Lecture Notes in Physics

The Lecture Notes in Physics

The series Lecture Notes in Physics (LNP), founded in 1969, reports new developments in physics research and teaching – quickly and informally, but with a high quality and the explicit aim to summarize and communicate current knowledge in an accessible way. Books published in this series are conceived as bridging material between advanced graduate textbooks and the forefront of research to serve the following purposes:

- to be a compact and modern up-to-date source of reference on a well-defined topic;
- to serve as an accessible introduction to the field to postgraduate students and nonspecialist researchers from related areas;
- to be a source of advanced teaching material for specialized seminars, courses and schools.

Both monographs and multi-author volumes will be considered for publication. Edited volumes should, however, consist of a very limited number of contributions only. Proceedings will not be considered for LNP.

Volumes published in LNP are disseminated both in print and in electronic formats, the electronic archive is available at springerlink.com. The series content is indexed, abstracted and referenced by many abstracting and information services, bibliographic networks, subscription agencies, library networks, and consortia.

Proposals should be sent to a member of the Editorial Board, or directly to the managing editor at Springer:

Dr. Christian Caron
Springer Heidelberg
Physics Editorial Department I
Tiergartenstrasse 17
69121 Heidelberg/Germany
christian.caron@springer.com

Walter Pötz Jaroslav Fabian
Ulrich Hohenester (Eds.)

Modern Aspects of Spin Physics

Springer

Editors

Walter Pötz
Ulrich Hohenester
Institut für Theoretische Physik
Universität Graz
Universitätsplatz 5
8010 Graz, Austria
E-mail: walter.poetz@uni-graz.at
ulrich.hohenester@uni-graz.at

Jaroslav Fabian
Institut für Theoretische Physik
Universität Regensburg
Universitätsstr. 31
93040 Regensburg, Germany
E-mail: jaroslav.fabian@physik.uni-regensburg.de

Walter Pötz et al., *Modern Aspects of Spin Physics*, Lect. Notes Phys. 712 (Springer, Berlin Heidelberg 2007), DOI 10.1007/b11824190

Library of Congress Control Number: 2006931572

ISSN 0075-8450
ISBN-10 3-540-38590-8 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-38590-5 Springer Berlin Heidelberg New York

Typesetting: by the authors and techbooks using a Springer LaTeX macro package
Cover design: WMXDesign GmbH, Heidelberg

Printed on acid-free paper SPIN: 11824190 54/techbooks 5 4 3 2 1 0

Preface

This volume contains a collection of lecture notes provided by the key speakers of the Schladming Winter School in Theoretical Physics, "43. Internationale Universitätswochen für Theoretische Physik", held in Schladming, Austria. This school took place from February 26 till March 4, 2005, and was titled *Spin Physics, Spintronics, and Spin-Offs*.

Until 2003 the Schladming Winter School, which is organized by the Division for Theoretical Physics of the University of Graz, Austria, has been devoted primarily to topics in subatomic physics. A few years ago, however, it was decided to broaden the scope of this school and, in particular, to incorporate hot topics in condensed matter physics. This was done in an effort to better represent the scientific activities of the theory group of the Physics Department at the University of Graz, resulting in the 42nd Winter School on "Quantum Coherence in Matter: From Quarks to Solids," held in 2004, and the 43rd Winter School on "Spin Physics, Spintronics, and Spin-Offs" in 2005. A compilation of lecture notes from the 2004 event have been released in the Springer series Lecture Notes in Physics LNP689 titled *Quantum Coherence: From Quarks to Solids*.

Spin is a fundamental property of elementary particles with important consequences on the macroscopic world. Beginning with the famous Stern–Gerlach experiment, research has been conducted to provide a sound microscopic understanding of this intriguing physical property. Indeed, the spin degree of freedom has physical implications on practically all areas of physics and beyond: from elementary particle physics, atomic-molecular physics, condensed matter physics, optics, to chemistry and biology.

Recently, the spin degree of freedom has been "rediscovered" in the context of quantum information storage and processing, colloquially summarized as "quantum computation." In addition, a relatively young field of solid-state device physics termed "spintronics," with the attempt to utilize the spin- rather than the charge-degree of freedom, has emerged. Each of these two topics is well worthy of its own school; however, in an attempt to provide an even broader perspective and to also attract students from elementary particle physics this winter school has included not only lectures and talks from both fields, but topics from elementary particle physics as well. As in past years, the Schladming Winter School and this compilation of lecture

notes is intended for advanced undergraduate and graduate students up to senior scientists who want to learn about or even get into this exciting field of physics. Research in this area is interdisciplinary and has both fundamental and applied aspects.

Listed below, in alphabetical order, are the speakers and titles of their lectures:

Enrico Arrigoni, Technical University of Graz, "Spin Pairing and High-Temperature Superconductors"

Tomasz Dietl, Polish Academy of Sciences, Warsaw, "Semiconductor Spintronics"

Stefano Forte, University of Milano, "Spin in Quantum Field Theories"

Elliot Leader, Imperial College, London, "Nucleon Spin"

Yuli V. Nazarov, Delft University, "Spin Currents and Spin Counting"

Igor Žutić, NRL, Washington DC, "Spin-Polarized Transport in Semiconductor Junctions: From Superconductors to Magnetic Bipolar Transistors"

Next to these lectures, there were a number of invited and contributed talks. For details, we refer to our Schladming Winter School web page http://physik.uni-graz.at/itp/iutp/index-iutp.html. This volume contains the lecture notes presented by T. Dietl, E. Arrigoni, S. Forte, and E. Leader. What has been said before about the flavor of the lectures also applies to the lecture notes presented in this volume.

In "Semiconductor Spintronics," Tomasz Dietl gives an overview of the modern field of spintronics, containing a brief history, motivation behind the field, past achievements, and future challenges. It should be mentioned that Prof. Dietl's Award of the Agilent Technologies Europhysics Prize 2005 (with David D. Awschalom and Hideo Ohno) was announced during the Winter School.

In "Lectures on Spin Pairing Mechanism in High-Temperature Superconductors," Enrico Arrigoni first reviews the essentials of conventional phonon-based superconductivity and then discusses alternative pairing mechanisms based on the Hubbard model, which may play a role in high-temperature superconducting materials with an antiferromagnetic phase.

In "Spin in Quantum Field Theories," Stefano Forte gives a pedagogical introduction to spin in quantum field theory, largely avoiding the usual framework of relativistic quantum field theory. This paper is intended as a bridge between elementary particle (relativistic quantum field theory) physics and condensed matter physics (nonrelativistic quantum field theory).

In "Nucleon Spin," Elliot Leader discusses proton (nucleon) spin and pitfalls encountered in the interpretation of its origin from the nucleon's constituents.

We are grateful to the lecturers for presenting their lectures in a very pedagogical way at the school and for taking the time for preparing the

manuscripts for publication in this book. We feel that this volume represents a good overview of current research on spin-related physics.

We acknowledge financial support from the main sponsors of the school: the Austrian Federal Ministry of Education, Science, and Culture, as well as the Government of Styria. We have received financial, material, and technical support from the University of Graz, the town of Schladming, RICOH Austria, and Hornig Graz. We also thank our colleagues, staff, and students at the Physics Department for their valuable technical assistance, as well as all participants and speakers for making the 43rd Schladming Winter School a great success.

Graz,
July 2006

Walter Pötz
Jaroslav Fabian
Ulrich Hohenester

Contents

Semiconductor Spintronics
T. Dietl 1

1 Why Spintronics? 1
2 Non-magnetic Semiconductors 4
2.1 Overview 4
2.2 Spin Relaxation and Dephasing 5
2.3 An Example of Spin Filter 6
3 Hybrid Structures 7
3.1 Overview 7
3.2 Spin Injection 8
3.3 Search for Solid-state Stern-Gerlach Effect 9
4 Diluted Magnetic Semiconductors 11
4.1 Overview 11
4.2 Magnetic Impurities in Semiconductors 12
4.3 Exchange Interaction Between Band and Localized Spins 14
4.4 Electronic Properties 15
4.5 Magnetic Polarons 17
4.6 Exchange Interactions between Localized Spins 17
4.7 Magnetic Collective Phenomena 18
5 Properties of Ferromagnetic Semiconductors 19
5.1 Overview 19
5.2 p-d Zener Model 20
5.3 Curie Temperature – Chemical Trends 22
5.4 Micromagnetic Properties 23
5.5 Optical Properties 26
5.6 Charge Transport Phenomena 28
5.7 Spin Transport Phenomena 36
5.8 Methods of Magnetization Manipulation 37
6 Summary and Outlook 37
References 40

Lectures on the Spin Pairing Mechanism in High-Temperature Superconductors
E. Arrigoni 47

1 Introduction 47
2 Superconductivity 48
3 Phonon-Mediated Effective Attraction between Electrons 50
4 BCS Theory 53
5 High-Temperature Superconductors 55
6 Pairing Mediated by Spin Fluctuations:
Linear Response to Magnetic Excitations 59
References 65

Spin in Quantum Field Theory
S. Forte 67

1 From Quantum Mechanics to Field Theory 67
2 Spin and Statistics 68
2.1 The Galilei Group and the Lorentz Group 68
2.2 Statistics and Topology 70
2.3 Bosons, Fermions and Anyons 74
3 A Path Integral for Spin 79
3.1 The Spin Action 79
3.2 Classical Dynamics 81
3.3 Geometric Quantization 82
4 Relativistic Spinning Particles 86
4.1 Path Integral for Spinless Particles 86
4.2 The Classical Spinning Particle 88
4.3 Quantum Spinning Particles and Fermions 90
5 Conclusion 93
References 94

Nucleon Spin
E. Leader 95

1 Introduction 95
2 Polarized Lepton-Nucleon Deep Inelastic Scattering 97
3 The Spin Crisis in the Parton Model 101
4 Resolution of the Spin Crisis: The Axial Anomaly 105
5 Matrix Elements of Angular Momentum Operators: The Problem 108
6 Relativistic Spin States 111
7 Matrix Elements of Angular Momentum Operators: The Results 113
7.1 Canonical Spin State Matrix Elements 114
7.2 Helicity State Matrix Elements 116
8 Angular Momentum Sum Rules 118
8.1 General Structure of Sum Rules:
Parton Transverse Momentum 118

8.2 The Longitudinal Sum Rule 122
8.3 The Transverse Case: The New Sum Rulexs 123
8.4 Comparison with Results in the Literature 125
9 Interpretation of the Sum Rules 126
References 127

List of Contributors

Enrico Arrigoni
Institute for Theoretical Physics
and Computational Physics
Graz University of Technology
Petersgasse 16
A-8010 Graz
Austria
arrigoni@itp.tu-graz.ac.at

Tomasz Dietl
Institute of Physics
Polish Academy of Sciences
and
ERATO Semiconductor Spintronics
Project
Japan Science and Technology
Agency
al. Lotnikow 32/46
PL-02668 Warszawa
Poland
and
Institute of Theoretical Physics
Warsaw University
Poland
dietl@ifpan.edu.pl

Stefano Forte
Dipartimento di Fisica
Universita di Milano
and
INFN Sezione di Milano
via Celoria 16
I-10129 Milano
Italy
forte@mi.infn.it

Elliot Leader
Prince Consort Road
London SW7 2BW
England
e.leader@imperial.ac.uk

Semiconductor Spintronics

T. Dietl

Institute of Physics, Polish Academy of Sciences and ERATO Semiconductor Spintronics Project, Japan Science and Technology Agency, al. Lotników 32/46, 02668 Warszawa, Poland and Institute of Theoretical Physics, Warsaw University, Poland
dietl@ifpan.edu.pl

Abstract. These informal lecture notes describe the recent progress in semiconductor spintronics in a historic perspective as well as in comparison to achievements of spintronics of ferromagnetic metals. After outlining motivations behind spintronic research, selected results of investigations on three groups of materials are presented. These include non-magnetic semiconductors, hybrid structures involving semiconductors and ferromagnetic metals, and diluted magnetic semiconductors either in paramagnetic or ferromagnetic phase. Particular attention is paid to the hole-controlled ferromagnetic systems whose thermodynamic, micromagnetic, transport, and optical properties are described in detail together with relevant theoretical models.

1 Why Spintronics?

The well-known questions fuelling a broad interest in nanoscience are: will it still be possible to achieve further progress in information and communication technologies simply by continuing to miniaturize the transistors in microprocessors and the memory cells in magnetic and optical discs? How to reduce power consumption of components in order to save energy and to increase battery operation time? How to integrate nowadays devices with biological molecules and functionalities?

Since 70s, the miniaturization by obeying Moore's law has persistently lead to an exponential increase in the quantity of information that can be processed, stored, and transmitted per unit area of microprocessor, memory, and fiberglass, respectively. A modern integrated circuit contains now one billion transistors, each smaller than 100 nm in size, i.e., a five hundred times smaller than the diameter of a human hair. The crossing of this symbolic 100 nm threshold at the outset of the 21st century ushered in the era of nanotechnology. As the size of transistors decreases, their speed increases, and their price falls. Today it is much less expensive to manufacture one transistor than to print a single letter. Despite the series of successes that industrial laboratories have scored over the past 40 years in surmounting one technical and physical barrier after another, there is a prevalent sense that in the near future a qualitative change is now in store for us in terms of the methods of data processing, storing, encoding, and transmission. For this reason, governments in many countries are financing ambitious interdisciplinary

T. Dietl: *Semiconductor Spintronics*, Lect. Notes Phys. **712**, 1–46 (2007)
DOI 10.1007/3-540-38592-4_1

programs aimed at insuring active participation in the future development of nanotechnology.

Among the many proposals for where to take such research, the field of spintronics, i.e., electronics aimed at understanding electron spin phenomena and at proposing, designing, and developing devices to harness these phenomena, is playing a major role. The hopes placed in spintronics are founded on the well-known fact that since magnetic monopoles do not exist, random magnetic fields are significantly weaker than random electric fields. For these reasons, magnetic memories are non-volatile, while memories based on an accumulated electric charge (dynamic random access memory, or DRAM) require frequent refreshing.

One of the ambitious goals in the spintronics field is to create magnetic random access memory (MRAM), a type of device that would combine the advantages of both magnetic memory and dynamic random access memory. This requires novel methods of magnetizing memory cells and reading back the direction of such magnetization, which would not involve any mechanical systems. Another important step along this path would be the ability to control magnetization isothermally, by means of light or electric field. Modern devices expend relatively large amounts of energy on controlling magnetization (i.e., storing data), as they employ Oersted magnetic fields generated by electric currents.

The development of more "intelligent" magnetization control methods would also make it possible to build spin transistors, devices composed of two layers of ferromagnetic conductors separated by non-magnetic material. It stands to reason that if carriers injected into the non-magnetic layer preserve their spin direction, then the electric conductivity depends on the relative direction of the magnetization vectors in the ferromagnetic layers. This could offer a means of producing an energy-conserving and fast switching device, as it would allow current to be controlled without changing the carrier concentration. An obvious prerequisite for such a transistor to operate is the efficient injection of spin-polarized carriers made of ferromagnetic material into the non-magnetic area. Also, there should be no processes that could disrupt the spin polarization. Simultaneously, researchers are seeking ways of generating, amplifying, and detecting spin currents: here, the underlying conviction is that the movement of electrons with opposite spins does not entail any losses, yet can carry information. This would lay the foundations for the development of low-power devices, characterized by significantly reduced heat dissipation. Another important issue is to develop methods for injecting spin-polarized carriers into semiconductors. Apart from the possibility of designing the magnetization sensors and spin transistors, polarized carrier injection could prove to be useful as a method for the fast modulation of semiconductor lasers and would allow surface-emission lasers to work in a single mode fashion.

Perhaps the most important intellectual challenge to be faced in spintronics is to create a hardware for quantum information science. Researchers over the world have joined efforts to lay the theoretical foundations for this new discipline [1], one notable example being the Horodecki family from Gdańsk [2]. Experiments conducted by David Awschalom's group in Santa Barbara show that spin degrees of freedom are of particular importance as they maintain their phase coherence significantly longer than orbital degrees of freedom do [3]. Electron spin is therefore much more suitable than electron charge for putting into practice modern ideas for performing numerical computations using the superposition and entanglement of quantum states. Spin nanostructures might consequently alter the basic principles not only in the design of electronic elements, but also in the very computer architecture that has been in use for half a century. It is noteworthy that quantum encoders are already now being sold and installed: such devices use the polarization of light to encode the transmitted information, and the unauthorized interception and reading of this information appears to be impossible.

Today's research on spin electronics involves virtually all material families. The most advanced are studies on magnetic multilayers. As demonstrated in 80s by groups of Albert Fert [4] in Orsay and Peter Grünberg [5] in Jülich, these systems exhibit giant magnetoresistance (GMR). According to theory triggered by these discoveries and developed by Józef Barnaś from Poznań and co-workers [6], GMR results from spin-dependent scattering at adjacent interfaces between non-magnetic and magnetic metals, which changes when the magnetic field aligns magnetization of particular layers. Since 90s, the GMR devices have been successfully applied in reading heads of high-density hard-discs. Recent works focuss also on spin-dependent tunnelling via an oxide film. Remarkably, for the case of crystalline MgO sandwiched between contacts of amorphous Fe-Co-B layers, the difference between tunnelling resistance for anti-parallel and parallel orientations of magnetization, the TMR, reaches a factor of three at 300 K [7, 8, 9]. Moreover, the magnetization direction can be switched by an electric current below 10^6 A cm^{-2}[10], opening the doors for a direct magnetization writing by current pulses. Last but not least such structures can be used for injecting highly polarized spin currents to semiconductors, such as GaAs [11].

These informal lecture notes on semiconductor spintronics exploit and update author's earlier reviews [12, 13, 14, 15, 16, 17, 18], where more systematic references to original papers can be found. Particular attention is paid here to those results of research on spin properties of semiconductors, which appear relevant in the context of disruptive classical and quantum information and communication technologies. First part of the paper shows briefly how spin effects specific to non-magnetic semiconductors can be exploited in spintronic devices. This is followed by a presentation of chosen properties of hybrid semiconductor/ferromagnetic metal structures. The main body of the paper is devoted to diluted magnetic semiconductors (DMS),

especially to materials exhibiting the ferromagnetic order, as they combine complementary resources of semiconductor materials and ferromagnetic metals. Here, the fundamental research problem is to identify the extent to which the methods that have been so successfully applied to controlling the density and degree of spin polarization of carriers in semiconductor structures might be employed to control the magnetization magnitude and direction. Apart from the possibility of designing the aforementioned magnetoresistive sensors and spin aligners, ferromagnetic semiconductors are the materials of choice for spin current amplification and detection. Furthermore, their outstanding magnetooptical properties can be exploited for fast light modulation as well as optical isolators, perhaps replacing hybrid structures consisting of paramagnetic DMS, such as (Cd,Mn)Te, and a permanent magnet.

In the course of the years semiconductor spintronics has evolved into a rather broad research field. These notes are by no means exhaustive and, moreover, they are biased by author's own expertise. Fortunately, however, in a number of excellent reviews the issues either omitted or only touched upon here has been thoroughly elaborated in terms of content and references to the original papers. For instance, the progress in fabrication and studies of spin quantum gates of double quantum dots has been described by van Viel et al. [19]. A comprehensive survey on spin-orbit effects and the present status of spin semiconductor transistors has been completed by Žutić, Fabian, and Das Sarma [20]. Finally, Jungwirth et al. [21] have reviewed various aspects of theory of (Ga,Mn)As and related materials. Excellent reviews on the entire semiconductor spintronics are also available [22, 23].

2 Non-magnetic Semiconductors

2.1 Overview

The beginning of spintronic research on non-magnetic semiconductors can be traced back to the detection of nuclear spin polarization in Si illuminated by circularly polarized light reported in late 60s by Georges Lampel at Ecole Polytechnique [24]. Already this pioneering experiment involved phenomena crucial for semiconductor spintronics: (i) the spin-orbit interaction that allows for transfer of orbital (light) momentum to spin degrees of freedom and (ii) the hyperfine interaction between electronic and nuclear spins. Subsequent experimental and theoretical works on spin orientation in semiconductors, carried out in 70s mostly by researchers around Ionel Solomon in Ecole Polytechnique and late Boris P. Zakharchenya in Ioffe Institute, were summarized in a by now classic volume [25].

More recently, notably David Awschalom and his co-workers first at IBM and then at Santa Barbara, initiated the use of time resolved optical magnetospectroscopies that have made it possible to both temporally and spatially

explore the spin degrees of freedom in a wide variety of semiconductor materials and nanostructures [26]. The starting point of this experimentally demanding technique is the preparation of spins in a particular orientation by optically pumping into selected electronic states. The electron spin then precesses in an applied or molecular magnetic field produced by electronic or nuclear spins. The precessing magnetic moment creates a time dependent Faraday rotation of the femtosecond optical probe. The oscillation and decay measure the effective Landé g-factor, the local magnetic fields, and coherence time describing the temporal dynamics of the optically injected spins.

Present spintronic activities focuss on two interrelated topics. The first is to exploit Zeeman splitting and spin-orbit interactions for spin manipulation. To this category belongs, in particular, research on spin filters and detectors, on the Datta-Das transistor [20], on optical generation of spin currents [27] and on the spin Hall effect [28]. The other topic is the quest for solid-state spin quantum gates that would operate making use of spin-spin exchange [29] and/or hyperfine interactions [30]. An important aspect of the field is a dual role of the interactions in question in non-magnetic semiconductors: from one hand they allow for spin functionalities, on the other they account for spin decoherence and relaxation, usually detrimental for spin device performance. This, together with isotope characteristics, narrows rather severely a window of material parameters at which semiconductor spin devices might operate.

2.2 Spin Relaxation and Dephasing

Owing to a large energy gap and the weakness of spin-orbit interactions, especially long spin life times are to be expected in the nitrides and oxides. Figure 1 depicts results of time-resolved Faraday rotation, which has been used to measure electron spin coherence in n-type GaN epilayers [31]. Despite densities of charged threading dislocations of 5×10^8 cm^{-2}, this coherence yields spin lifetimes of about 20 ns at temperatures of 5 K, and persists up to room temperature.

Figure 2 presents a comparison of experimental and calculated magnetoresistance (MR) of a ZnO:Al thin film containing $1.8 \cdot 10^{20}$ electrons per cm^3 [32]. Here, spin effects control quantum interference corrections to the classical Drude-Boltzmann conductivity. A characteristic positive component of MR, signalizing the presence of spin-orbit scattering, is detected below 1 mT at low temperatures. This scattering is linked to the presence of a Rashba-like term $\lambda_{\rm so}\boldsymbol{c}(\boldsymbol{s}\times\boldsymbol{k})$ in the kp hamiltonian of the wurzite structure, first detected in n-CdSe in the group of the present author [33]. As shown in Fig. 2, a quite good description of the findings is obtained with $\lambda_{\rm so} = 4.4 \cdot 10^{-11}$ eV cm, resulting in the spin coherence time 1 ns, more than 10^4 times longer than the momentum relaxation time. Importantly, this low decoherence rate of wide-band gap semiconductors is often coupled with a small value of the dielectric constant that enhances characteristic energy scales for quantum dot charging as well as for the exchange interaction of the electrons residing on the

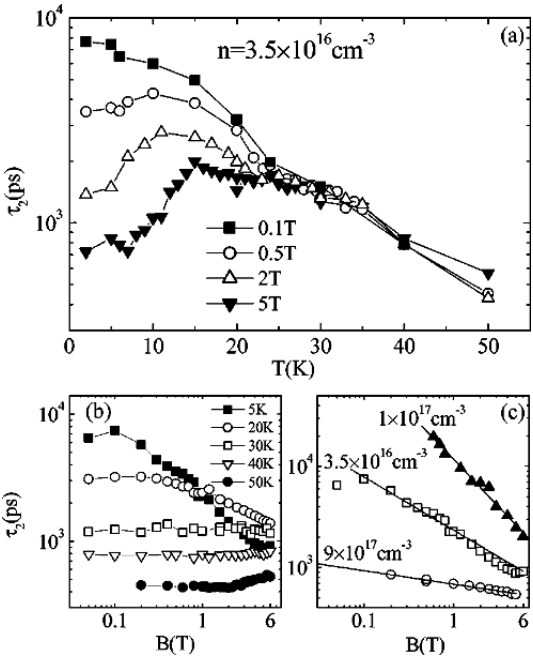

Fig. 1: Spin scattering time τ_2 of n-GaN at various magnetic fields (**a**), temperatures (**b**) ($n = 3.5 \times 10^{16}$ cm^{-3}), and electron concentrations at 5 K (**c**) (after Beschoten et al. [31])

neighboring dots. This may suggest some advantages of these compounds for fabrication of spin quantum gates. Another material appealing in this context is obviously Si, and related quantum structures, in which the interfacial electric field controls the magnitude of the Rashba term [34] and material containing no nuclear spins can be obtained.

2.3 An Example of Spin Filter

Turning to the case of narrow-gap semiconductors we note that strong spin-orbit effects specific to these systems results, among other things, in a large Zeeman splitting of the carrier states, which can be exploited for fabrication of efficient spin filters. As an example, we consider quantum point contacts patterned of PbTe quantum wells embedded by Bi-doped $Pb_{0.92}Eu_{0.08}Te$ barriers [35, 36]. Owing to biaxial strain, the fourfold L-valley degeneracy of the conduction band in PbTe is lifted, so that the relevant ground-state 2D subband is formed of a single valley with the long axis parallel to the [111] growth direction. As discussed recently [36], the paraelectric character of PbTe results in efficient screening of Coulomb scattering potentials, so that signatures of ballistic transport can be observed despite of significant amount of charged

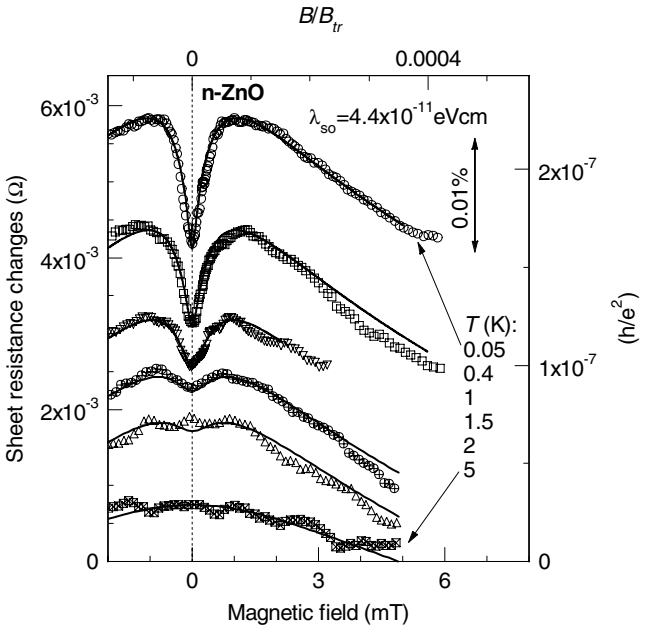

Fig. 2: Resistance changes in the magnetic field for n-ZnO (*symbols*) compared to calculations (*solid lines*) within the weak localization theory for the 2D case. Curves are vertically shifted for clarity (after Andrearczyk et al. [32])

defects in the vicinity of the channel. At the same time, the electron density can be tuned over a wide range by biasing a p-n junction that is formed between the p^+ interfacial layer and the n-type quantum well [36]. Furthermore, a rather large magnitude of electron spin splitting for the magnetic field along the growth direction, corresponding to the Landé factor $|g*| \approx 66$, can serve to produce a highly spin-selective barrier. According to results displayed in Fig. 3, spin-degeneracy of the quantized conductance steps starts to be removed well below 1 T, so that it has become possible to generate entirely polarized spin current carried by a number of 1D subbands [35].

3 Hybrid Structures

3.1 Overview

The hybrid nanostructures, in which both electric and magnetic field are spatially modulated, are usually fabricated by patterning of a ferromagnetic metal on the top of a semiconductor or by inserting ferromagnetic nanoparticles or layers into a semiconductor matrix. In such devices, the stray fields

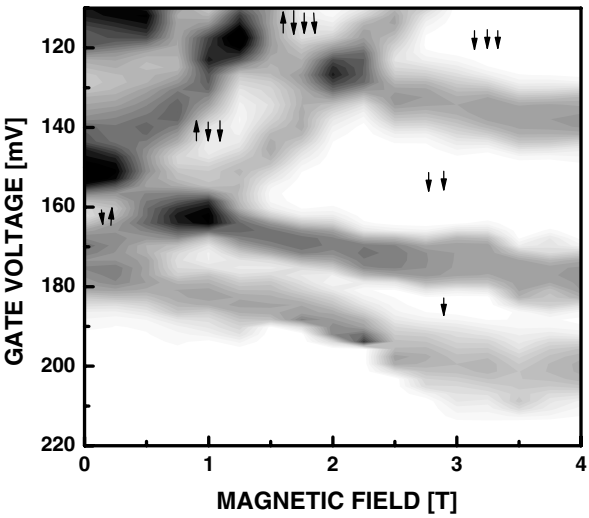

Fig. 3: Transconductance dG/dV_g (*gray scale*) showing dependence of 1D subbands on the magnetic field and gate voltage for PbTe nanoconstriction of a wide (Pb, Eu)Te/PbTe/(Pb, Eu)Te quantum well (after Grabecki et al. [35])

can control charge and spin dynamics in the semiconductor. At the same time, spin-polarized electrons in the metal can be injected into or across the semiconductor [37, 38]. Furthermore, the ferromagnetic neighbors may affect semiconductor electronic states by the ferromagnetic proximity effect even under thermal equilibrium conditions. Particularly perspective materials in the context of hybrid structures appear to be those elemental or compound ferromagnets which can be grown in the same reactor as the semiconductor counterpart.

3.2 Spin Injection

It is now well established that efficient spin injection from a ferromagnetic metal to a semiconductor is possible provided that semiconductor Sharvin resistance is comparable or smaller than the difference in interface resistances for two spin orientations. Often, to enhance the latter, a heavily doped or oxide layer is inserted between the metal and as-grown semiconductor. In this way, spin current reaching polarization tens percents has been injected form Fe into GaAs [11, 39]. At the same time, it is still hard to achieve TMR above 10% in Fe/GaAs/Fe trilayer structures without interfacial layer [40], which may suggest that the relevant Schottky barriers are only weakly spin selective.

The mastering of spin injection is a necessary condition for the demonstration of the Datta-Das transistor [41], often regarded as a flag spintronic device. In this spin FET, the orientation of the spins flowing between ferromagnetic contacts, and thus the device resistance, is controlled by the Rashba field generated in the semiconductor by an electrostatic gate. Recently, a current modulation up to 30% by the gate voltage was achieved in a Fe/(In,Ga)As/Fe FET at room temperature [42]. This important finding was obtained for a 1 μm channel of narrow gap $In_{0.81}Ga_{0.19}As$, in which TMR achieved 200%, indicating that the destructive role of the Schottky barriers got reduced. Furthermore, an engineered interplay between the Rashba and Dresselhaus effects [43, 44] resulted in a spin relaxation time long comparing to spin precession period and the dwell time.

3.3 Search for Solid-state Stern-Gerlach Effect

The ferromagnetic component of hybrid structures can also serve for the generation of a magnetic field. This field, if uniform, produces a spin selective barrier that can serve as a local spin filter and detector. A non-homogenous field, in turn, might induce spatial spin separation via the Stern-Gerlach (S-G) mechanism. Figure 4(a) presents a micrograph of a Stern-Gerlach device, whose design results from an elaborated optimization process [45]. A local magnetic field was produced by NiFe (permalloy, Py) and cobalt (Co) films. The micromagnets resided in deep groves on the two sides of the wire, so that the 2D electron gas in the modulation-doped GaAs/AlGaAs heterostructure was approximately at the center of the field, and the influence of the competing Lorentz force was largely reduced. Hall magnetometry was applied in order to visualize directly the magnetizing process of the two micromagnets in question.

As shown in Fig. 5, a current increase in counters was detected when a field gradient was produced by an appropriate cycle of the external magnetic field at 100 mK. The range of magnetic fields where the enhancement was observed corresponded to the the presence of the field gradient according to the Hall magnetometry, which also showed that Py magnetization diminished almost twofold prior to a change in the direction of the external magnetic field. This effect, associated with the formation of closure domains in soft magnets, explained why the current changes appeared before the field reversal. The relative change ΔI of counter current depended on V_G, $\Delta I/I$ increased from 0.5% at zero gate voltage to 50% close to the threshold. Furthermore, for V_G about -0.8 V ΔI was negative. It was checked that results presented in Fig. 5 were unaltered by increasing the temperature up to 200 mK and independent of the magnetic field sweep rate.

Theoretical studies [45] of the results shown in Fig. 5 demonstrated that semiconductor nanostructures of the kind shown in Fig. 4 can indeed serve to generate and detect spin polarized currents in the absence of an external magnetic field. Moreover, the degree and direction of spin polarization at

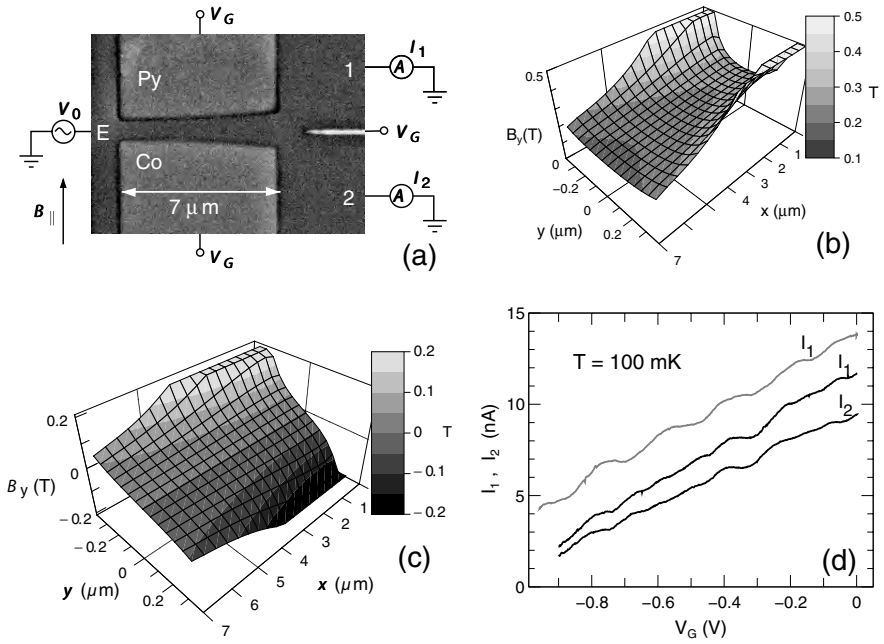

Fig. 4: (**a**) Scanning electron micrograph of the spin-filter device. Fixed AC voltage V_0 is applied between emitter (E) and "counters" (1), (2); V_G is the DC gate voltage. The external in-plane magnetizing field ($B_{\|}$) is oriented as shown. (**b**) The in-plane magnetic field B_y (wider part of the channel is in front) calculated for half-plane, 0.1 μm thick magnetic films separated by a position dependent gap $W(x)$ and magnetized in the same directions (saturation magnetization as for Co). (**c**) B_y calculated for antiparallel directions of micromagnet magnetizations. (**d**) Counter currents I_1 and I_2 as a function of the gate voltage at $V_0 = 100$ μV and $B_{\|} = 0$; upper curve (shown in gray) was collected during a different thermal cycle and after longer infra-red illumination (after Wróbel et al. [45])

low electron densities can easily be manipulated by gate voltage or a weak external magnetic field. While the results of the performed computations suggest that the spin separation and thus Stern-Gerlach effect occurs under experimental conditions in question, its direct experimental observation would require incorporation of spatially resolved spin detection.

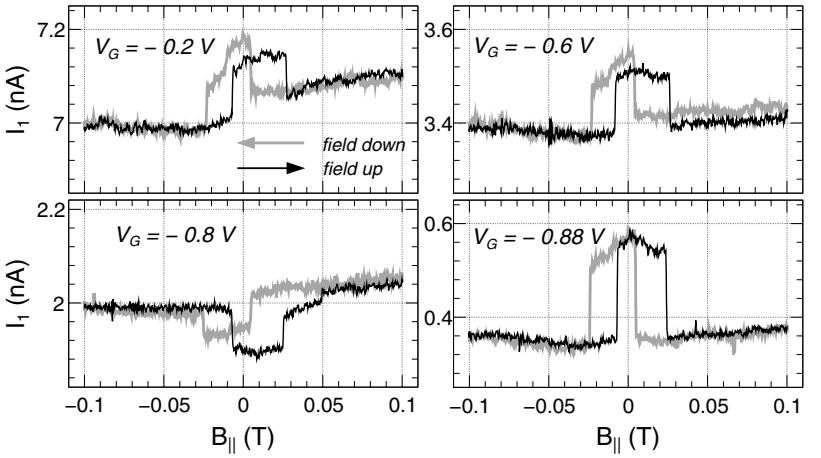

Fig. 5: The counter current I_1 of as a function of the in-plane magnetic field for various gate voltages for the device shown in Fig. 4. After Wrobel et al. [45]

4 Diluted Magnetic Semiconductors

4.1 Overview

This family of materials encompasses standard semiconductors, in which a sizable portion of atoms is substituted by such elements, which produce localized magnetic moments in the semiconductor matrix. Usually, magnetic moments originate from 3d or 4f open shells of transition metals or rare earths (lanthanides), respectively, so that typical examples of diluted magnetic semiconductors (DMS) are $Cd_{1-x}Co_xSe$, $Ga_{1-x}Mn_xAs$, $Pb_{1-x}Eu_xTe$ and, in a sense, Si:Er. A strong spin-dependent coupling between the band and localized states accounts for outstanding properties of DMS. This coupling gives rise to spin-disorder scattering, giant spin-splittings of the electronic states, formation of magnetic polarons, and strong indirect exchange interactions between the magnetic moments, the latter leading to collective spin-glass, antiferromagnetic or ferromagnetic spin ordering. Owing to the possibility of controlling and probing magnetic properties by the electronic subsystem or vice versa, DMS have successfully been employed to address a number of important questions concerning the nature of various spin effects in various environments and at various length and time scales. At the same time, DMS exhibit a strong sensitivity to the magnetic field and temperature as well as constitute important media for generation of spin currents and for manipulation of localized or itinerant spins by, e.g., strain, light, electrostatic or ferromagnetic gates. These properties, complementary to both non-magnetic

semiconductors and magnetic metals, open doors for application of DMS as functional materials in spintronic devices.

Extensive studies of DMS started in 70s, particularly in the group of Robert R. Gałązka in Warsaw, when appropriately purified Mn was employed to grow bulk II-VI Mn-based alloys by various modifications of the Bridgman method [46]. Comparing to magnetic semiconductors, such as Eu chalcogenides (e.g., EuS) and Cr spinels (e.g., $CdCr_2Se_4$) investigated earlier [47], DMS exhibited smaller defect concentrations and were easier to dope by shallow impurities. Accordingly, it was possible to examine their properties by powerful magnetooptical and magnetotransport techniques [12, 46, 48, 49]. Since, in contrast to magnetic semiconductors, neither narrow magnetic bands nor long-range magnetic ordering affected low-energy excitations, DMS were named semimagnetic semiconductors. More recently, research on DMS have been extended toward materials containing magnetic elements other than Mn as well as to III-VI, IV-VI [50] and III-V [51] compounds as well as group IV elemental semiconductors and various oxides [52]. In consequence, a variety of novel phenomena has been discovered, including effects associated with narrow-bands and magnetic phase transformations, making the borderline between properties of DMS and magnetic semiconductors more and more elusive.

A rapid progress of DMS research in 90s stemmed, to a large extend, from the development of methods of crystal growth far from thermal equilibrium, primarily by molecular beam epitaxy (MBE), but also by laser ablation. These methods have made it possible to obtain DMS with the content of the magnetic constituent beyond thermal equilibrium solubility limits [53]. Similarly, the doping during MBE process allows one to increase substantially the electrical activity of shallow impurities [54, 55]. In the case of III-V DMS [51], in which divalent magnetic atoms supply both spins and holes, the use of the low-temperature MBE (LT MBE) provides thin films of, e.g., $Ga_{1-x}Mn_xAs$ with x up to 0.07 and the hole concentration in excess of 10^{20} cm^{-3}, in which ferromagnetic ordering is observed above 170 K [56]. Remarkably, MBE and processes of nanostructure fabrication, make it possible to add magnetism to the physics of semiconductor quantum structures. Particularly important are DMS, in which ferromagnetic ordering was discovered, as discussed in some details later on.

4.2 Magnetic Impurities in Semiconductors

A good starting point for the description of DMS is the Vonsovskii model, according to which the electron states can be divided into two categories: (i) localized magnetic d or f shells and (ii) extended band states built up of s, p, and sometimes d atomic orbitals. The former give rise to the presence of local magnetic moments and intra-center optical transitions. The latter form bands, much alike as in the case of non-magnetic semiconductor alloys. Indeed, the lattice constant of DMS obeys the Vegard low, and the energy

gap E_g between the valence and the conduction band depends on x in a manner qualitatively similar to non-magnetic counterparts. According to the Anderson model, the character of magnetic impurities in solids results from a competition between (i) hybridization of local and extended states, which tends to delocalized magnetic electrons and (ii) the on-site Coulomb interactions among the localized electrons, which stabilizes the magnetic moment in agreement with Hund's rule.

Figure 6 shows positions of local states derived from 3d shells of transition metal (TM) impurities in respect to the band energies of the host II-VI and III-V compounds. In figure the levels labelled "donors" denote the ionization energy of the magnetic electrons ($TM^{2+} \rightarrow TM^{3+}$ or $d^n \rightarrow d^{n-1}$), whereas the "acceptors" correspond to their affinity energy ($TM^{2+} \rightarrow TM^{1+}$ or $d^n \rightarrow d^{n+1}$). The difference between the two is the on-d-shell Coulomb (Hubbard) repulsion energy U in the semiconductor matrix. In addition, the potential introduced by either neutral or charged TM can bind a band carrier in a Zhang-Rice-type singlet or hydrogenic-like state, respectively. Such bound states are often experimentally important, particularly in III-V compounds, as they correspond to lower energies than the competing d-like states, such as presented in Fig. 6.

In the case of Mn, in which the d shell is half-filled, the d-like donor state lies deep in the valence band, whereas the acceptor level resides high in the conduction band, so that $U \approx 7$ eV according to photoemission and inverse photoemission studies. Thus, Mn-based DMS can be classified as charge transfer insulators, $E_g < U$. The Mn ion remains in the 2+ charge state, which means that it does not supply any carriers in II-VI materials. However, it acts as a hydrogenic-like acceptor in the case of III-V antimonides and arsenides, while the corresponding Mn-related state is deep, presumably due to a stronger p-d hybridization, in the case of phosphides and nitrides. According to Hund's rule the total spin $S = 5/2$ and the total orbital momentum $L = 0$ for the d^5 shell in the ground state. The lowest excited state d^{*5} corresponds to $S = 3/2$ and its optical excitation energy is about 2 eV. Thus, if there is no interaction between the spins, their magnetization is described by the paramagnetic Brillouin function. In the case of other transition metals, the impurity-induced levels may appear in the gap, and then compensate shallow impurities, or even act as resonant dopant, e.g., Sc in CdSe, Fe in HgSe or Cu in HgTe. Transport studies of such systems have demonstrated that inter-site Coulomb interactions between charged ions lead to the Efros-Shklovskii gap in the density of the impurity states, which makes resonant scattering to be inefficient in semiconductors [59]. Furthermore, spin-orbit interaction and Jahn-Teller effect control positions and splittings of the levels in the case of ions with $L \neq 0$. If the resulting ground state is a magnetically inactive singlet there is no permanent magnetic moment associated with the ion, the case of Fe^{2+}, whose magnetization is of the Van Vleck-type at low temperatures.

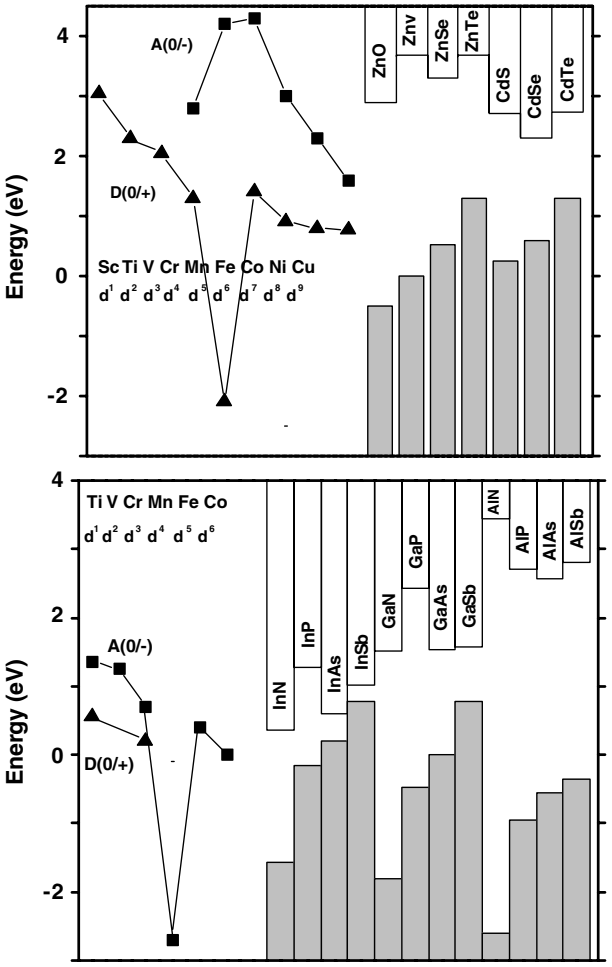

Fig. 6: Approximate positions of transition metals levels relative to the conduction and valence band edges of II-VI (*left panel*) and III-V (*right panel*) compounds. By triangles the d^N/d^{N-1} donor and by squares the d^N/d^{N+1} acceptor states are denoted (adapted from Langer et al. [57] and Zunger [58])

4.3 Exchange Interaction Between Band and Localized Spins

The important aspect of DMS is a strong spin-dependent coupling of the effective mass carriers to the localized d electrons, first discovered in (Cd,Mn)Te [60, 61] and (Hg,Mn)Te [62, 63]. Neglecting non-scalar corrections that can appear for ions with $L \neq 0$, this interaction assumes the Kondo form,

$$H_K = -I(\boldsymbol{r} - \boldsymbol{R}^{(i)})\boldsymbol{s}\boldsymbol{S}^{(i)} \ , \tag{1}$$

where $I(\boldsymbol{r} - \boldsymbol{R}^{(i)})$ is a short-range exchange energy operator between the carrier spin $\boldsymbol{s}$ and the TM spin localized at $\boldsymbol{R}^{(i)}$. When incorporated to the kp scheme, the effect of H_K is described by matrix elements $\langle u_i|I|u_i\rangle$, where u_i are the Kohn-Luttinger amplitudes of the corresponding band extreme. In the case of carriers at the Γ point of the Brillouin zone in zinc-blende DMS, the two relevant matrix elements $\alpha = \langle u_c|I|u_c\rangle$ and $\beta = \langle u_v|I|u_v\rangle$ involve s-type and p-types wave functions, respectively. There are two mechanisms contributing to the Kondo coupling [48, 64, 65]: (i) the exchange part of the Coulomb interaction between the effective mass and localized electrons; (ii) the spin-dependent hybridization between the band and local states. Since there is no hybridization between Γ_6 and d-derived (e_g and t_{2g}) states in zinc-blende structure, the s-d coupling is determined by the direct exchange. The experimentally determined values are of the order of $\alpha N_o \approx 0.25$ eV, where N_o is the cation concentration, somewhat reduced comparing to the value deduced from the energy difference between $S \pm 1$ states of the free singly ionized Mn atom $3d^5 4s^1$, $\alpha N_o = 0.39$ eV. In contrast, there is a strong hybridization between Γ_8 and t_{2g} states, which affects their relative position, and leads to a large magnitude of $|\beta N_o| \approx 1$ eV. If the relevant effective mass state is above the t_{2g} level (the case of, e.g., Mn-based DMS), $\beta < 0$ but otherwise β can be positive (the case of, e.g., $Zn_{1-x}Cr_xSe$ [66]).

4.4 Electronic Properties

Effects of Giant Spin Splitting

In the virtual-crystal and molecular-field approximations, the effect of the Kondo coupling is described by $H_K = I\boldsymbol{M}(\boldsymbol{r})\boldsymbol{s}/g\mu_B$, where $\boldsymbol{M}(\boldsymbol{r})$ is magnetization (averaged over a microscopic region around $\boldsymbol{r}$) of the localized spins, and g is their Landé factor. Neglecting thermodynamic fluctuations of magnetization (the mean-field approximation) $\boldsymbol{M}(\boldsymbol{r})$ can be replaced by $M_o(T, H)$, the temperature and magnetic field dependent macroscopic magnetization of the localized spins available experimentally. The resulting spin-splitting of s-type electron states is given by

$$\hbar\omega_s = g^*\mu_B B + \alpha M_o(T, H)/g\mu_B \ , \tag{2}$$

where g^* is the band Landé factor. The exchange contribution is known as the giant Zeeman splitting, as in moderately high magnetic fields and low temperatures it attains values comparable to the Fermi energy or to the binding energy of excitons and shallow impurities. For effective mass states, whose periodic part of the Bloch function contains spin components mixed up by a spin-orbit interaction, the exchange splitting does not depend only on the product of M_o and the relevant exchange integral, say β, but usually also on the magnitude and direction of $\boldsymbol{M}_o$, confinement, and strain. Furthermore, because of confinement or non-zero k the Bloch wave function

contains contributions from both conduction and valence band, which affects the magnitude and even the sign of the spin splitting [49, 62, 63, 67]. The giant Zeeman splitting is clearly visible in magnetooptical phenomena as well as in the Shubnikov-de Haas effect, making an accurate determination of the exchange integrals possible, particularly in wide-gap materials, in which competing Landau and ordinary spin splittings are small.

The possibility of tailoring the magnitude of spin splitting in DMS structures offers a powerful tool to examine various phenomena. For instance, spin engineering was explored to control by the magnetic field the confinement of carriers and photons [68], to map atom distributions at interfaces [69] as well as to identify the nature of optical transitions and excitonic states. Furthermore, a subtle influence of spin splitting on quantum scattering amplitude of interacting electrons with opposite spins was put into evidence in DMS in the weakly localized regime in 3D [33], 2D [70, 71], and 1D systems [72]. The redistribution of carriers between spin levels induced by spin splitting was found to drive an insulator-to-metal transition [73] as well as to generate universal conductance fluctuations in DMS quantum wires [72]. Since the spin splitting is greater than the cyclotron energy, there are no overlapping Landau levels in modulation-doped heterostructures of DMS in the quantum Hall regime in moderately strong magnetic fields. This made it possible to test a scaling behavior of wave functions at the center of Landau levels [74]. At higher fields, a crossing of Landau levels occurs, so that quantum Hall ferromagnet could be evidenced and studied [75]. At the same time, it has been confirmed that in the presence of a strong spin-orbit coupling (e.g., in the case of p-type wave functions) the spin polarization can generate a large extraordinary (anomalous) Hall voltage [76]. Last but not least, optically [77] and electrically controlled spin-injection [78] and filtering [79] were observed in all-semiconductor structures containing DMS.

Spin-disorder Scattering

Spatial fluctuations of magnetization, disregarded in the mean-field approximation, lead to spin disorder scattering. According to the fluctuation-dissipation theorem, the corresponding scattering rate in the paramagnetic phase is proportional to $T\chi(T)$, where $\chi(T)$ is the magnetic susceptibility of the localized spins [12, 80]. Except to the vicinity of ferromagnetic phase transitions, a direct contribution of spin-disorder scattering to momentum relaxation turns out to be small. In contrast, this scattering mechanism controls the spin lifetime of effective mass carriers in DMS, as evidenced by studies of universal conductance fluctuations [81], line-width of spin-flip Raman scattering [80], and optical pumping efficiency [82]. Furthermore, thermodynamic fluctuations contribute to the temperature dependence of the band gap and band off-set. In the case when the total potential introduced by a magnetic ion is grater than the width of the carrier band, the virtual crystal and molecular field approximations break down, a case of the holes in $Cd_{1-x}Mn_xS$.

A non-perturbative scheme was developed [83, 84] to describe nonlinear dependencies of the band gap on x and of the spin splitting on magnetization observed in such situations.

4.5 Magnetic Polarons

Bound magnetic polaron (BMP), that is a bubble of spins ordered ferromagnetically by the exchange interaction with an effective mass carrier in a localized state, modifies optical, transport, and thermodynamic properties of DMS. BMP is formed inside the localization radius of an occupied impurity or quantum dot state but also around a trapped exciton, as the polaron formation time is typically shorter than the exciton lifetime [85]. The BMP binding energy and spontaneous carrier spin-splitting are proportional to the magnitude of local magnetization, which is built up by two effects: the molecular field of the localized carrier and thermodynamic fluctuations of magnetization [86, 87, 88, 12]. The fluctuating magnetization leads to dephasing and enlarges width of optical lines. Typically, in 2D and 3D systems, the spins alone cannot localized itinerant carriers but in the 1D case the polaron is stable even without any pre-localizing potential [83]. In contrast, a free magnetic polaron – a delocalized carrier accompanied by a travelling cloud of polarized spins – is expected to exist only in magnetically ordered phases. This is because coherent tunnelling of quasi-particles dressed by spin polarization is hampered, in disordered magnetic systems, by a smallness of quantum overlap between magnetizations in neighboring space regions. Interestingly, theory of BMP can readily be applied for examining effects of the hyperfine coupling between nuclear spins and carriers in localized states.

4.6 Exchange Interactions between Localized Spins

As in most magnetic materials, classical dipole-dipole interactions between magnetic moments are weaker than exchange couplings in DMS. Direct d-d or f-f exchange interactions, known from properties of magnetic dimmers, are thought to be less important than indirect exchange channels. The latter involve a transfer of magnetic information via spin polarization of bands, which is produced by the exchange interaction or spin-dependent hybridization of magnetic impurity and band states. If magnetic orbitals are involved in the polarization process, the mechanism is known as superexchange, which is merely antiferromagnetic and dominates, except for p-type DMS. If fully occupied band states are polarized by the sp-d exchange interaction, the resulting indirect d-d coupling is known as the Bloembergen-Rowland mechanism. In the case of Rudermann-Kittel-Kasuya-Yosida (RKKY) interaction, the d-d coupling proceeds via spin polarization of partly filled bands, that is by free carriers. Since in DMS the sp-d is usually smaller than the width of the relevant band (weak coupling limit) as well as the carrier concentration is

usually smaller than those of localized spins, the energetics of the latter can be treated in the continuous medium approximation, an approach referred here to as the Zener model. Within this model the RKKY interaction is ferromagnetic, and particularly strong in p-type materials, because of a large magnitudes of the hole mass and exchange integral β. It worth emphasizing that the Zener model is valid for any ratio of the sp-d exchange energy to the Fermi energy. Finally, in the case of systems in which magnetic ions in different charge states coexist, hopping of an electron between magnetic orbitals of neighboring ions in differing charge states tends to order them ferromagnetically. This mechanism, doubted the double exchange, operates in manganites but its relevance in DMS has not yet been found.

In general, the bilinear part of the interaction Hamiltonian for a pair of spins i and j is described by a tensor $\hat{J}$,

$$H_{ij} = -2\boldsymbol{S}^{(i)}\hat{J}^{(ij)}\boldsymbol{S}^{(j)} \ , \tag{3}$$

which in the case of the coupling between nearest neighbor cation sites in the unperturbed zinc-blende lattice contains four independent components. Thus, in addition to the scalar Heisenberg-type coupling, $H_{ij} = -2J^{(ij)}\boldsymbol{S}^{(i)}\boldsymbol{S}^{(j)}$, there are non-scalar terms (e.g., Dzialoshinskii-Moriya or pseudo-dipole). These terms are induced by the spin-orbit interaction within the magnetic ions or within non-magnetic atoms mediating the spin-spin exchange. The non-scalar terms, while smaller than the scalar ones, control spin-coherence time and magnetic anisotropy. Typically, $J^{(ij)} \approx -1$ meV for nearest-neighbor pairs coupled by the superexchange, and the interaction strength decays fast with the pair distance. Thus, with lowering temperature more and more distant pairs become magnetically neutral, $S_{tot} = 0$. Accordingly, the temperature dependence of magnetic susceptibility assumes a modified Curie form, $\chi(T) = C/T^{\gamma}$, where $\gamma < 1$ and both C and γ depend on the content of the magnetic constituent x. Similarly, the field dependence of magnetization is conveniently parameterized by a modified Brillouin function B_S [89],

$$M_o(T,H) = Sg\mu_B N_o x_{eff} \mathrm{B}_S[Sg\mu_B H/k_B(T+T_{AF})] \ , \tag{4}$$

in which two x- and T-dependent empirical parameters, $x_{eff} < x$ and $T_{AF} > 0$, describe the presence of antiferromagnetic interactions.

4.7 Magnetic Collective Phenomena

In addition to magnetic and neutron techniques [90], a variety of optical and transport methods, including $1/f$ noise study of nanostructures [81], have successfully been employed to characterize collective spin phenomena in DMS. Undoped DMS belong to a rare class of systems, in which spin-glass freezing is driven by purely antiferromagnetic interactions, an effect of spin frustration inherent to the randomly occupied fcc sublattice. Typically, in II-VI DMS, the spin-glass freezing temperature T_g increases from 0.1 K for

$x = 0.05$ to 20 K at $x = 0.5$ according to $T_g \sim x^{\delta}$, where $\delta \approx 2$, which reflects a short-range character of the superexchange. For x approaching 1, antiferromagnetic type III ordering develops, according to neutron studies. Here, strain imposed by the substrate material–the strain engineering–can serve to select domain orientations as well as to produce spiral structures with a tailored period [91]. Particularly important is, however, the carrier-density controlled ferromagnetism of bulk and modulation-doped p-type DMS described next.

5 Properties of Ferromagnetic Semiconductors

5.1 Overview

Since for decades III-V semiconductor compounds have been applied as photonic and microwave devices, the discovery of ferromagnetism first in $In_{1-x}Mn_xAs$ [92] and then in $Ga_{1-x}Mn_xAs$ by Hideo Ohno and collaborators in Sendai [93] came as a landmark achievement. In these materials, substitutional divalent Mn ions provide localized spins and function as acceptor centers that provide holes which mediate the ferromagnetic coupling between the parent Mn spins [94, 95, 96]. In another technologically important group of semiconductors, in II-VI compounds, the densities of spins and carriers can be controlled independently, similarly to the case of IV-VI materials, in which hole-mediated ferromagnetism was discovered by Tomasz Story et al. in Warsaw already in the 80s [97]. Stimulated by the theoretical predictions of the present author [94], laboratories in Grenoble and Warsaw, led by late Yves Merle d'Aubigné and the present author, joined efforts to undertake comprehensive research dealing with carrier-induced ferromagnetism in II-IV materials containing Mn. Experimental studies conducted with the use of magnetooptical and magnetic methods led to the discovery of ferromagnetism in 2D and [54] 3D II-VI materials [55] doped by nitrogen acceptors.

Guided by the growing amount of experimental results, the present author and co-workers proposed a theoretical model of the hole-controlled ferromagnetism in III-V, II-VI, and group IV semiconductors containing Mn [98, 99]. In these materials conceptual difficulties of charge transfer insulators and strongly correlated disordered metals are combined with intricate properties of heavily doped semiconductors, such as Anderson-Mott localization and defect generation by self-compensation mechanisms. Nevertheless, the theory built on Zener's model of ferromagnetism and the Kohn-Luttinger kp theory of the valence band in tetrahedrally coordinated semiconductors has quantitatively described thermodynamic, micromagnetic, transport, and optical properties of DMS with delocalized or weakly localized holes [21, 98, 99, 100], challenging competing theories. It is often argued that owing to these studies $Ga_{1-x}Mn_xAs$ has become one of the best-understood ferromagnets. Accordingly, this material is now employed as a testing ground for various

ab initio computation approaches to strongly correlated and disordered systems. Moreover, the understanding of the carrier-controlled ferromagnetic DMS has provided a basis for the development of novel methods enabling magnetization manipulation and switching.

5.2 p-d Zener Model

It is convenient to apply the Zener model of carrier-controlled ferromagnetism by introducing the functional of free energy density, $\mathcal{F}[\boldsymbol{M}(\boldsymbol{r})]$. The choice of the local magnetization $\boldsymbol{M}(\boldsymbol{r})$ as an order parameter means that the spins are treated as classical vectors, and that spatial disorder inherent to magnetic alloys is neglected. In the case of magnetic semiconductors $\mathcal{F}[\boldsymbol{M}(\boldsymbol{r})]$ consists of two terms, $\mathcal{F}[\boldsymbol{M}(\boldsymbol{r})] = \mathcal{F}_S[\boldsymbol{M}(\boldsymbol{r})] + \mathcal{F}_c[\boldsymbol{M}(\boldsymbol{r})]$, which describe, for a given magnetization profile $\boldsymbol{M}(\boldsymbol{r})$, the free energy densities of the Mn spins in the absence of any carriers and of the carriers in the presence of the Mn spins, respectively. A visible asymmetry in the treatment of the carries and of the spins corresponds to an adiabatic approximation: the dynamics of the spins in the absence of the carriers is assumed to be much slower than that of the carriers. Furthermore, in the spirit of the virtual-crystal and molecular-field approximations, the classical continuous field $\boldsymbol{M}(\boldsymbol{r})$ controls the effect of the spins upon the carriers. Now, the thermodynamics of the system is described by the partition function Z, which can be obtained by a functional integration of the Boltzmann factor $\exp(-\int d\boldsymbol{r}\mathcal{F}[\boldsymbol{M}(\boldsymbol{r})]/k_BT)$ over all magnetization profiles $\boldsymbol{M}(\boldsymbol{r})$ [87, 88]. In the mean-field approximation (MFA), a term corresponding to the minimum of $\mathcal{F}[\boldsymbol{M}(\boldsymbol{r})]$ is assumed to determine Z with a sufficient accuracy.

If energetics is dominated by spatially uniform magnetization $\boldsymbol{M}$, the spin part of the free energy density in the magnetic field $\boldsymbol{H}$ can be written in the form [101]

$$\mathcal{F}_S[\boldsymbol{M}] = \int_0^{\boldsymbol{M}} d\boldsymbol{M}_o \boldsymbol{h}(\boldsymbol{M}_o) - \boldsymbol{M}\boldsymbol{H} \ . \tag{5}$$

Here, $\boldsymbol{h}(\boldsymbol{M}_o)$ denotes the inverse function to $\boldsymbol{M}_o(\boldsymbol{h})$, where $\boldsymbol{M}_o$ is the available experimentally macroscopic magnetization of the spins in the absence of carriers in the field h and temperature T. In DMS, it is usually possible to parameterize $M_o(h)$ by the Brillouin function that, according to Eq. (4), takes the presence of intrinsic short-range antiferromagnetic interactions into account. Near T_C and for $H = 0$, M is sufficiently small to take $M_o(T, h) = \chi(T)h$, where $\chi(T)$ is the magnetic susceptibility of localized spins in the absence of carriers. Under these conditions,

$$\mathcal{F}_S[M] = M^2/2\chi(T) \ , \tag{6}$$

which shows that the increase of $\mathcal{F}_S$ with M slows down with lowering temperature, where $\chi(T)$ grows. Turning to $\mathcal{F}_c[M]$ we note that owing to the

giant Zeeman splitting of the bands proportional to M, the energy of the carriers, and thus $\mathcal{F}_c[M]$, decreases with $|M|$, $\mathcal{F}_c[M] - \mathcal{F}_c[0] \sim -M^2$. Accordingly, a minimum of $\mathcal{F}[M]$ at non-zero M may develop in $H = 0$ at sufficiently low temperatures signalizing the appearance of a ferromagnetic order.

The present authors and co-workers [98] found that the minimal hamiltonian necessary to describe properly effects of the complex structure of the valence band in tetrahedrally coordinated semiconductors upon $\mathcal{F}_c[M]$ is the Luttinger 6×6 kp model supplemented by the p-d exchange contribution taken in the virtual crystal and molecular field approximations,

$$H_{pd} = \beta \boldsymbol{s}\boldsymbol{M}/g\mu_B \ . \tag{7}$$

This term leads to spin splittings of the valence subbands, whose magnitudes – owing to the spin-orbit coupling – depend on the hole wave vectors $\boldsymbol{k}$ in a complex way even for spatially uniform magnetization $\boldsymbol{M}$. It would be technically difficult to incorporate such effects to the RKKY model, as the spin-orbit coupling leads to non-scalar terms in the spin-spin Hamiltonian. At the same time, the indirect exchange associated with the virtual spin excitations between the valence subbands, the Bloembergen-Rowland mechanism, is automatically included. The model allows for biaxial strain, confinement, and was developed for both zinc blende and wurzite materials [99]. Furthermore, the direct influence of the magnetic field on the hole spectrum was taken into account. Carrier-carrier spin correlation was described by introducing a Fermi-liquid-like parameter A_F [54, 94, 96], which enlarges the Pauli susceptibility of the hole liquid. No disorder effects were taken into account on the ground that their influence on thermodynamic properties is relatively weak except for strongly localized regime. Having the hole energies, the free energy density $\mathcal{F}_c[\boldsymbol{M}]$ was evaluated according to the procedure suitable for Fermi liquids of arbitrary degeneracy. By minimizing $\mathcal{F}[\boldsymbol{M}] = \mathcal{F}_S[\boldsymbol{M}] + \mathcal{F}_c[\boldsymbol{M}]$ with respect to $\boldsymbol{M}$ at given T, H, and hole concentration p, Mn spin magnetization $M(T, H)$ was obtained as a solution of the mean-field equation,

$$\boldsymbol{M}(T, H) = x_{eff} N_o g \mu_B S \mathrm{B}_S[g\mu_B(-\partial \mathcal{F}_c[\boldsymbol{M}]/\partial \boldsymbol{M} + \boldsymbol{H})/k_B(T + T_{AF})] \ , \tag{8}$$

where peculiarities of the valence band structure, such as the presence of various hole subbands, anisotropy, and spin-orbit coupling, are hidden in $F_c[\boldsymbol{M}]$. Near the Curie temperature T_C and at $H = 0$, where M is small, we expect $\mathcal{F}_c[M] - \mathcal{F}_c[0] \sim -M^2$. It is convenient to parameterize this dependence by a generalized carrier spin susceptibility $\tilde{\chi}_c$, which is related to the magnetic susceptibility of the carrier liquid according to $\tilde{\chi}_c = A_F(g * \mu_B)^2 \chi_c$. In terms of $\tilde{\chi}_c$,

$$\mathcal{F}_c[M] = \mathcal{F}_c[0] - A_F \tilde{\chi_c} \beta^2 M^2 / 2(g\mu_B)^2 \ . \tag{9}$$

By expanding $\mathrm{B}_S(M)$ for small M one arrives to the mean-field formula for $T_C = T_F - T_{AF}$, where T_F is given by

$$T_F = x_{eff} N_o S(S+1) A_F \tilde{\chi}_c(T_C) \beta^2 / 3k_B \ . \tag{10}$$

For a strongly degenerate carrier liquid $|\epsilon_F|/k_B T \gg 1$, as well as neglecting the spin-orbit interaction $\tilde{\chi}_c = \rho/4$, where ρ is the total density-of-states for intra-band charge excitations, which in the 3D case is given by $\rho = m^*_{DOS} k_F / \pi^2 \hbar^2$. In this case and for $A_F = 1$, T_F assumes the well-known form, derived already in 40s in the context of carrier-mediated nuclear ferromagnetism [102]. In general, however, $\tilde{\chi}_c$ has to be determined numerically by computing $\mathcal{F}_c[M]$ for a given band structure and degeneracy of the carrier liquid. The model can readily be generalized to various dimensions as well as to the case, when $\boldsymbol{M}$ is not spatially uniform in the ground state.

The same formalism, in addition to T_C and Mn magnetization $M(T, H)$, as discussed above, provides also quantitative information on spin polarization and magnetization of the hole liquid [99]. Furthermore, it can be exploited to describe chemical trends as well as micromagnetic, transport, and optical properties of ferromagnetic DMS, the topics discussed in the subsequent sections.

5.3 Curie Temperature – Chemical Trends

Large magnitudes of both density of states and exchange integral specific to the valence band make T_F to be much higher in p-type than in n-type materials with a comparable carrier concentration. Accordingly, in agreement with theoretical evaluations [94], no ferromagnetic order was detected above 1 K in n-(Zn,Mn)O:Al, even when the electron concentration exceeded 10^{20} cm^{-3} [103]. At the same time, theoretical calculations carried out with no adjustable parameters explained satisfactorily the magnitude of T_C in both (Ga,Mn)As [98, 104] and p-type (Zn,Mn)Te [55]. Furthermore, theoretical expectations within the p-d Zener model are consistent with chemical trends in T_C values observed experimentally in (Ga,Mn)Sb, (Ga,Mn)P, (In,Mn)As, (In,Mn)Sb, (Ge,Mn), and p-(Zn,Be)Te though effects of hole localization [99, 55] preclude the appearance of a uniform ferromagnetic order with a univocally defined T_C value in a number of cases. In addition to localization, a competition between long-range ferromagnetic interactions and intrinsic short-range antiferromagnetic interactions [100], as described by $T_{AF} > 0$ and $x_{eff} < x$, may affect the character of magnetic order [105]. It appears that the effect is more relevant in II-VI DMS than in III-V DMS where Mn centers are ionized, so that the enhanced hole density at closely lying Mn pairs may compensate antiferromagnetic interactions [98]. In both groups of materials the density of compensating donor defects appear to grow with the Mn concentration [95, 55]. In the case of (Ga,Mn)As the defect involved is the Mn interstitial [106], which can be driven and passivated at the surface be low temperature annealing [107].

According to evaluations carried out by the present author and co-workers [98] room temperature ferromagnetism could be observed in a weakly com-

pensated (Ga,Mn)As containing at least 10% of Mn. At the same time, because of stronger p-d hybridization in wide band-gap materials, such as (Ga,Mn)N and (Zn,Mn)O, $T_C > 300$ K is expected already for $x = 5\%$, provided that the hole concentration would be sufficiently high. However, it was clear from the beginning [98] that the enhancement of the hole binding energy by p-d hybridization as well as a limited solubility of magnetic constituent together with the effect of self-compensation may render the fabrication of high temperature ferromagnetic DMS challenging. Nevertheless, a number of group has started the growth of relevant systems, the effort stimulated even further by a number of positive results as well as by numerous theoretical papers suggesting, based on ab initio computations, that high temperature ferromagnetism is possible in a large variety of DMS even without band holes. Today, however, a view appears to prevail that the high temperature ferromagnetism, as evidenced by either magnetic, magnetotransport or magnetooptical phenomena, results actually from the presence of precipitates of known or so-far unknown ferromagnetic or ferrimagnetic nanocrystals containing a high density of magnetic ions. At the same time, it becomes more and more clear that the ab initio computations in question suffered from improper treatment of correlation and disorder, which led to an overestimation of tendency towards a ferromagnetic order. It seems at the end that, as argued initially [94, 98], the delocalized or weakly localized holes are necessary to stabilize a long-range ferromagnetic order in tetrahedrally coordinated DMS with a small concentration of randomly distributed magnetic ions.

5.4 Micromagnetic Properties

Magnetic Anisotropy

As the energy of dipole-dipole magnetic interactions depends on the dipole distribution, there exists the so-called shape anisotropy. In particular, for thin films, the difference in energy density corresponding to the perpendicular and in-plane orientation of magnetization M is given by

$$E = \mu_o M^2/2 \,, \tag{11}$$

which leads to the anisotropy field $\mu_o H_A = \mu_o M$ of about 60 mT for $Ga_{0.95}Mn_{0.05}As$.

Already early studies of the ferromagnetic phase in (In,Mn)As [108] and (Ga,Mn)As [109] demonstrated the existence of magnetic anisotropy, whose character and magnitude implied a sizable contribution of a microscopic origin. Magneto-crystalline anisotropy is usually associated with the interaction between spin and orbital degrees of freedom of the magnetic ion d-electrons. According to the model advocated here, these electrons are in the d^5 configuration. For such a case the orbital momentum $L = 0$, so that effects stemming from the spin-orbit coupling are expected to be rather weak. It was, however,

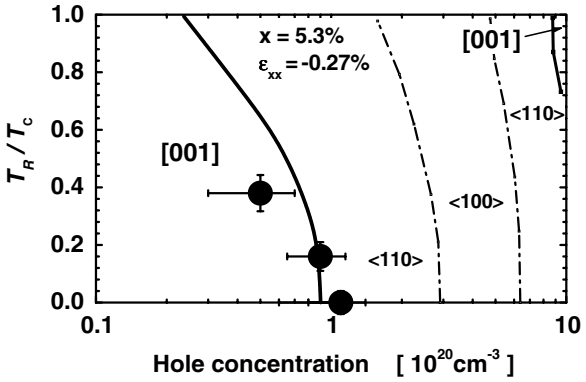

Fig. 7: Experimental (*full points*) and computed values (*thick lines*) of the ratio of the reorientation to Curie temperature for the transition from perpendicular to in-plane magnetic anisotropy. Dashed lines mark expected temperatures for the reorientation of the easy axis between ⟨100⟩ and ⟨110⟩ in-plane directions (after Sawicki et al. [111])

been noted that the interaction between the localized spins is mediated by the holes that have a non-zero orbital momentum $l = 1$ [98]. An important aspect of the p-d Zener model is that it does take into account the anisotropy of the carrier-mediated exchange interaction associated with the spin-orbit coupling in the host material [98, 99, 110].

A detail theoretical analysis of anisotropy energies and anisotropy fields in films of (Ga,Mn)As was carried out for a number of experimentally important cases within the p-d Zener model [99, 110]. In particular, the cubic anisotropy as well as uniaxial anisotropy under biaxial epitaxial strain were examined as a function of the hole concentration p. Both shape and magneto-crystalline anisotropies were taken into account. The perpendicular and in-plane orientation of the easy axis is expected for the compressive and tensile strain, respectively, provided that the hole concentration is sufficiently small. However, according to theory, a reorientation of the easy axis direction is expected at higher hole concentrations. Furthermore, in a certain concentration range the character of magnetic anisotropy is computed to depend on the magnitude of spontaneous magnetization, that is on the temperature. The computed phase diagram for the reorientation transition compared to the experimental results for a film is shown in Fig. 7. In view that theory is developed with no adjustable parameters the agreement between experimental and computed concentrations and temperature corresponding to the reorientation transition is very good. Furthermore, the computed magnitudes of the anisotropy field H_u [99] are consistent with the available findings for both compressive and tensile strain.

According to the discussion above, the easy axis assumes the in-plane orientation for typical carrier concentrations in the most thoroughly studied system (Ga,Mn)As/GaAs. In this case the easy axis is expected to switch between ⟨100⟩ and ⟨110⟩ in-plane cubic directions as a function of p [99, 110]. Surprisingly, however, only the ⟨100⟩ biaxial magnetic symmetry has so-far been observed in films of (Ga,Mn)As/GaAs at low temperatures. Nevertheless, the corresponding in-plane anisotropy field assumes the expected magnitude, of the order of 0.1 T, which is typically much smaller than that corresponding to the strain-induced energy of magnetic anisotropy. It is possible that anisotropy of the hole magnetic moment, neglected in the theoretical calculations [99, 110], stabilizes the ⟨100⟩ orientation of the easy axis.

In addition to the cubic in-plane anisotropy, the accumulated data for both (Ga,Mn)As/GaAs and (In,Mn)As/(In,Al)As point to a non-equivalence of [110] and [−110] directions, which leads to the in-plane uniaxial magnetic anisotropy. Such a uniaxial anisotropy is not expected for D_{2d} symmetry of a T_d crystal under epitaxial strain [112, 113]. Furthermore, the magnitude of the corresponding anisotropy field appears to be independent of the film thickness [114], which points to a puzzling symmetry breaking in the film body.

Magnetic Stiffness and Domain Structure

Another important characteristics of any ferromagnetic system is magnetic stiffness A, which describes the energy penalty associated with the local twisting of the direction of magnetization. Remarkably, A determines the magnitude and character of thermodynamic fluctuations of magnetization, the spectrum of spin excitations as well as the width and energy of domain walls. An important result is that the magnetic stiffness computed within the 6×6 Luttinger model is almost by a factor of 10 greater than that expected for a simple spin degenerate band with the heave-hole band-edge mass [115]. This enhancement, which stabilizes strongly the spatially uniform spin ordering, stems presumably from p-like symmetry of the valence band wave functions, as for such a case the carrier susceptibility (the Lindhard function) decreases strongly with q [116].

The structure of magnetic domains in (Ga,Mn)As under tensile strain has been determined by micro-Hall probe imaging [117]. The regions with magnetization oriented along the [001] and [00-1] easy axis form alternating stripes extending in the [110] direction. As shown in Fig. 8, the experimentally determined stripe width is $W = 1.5$ μm at 5 K for 0.2 μm film of $Ga_{0.957}Mn_{0.043}As$ on $Ga_{0.84}In_{0.16}As$, for which tensile strain of $\epsilon_{xx} = 0.9\%$ is expected. According to micromagnetic theory, W is determined by the ratio of the domain wall energy to the stray field energy. As shown in Fig. 8, the computed value with no adjustable parameters $W = 1.1$ μm [118] compares favorably with the experimental finding, $W = 1.5$ μm at low temperatures. However, the model

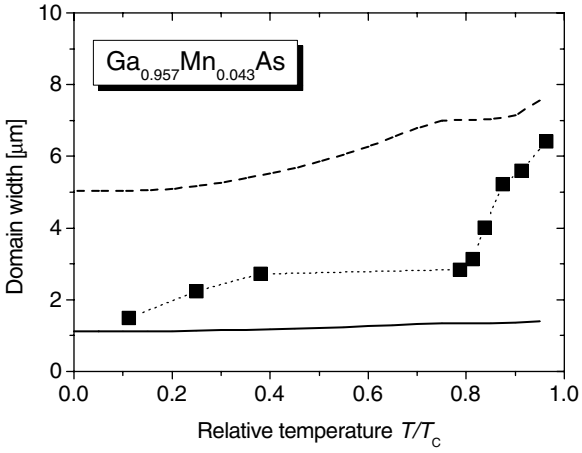

Fig. 8: Temperature dependence of the width of domain stripes as measured by Shono et al. [117] for the $Ga_{0.957}Mn_{0.043}As$ film with the easy axis along the growth direction (*full squares*). Computed domain width is shown by the *solid line* (after Dietl et al. [118])

predicts much weaker temperature dependence of W than that observed experimentally, which was linked [118] to critical fluctuations, disregarded in the mean-field approach.

5.5 Optical Properties

Magnetic Circular Dichroism

Within the Zener model, the strength of the ferromagnetic spin-spin interaction is controlled by the $k \cdot p$ parameters of the host semiconductor and by the magnitude of the spin-dependent coupling between the effective mass carriers and localized spins. In the case of II-VI DMS, detailed information on the exchange-induced spin-splitting of the bands, and thus on the coupling between the effective mass electrons and the localized spins has been obtained from magnetooptical studies [12]. A similar work on (Ga,Mn)As [119, 120, 121] led to a number of surprises. The most striking was the opposite order of the absorption edges corresponding to the two circular photon polarizations in (Ga,Mn)As comparing to II-VI materials. This behavior of circular magnetic dichroism (MCD) suggested the opposite order of the exchange-split spin subbands, and thus a different origin of the sp-d interaction in these two families of DMS. A new light on the issue was shed by studies of photoluminescence (PL) and its excitation spectra (PLE) in p-type (Cd,Mn)Te quantum wells [54]. As shown schematically in Fig. 9, the reversal of the order of PLE edges corresponding to the two circular polarizations results from the

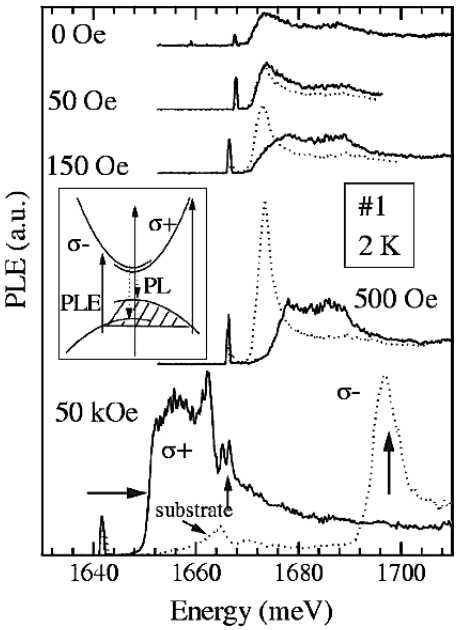

Fig. 9: Photoluminescence excitation spectra (PLE), that is the photoluminescence (PL) intensity as a function of the excitation photon energy intensity, for σ^+ (*solid lines*) and σ^- (*dotted lines*) circular polarizations at selected values of the magnetic field in a modulation-doped p-type quantum well of $Cd_{0.976}Mn_{0.024}Te$ at 2 K. The photoluminescence was collected in σ^+ polarization at energies marked by the narrowest features. The sharp maximum (*vertical arrow*) and step-like form (*horizontal arrow*) correspond to quasi-free exciton and transitions starting at the Fermi level, respectively. Note reverse ordering of transition energies at σ^+ and σ^- for PL and PLE (the latter is equivalent to optical absorption). The band arrangement at 150 Oe is sketched in the inset (after Haury et al. [54])

Moss-Burstein effect, that is from the shifts of the absorption edges associated with the empty portion of the valence subbands in the p-type material.

The above model was subsequently applied to interpret the magnetoabsorption data for metallic (Ga,Mn)As [99, 120]. More recently, the theory was extended by taking into account the effect of scattering-induced mixing of k states [122]. As shown in Fig. 10, this approach explains the slop of the absorption edge as well as its field-induced splitting assuming the value of the p-d exchange energy $\beta N_0 = -1$ eV.

Recently, the formalisms suitable for description of either interband [99] or intraband [123] optical absorption were combined [124] in order to examine theoretically optical (dynamic) conductivity in the whole spectral range up to

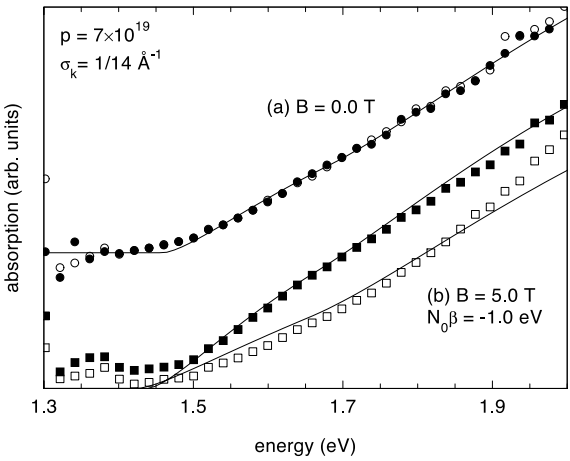

Fig. 10: Transmission of $Ga_{0.968}Mn_{0.032}As$ film for two circular light polarizations in the Faraday configuration in the absence of the magnetic field (data shifted up for clarity) and in 5 T at 2 K (points) [120]. *Solid lines* are calculated for the hole concentration $p = 7 \times 10^{19}$ cm^{-3}, exchange energy $N_0\beta = -1$ eV, and allowing for scattering-induced breaking of the k selection rules [122]

2 eV. Furthermore a possible presence of optical absorption involving defect states was taken into account. In this way, the most general quantitative theory of optical and magnetoptical effects in magnetic semiconductors available to date was worked out. A good quantitative description of experimental data [125, 126] was obtained verifying the model. However, some discrepancies in the low photon energy range were detected, which confirmed the presence of quantum localization effects. At the same time, a disagreement in the high energy region pointed to the onset of intra-d band transitions. The Faraday and Kerr rotations were also computed showing a large magnitude and a complex spectral dependence in the virtually whole studied photon energy range up to 2 eV, which suggests a suitability of this material family for magnetooptical applications.

5.6 Charge Transport Phenomena

Hall Effect in Ferromagnetic Semiconductors – Theory

The assessment of magnetic characteristics by means of magnetotransport studies is of particular importance in the case of thin films of diluted magnets, in which the magnitude of the total magnetic moment is typically small. For this reason, recent years have witnessed a renewed interest in the nature of the anomalous Hall effect (AHE), which–if understood theoretically–can

serve to determine the magnitude of magnetization. Also magnetoresistance, to be discussed later on, provides information on the magnetism and on the interplay between electronic and magnetic degrees of freedom.

The Hall resistance $R_{Hall} \equiv \rho_{yx}/d$ of a film of the thickness d is empirically known to be a sum of ordinary and anomalous Hall terms in magnetic materials [127],

$$R_{Hall} = R_0 \mu_o H/d + R_S \mu_o M/d \, . \tag{12}$$

Here, R_0 and R_S are the ordinary and anomalous Hall coefficients, respectively ($R_0 > 0$ for the holes), and $M(T, H)$ is the component of the magnetization vector perpendicular to the sample surface. While the ordinary Hall effect serves to determine the carrier density, the anomalous Hall effect (known also as the extraordinary Hall effect) provides valuable information on magnetic properties of thin films. The coefficient R_S is usually assumed to be proportional to R_{sheet}^{α}, where $R_{sheet}(T, H)$ is the sheet resistance and the exponent α depends on the mechanisms accounting for the AHE.

If the effect of stray magnetic fields produced by localized magnetic moments were been dominating, R_S would scale with magnetization M but would be rather proportional to R_0 than to R_{sheet}. There is no demagnetization effect in the magnetic field perpendicular to the surface of a uniformly magnetized film, $B = \mu_o H$. However, this is no longer the case in the presence of magnetic precipitates, whose stray fields and AHE may produce an apparent magnetization-dependent contribution the host Hall resistance.

When effects of stray fields can be disregarded, spin-orbit interactions control totally R_S. In such a situation α is either 1 or 2 depending on the origin of the effect: the skew-scattering (extrinsic) mechanism, for which the Hall conductivity is proportional to momentum relaxation time τ, results in $\alpha \approx 1$ [127]. From the theory point of view particularly interesting is the intrinsic mechanism for the Hall conductivity $\sigma_{AH} = R_S M/(R_{sheet} d)^2]$ does not depend explicitly on scattering efficiency but only on the band structure parameters [21, 128, 129].

For both extrinsic and intrinsic mechanisms, the overall magnitude of the anomalous Hall resistance depends on the strength of the spin-orbit interaction and spin polarization of the carriers at the Fermi surface. Accordingly, at given magnetization M, the effect is expected to be much stronger for the holes than for the electrons in tetrahedrally coordinated semiconductors. For the carrier-mediated ferromagnetism, the latter is proportional to the exchange coupling of the carriers to the spins, and varies – not necessarily linearly – with the magnitude of spin magnetization M. Additionally, the skew-scattering contribution depends on the asymmetry of scattering rates for particular spin subbands, an effect which can depend on M in a highly nontrivial way. Importantly, the sign of either of the two contributions can be positive or negative depending on a subtle interplay between the orientations of orbital and spin momenta as well as on the character (repulsive vs. attractive) of scattering potentials.

Recently, Jungwirth et al. [129] developed a theory of the intrinsic AHE in p-type zinc-blende magnetic semiconductors, and presented numerical results for the case of (Ga,Mn)As, (In,Mn)As, and (Al,Mn)As. The derived formula for σ_{AH} corresponds to that given earlier [128, 130, 131] in the weak scattering limit. The intrinsic AHE can also be regarded as a zero-frequency limit of $\sigma_{xy}(\omega)$, where $\sigma(\omega)$ is the dynamic (optical) conductivity tensor, related directly the Kerr effect, widely studied in experimentally and theoretically in ferromagnetic metals [132]. For the hole concentration p such that the Fermi energy is much smaller than the spin-orbit splitting Δ_o but larger than the exchange splitting h between the majority $j_z = -3/2$ and minority $j_z = +3/2$ bands at $k = 0$, $\Delta_o \gg |\epsilon_F| \gg h$, Jungwirth et al. [129] predict within the 4×4 spherical Luttinger model

$$\sigma_{AH}^{in} = e^2 h m_{hh}/[4\pi^2\hbar^3(3\pi p)^{1/3}] \, . \tag{13}$$

Here the heavy hole mass m_{hh} is assumed to be much larger than the light hole mass m_{lh}, whereas σ_{AH}^{in} becomes by the factor of $2^{4/3}$ greater in the opposite limit $m_{hh} = m_{lh}$. In the range $h \ll |\epsilon_F| \ll \Delta_o$ the determined value of σ_{AH}^{in} is positive, that is the coefficients of the normal and anomalous Hall effects are expected to have the same sign. However, if the Fermi level were approached the split-off Γ_7 band, a change of sign would occur.

A formula for σ_{AH}^{in} was also derived [133] from 4 of Jungwirth et al. [129]), employing the known form of the heavy hole Bloch wave functions $u_{\boldsymbol{k},j_z}$ [116]. Neglecting a small effect of the spin splitting on the heavy hole wave functions, σ_{AH}^{in} was found to be given by the right hand side of 14 multiplied by the factor $(16/9)\ln 2 - 1/6 \approx 1.066$ [133].

In order to evaluate the ratio of intrinsic and skew-scattering mechanisms, the general theory of the AHE effect in semiconductors [128, 129, 130, 131] was applied [15]. Assuming that scattering by ionized impurities dominates, this ratio is then given by [131, 134, 135],

$$\frac{\sigma_{AH}^{in}}{\sigma_{AH}^{ss}} = \pm f(\xi)(N_A + N_D)/(p r_s k_F \ell) \, . \tag{14}$$

Here, $f(\xi) \approx 10$ is a function that depends weakly on the screening dimensionless parameter ξ; $(N_A + N_D)/p$ is the ratio of the ionized impurity and carrier concentrations; r_s is the average distance between the carriers in the units of the effective Bohr radius, and ℓ is the mean free path. Similarly, for spin-independent scattering by short range potentials, $V(\boldsymbol{r}) = V\delta(\boldsymbol{r} - \boldsymbol{r}_i)$ [130] was applied [133]. Assuming that scattering by ionized impurities is negligible,

$$\frac{\sigma_{AH}^{in}}{\sigma_{AH}^{ss}} = -3/[\pi V \rho(\varepsilon_F) k_F \ell] \, , \tag{15}$$

where $\rho(\varepsilon_F)$ is the density of states at the Fermi level. Of course, the overall sign depends on the sign of the scattering potential V.

In order to find out which of the two AHE mechanisms operates predominantly in p-type tetrahedrally coordinated ferromagnetic semiconductors, we note that scattering by ionized impurities appears to dominate in these heavily doped and compensated materials. This scattering mechanism, together with alloy and spin disorder scattering, limits presumably the hole mobility and leads ultimately to the metal-to-insulator transition (MIT). Since at the MIT $r_s \approx 2$ and $k_F\ell \approx 1$ one expects from 15 that as long as the holes remain close to the localization boundary the intrinsic mechanism accounts for the AHE. It would be interesting to know how quantum localization corrections affect the anomalous Hall conductivity as well as how to extend theory towards the insulator side of the MIT. A work in this direction was reported [136].

Obviously, the presence of the AHE makes a meaningful determination of the carrier type and density difficult in ferromagnetic semiconductors. Usually, the ordinary Hall effect dominates only in rather high magnetic fields or at temperatures several times larger than T_C. It appears, therefore, that a careful experimental and theoretical examination of the resistivity tensor in wide field and temperature ranges is necessary to separate characteristics of the spin and carrier subsystems.

Comparison Between Theoretical and Experimental Results

As mentioned above, because of the dominance of the anomalous Hall term in wide temperature and field ranges, it is not straightforward to determine the carrier type and concentration in ferromagnetic semiconductors. Only at low temperatures and under very high fields, the anomalous Hall term saturates, so that the ordinary Hall coefficient can be determined from the remaining linear change of the Hall resistance in the magnetic field. Note that although magnetization saturates in relatively low magnetic fields, the negative MR usually persists, and generates the field dependence of the anomalous Hall coefficient.

Magnetotransport data collected for (Ga,Mn)As in a wide temperature and field ranges [95, 137] were exploited to test the theory of the AHE [129]. The results of such a comparison are shown in Fig. 11. There is a good agreement between the theoretical and experimental magnitude of the Hall conductivity. Importantly, no significant contribution from skew scattering is expected for the (Ga,Mn)As sample in question [137], for which $(N_A + N_D)/p \approx 5$, $r_s \approx 1.1$, and $k_F\ell \approx 0.8$, so that $\sigma_{AH}^{in}/\sigma_{AH}^{ss} \approx 57$.

Another material for which various contributions to Hall resistance were analyzed is $Zn_{0.981}Mn_{0.019}Te$:N containing 1.2×10^{20} holes per cm^3 [133]. In Fig. 12, $\rho_{yx}/\rho_{xx} - \mu B$, i.e., the spin dependent Hall angle, is compared to the magnetization measured in a vibrating sample magnetometer [138] for this film. The normal Hall angle $\mu B = \mu\mu_o H$ was subtracted assuming a constant hole mobility μ i.e., assigning the conductivity changes entirely to variations in the hole concentration. This assumption is not crucial for the

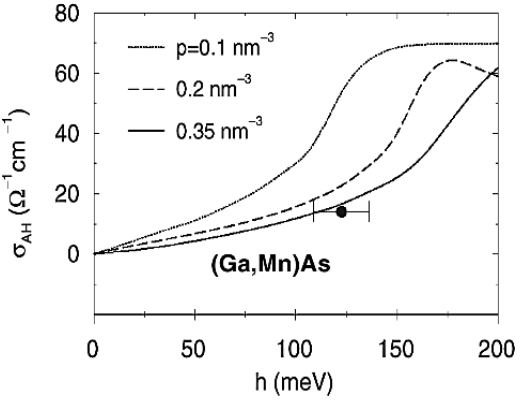

Fig. 11: Full numerical simulations of the anomalous Hall conductivity σ_{AH} for GaAs host with hole densities $p = 10^{20}$, (*dotted lines*), 2×10^{20} (*dashed lines*), and 3.5×10^{20} cm^{-3} (*solid lines*). Filled circle represents measured Hall conductivity (Fig. 2). The saturation mean-field value of the splitting h between Γ_8 heavy hole subbands was estimated from nominal sample parameters. Horizontal error bar corresponds to the experimental uncertainty of the p−d exchange integral. Experimental hole density in the (Ga,Mn)As sample is 3.5×10^{20} cm^{-3} (after Jungwirth et al. [129])

present highly doped sample, but it proves to be less satisfactory for the less doped samples. As shown in Fig. 12, a reasonable agreement is found by taking,

$$\rho_{yx}/\rho_{xx} = \mu B + \Theta M/M_S \ , \tag{16}$$

where M_S is the saturation value of magnetization and $\Theta = 0.04$ is the adjustable parameter. For the sample in question, the maximum value of hole polarization, $(p^{up} - p^{down})/(p^{up} + p^{down})$, has been estimated to be of the order of 10% [138].

Here, similarly to the case of (Ga,Mn)As, the sign and magnitude of the anomalous Hall coefficient indicated that the intrinsic mechanism is involved. The value of Θ was evaluated theoretically from 13 by adopting parameters suitable for the sample in question, $m_{hh} = 0.6m_o$, $\rho_{xx} = 5 \times 10^{-3}$ Ωcm and the saturation value of the splitting $h = 41$ meV. This leads to $\sigma^{in}_{AH} = 13.1$ $(\Omega\text{cm})^{-1}$ and $\Theta^{in} = 0.065$ [133], in a reasonable agreement with the experimental value $\Theta = 0.04$. Since a contribution from the light hole band will enhance the theoretical value, it was concluded [133] that the present theory describes the anomalous hole effect within the factor of about two.

It is important to note that there exist several reasons causing that the Hall effect and direct magnetometry can provide different information on magnetization. Indeed, contrary to the standard magnetometry, the AHE does not provide information about the magnetization of the whole samples

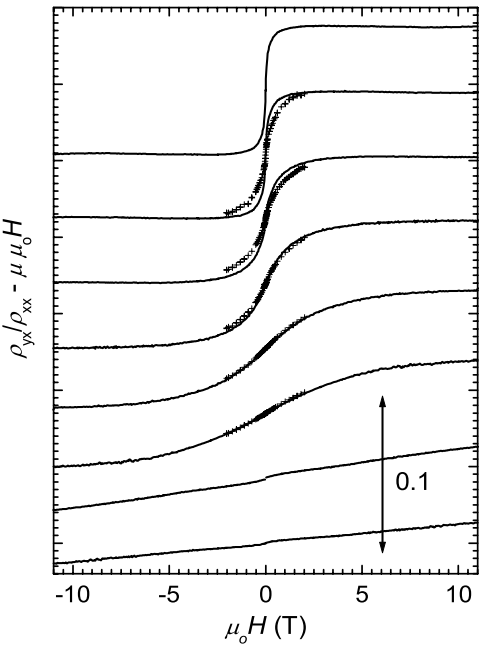

Fig. 12: Comparison of the normalized anomalous Hall effect (*lines*) with the normalized magnetization M/M_S (*crosses*); from *top* to *bottom*: 1.7, 2.8, 4.2, 7, 10, 30, and 50 K; the data are shifted for clarity (after [138])

but only about its value in regions visited by the carriers. Near the metal-insulator boundary, especially when the compensation is appreciable, the carrier distribution is highly non-uniform. In the regions visited by the carriers the ferromagnetic interactions are strong, whereas the remaining regions may remain paramagnetic. Under such conditions, magnetotransport and direct magnetic measurements will provide different magnetization values [98]. In particular, M_S at $T \to 0$, as seen by a direct magnetometry, can be much lower than that expected for a given value of the magnetic ion concentration. High magnetic fields are then necessary to magnetize all localized spins. The corresponding field magnitude is expected to grow with the temperature and strength of antiferromagnetic interactions that dominate in the absence of the holes.

Anisotropic Magnetoresistance and Planar Hall Effect

In cubic materials the conductivity tensor is diagonal in the absence of an external magnetic field. However, non-zero values of strain make the resistance to depend on the orientation of current in respect to crystallographic axes. Furthermore, the spin-orbit interaction accounts for anisotropic magnetore-

sistance (AMR), that is the dependence of resistance on the angle between the current and magnetization, an effect particularly useful for position sensing in engines. Under these conditions, even if magnetization remains in-plane the resistivity tensor may assume a non-diagonal form. This leads to the appearance of a Hall voltage, a phenomenon known as the planar Hall effect (PHE) [139]. The information on the orientation of in-plane magnetization, which can be obtained from AMR and PHE is, thus, complementary to that provided by AHE which is sensitive only to the perpendicular component of magnetization. In particular, AMR and PHE can trace the character of in-plane magnetization reorientation at the coercive field and serve to determine the corresponding anisotropy fields. Last but not least, AMR and PHE are sensitive probe of spin anisotropy at the Fermi surface associated with the strain and spin-orbit interaction for non-zero magnetization. The corresponding theory of AMR was developed by Jungwirth et al. [140] within the Drude-Boltzmann formulation of charge transport in solids.

To test the theoretical predictions concerning effects of biaxial strain upon AMR, (Ga,Mn)As samples under compressive and tensile strain were studied for longitudinal and two perpendicular orientations of the magnetic field in respect to electric current [141]. As show in Fig. 13, above 0.5 T, negative magnetoresistance is observed, whose magnitude is virtually independent of experimental configuration. However, the absolute value of resistance ρ in this range depends on the field direction, which is the signature of AMR. These data provide information on processes of the field-induced rotation of magnetization for various orientations of the field in respect to crystal and easy axes. In particular, the values of the field corresponding to the resistance maxima are expected to be of the order of the anisotropy field.

If only spin-orbit effects were controlled AMR, its magnitude would depend only on the angle between the current and field directions. According to Fig. 13, this is not the case since AMR depends also on the directions of the field and current in respect to crystal axes. It is convenient to introduce $\mathrm{AMR}_{op} = [\rho_{xx}(H \parallel x) - \rho_{xx}(H \parallel y)]/\rho_{xx}(H \parallel y)$ and $\mathrm{AMR}_{op} = [\rho_{xx}(H \parallel x) - \rho_{xx}(H \parallel z)]/\rho_{xx}(H \parallel z)$, where the current and growth directions are denoted by x and z, respectively, and ρ_{xx} is the longitudinal resistivity. Importantly, the sign and order of magnitude of AMR is consistent with theoretical expectations [140, 142]. In particular, the predicted difference in sign of $\mathrm{AMR}_{op}-\mathrm{AMR}_{ip}$ in the case of compressive and tensile strain is corroborated by the data. On the other hand, the dependence of AMR_{op} and AMR_{ip} at given strain on the current direction appears as challenging. It may result from the lowering of the symmetry of the (Ga,Mn)As films from the expected D_{2d} to C_{2v}, as discussed in Sect. 3.

Particularly intriguing are hysteretic resistance jumps observed for samples under compressive strain and for the field pointing along the growth direction. We assign this effect to a large ratio of the anisotropy and coercive fields, which makes that even a rather small misalignment, and thus a minute

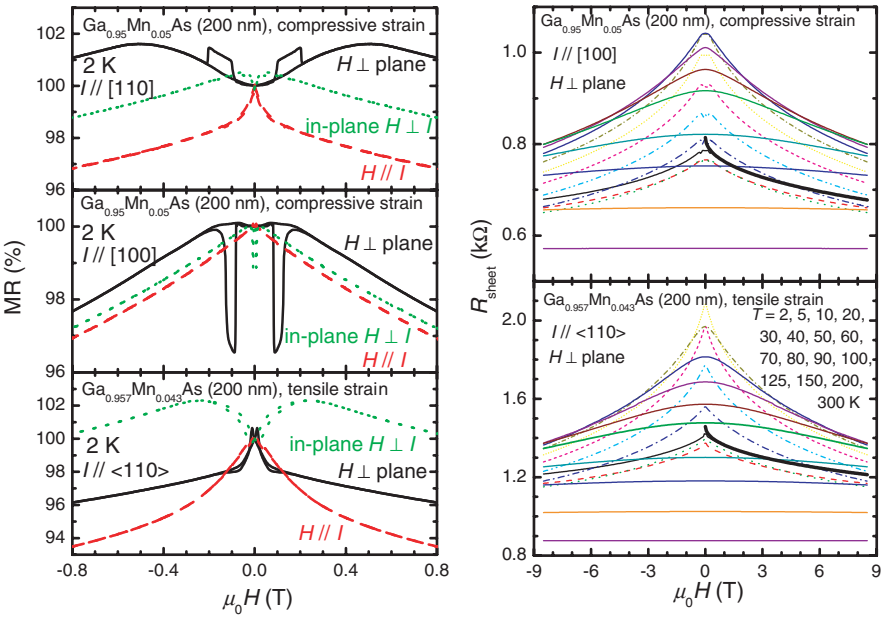

Fig. 13: *Left panel*: Field-induced changes in resistance of $Ga_{0.95}Mn_{0.05}As$/GaAs (compressive strain) *lower panel*: current along [110]; *middle panel*: current along [100]) and of $Ga_{0.957}Mn_{0.043}As$/(In,Ga)As under tensile strain (*lower panel*, current along ⟨110⟩) for three orientations of the magnetic field in respect to current direction at 2 K. *Right panel*: field and temperature dependencies of resistance in $Ga_{0.95}Mn_{0.05}As$/GaAs (compressive strain, *upper panel*) and in tensile strained $Ga_{0.957}Mn_{0.043}As$/(In,Ga)As (*lower panel*) for magnetic field perpendicular to the film plane. Starting from up, subsequent curves at $H = 0$ correspond to temperatures in K: 70, 60, 80, 50, 90, 40, 100, 30, 125, 20, 2, 5, 10, 150, 200, 300 (*upper panel*) and to 50, 60, 40, 70, 30, 80, 90, 20, 100, 2, 10, 5, 125, 150, 200, 300 (*lower panel*). The thick *solid lines* superimposed on 2 K data in positive magnetic field side show Kawabata's theory predictions (after Matsukura et al. [141])

in-plane field, can result in magnetization switching between in-plane easy directions. These results provide, therefore, information on resistance values in a demagnetized state for the studied current directions.

Low and High Field Magnetoresistance

Apart from AMR, there is a number of other effects that can produce a sizable magnetoresistance in magnetic semiconductors, especially in the vicinity of the localization boundary [12], where quantum corrections to Drude-Boltzmann conductivity become important. In particular, carrier diffusion in

the molecular field of randomly oriented spin clusters that form above T_C shifts the metal-to-insulator transition towards higher carrier concentrations [71]. The resulting temperature dependent localization may lead to a resistance maximum at T_C, which will be destroyed by the magnetic field. This accounts presumably for the field and temperature dependence of resistivity near T_C visible clearly in Fig. 13.

However, the negative magnetoresistance hardly saturates even in rather strong magnetic fields, and occurs also at low temperatures, where the spins are fully ordered ferromagnetically according to the Hall effect data. This surprising observation was explained by the present author and co-workers [133, 141] in terms of weak localization orbital magnetoresistance. Indeed, in the regime in question the giant splitting of the valence band makes both spin-disorder and spin-orbit scattering relatively inefficient. Under such conditions, weak localization magnetoresistance can show up at low temperatures, where phase breaking scattering ceases to operate. According to Kawabata [143],

$$\triangle\rho/\rho \approx -\triangle\sigma/\sigma = -n_v e^2 C_o (e/\hbar B)^{1/2}/(2\pi^2\hbar) \; , \tag{17}$$

where $C_o = 0.605$, σ is the conductivity, and $1/2 \leq n_v \leq 2$ depending on whether one or all four hole subbands contribute to the charge transport. For the samples under compressive and tensile strain, the above formula gives $\triangle\rho/\rho = -0.13 n_v$ and $-0.25 n_v$, respectively at $B = 9$ T. These values are to be compared to experimental data of Fig. 13, $\triangle\rho/\rho = -0.09$ and -0.14 at 2 K. The fitting to Eq. (17) reproduces the data at 2 K quite well (thin *solid lines* in Fig. 13 and gives $n_v = 1.46$ and 0.82 for the compressive and tensile samples, respectively, as could be expected for ferromagnetic films of (Ga,Mn)As. Since negative magnetoresistance takes over above $B_i \approx 1$ T, we can evaluate a lower limit for the spin-disorder scattering time, $\tau_s = m*/(eB_i k_F l) = 8$ ps for the hole effective mass $m* = 0.7 m_o$ and $k_F l = 0.5$, where k_F the Fermi momentum and l the mean free path.

5.7 Spin Transport Phenomena

To this category belongs a number of effects observed in heterostructures of (Ga,Mn)As, and important for perspective spintronic devices, such as spin injection of holes [144, 145] and electrons in the Zener diode [146, 147], giant magnetoresistance (GMR) [148], tunnelling magnetoresistance (TMR) [149, 150, 151], tunnelling anisotropic magnetoresistance (TAMR) [152, 153], and domain wall resistance [154, 155].

Since in most semiconductor spin transport devices the relevant length scale is shorter that the phase coherence length, a formulation of theory in terms of the Boltzmann distribution function f is not valid. Recently, theory that combines an empirical tight-binding approach with a Landauer-Büttiker formalism was developed [156, 157]. In contrast to the standard kp method,

this theory describes properly the interfaces and inversion symmetry breaking as well as the band dispersion in the entire Brillouin zone, so that the essential for the spin-dependent tunnelling Rashba and Dresselhaus terms as well as the tunnelling *via* $\boldsymbol{k}$ points away from the zone center are taken into account. This approach [156, 157], developed with no adjustable parameters, explained experimentally observed large magnitudes of both electron current spin polarization up to 70% in the (Ga,Mn)As/n-GaAs Zener diode [158] and TMR of the order of 300% in a (Ga,Mn)As/GaAs/(Ga,Mn)As trilayer structure [151]. Furthermore, theory reproduced a fast decrease of these figures with the device bias as well as it indicated that the magnitude of TAMR should not exceed 10% under usual strain conditions.

5.8 Methods of Magnetization Manipulation

Since magnetic properties are controlled by band holes, an appealing possibility is to influence the magnetic ordering isothermally, by light or by the electric field, which affect the carrier concentration in semiconductor structures. Such tuning capabilities of the materials systems in question were put into the evidence in (In,Mn)As/(Al,Ga)Sb [159, 160] and modulation doped p-(Cd,Mn)Te/(Cd,Mg,Zn)Te [54, 161] heterostructures, as depicted in Figs. 14 and 15. Actually, these findings can be quantitatively interpreted by considering the effect of the electric field or illumination on the hole density under stationary conditions and, therefore, on the Curie temperature in the relevant magnetic layers. Interestingly, according to experimental findings and theoretical modelling, photocarriers generated in II-VI systems by above barrier illumination destroy ferromagnetic order in the magnetic quantum well residing in an undoped (intrinsic) region of a p-i-p structure [54, 161] but they enhance the magnitude of spontaneous magnetization in the case of a p-i-n diode [161], as shown in Fig. 15.

Another method of magnetization manipulation, suitable for low-power switching of bits in magnetic memories, was invoked by Luc Berger [162] and John Slonczewski [163], who considered since dozen of years magnetization reversal by a transfer of spin momentum from the current of spin polarized carriers to localized magnetic moments in ferromagnetic metals. In the case of semiconductors, the current-induced magnetization reversal was demonstrated in submicron pillars of (Ga,Mn)As/GaAs/(Ga,Mn)As [164]. Furthermore, spin-polarized current was shown to displace magnetic domain walls in (Ga,Mn)As with the easy axis perpendicular to the film plane [165, 166].

6 Summary and Outlook

As an outcome of the great progress made in the field of semiconductor spintronics in the past few years as reviewed above, spin transistors were for the

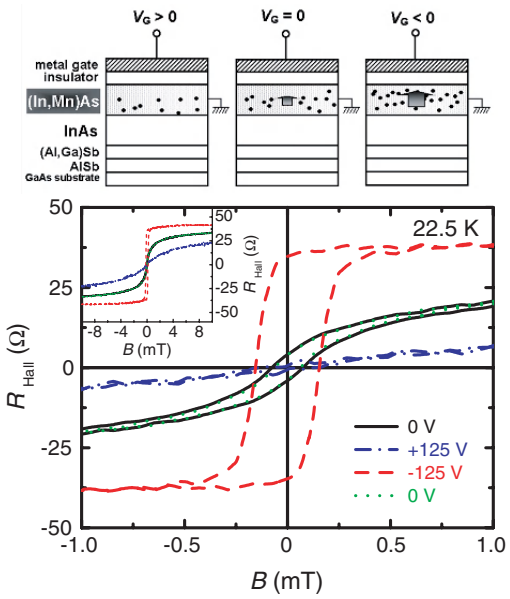

Fig. 14: Magnetization hysteresis loops determined by measurements of anomalous Hall effect at constant temperature of 22.5 K for various gate voltages in field-effect transistor with (In,Mn)As channel (after Ohno et al. [160])

first time described in The International Technology Roadmap for Semiconductors: Update 2004 Emerging Research Devices. Here it was also suggested that spin transistors might replace unipolar silicon transistors, which have been so successfully employed since the 1960s. It is, however, also obvious from this review that a number of challenges are ahead, so that semiconductor spintronics will attract a lot of attentions of the research community in the years to come.

From the device physics perspective, further works on magnetooptical isolators and modulators as well as on electrically controlled spin current generation, injection, detection, filtering, and amplification, particularly in spin bipolar devices [20] are expected. At the same time, further advancement in low-power magnetization switching will allow the development of new generation magnetic random access memories (MRAM) and, perhaps, extend the use of magnetism towards logics. Last but not least a progress in manipulation of single electron or nuclear spins in scalable solid-state devices can be envisaged, though a time scale in question is hard to predict.

Similarly to other branches of condensed matter physics, breakthrough achievements will be triggered by developments of new materials. Further progress in p-type doping and magnetic ion incorporation to standard semiconductors will make it possible to synthesize functional high temperature

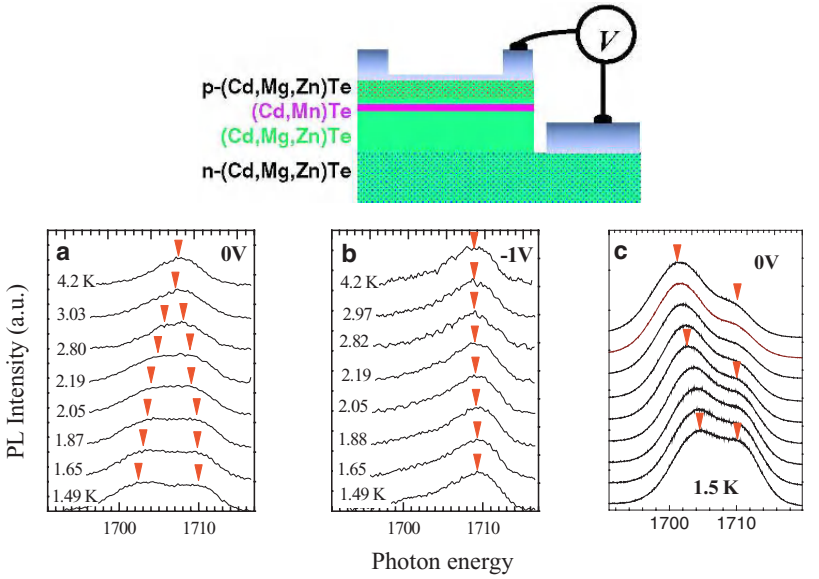

Fig. 15: Effect of temperature (**a**), bias voltage (**b**), and illumination (**c**) on photoluminescence of structure consisting of modulation doped p-(Cd,Mn)Te quantum well and n-type barrier. Zero-field line splitting (marked by arrows) witnesses the appearance of a ferromagnetic ordering (a) which does not show up if the quantum well is depleted from the holes by reverse bias of p-i-n diode (b). Low-temperature splitting is enhanced by additional illumination by white light (c), which increases hole concentration in quantum well (after Boukari et al. [161])

ferromagnetic DMS. At the same time, a control over ferromagnetic precipitates in various semiconductors will result in composite materials that will be useful as magnetooptical media and for high density memories. Particular attention will be paid to insulating ferrimagnetic oxides and nitrides, which could serve as spin selective barriers up to well above room temperature. Moreover, efforts will be undertaken to convert them into functional magnetic semiconductors by elaboration of purification methods and mastering doping protocols that will produce high mobility electrons and holes in these systems. Another line of research will be devoted to search for nonmagnetic barrier materials in which transmission coefficients could be electrically adjusted to optimize either reading or writing process in MRAM cells. Particularly prospective appear multiferroic systems, as in these multifunctional materials the coupling between magnetic and electric polarizations offers new device paradigms.

The important aspect of extensive studies on ferromagnetism in semiconductors discussed in this review is the demonstration of suitability of

empirically-constrained theoretical methods in quantitative description of a large body of thermodynamic, micromagnetic, transport, and optical properties of ferromagnetic semiconductors. In particular, a successful description of spintronic effects in both nonmagnetic and magnetic semiconductors is possible provided that all peculiarities of the host band structure, especially those associated with the spin-orbit interaction, are carefully taken into account. Indeed, as a result of such an effort (Ga,Mn)As has reached the status of the best understood ferromagnet. At the same time, research on DMS has disclosed shortcomings of today's computational materials science in predicting and elucidating magnetic properties of solids. It appears that this failure of ab initio methods (prediction of ferromagnetism in systems where it is absent and inability to explain its nature in materials where it does exist) originates from the co-existence of strong correlation with electronic and magnetic disorder in DMS. This calls for novel computation protocols that will be able to handle randomness and correlation on equal footing, also at non-zero temperatures, and will allow for the existence of electronic and/or chemical nanoscale phase separations. Such computational tools, together with advanced methods of spatially resolved material characterization, will in particular answer a persistently raised question on whether a high temperature ferromagnetism is possible in materials containing no magnetic ions.

With no doubt, in course of the years semiconductor spintronics has evolved into an important branch of today's materials science, condensed matter physics, and device engineering.

Acknowledgements

The author would like to thank his numerous Warsaw's co-workers and Hideo Ohno and his co-workers, as indicated in the reference list, for many years of stimulating discussions and fruitful collaboration.

References

1. C. Benett, D. DiVincenzo: Nature **404**, 247 (2000)
2. M. Horodecki, K. Horodecki, P. Horodecki, R. Horodecki, J. Oppenheim, A. Sen(De), U. Sen: Phys. Rev. Lett. **90**, 100 402 (2003)
3. J.M. Kikkawa, D.D. Awschalom, I.P. Smorchkova, N. Samarth:
4. M.N. Baibich, J.M. Broto, A. Fert, F. Nguyen Van Dau, F. Petroff, P. Eitenne, G. Creuzet, A. Friederich, J. Chazelas: Phys. Rev. Lett. **61**, 2472 (1988)
5. G. Binasch, P. Grünberg, F. Saurenbach, W. Zinn: Phys. Rev. **B 39**, 4828 (1989)
6. R.E. Camley, J. Barnaś: Phys. Rev. Lett. **63**, 664 (1989)
7. S. Parkin, C. Kaiser, A. Panchula, P. Rice, B.H.M. Samant, S.H. Yang: Nature Mat. **3**, 862 (2004)

8. S. Yuasa, T. Nagahama, A. Fukushima, Y. Suzuki, K. Ando: Nature Mat. **3**, 868 (2004)
9. S. Ikeda, J. Hayakawa, Y.M. Lee, R. Sasaki, T. Meguro, F. Matsukura, H. Ohno: Japan. J. Appl. Phys. **44**, L1442
10. J. Hayakawa, S. Ikeda, Y.M. Lee, R. Sasaki, T. Meguro, F. Matsukura, H.T.H. Ohno: Japan. J. Appl. Phys. **44**, L1267
11. X. Jiang, R. Wang, R.M. Shelby, R.M. Macfarlane, S.R. Bank, J.S. Harris, S.S.P. Parkin: Phys. Rev. Lett. **94**, 056 601
12. T. Dietl: "Diluted magnetic semiconductors", in *Handbook of Semiconductors*, ed. by S. Mahajan, Vol. 3B (North Holland, Amsterdam, 1994), p. 1251
13. T. Dietl: Acta Phys. Polon. **A 100**, 139 (2001)
14. T. Dietl: Semicond. Sci. Technol. **17**, 377 (2002)
15. T. Dietl: "III-V and II-VI Mn-based ferromagnetic semiconductors", in *Advances in Solid State Physics*, ed. by B. Kramer (Springer, Berlin, 2003), p. 413
16. T. Dietl: J. Phys.: Condens. Matter **16**, 5471 (2004)
17. T. Dietl: J. Mag. Mag Mat. **290-291**, 14 (2005)
18. T. Dietl: Proceedings 27th International Conference on Physics of Semiconductors, Flagstaff, USA, 2004 (AIP, Melville) p. 56 (2005)
19. W.G. van der Wiel, S. De Franceschi, J.M. Elzerman, T. Fujisawa, S. Tarucha, L.P. Kouwenhoven: Rev. Mod. Phys. **75**, 1 (2003)
20. I. Žutić, J. Fabian, S. Das Sarma: Rev. Mod. Phys. **76**, 323 (2004)
21. T. Jungwirth, J. Sinova, J. Mašek, J. Kučera, A. MacDonald: Rev. Mod. Phys **96**, in press (2006)
22. S.A. Wolf, D.D. Awschalom, R.A. Buhrman, J.M. Daughton, S. von Molnár, M.L. Roukes, A.Y. Chtchelkanova, D.M. Treger: Science **294**, 1488 (2001)
23. H. Ohno, F. Matsukura, Y. Ohno: JSAP International **5**, 4 (2002)
24. G. Lampel: Phys. Rev. Lett. **20**, 491 (1968)
25. In *Optical Orientation*, ed. by F. Meyer, B. Zacharchenya (North Holland, Amsterdam, 1986), Vol. 8 of Modern Problems in Condensed Matter Sciences
26. D. Awschalom, N. Samarth: "Optical manipulation, transport and storage of spin coherence in semiconductors", in *Semiconductor Spintronics and Quantum Computation*, ed. by D. Awschalom, D. Loss, N. Samarth (Springer, Berlin, 2002), pp. 147–194
27. S. Ganichev, W. Prettl: J. Phys.: Condens. Matter **15**, R935 (2003)
28. V. Sih, Y. Kato, D. Awschalom: Phys. World **18**, 33 (November 2005)
29. D. Loss, D.P. DiVincenzo: Phys. Rev. **A 57**, 120 (1998)
30. B. Kane: Nature **393**, 133 (1998)
31. B. Beschoten, E. Johnston-Halperin, D.K. Young, M. Poggio, J.E. Grimaldi, S. Keller, S.P. DenBaars, U.K. Mishra, E.L. Hu, D.D. Awschalom: Phys. Rev. **B 63**, 121 202 (2001)
32. T. Andrearczyk, J. Jaroszyński, G. Grabecki, T. Dietl, T. Fukumura, M. Kawasaki: Phys. Rev. **B. 72**, 121 309(R) (2005)
33. M. Sawicki, T. Dietl, J. Kossut, J. Igalson, T. Wojtowicz, W. Plesiewicz: Phys. Rev. Lett. **56**, 508 (1986)
34. Z. Wilamowski, W. Jantsch, N. Sandersfeld, M. Mühlberger, F. Schäffler, S. Lyon: Physica **E 16**, 111 (2003)
35. G. Grabecki, J. Wróbel, T. Dietl, E. Papis, E. Kamińska, A. Piotrowska, G. Springholz, G. Bauer: Physica E **13**, 649 (2002)

36. G. Grabecki, J. Wróbel, T. Dietl, E. Janik, M. Aleszkiewicz, E. Papis, E. Kamińska, A. Piotrowska, G. Springholz, G. Bauer: Phys. Rev. **B 72**, 125 332 (2005)
37. M. Johnson: Semicond. Sci. Technol. **17**, 298 (2002)
38. G. Prinz: Science **282**, 1660 (1998)
39. A. Hanbicki, O. van t Erve, R. Magno, G. Kioseoglou, C. Li, B. Jonker, G. Itskos, R. Mallory, M. Yasar, A. Petrou: Appl. Phys. Lett. **82**, 4092 (2003)
40. M. Zenger, J. Moser, W. Wegscheider, D. Weiss, T. Dietl: J. Appl. Phys. **96**, 2400 (2004)
41. S. Datta, B. Das: Appl. Phys. Lett. **56**, 665 (1990)
42. K. Yoh, M. Ferhat, S. Kashiwada, T. Tsuchiyam, A. Riposan, J. Mirecki-Millunchick: "Electrical characterization of an Fe/InGaAs spin FET", in *Abstracts – SpinTech III Conference, Awaji Island, Japan 2005* (unpublished, 2005)
43. A. Łusakowski, J. Wróbel, T. Dietl: Phys. Rev. B **68**, 081 201(R) (2003)
44. R. Winkler: Phys. Rev. **B 69**, 045 317 (2004)
45. J. Wróbel, T. Dietl, A. Łusakowski, G. Grabecki, K. Fronc, R. Hey, K. Ploog, H. Shtrikman: Phys. Rev. Lett. **93**, 246 601 (2004)
46. R. Gałązka: "Semimagnetic semiconductors", in *Proceedings 14th International Conference on Physics of Semiconductors, Edinburgh 1978*, ed. by B. Wilson (IoP, Bristol, 1978), p. 133
47. R. Nagaev: *Physics of Magnetic Semiconductors* (Mir, Moscow, 1983)
48. T. Dietl: "Semimagnetic semiconductors in high magnetic fields", in *Physics in High Magnetic Fields*, ed. by S. Chikazumi, N. Miura (Springer, Berlin)
49. J. Furdyna, J. Kossut: *Diluted Magnetic Semiconductors*, Vol. 25 of Semiconductor and Semimetals (Academic Press, New York, 1988)
50. G. Bauer, W. Pascher, W. Zawadzki: Semicond. Sci. Technol. **7**, 703 (1992)
51. F. Matsukura, H. Ohno, T. Dietl: Handbook of Magnetic Materials **14**, 1–87 (2002)
52. W. Prellier, A. Fouchet, B. Mercey: J. Phys.: Condens. Matter **15**, R1583 (2003)
53. H. Ohno: Science **281**, 951 (1998)
54. A. Haury, A. Wasiela, A. Arnoult, J. Cibert, S. Tatarenko, T. Dietl, Y. Merle d'Aubigné: Phys. Rev. Lett. **79**, 511 (1997)
55. D. Ferrand, J. Cibert, A. Wasiela, C. Bourgognon, S. Tatarenko, G. Fishman, T. Andrearczyk, J. Jaroszynski, S. Kolesnik, T. Dietl, B. Barbara, D. Dufeu: Phys. Rev. **B 63**, 085 201 (2001)
56. K. Wang, K. Edmonds, R. Campion, B. Gallagher, T. Foxon, M. Sawicki, T. Dietl, P. Boguslawski, T. Jungwirth: Flagstaff, Arizona, USA, July 2004, ed. J. Mendez, AIP Melville, 2005 p. 333 (2005)
57. J. Langer, C. Delerue, M. Lannoo, H. Heinrich: Phys. Rev. **B 38**, 7723 (1988)
58. A. Zunger: "Electronic structure of 3d transition-atom impurities in semiconductors", in *Solid State Physics*, ed. by F. Seitz, D. Turnbull, Vol. 39 (Academic Press, New York, 1986), pp. 275–464
59. Z. Wilamowski, K. Świątek, T. Dietl, J. Kossut: Solid State Commun. **74**, 833 (1990)
60. A. Komarov, S. Ryabchenko, O. Terletskii, I. Zheru, R. Ivanchuk: Sov. Phys. JETP **46**, 318 (1977)
61. J.A. Gaj, R.R. Gałązka, M. Nawrocki: Solid State Commun. **25**, 193 (1978)

62. G. Bastard, C. Rigaux, Y. Guldner, J. Mycielski, A. Mycielski: J. de Physique (Paris) **39**, 87 (1978)
63. M. Jaczyński, J. Kossut, R.R. Gałązka: Phys. Stat. Sol. (b) **88**, 73 (1978)
64. A.K. Bhattacharjee, G. Fishman, B. Coqblin: Physica **B+C 117-118**, 449 (1983)
65. P. Kacman: Semicond. Sci. Technol. **16**, R25 R39 (2001)
66. W. Mac, Nguyen The Khoi, A. Twardowski, J.A. Gaj, M. Demianiuk: Phys. Rev. Lett. **71**, 2327 (1993)
67. I.A. Merkulov, D.R. Yakovlev, A. Keller, J.G. W. Ossau, A. Waag, G. Landwehr, G. Karczewski, T. Wojtowicz, J. Kossut: Phys. Rev. Lett. **83**, 1431 (1999)
68. J. Sadowski, H. Mariette, A. Wasiela, R. André, Y. Merle d'Aubigné, T. Dietl: Phys. Rev. **B 56**, 1664(R) (1997)
69. J.A. Gaj, W. Grieshaber, C. Bodin-Deshayes, J. Cibert, G. Feuillet, Y. Merle d'Aubigné, A. Wasiela: Phys. Rev. **B 50**, 5512 (1994)
70. I. Smorchkova, N. Samarth, J. Kikkawa, D. Awschalom: Phys. Rev. Lett. **78**, 3571 (1997)
71. J. Jaroszyński, T. Andrearczyk, G. Karczewski, J. Wróbel, T. Wojtowicz, D. Popoviæ, T. Dietl: cond-mat/0509189
72. J. Jaroszyński, J. Wróbel, M. Sawicki, E. Kamińska, T.Skośkiewicz, G. Karczewski, T. Wojtowicz, A. Piotrowska, J.Kossut, T. Dietl: Phys. Rev. Lett. **75**, 3170 (1995)
73. T. Wojtowicz, T. Dietl, M. Sawicki, W. Plesiewicz, J. Jaroszyński: Phys. Rev. Lett. **56**, 2419 (1986)
74. J. Jaroszyński, G. Karczewski, J. Wróbel, T. Andrearczyk, T.W. A. Strycharczuk, G. Grabecki, E. Papis, E. Kamińska, A. Piotrowska, T. Dietl: Physica E **6**, 790 (2000)
75. J. Jaroszyński, T. Andrearczyk, G. Karczewski, J. Wróbel, T. Wojtowicz, E. Papis, E. Kamińska, A. Piotrowska, D. Popovic, T. Dietl: Phys. Rev. Lett. **89**, 266 802 (2002)
76. N. Brandt, V. Moshchalkov: Adv. Phys. **33**, 193 (1984)
77. M. Oestreich, J. Hübner, D. Hägele, P.J. Klar, W. Heimbrodt, W.W. Rühle, D.E. Ashenford, B. Lunn: Appl. Phys. Lett. **74**, 1251 (1999)
78. R. Fiederling, M. Keim, G. Reuscher, W. Ossau, G. Schmidt, A. Waag, L.W. Molenkamp: Nature **402**, 787 (1999)
79. A. Slobodskyy, C. Gould, T. Slobodskyy, C.R. Becker, G. Schmidt, L.W. Molenkamp: Phys. Rev. Lett. **90**, 246 601 (2003)
80. T. Dietl, M. Sawicki, E. Isaacs, M. Dahl, D. Heiman, M. Graf, S. Gubarev, D.L. Alov: Phys. Rev. B **43**, 3154 (1991)
81. J. Jaroszyński, J. Wróbel, G. Karczewski, T. Wojtowicz, T. Dietl: Phys. Rev. Lett. **80**, 5635 (1998)
82. H. Krenn, K. Kaltenegger, T. Dietl, J. Spałek, G. Bauer: Phys. Rev. **B 39**, 10 918 (1989)
83. C. Benoit à la Guillaume: Phys. Stat. Solidi (b) **175**, 369 (1993)
84. T. Dietl: Acta Phys. Polon. A **94**, 111 (1998)
85. T. Dietl, P. Peyla, W. Grieshaber, Y. Merle d'Aubigné: Phys. Rev. Lett. **74**, 474 (1995)
86. T. Dietl, J. Spałek: Phys. Rev. Lett. **48**, 355 (1982)
87. T. Dietl, J. Spałek: Phys. Rev. **B 28**, 1548 (1983)

88. T. Dietl: J. Mag. Mag. Mat. **38**, 34 (1983)
89. J.A. Gaj, R. Planel, G. Fishman: Solid State Commun. **29**, 435 (1979)
90. R. Gałązka: J. Magn. Magn. Mat. **140-144**, 13 (1995)
91. T.M. Giebułtowicz, N. Samarth, H. Luo, J.K. Furdyna, P. Kłosowski, J.J. Rhyne: Phys. Rev. B **46**, 12 076(R) (1992)
92. H. Ohno, H. Munekata, T. Penney, S. von Molnár, L.L. Chang: Phys. Rev. Lett. **68**, 2664 (1992)
93. H. Ohno, A. Shen, F. Matsukura, A. Oiwa, A. Endo, S. Katsumoto, Y. Iye: Appl. Phys. Lett. **69**, 363 (1996)
94. T. Dietl, A. Haury, Y. Merle d'Aubigné: Phys. Rev. **B 55**, 3347(R) (1997)
95. F. Matsukura, H. Ohno, A. Shen, Y. Sugawara: Phys. Rev. **B 57**, R2037 (1998)
96. T. Jungwirth, W. Atkinson, B. Lee, A. MacDonald: Phys. Rev. **B 59**, 9818 (1999)
97. T. Story, R.R. Gałązka, R.B. Frankel, P.A. Wolff: Phys. Rev. Lett. **56**, 777 (1986)
98. T. Dietl, H. Ohno, F. Matsukura, J. Cibert, D. Ferrand: Science **287**, 1019 (2000)
99. T. Dietl, H. Ohno, F. Matsukura: Phys. Rev. **B 63**, 195 205 (2001)
100. H. Kępa, Le Van Khoi, C. Brown, M. Sawicki, J. Furdyna, T. Giebułtowicz, T. Dietl: Phys. Rev. Lett. **91**, 087 205 (2003)
101. L. Świerkowski, T. Dietl: Acta Phys. Polon. **A 73**, 431 (1988)
102. F. Fröhlich, F. Nabarro: Proc. Roy. Soc. London **A 175**, 382 (1940)
103. T. Andrearczyk, J. Jaroszyński, M. Sawicki, Le Van Khoi, T. Dietl, D. Ferrand, C. Bourgognon, J. Cibert, S. Tatarenko, T. Fukumura, Z. Jin, H. Koinuma, M. Kawasaki: in *Proceedings 25th International Conference on Physics of Semiconductors, Osaka, Japan, 2000*, ed. by N. Miura, T. Ando (Spriger, Berlin)
104. T. Jungwirth, K. Wang, J. Mašek, K. Edmonds, J. König, J. Sinova, M. Polini, N. Goncharuk, A. MacDonald, M. Sawicki, R. Campion, L. Zhao, C. Foxon, B. Gallagher: Phys. Rev. **B 72**, 165 204
105. D. Kechrakos, N. Papanikolaou, K. Trohidou, T. Dietl: Phys. Rev. Lett **94**, 127 201
106. K.M. Yu, W. Walukiewicz, T. Wojtowicz, I. Kuryliszyn, X. Liu, Y. Sasaki, J.K. Furdyna: Phys. Rev. **B 65**, 201 303(R) (2002)
107. K. Edmonds, P. Boguslawski, K. Wang, R. Campion, N. Farley, B. Gallagher, C. Foxon, M. Sawicki, T. Dietl, M. Nardelli, J. Bernholc: Phys. Rev. Lett. **92**, 037 201 (2004)
108. H. Munekata, A. Zaslavsky, P. Fumagalli, R.J. Gambino: Appl. Phys. Lett. **63**, 2929 (1993)
109. A. Shen, H. Ohno, F. Matsukura, Y. Sugawara, N. Akiba, T. Kuroiwa, A. Oiwa, A. Endo, S. Katsumoto, Y. Iye: J. Cryst. Growth **175/176**, 1069 (1997)
110. M. Abolfath, T. Jungwirth, J. Brum, A. MacDonald: Phys. Rev. **B 63**, 054 418 (2001)
111. M. Sawicki, K.Y. Wang, K.W. Edmonds, R. Campion, C. Staddon, N. Farley, C. Foxon, E. Papis, E. Kamińska, A. Piotrowska, T. Dietl, B. Gallagher: Phys. Rev. **B 71**, 121 302 (2005)

112. M. Sawicki, K.Y. Wang, K.W. Edmonds, R.P. Campion, N.R.S.F. C. R. Staddon, C.T. Foxon, E. Papis, E. Kamińska, A. Piotrowska, T. Dietl, B.L. Gallagher: Phys. Rev. **B 71**, 121 302(R)
113. K.Y. Wang, M. Sawicki, K. Edmonds, R. Campion, S. Maat, C. Foxon, B. Gallagher, T. Dietl: Phys. Rev. Lett. **95** (2005)
114. U. Welp, V.K. Vlasko-Vlasov, A. Menzel, H.D. You, X. Liu, J.K. Furdyna, T. Wojtowicz: Appl. Phys. Lett. **85**, 260 (2004)
115. J. König, T. Jungwirth, A. MacDonald: Phys. Rev. **B 64**, 184 423 (2001)
116. W. Szymańska, T. Dietl: J. Phys. Chem. Solids **39**, 1025 (1978)
117. T. Shono, T. Hasegawa, T. Fukumura, F. Matsukura, H. Ohno: Appl. Phys. Lett. **77**, 1363 (2000)
118. T. Dietl, J. König, A.H. MacDonald: Phys. Rev. **B 64**, 241 201 (2001)
119. K. Ando, T. Hayashi, M. Tanaka, A. Twardowski: J. Appl. Phys. **83**, 6548 (1998)
120. J. Szczytko, W. Mac, A. Twardowski, F. Matsukura, H. Ohno: Phys. Rev. **B 59**, 12 935 (1999)
121. B. Beschoten, P. Crowell, I. Malajovich, D. Awschalom, F. Fumihiro, A. Shen, H. Ohno: Phys. Rev. Lett. **83**, 3073 (1999)
122. J. Szczytko, W. Bardyszewski, A. Twardowski: Phys. Rev. **B 64**, 075 306 (2001)
123. J. Sinova, T. Jungwirth, S.R.E. Yang, J. Kučera, A. MacDonald: Phys. Rev. **B 66**, 041 202 (2002)
124. E.M. Hankiewicz, T. Jungwirth, T. Dietl, C. Timm, J. Sinova: Phys. Rev. **B 70**, 245 211 (2004)
125. Y. Nagai, T. Junimoto, K. Ngasaka, H. Nojiri, M. Motokawa, F. Matsujura, T. Dietl, H. Ohno: Jpn. J. Appl. Phys. **40**, 6231 (2001)
126. E.J. Singley, K.S. Burch, R. Kawakami, J. Stephens, D.D. Awschalom, D.N. Basov: Phys. Rev. **B 68**, 165 204 (2003)
127. L. Chien, C. Westgate: *The Hall Effect and Its Applications* (Plenum, New York, 1980)
128. J.M. Luttinger: Phys. Rev. **112**, 739 (1958)
129. T. Jungwirth, Q. Niu, A. MacDonald: Phys. Rev. Lett. **88**, 207 208 (2002)
130. P. Nozieres, C. Lewiner: Le Journal de Physique **34**, 901 (1973)
131. J.N. Chazalviel: Phys. Rev. **B 11**, 3918 (1975)
132. P. Oppener: "Ferromagnetic semiconductors", in *Handbook of Magnetic Materials*, ed. by K. Buschow, Vol. 13 (Elsevier, Amsterdam, 2001), p. 229
133. T. Dietl, F. Matsukura, H. Ohno, J. Cibert, D. Ferrand: "Hall effect and magnetoresistance in p-type ferromagnetic semiconductors", in *Recent Trends in Theory of Physical Phenomena in High Magnetic Fields*, ed. by I. Vagner (Kluwer, Dordrecht, 2003), p. 197
134. P. Leroux-Hugon, A. Ghazali: J. Phys. C: Solid State Phys. **5**, 1072 (1972)
135. J.N. Chazalviel: Phys. Rev. **B 10**, 3018 (1974)
136. V.K. Dugaev, A. Crépieux, P. Bruno: Phys. Rev. **B 64**, 104 411 (2001)
137. T. Omiya, F. Matsukura, T. Dietl, Y. Ohno, T. Sakon, M. Motokawa, H. Ohno: Physica **E 7**, 976 (2000)
138. D. Ferrand, J. Cibert, A. Wasiela, C. Bourgognon, S. Tatarenko, G. Fishman, S. Koleśnik, J. Jaroszyński, T. Dietl, B. Barbara, D. Dufeu: J. Appl. Phys. **87**, 6451 (2000)
139. H. Tang, R. Kawakami, D. Awschalom, M. Roukes: Phys. Rev. Lett. **90**, 107 201 (2003)

140. T. Jungwirth, M. Abolfath, J. Sinova, J. Kučera, A. MacDonald: Appl. Phys. Lett. **81**, 4029 (2002)
141. F. Matsukura, M. Sawicki, T. Dietl, D. Chiba, H. Ohno: Physica **E 21**, 1032 (2004)
142. T. Jungwirth, J. Sinova, K. Wang, K.W. Edmonds, R. Campion, B. Gallagher, C. Foxon, Q. Niu, A. MacDonald: Appl. Phys. Lett. **83**, 320 (2003)
143. A. Kawabata: Solid State Commun. **34**, 432 (1980)
144. Y. Ohno, D.K. Young, B. Beschoten, F. Matsukura, H. Ohno, D.D. Awschalom: Nature **402**, 790 (1999)
145. K. Young, J.A. Gupta, E. Johnston-Halperin, R. Epstein, Y. Kato, D.D. Awschalom: Semicond. Sci. Technol. **17**, 275 (2002)
146. M. Kohda, Y. Ohno, K. Takamura, F. Matsukura, H. Ohno: Jpn. J. Appl. Phys. **40**, L1274 (2001)
147. E. Johnston-Halperin, D. Lofgreen, R. Kawakami, D. Young, L. Coldren, A. Gossard, D. Awschalom: Phys. Rev. B **65**, 041 306(R) (2002)
148. D. Chiba, N. Akiba, F. Matsukura, Y. Ohno, H. Ohno: Appl. Phys. Lett. **77**, 1873 (2000)
149. M. Tanaka, Y. Higo: Phys. Rev. Lett. **87**, 026 602 (2001)
150. R. Mattana, J.M. George, H. Jaffrès, F.N.V. Dau, A. Fert, B. Lépine, A. Guivarc'h, G. Jézéquel: Phys. Rev. Lett. **90**, 166 601 (2003)
151. D. Chiba, F. Matsukura, H. Ohno: Physica **E 21**, 966 (2004)
152. C. Ruster, C. Gould, T. Jungwirth, E. Girgis, G.M. Schott, R. Giraud, K. Brunner, G. Schmidt, L.W. Molenkamp: J. Appl. Phys. **97**, 10C506 (2005)
153. A. Giddings, M. Khalid, T. Jungwirth, J. Wunderlich, S. Yasin, R. Campion, K. Edmonds, J. Sinova, K. Ito, K.Y. Wang, D. Williams, B. Gallagher, C. Foxon: Phys. Rev. Lett. **94**, 127 202 (2005)
154. H. Tang, S. Masmanidis, R. Kawakami, D. Awschalom, M. Roukes: Nature **431**, 52 (2004)
155. D. Chiba, M. Yamanouchi, F. Matsukura, T. Dietl, H. Ohno: Phys. Rev. Lett. **96**, 096602 (2006)
156. P. Van Dorpe, W. Van Roy, J. De Boeck, G. Borghs, P. Sankowski, P. Kacman, J.A. Majewski, T. Dietl: Phys. Rev. B **72**, 205 322 (2005)
157. P. Sankowski, P. Kacman, J. Majewski, T. Dietl: cond-mat/0607206, Phys. Rev. B, submitted
158. P. Van Dorpe, W. Van Roy, V. Motsnyi, M. Sawicki, G. Borghs, J. De Boeck: Appl. Phys. Lett. **84**, 3495 (2004)
159. S. Koshihara, A. Oiwa, M. Hirasawa, S. Katsumoto, Y. Iye, C. Urano, H. Takagi, H. Munekata: Phys. Rev. Lett. **78**, 4617 (1997)
160. H. Ohno, D. Chiba, F. Matsukura, T. Omiya, E. Abe, T. Dietl, Y. Ohno, K. Ohtani: Nature **408**, 944 (2000)
161. H. Boukari, P. Kossacki, M. Bertolini, D. Ferrand, J. Cibert, S. Tatarenko, A. Wasiela, J.A. Gaj, T. Dietl: Phys. Rev. Lett. **88**, 207 204 (2002)
162. L. Berger: J. Appl. Phys. **55**, 1954 (1984)
163. J.C. Slonczewski: J. Magn. Magn. Mater. **159**, L1 (1996)
164. D. Chiba, Y. Sato, T. Kita, F. Matsukura, H. Ohno: Phys. Rev. Lett. **93**, 216 602 (2004)
165. M. Yamanouchi, D. Chiba, F. Matsukura, H. Ohno: Nature **428**, 539 (2004)
166. M. Yamanouchi, D. Chiba, F. Matsukura, T. Dietl, H. Ohno: Phys. Rev. Lett. **96**, 096601 (2006)

Lectures on the Spin Pairing Mechanism in High-Temperature Superconductors

E. Arrigoni

Institute of Theoretical Physics and Computational Physics,
Graz University of Technology, Petersgasse 16, 8010 Graz, Austria
arrigoni@itp.tu-graz.ac.at

Abstract. The aim of these lectures is to present a brief introduction, appropriate for non-experts in the field, on the physics of high-temperature superconductors, focusing in particular on their pairing mechanism. These notes are a summary of the short lecture held at the Schladming Winter School, and are certainly not a complete review on the topic. For a more complete and appropriate overview, the reader is referred to the extended literature on the subject.

After a short introduction about "conventional" superconductors and about the BCS phonon-mediated pairing mechanism, I will present one of the models which is commonly used to describe the physics of these materials, namely, the Hubbard model. Here, I will show how the Pauli principle naturally leads to the superexchange mechanism, i.e. to an antiferromagnetic coupling between electron spins on neighboring orbitals. In undoped compounds this coupling leads to an insulating antiferromagnetic phase. Upon doping, charge carriers (holes) are introduced into the copper-oxide layers and destroy long-range antiferromagnetism which is replaced by short-range antiferromagnetic fluctuations. In the last part of these lectures I will show how these magnetic fluctuations can lead to an effective attraction between charge carriers and to pairing, which is the main ingredient of superconductivity.

1 Introduction

Superconductivity was first discovered 1911 in mercury by H. Kamerling Omnes at Leiden University. However, the theoretical explanation for this phenomenon was given only 46 years later by Bardeen, Cooper and Schrieffer [1]. In order to become superconductor, Mercury was cooled down below a critical temperature $T_c \approx 4K$. In the years following Omnes' discovery, superconductivity was observed in a number of additional elements and compounds. However, until about 20 years ago, the maximum T_c obtained was around $20K$ (Fig. 2). In 1986, Bednorz and Müller from IBM in Zürich broke this barrier and first observed superconductivity with $T_c \approx 35K$ in a new compound, LaBaCuO [2]. A few months later P. Chu from the University of Huston, reported superconductivity at $T_c \approx 90K$ in YBaCuO. In few years, a large number of new superconducting materials were discovered, with T_c 's above the melting point of liquid nitrogen (Fig. 2).

It was quickly realized that these new materials were quite different from the binary and ternary "conventional" superconductors found before 1986.

E. Arrigoni: *Lectures on the Spin Pairing Mechanism in High-Temperature Superconductors*, Lect. Notes Phys. **712**, 47–65 (2007)
DOI 10.1007/3-540-38592-4_2

In particular, it was very soon clear that the superconducting mechanism was different. Up to now, there is not yet a complete agreement about the mechanism for these High-Temperature Superconductors (HTSC).

In these lectures, I will first present a short introduction about the "conventional" superconducting mechanism and then present a simple description of one of the most widely accepted theories for the HTSC mechanism, namely the one based on spin-mediated pairing.

2 Superconductivity

The term superconductivity is related to dissipationless current transport, which is reached upon cooling below the superconducting transition temperature. However, this is not the whole story. One additional important property of superconductors is the Meissner effect, namely, the fact that magnetic-field lines are (at least partially) expelled from the bulk of a superconductor. This effect can be observed by putting a magnet on a piece of superconducting YBaCuO: the repulsion of the magnetic lines holds the magnet floating on top of the superconductor, see Fig. 1.

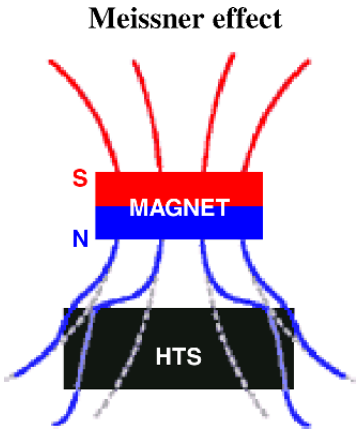

Fig. 1: Meissner effect. Magnetic-field lines are expelled from the bulk of a superconductor. Due to this effect, a magnet levitates on top of a superconductor

What is the microscopic origin of such a dissipationless state? Figures 3 and 4 schematically show how the pairing between charge carriers (electrons) taking place in a superconductor, produce a "rigid" (coherent) state consisting of a coherent superposition of electron pairs (Cooper pairs). Current-

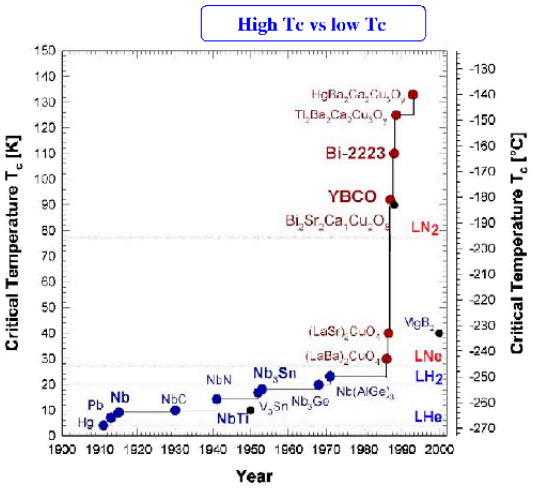

Fig. 2: The discovery of superconductivity: T_c versus years

dissipative scattering processes, e.g. with impurities, cannot take place, because they first need to break a Cooper pair, which costs an energy of the order of the Cooper-pair binding energy $\Delta(\propto K_B T_c)$.

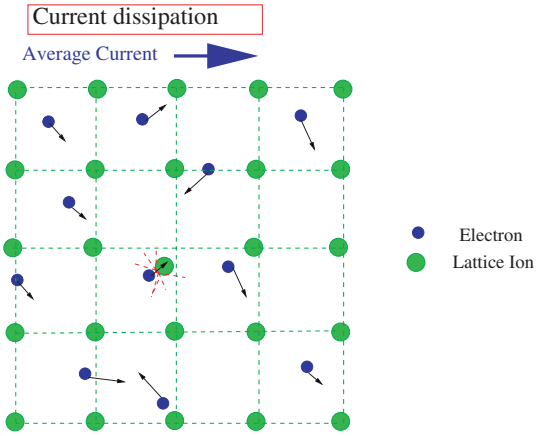

Fig. 3: Current dissipation in a non-superconducting state

The fundamental question is, how can electrons, which normally repel each other due to the Coulomb interaction, form bound pairs? As explained in Sect. 3, in conventional superconductors, the attractive force is provided by so-called phonons, i.e. by the lattice which is made of positively charged

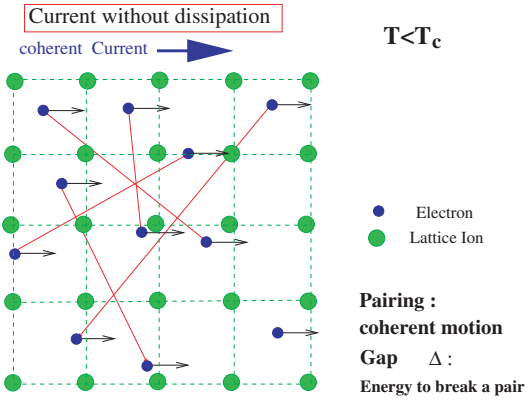

Fig. 4: Superconducting state made of a rigid coherent state of Cooper pairs

ions oscillating around their equilibrium position. On the other hand, the pairing mechanism is probably different in HTSC, as discussed in Sects. 5 and 6. One of the most widespread idea, which will be discussed in Sect. 6 is that pairing in HTSC is mediated by magnetic fluctuations.

3 Phonon-Mediated Effective Attraction between Electrons

Let us first consider the simple picture of Fig. 5. Here the lattice is schematically represented as an elastic layer, while electrons are "billiard balls" which locally deform the layer with their "weight". A second ball is attracted by the deformation produced by the first one, so that putting two balls close together reduces the elastic deformation energy. The elastic layer, thus, mediates an effective attraction between the balls.

A more realistic picture is shown in Fig. 6. Here, a negatively charged electron locally polarizes the positively charged lattice. Since ions are heavier than electrons, the local lattice distortion and associated positive polarization survives for some time after the electron has gone away (retardation effect). A second electron then feels the attraction of the local positive polarization and follows the path of the first electron. This picture already shows how this relatively weak "elastic" effect can overwhelm the repulsive Coulomb interaction thanks to the retardation effect.

We want now to describe the effective attraction between electrons as an interaction mediated by the exchange of "elastic waves" (see Fig. 5). These waves are the quanta of elastic lattice deformation and are referred to as phonons because they are responsible for the transmission of sound in solids. Phonons play a role similar to "mesons" in high-energy physics by mediating a force between electrons.

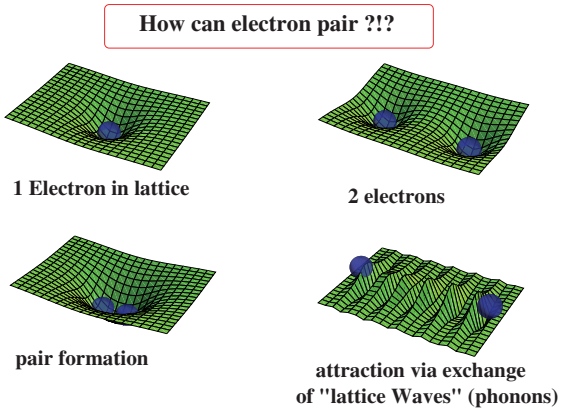

Fig. 5: Simple picture of the attraction between two "particles" due to a minimization of the elastic energy

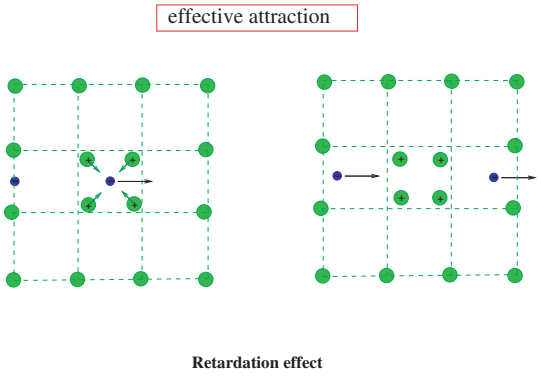

Fig. 6: Illustration of the retardation effect

As a starting point for this description we shall adopt a tight-binding Hamiltonian, containing creation ($c^{\dagger}_{R,\sigma}$) and destruction ($c_{R,\sigma}$) operators for an electron at lattice site R with spin projection σ. These lattice electrons are coupled via their density $n(R) \equiv \sum_{\sigma} c^{\dagger}_{R,\sigma} c_{R,\sigma}$ to a local elastic displacement (phonon) described by the field $u(R)$ with conjugate momentum $P(R)$. The total Hamiltonian for this system is given by

$$\sum_{R,R'} T(R-R')\; c^{\dagger}_{R,\sigma} c_{R',\sigma} + g \sum_{R} \sqrt{\Omega}\; n(R)\; u(R) + \sum_{R} \frac{1}{2} \left[P(R)^2 + \Omega^2 u(R)^2 \right] \; . \tag{1}$$

Here, $T(R-R')$ is the hopping matrix element for electrons, g is the electron-phonon coupling constant, and Ω is the phonon frequency. For simplicity, we have considered optical (i.e. dispersionless) phonons, and we have

conveniently absorbed the ionic mass M into the displacement field u, its momentum P, and g.

In the so-called antiadiabatic limit, one can neglect the P term, so that one can complete the square in the last two terms of Eq. (1):

$$g\sum_R n(R)\sqrt{\Omega}u(R) + \sum_R \frac{1}{2}\Omega^2 u(R)^2 = \sum_R \frac{1}{2}\Omega^2 \bar{u}(R)^2 - \sum_R \frac{1}{2}\frac{g^2}{\Omega}n(R)^2 \quad (2)$$

with

$$\bar{u} \equiv u + \frac{gn(R)}{\Omega} \; . \quad (3)$$

The new phonon field $\bar{u}$ is now decoupled from the electrons, which, however, feel an effective *attractive* interaction

$$V_{eff,ph} = -\frac{1}{2}\sum_R \frac{g^2}{\Omega}n(R)^2 \; . \quad (4)$$

From the discussion above, it is clear that the antiadiabatic limit is unphysical: it means that phonons move much faster than electrons, which is precisely the opposite to what happens in reality. For this reason, we need to introduce *retardation* effects. This is done via the introduction of a Matsubara-frequency (ω_λ) term, which modifies Eq. (4) to

$$V_{eff,ph} = -\frac{1}{2}\sum_R \frac{g^2\ \Omega}{\omega_\lambda^2 + \Omega^2}n(R)^2 \; . \quad (5)$$

In momentum space this becomes

$$-\frac{1}{2N}\sum_{k,k',q,\sigma,\sigma'} g_q^2\ D(q,\omega_\lambda)\ c^\dagger_{k+q,\sigma}c^\dagger_{k'-q,\sigma'}c_{k',\sigma'}c_{k,\sigma} \; , \quad (6)$$

where N is the number of lattice sites, the electron operators should be now considered as Fermi (Grassmann) fields with an implicit Matsubara-frequency dependence, and we have introduced the phonon propagator

$$D(q,\omega_\lambda) \equiv \frac{2\ \Omega_q}{\omega_\lambda^2 + \Omega_q^2} \; . \quad (7)$$

Here, we have introduced a q dependence of the phonon frequency Ω and of the coupling constant g.

The expression Eq. (6) can be also obtained via a diagrammatic expansion, which is represented in Fig. 7:

As discussed above, the retardation term ω_λ is important when taking into account the Coulomb *repulsion* between the electrons. The latter is much stronger than the attraction Eq. (6). However, fortunately pairing can take advantage of the retardation of the phonon-mediated attractive term (as already discussed in Fig. 6).

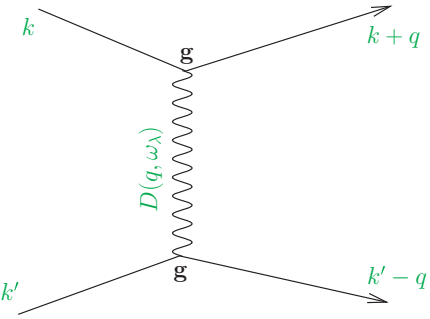

Fig. 7: Diagrammatic representation of the effective phonon-mediated electron-electron interaction. *Solid lines* are ingoing and outgoing electron lines (whose propagators don't contribute to the diagram). The *wavy line* is the phononic propagator $D(q, \omega_\lambda)$

4 BCS Theory

How does one exploit the attractive term Eq. (6) to get the superconducting state? Within the theory developed by Bardeen, Cooper and Schrieffer [1], which was termed accordingly BCS theory, one takes into account retardation effects by restricting the effective attraction Eq. (6) to particles within a typical phonon energy $\hbar\Omega_D$ around the Fermi energy. For larger energies ($\omega_\lambda \gg \Omega_d$), the interaction term (originating from the propagator Eq. (7)) is cut off. This is equivalent to say that higher electronic energies correspond to time scales that are too fast to see the phononic-mediated attraction.

After having taken for granted that in this energy range the attractive part of the interaction "wins" over the repulsion, as suggested above (this is proven in more detail within Eliashberg's theory [3]), one can consider the purely fermionic attractive model:

$$H = \sum_{k,\sigma} \xi_k \; c^\dagger_{k,\sigma} c_{k,\sigma} - \frac{1}{2N} \sum_{k,k',q,\sigma,\sigma'} V(q) c^\dagger_{k+q,\sigma} c^\dagger_{k'-q,\sigma'} c_{k',\sigma'} c_{k,\sigma} \; , \qquad (8)$$

where $-V(q)$ is the (attractive) electron-electron interaction, and the fermionic (electron) operators are written in momentum space, in which the "kinetic"-energy term $\xi_k = \epsilon_k - \mu$ (here, the chemical potential μ is included) is diagonal.

This model is solved by decoupling the interaction part in an unusual way, namely, one assumes that the operator $c_{k',\sigma'}\; c_{k,\sigma}$ has a finite expectation value ϕ_k for $k' = -k$ and $\sigma' = -\sigma$:

$$\phi_k \equiv \langle c_{-k,\downarrow} \; c_{k,\uparrow} \rangle \; . \qquad (9)$$

Equation (8) then yields the following form

$$H = \sum_{k,\sigma} \xi_k \; c^\dagger_{k,\sigma} c_{k,\sigma} - \frac{1}{N} \sum_{k,q,\sigma} V(q) \left[c^\dagger_{k+q,\uparrow} c^\dagger_{-k-q,\downarrow} \phi_k + H.C. \right] . \qquad (10)$$

In order to avoid introducing Bogolubov transformations, we carry out a particle-hole transformation for the spin-down particles $c^\dagger_{-k,\downarrow} = d_{k,\downarrow}$, and obtain the following expression (in matrix notation)

$$H = \sum_k \left(c^\dagger_{k,\uparrow}, d^\dagger_{k,\downarrow} \right) \begin{pmatrix} \xi(k) & -\Delta^*_k \\ -\Delta_k & -\xi_k \end{pmatrix} \begin{pmatrix} c_{k,\uparrow} \\ d_{k,\downarrow} \end{pmatrix} + \text{const.} , \qquad (11)$$

where

$$\Delta_k \equiv \frac{1}{N} \sum_q \phi_{k-q} V(q) . \qquad (12)$$

The eigenvalues of the matrix in Eq. (11) are

$$\pm E(k) \equiv \pm \sqrt{\xi_k^2 + \Delta_k^2} . \qquad (13)$$

The ground-state energy of Eq. (11) is obtained by summing over negative eigenvalues, yielding $E_{GS} = -\sum_k E(k)$. Evaluating the expectation value

$$\langle d^\dagger_{k,\downarrow} c_{k,\uparrow} \rangle = \phi_k = -\frac{\partial E_{GS}}{\partial \Delta_k} = \frac{\Delta_k}{E(k)} \qquad (14)$$

we obtain the BCS self-consistent gap equation [1]

$$\Delta_k = \frac{1}{N} \sum_{k'} V(k-k') \frac{\Delta_{k'}}{E_{k'}} . \qquad (15)$$

Retardation effects included in Eq. (6) can be considered in a rough way by setting $V(k,k')$ to a constant V whenever the electron momenta k and k' lie within an energy shell δ_D of width $\hbar\Omega_D$ from the FS. From Eq. (12) one can then approximately take $\Delta_k = \Delta$ independent of k in this region. In this way, one can rewrite the self-consistent equation Eq. (15) as

$$\frac{1}{V} = \frac{1}{N} \sum_k^{\delta_D} \frac{1}{\sqrt{\xi(k)^2 + \Delta^2}} \approx \frac{n_0}{2} \int_0^{\hbar\Omega_D} d\xi \; \frac{1}{\sqrt{\xi^2 + \Delta^2}} \approx n_0 \; \log \frac{\hbar\Omega_D}{\Delta} , \qquad (16)$$

where n_0 is the density of states at the Fermi level (summed over spins). The solution for $n_0 V \ll 1$ can be written as

$$\Delta \approx \hbar\Omega_D e^{-\frac{1}{n_0 V}} . \qquad (17)$$

It is clear from Eq. (16) that a repulsive interaction ($V < 0$) only gives a solution with a vanishing gap $\Delta = 0$. We will see below for the case of High-T_c superconductors that this is not always true.

An important effect contained in the solution Eq. (17) is the fact that the gap Δ is proportional to the phononic frequency $\Omega_D \propto 1/\sqrt{M_{Ion}}$, where M_{Ion} is the ionic mass. Since the transition temperature T_c is in turn proportional to Δ, one obtains the well-known *isotope effect*. This means that by replacing the nuclei by isotopes with different masses, T_c becomes proportional to the inverse squared isotopic mass. This experimental fact was one of the key observation suggesting that phonons could be at the origin of the superconducting mechanism. In HTSC the isotope effect is much smaller, especially at the doping where T_c is maximum (top of the "dome" in Fig. 8), thus questioning a possible phononic (-only) mechanism in these materials.

5 High-Temperature Superconductors

High-temperature superconducting materials have, in contrast to conventional superconductors, a quite complicated crystal structure (Fig. 9). Moreover, these materials, when undoped, are in fact insulators. A metallic (and superconducting) state is obtained upon replacing some of the rare earth (e. g. La with Sr in LaCuO) or adding oxygen (in YBaCuO). There is an optimal value of the doping x for which the superconducting T_c is maximum (Fig. 8).

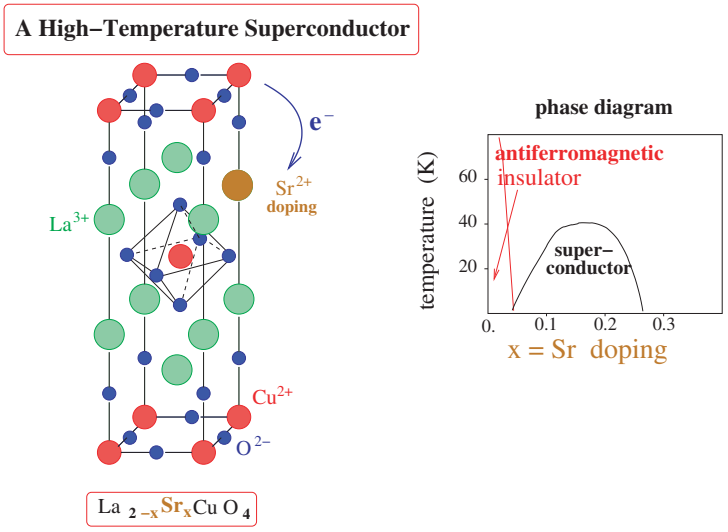

Fig. 8: Crystal structure of $La_{2-x}Sr_xCuO_4$, and the effect of replacing (doping) La with Sr. Phase diagram as a function of doping x and temperature T

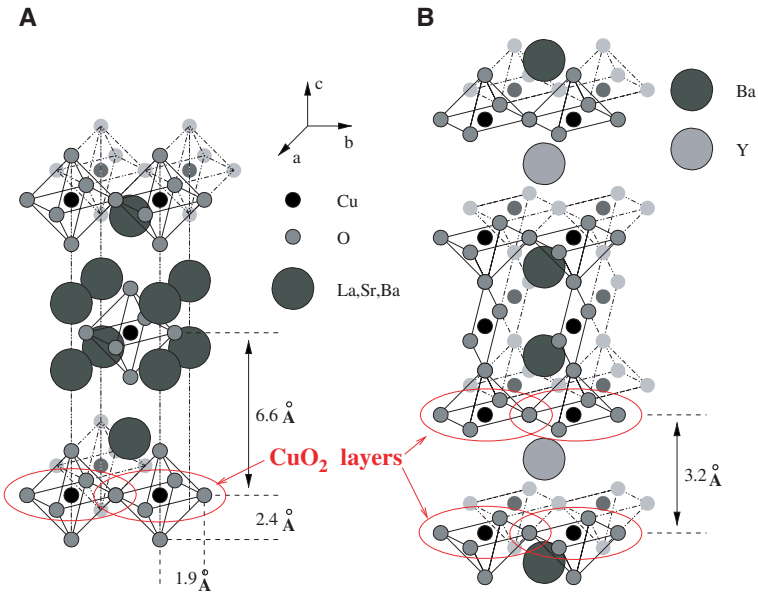

Fig. 9: Crystal structure of La_2CuO_4 (**A**) and $YBa_2Cu_3O_7$ (**B**)

Despite the complicated structure, it has soon become clear that the important electron dynamics takes place on a subset of this structure, namely on the ubiquitous copper-oxide layers. Moreover, only three orbitals per unit cell are important to describe the electronic structure, namely the Cu-$d_{x^2-y^2}$ and the O p_x and p_y orbitals (Fig. 10).

The electron dynamics in these orbitals is described by the three-band (Emery) model [4]

$$\begin{aligned} H_{3b} = & -t_{pd} \sum_{\delta=x,y} \sum_{\langle RR' \rangle \sigma} \alpha_{RR'\delta} \left(p^\dagger_{\delta R\sigma} d_{R'\sigma} + H.C. \right) \\ & -t_{pp} \sum_{\langle\langle RR' \rangle\rangle \sigma} \alpha'_{RR'} \left(p^\dagger_{xR\sigma} p_{yR'\sigma} + H.C. \right) \\ & + \Delta \sum_{R\sigma} \sum_{\delta=x,y} p^\dagger_{\delta R\sigma} p_{\delta R\sigma} + \frac{U_d}{2} \sum_{R\sigma} n^d_{R\sigma} n^d_{R-\sigma} \end{aligned}$$

(Fig. 11). Here, $d^\dagger_{R\sigma}$ and $p^\dagger_{\delta R\sigma}$ create a hole (not an electron!) in the copper $3d_{x^2-y^2}$ and in the oxygen $2p_\delta$ ($\delta = x, y$) orbital, respectively. $\langle \ldots \rangle$ denotes next neighbors Cu-O orbitals and $\langle\langle \ldots \rangle\rangle$ next neighbors O-O orbitals. ϵ_d and ϵ_p are the on-site energies of the copper and the oxygen orbitals, respectively. We have set the zero of the energy on the copper site $\epsilon_d = 0$, and introduced the charge-transfer gap $\Delta = \epsilon_p - \epsilon_d$. The Coulomb repulsion between two holes is taken into account by the term U. $\alpha_{RR'\delta}$ and $\alpha'_{RR'}$ contain the sign

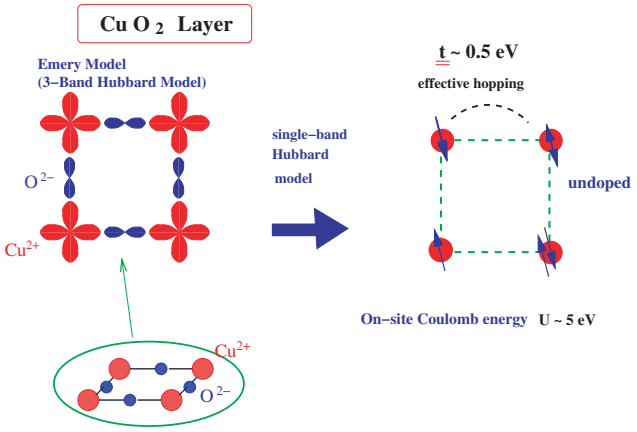

Fig. 10: The copper-oxide planes with their relevant orbitals. Modeling of these planes via a three- (*left*) and a single-band (*right*) Hubbard model

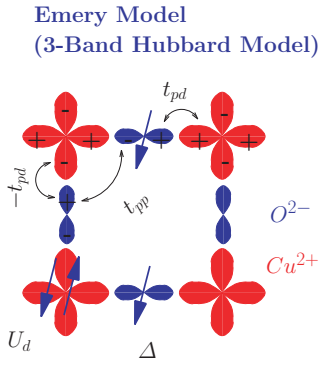

Fig. 11: Parameters of the three-band Hubbard model describing the most relevant orbitals in the CuO_2 layers

of the phases of the Cu-O and O-O hopping due to the d and p symmetries of the Cu and O orbitals, respectively, according to the convention in Fig. 11.

This model can be further simplified by restricting only to the Cu orbitals and considering the O orbitals as an effective "bridge" between them (this is not completely correct in hole-doped HTSC, however, we won't go into detail here. For a more rigorous treatment see [5]). This is the single-band Hubbard model:

$$H_{1\mathrm{b}} = -t \sum_{\sigma} \sum_{\langle RR' \rangle} \left(c^{\dagger}_{R\sigma} c_{R'\sigma} + h.c. \right) + U \sum_{R} n_{R\uparrow} n_{R\downarrow} \tag{18}$$

with the same conventions as in Eq. (1), and $\langle RR' \rangle$ means that the lattice sum is restricted to nearest-neighbor Cu orbitals. The single-band Hubbard

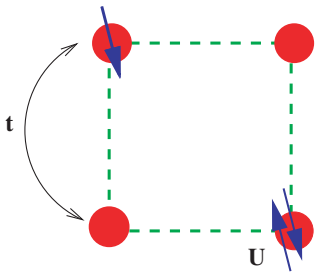

Fig. 12: Single-band Hubbard model

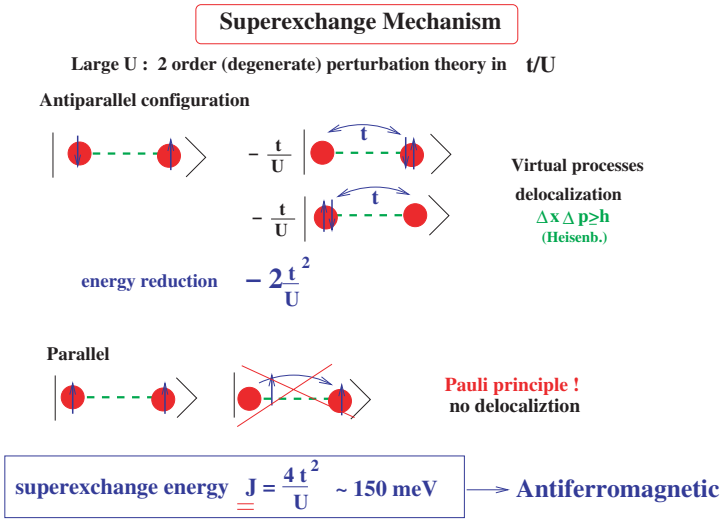

Fig. 13: Superexchange mechanism: due to Pauli principle, the electron spins on two neighboring site prefer to stay antiparallel in order to gain the energy $\propto -t^2/U$

Hamiltonian was independently introduced by Hubbard [6] and Gutzwiller [7] in an attempt to describe the effect of electron correlation in transition metal compounds.

At half filling and for large U/t, an important effect takes place (Fig. 13). Two particles on neighboring sites prefer to have their spin antiparallel in order to be able to delocalize to the neighboring site. The amplitude for this process is $\propto t/U$ as obtained from second-order perturbation theory. This gives an energy gain $\propto -J/2 \approx -2t^4/U$ for antiparallel configurations with respect to parallel ones. Therefore, in a square lattice, particles tend to order

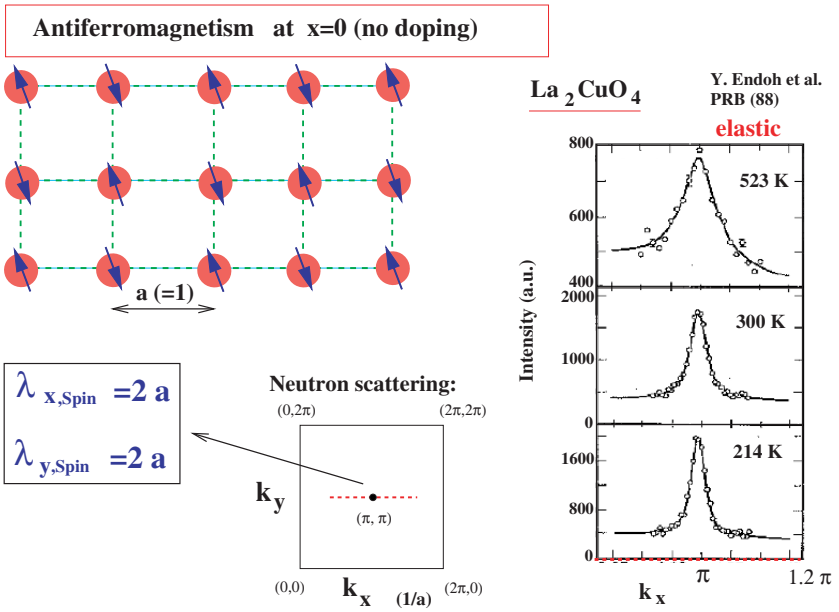

Fig. 14: Detecting antiferromagnetic order via neutron scattering

antiferromagnetically. This ordered phase can be detected, for example, by neutron scattering (Fig. 14).

Doped holes frustrate and eventually destroy the antiferromagnetic state. However, strong antiferromagnetic fluctuations survive deep into the superconducting phase, as shown by neutron-scattering experiments. Many theories consider magnetic fluctuations responsible for the pairing mechanism in HTSC. A simple view on how this can take place will be presented in Sect. 6.

6 Pairing Mediated by Spin Fluctuations: Linear Response to Magnetic Excitations

The magnetic-mediated pairing is similar to the one produced by phonons discussed in Sect. 3. In the former case, however, the bosons responsible for the interaction between the charge carriers are magnetic excitations, also called spin waves. A spin wave can be considered as a local perturbation of the antiferromagnetic (AF) state which can then propagate in space. In a pure AF state, long-wave spin waves have a very long lifetime (Goldstone modes), while in a state with short-range AF order, spin waves are strongly damped.

We now consider a simple approximation to the Hubbard Hamiltonian Eq. (18), which will help us understanding how spin waves propagate in a

non-antiferromagnetic state and how these can induce an effective interaction between electrons. Starting from Eq. (18), we add an external spin-dependent potential $V_{ext}(R,\sigma)$ and consider the linear response to this potential. We approximate the total Hamiltonian within a space-dependent mean field-approach, i.e. each particle is affected only by the "average" interaction with all other particles. This mean-field Hamiltonian reads

$$H = H_{kin} + \sum_{R,\sigma} (V_{ext}(R,\sigma) + V_{int}(R,\sigma))\ n(R,\sigma) \tag{19}$$

where H_{kin} contains the kinetic part of the Hubbard Hamiltonian Eq. (18) (including the chemical potential), while the (mean-field) interaction potential can be obtained from Eq. (18)

$$V_{int}(R,\sigma) \equiv U\rho(R,-\sigma)\ , \tag{20}$$

where $\rho(R,\sigma) \equiv < n(R,\sigma) >$ is the mean density. We consider the fluctuation of ρ in the presence of a total potential

$$V_{tot}(R,\sigma) \equiv V_{ext}(R,\sigma) + V_{int}(R,\sigma) - U\ \rho_0/2 \tag{21}$$

up to linear order in V_{tot}. In Eq. (21), for convenience, we have subtracted the homogeneous (constant) part of the interaction potential, ρ_0 being the average particle density.

In a non magnetic homogeneous state $\rho(R,\sigma) = \rho_0/2 + \delta\rho(R,\sigma)$, where $\delta(R,\sigma)$ is induced by the non-homogeneous potential V_{tot}, we can write, to lowest order in V_{tot},

$$\delta\rho(R,\sigma) \approx -\sum_{R'} \chi_0(R-R')V_{tot}(R',\sigma)\ , \tag{22}$$

where we have introduced the polarisability χ_0, and we have exploited translation invariance by a lattice vector. The choice of the minus sign in the convention for χ_0 in Eq. (22) reflects the fact that we expect a *negative* $\delta\rho$ in the presence of a *positive* V_{tot} (at least at short distances), so that we choose the sign in Eq. (22) in order to have a positive χ_0. By inserting Eq. (20) into Eq. (21) and Eq. (22), and by going over to momentum space q (in which the convolution becomes a product), we obtain a self-consistent equation for $\delta\rho$

$$\delta\rho(q,\sigma) = -\chi_0(q)\left(V_{ext}(q,\sigma) + U\delta\rho(q,-\sigma)\right)\ . \tag{23}$$

Equation (23) can be easily decoupled by introducing the total charge and spin densities:

$$\delta\rho(q) \equiv \delta\rho(q,\uparrow) + \delta\rho(q,\downarrow) \qquad S(q) \equiv \delta\rho(q,\uparrow) - \delta\rho(q,\downarrow)\ , \tag{24}$$

as well as the external charge potential and magnetic field

$$V_{ext}(q) = V_{ext}(q,\uparrow) + V_{ext}(q,\downarrow) \qquad h(q) = V_{ext}(q,\uparrow) - V_{ext}(q,\downarrow)\ . \tag{25}$$

In this way, Eq. (23) becomes

$$\delta\rho(q) = -\chi_0(q)(V_{ext}(q) + U\delta\rho(q)) = -\frac{\chi_0(q)}{1+U\chi_0(q)}V_{ext}(q)\ , \tag{26}$$

and

$$S(q) = -\chi_0(q)(h(q) - US(q)) = -\frac{\chi_0(q)}{1-U\chi_0(q)}h(q)\ . \tag{27}$$

Equation (27) already suggests the possibility of a magnetic instability: if $U\chi_0(q)$ approaches 1 for some q, then the corresponding Fourier component $S(q)$ of the spin density will become nonzero *even for vanishing external magnetic field* h. Clearly, $q = 0$ corresponds to a ferromagnetic, $q = (\pi, \pi)$ to an antiferromagnetic instability[1].

Close to such an instability, the static *spin susceptibility*

$$\chi_s(q) \equiv \frac{\chi_0(q)}{1-U\chi_0(q)} \tag{28}$$

will be strongly peaked at this wave vector.

For the sake of simplicity, we now assume to be close to such an instability, so that we can take $\chi_s U \gg 1$ and neglect the charge susceptibility $\frac{\chi_0(q)}{1+U\chi_0(q)}$ (although the calculation with inclusion of this term is straightforward). Our goal is to evaluate the total potential

$$V_{tot}(q,\sigma) = V_{ext}(q,\sigma) + U\delta\rho(q,-\sigma) \approx V_{ext}(q,\sigma) - \frac{U}{2}\sigma S(q) \tag{29}$$

$$= V_{ext}(q,\sigma) - \frac{U}{2}\sigma[-\chi_s(q)h(q)] \tag{30}$$

$$= V_{ext}(q,\sigma) + \frac{U}{2}\chi_s(q)(V_{ext}(q,\sigma) - V_{ext}(q,-\sigma)) \tag{31}$$

$$= V_{ext}(q,\sigma)\left(1 + \frac{U}{2}\chi_s(q)\right) - V_{ext}(q,-\sigma)\frac{U}{2}\chi_s(q)\ . \tag{32}$$

If we take V_{ext} as the Hubbard interaction originating from a density perturbation $n(q,\sigma)$, then

$$V_{ext}(q,\sigma) = Un(q,-\sigma) \tag{33}$$

$V_{tot}(q,\sigma)$ is thus the total potential acting on a charge density with momentum q and spin σ in the presence of the density Eq. (33). One can interpret this in terms of an effective interaction

$$V_{eff,\uparrow\uparrow}(q) = -\frac{U^2}{2}\chi_s(q) \tag{34}$$

[1] From now on, we will work in units of the lattice constant (Cu-Cu distance)$a = 1$

and

$$V_{eff,\uparrow\downarrow}(q) = U\left(1 + \frac{U}{2}\chi_s(q)\right) \approx \frac{U^2}{2}\chi_s(q)\ . \tag{35}$$

This means that particles with parallel spin experience an effective spin-mediated *attractive interaction*, while particles with antiparallel spin a *repulsive* one.

In Eq. (35) we have neglected the contribution from transverse spin fluctuations, as well as time-dependent effects. These can be obtained more appropriately within a diagrammatic approach. The diagrams leading to the effective interactions are plotted in Fig. 15. The diagrams in (A) yield a contribution

$$U(1 + (U\chi_0)^2 + (U\chi_0)^4 + \cdots) = \frac{U}{1 - (U\chi_0)^2} \tag{36}$$

which is equal to Eq. (35) close to a magnetic instability for which $1-(U\chi_0) \ll 1$. In addition, one obtains a contribution from transverse spin fluctuations, plotted in (B):

$$U^2\chi_0(1 + (U\chi_0) + (U\chi_0)^2 + \cdots) = \frac{U^2\chi_0}{1 - (U\chi_0)} \tag{37}$$

notice, however, that due to particle exchange, the polarisability χ_0 is evaluated not at the momentum transfer q but at $k - k' - q$.

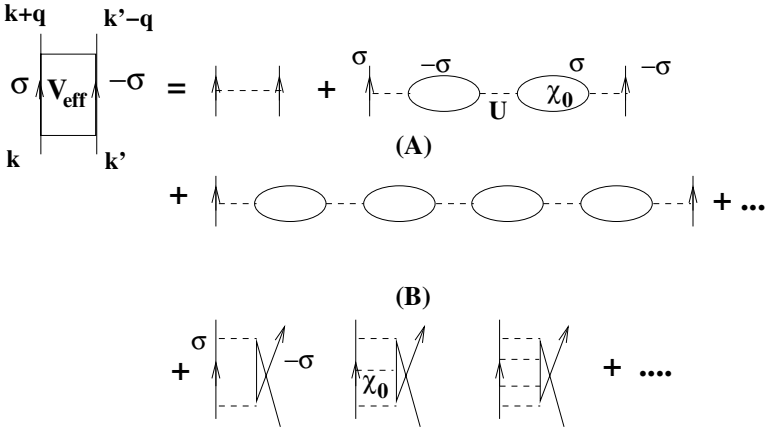

Fig. 15: Diagrams contributing to the effective interaction $V_{eff,\uparrow\downarrow}$ between opposite spins. Diagrams in (**B**) contain the contribution from transverse spin fluctuations, not taken into account in Eq. (35). *Dashed lines* represent the Hubbard interaction U, *solid lines* are Fermion propagators. One closed fermionic loop (bubble) corresponds to the polarisability χ_0

The idea of a superconducting pairing mechanism originating from magnetic fluctuations was originally suggested by Emery [8] and by Berk and Schrieffer [9]. They were interested in a mechanism for superconductivity in nearly ferromagnetic materials in which $\chi_s(q)$ is strongly peaked at $q = 0$.

As discussed in Sect. 4, a repulsive interaction seems to suggest that it is impossible to get a superconducting state with antiparallel spin (singlet state). This is not true. Here, the q dependence of the interaction plays a crucial role. As a matter of fact, close to an AF instability, $\chi_s(q)$, and thus $V_{eff,\uparrow\downarrow}(q)$ is peaked around (π, π). In order to illustrate the consequences of this strong q dependence, we take a model V_{eff} with this property [10, 11]:

$$V_{eff,\uparrow\downarrow}(q) \approx \frac{V_0}{1 + \xi^2 \left[(q_x - \pi)^2 + (q_y - \pi)^2\right]} \,. \tag{38}$$

The Fourier transform in real space of Eq. (38) will decay on a typical distance (AF correlation length) ξ, will be positive at $R = 0$ so that particles on the same site will repel each other, and, most importantly, Eq. (38) will be negative at nearest-neighbor sites $|R| = 1$. This can be seen from the most important contribution to the Fourier integral at (π, π) yielding a phase $e^{i\pi R_x + i\pi R_y}$, which is -1 at $R = (1, 0)$ and equivalent points. This means that particles on nearest-neighbor sites will attract each other and will be able to form Cooper pairs.

In momentum space this can be seen by invoking again the BCS gap equation Eq. (15) with V replaced with $-V_{eff,\uparrow\downarrow}$ (here, $V_{eff,\uparrow\downarrow} > 0$):

$$\Delta_k = -\frac{1}{N} \sum_{k'} V_{eff,\uparrow\downarrow}(k - k') \frac{\Delta_{k'}}{E_{k'}} \,. \tag{39}$$

Clearly, there is no solution for a constant (s-wave) Δ_k. On the other hand, if V_{eff} is strongly peaked at $k - k' = Q \equiv (\pi, \pi)$ one can look for a solution in which $\Delta_{k+Q} = -\Delta_k$, so that one obtains

$$\Delta_k = \frac{1}{N} \sum_{k'} V_{eff,\uparrow\downarrow}(k - k' - Q) \frac{\Delta_{k'}}{E_{k'-Q}} \,. \tag{40}$$

One possible solution is obtained by assuming an angular dependence of the gap function in the form

$$\Delta_k = \Delta_0(\cos k_x - \cos k_s) \,. \tag{41}$$

This form changes sign under a rotation by an angle $\pi/2$, and belongs to the so called $d_{x^2-y^2}$ irreducible representation of the square lattice symmetry [12]. The presence of a $d_{x^2-y^2}$ form of the superconducting gap in HTSC has been supported by many experiments [12, 13, 14, 15], and is by now a well accepted result. The present discussion suggests that a $d_{x^2-y^2}$ pairing gap is consistent with a magnetic mechanism for high-T_c superconductivity, while at

the same time probably excluding a (pure) phononic mechanism in which the effective interaction is not strongly q-dependent. In addition, the magnetic mechanism is supported by the observation of a strong magnetic resonance in the neutron-scattering spectrum of many HTSC in the superconducting phase [16, 17, 18].

Similar conclusions are reached within the Nearly Antiferromagnetic Fermi Liquid theory by Pines [10], in which the spin susceptibility is taken phenomenologically from experiments in the form Eq. (38). A related theory based on AF spin fluctuations was developed by Moriya and coworkers [19]. Schrieffer proposed the so-called "spin-bag" magnetic pairing mechanism based on the idea that local AF order is still present upon doping and that a particle (hole) is dressed with a spatially (typically a few lattice constants) extended spin cloud, which is due to the frustration of the local antiferromagnetic order. Two particles tend to pair because in this way they minimize the frustration of the AF order [20], similarly to the effect displayed in Fig. 5. On the other hand, Schrieffer criticizes the spin-wave mediated pairing discussed in Sect. 6, suggesting that vertex corrections to the diagrams shown in Fig. 15 are crucial and could dramatically reduce the effective interactions [21]. This is due to the fact that spin waves in an antiferromagnet are in fact Goldstone modes, i.e. they should become exact eigenstates of the Hamiltonian (i.e. decoupled from other excitations) in the long wavelength limit. Indeed, a strong reduction of the electron-spin-wave vertex has been confirmed by recent Quantum-Monte Carlo calculations [22].

Despite the fact that many experiments and theories support a magnetic-based pairing mechanism, recent measurements suggest that electron-phonon coupling could also play an important role in HTSC after all. There are strong indications for a superconductivity-induced phonon renormalization, for a large isotope coefficients away from optimal doping [23], and of a strong electron-phonon coupling in the electron dispersion, as observed by angle-

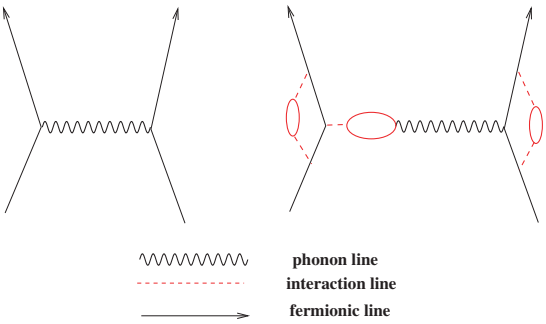

Fig. 16: A vertex correction to the phonon-mediated electron-electron interaction. *Dashed lines* describe the bare electron-electron interaction U, and *wavy lines* describe the phonon propagator

resolved photoemission spectroscopy [24]. In this respect, the combined effect of electron-phonon and electron-electron interaction may play an important role. Indeed, vertex corrections due to electron-electron interaction (see Fig. 16) introduce a strong q dependence in the electron-phonon coupling [23, 25, 26, 27], possibly making an electron-phonon mechanism consistent with $d_{x^2-y^2}$ pairing.

References

1. J. Bardeen, L.N. Cooper, J.R. Schrieffer: Phys. Rev. B **108**, 1175 (1957)
2. J.G. Bednorz, K.A. Müller: Zeit. Phys. B **64**, 189–193 (1986)
3. G.M. Eliashberg: Soviet Phys. JETP, **11**, 696 (1960)
4. V.J. Emery: Phys. Rev. Lett. **58**, 2794 (1987)
5. F.C. Zhang, T.M. Rice: Phys. Rev. B **37**, 3759–3761 (1988)
6. J. Hubbard: Proc. R. Soc. London **276**, 238 (1963)
7. M.C. Gutzwiller: Phys. Rev. Lett. **10**, 159 (1963)
8. V.J. Emery: Ann. Phys. (NY) **28**, 1 (1964)
9. N.F. Berk, J.R. Schrieffer: Phys. Rev. Lett. **17**, 433 (1966)
10. D. Pines: Phys. B **163**, 78 (1990)
11. A.J. Millis, H. Monien, D. Pines: Phys. Rev. B **42**, 167–178 (1990)
12. D.J. Scalapino: Phys. Rep. **250**, 329–365 (1995)
13. Z.X. Shen, D.S. Dessau: Phys. Rep. **253**, 1 (1995)
14. C.C. Tsuei, J.R. Kirtley: Phys. Rev. Lett. **73**, 593 (1994)
15. J.R. Kirtley, C.C. Tsuei: Scientific American August, 50–55 (1996)
16. H.F. Fong, B. Keimer, P.W. Anderson, D. Reznik, F. Dogan, I.A. Aksay: Phys. Rev. Lett. **75**, 316–319 (1995)
17. H.F. Fong, B. Keimer, D.L. Milius, I.A. Aksay: Phys. Rev. Lett. **78**, 713–716 (1997)
18. H. He, P. Bourges, Y. Sidis, C. Ulrich, L.P. Regnault, S. Pailhes, N.S. Berzigiarova, N.N. Kolesnikov, B. Keimer: Science **295**, 1045–1047 (2002)
19. T. Moriya, Y. Takahashi, K. Ueda: J. Phys. Soc. Japan **59**, 2905–2915 (1990)
20. J.R. Schrieffer, X.G. Wen, S.C. Zhang: Phys. Rev. B **39**, 11 663–11 679 (1989)
21. J.R. Schrieffer: J. Low Temp. Phys. **99**, 397 (1995)
22. Z.B. Huang, W. Hanke, E. Arrigoni: Europhys. Lett. **71**, 959 (2005)
23. M.L. Kulic: Phys. Rep. **388**, 1–264 (2000)
24. A. Lanzara, P.V. Bogdanov, X.J. Zhou, S.A. Kellar, D.L. Feng, E.D. Lu, T. Yoshida, H. Eisaki, A. Fujimori, K. Kishio, J.I. Shimoyama, T. Noda, S. Uchida, Z. Hussain, Z.X. Shen: Nature (London) **412**, 510–514 (2001)
25. R. Zeyher, M.L. Kulic: Phys. Rev. B **53**, 2850–2862 (1996)
26. M. Grilli, C. Castellani: Phys. Rev. B **50**, 16 880–16 898 (1994)
27. Z.B. Huang, W. Hanke, E. Arrigoni, D.J. Scalapino: Phys. Rev. B **68**, 220 507(R) (2003)

Spin in Quantum Field Theory

S. Forte

Dipartimento di Fisica, Università di Milano and INFN, Sezione di Milano
via Celoria 16, 10129 Milano, Italy
forte@mi.infn.it

Abstract. I introduce spin in field theory by emphasizing the close connection between quantum field theory and quantum mechanics. First, I show that the spin-statistics connection can be derived in quantum mechanics without relativity or field theory. Then, I discuss path integrals for spin without using spinors. Finally, I show how spin can be quantized in a path-integral approach, without introducing anticommuting variables.

1 From Quantum Mechanics to Field Theory

Even though everybody learns about spin in their childhood in the context of nonrelativistic quantum mechanics, many of the more interesting dynamical features of spin are only introduced in relativistc quantum field theory. In these lectures, which were originally addressed to an audience of (mostly) condensed-matter physicists, I discuss some relevant aspects of spin dynamics in quantum field theory by showing their origin in quantum mechanics. In the first lecture, after a brief discussion of the way spin appears in nonrelativistic (Galilei invariant) or relativistic (Lorentz invariant) dynamics, I show how the spin-statistics connection can be obtained with minimal assumptions in nonrelativistic quantum mechanics, without invoking relativity or field theory. In the second lecture I show how spin can be quantized in a path-integral approach with no need for introducing quantum fields. In the third lecture I discuss the dynamics of relativistic spinning particles and show that its quantization can be described without using anticommuting variables. A fourth lecture was devoted to the quantum breaking of chiral symmetry – the axial anomaly – and its origin in the structure of the spectrum of the Dirac operator, but since this subject is already covered in many classic lectures [1] it will not be covered here. We will see that even though the standard methods of quantum field theory are much more practical for actual calculations, a purely quantum-mechanical approach helps in understanding the meaning of field-theoretic concepts.

S. Forte: *Spin in Quantum Field Theory*, Lect. Notes Phys. **712**, 67–94 (2007)
DOI 10.1007/3-540-38592-4_3

2 Spin and Statistics

2.1 The Galilei Group and the Lorentz Group

In both relativistic and non-relativistic dynamics we can understand the meaning of quantum numbers in terms of the symmetries of the Hamiltonian and the Lagrangian and associated action. Indeed, the invariance of the Hamiltonian determines the spectrum of physical states: eigenstates of the Hamiltonian are classified by the eigenvalues of operators which commute with it, and this gives the set of observables which are conserved by time evolution. However, the invariance of the dynamics is defined by the invariance of the action. This is bigger than that of the Hamiltonian, because it also involves time-dependent transformations. For example, in a nonrelativistic theory the action must be invariant under Galilei boost: the change between two frames that move at constant velocity with respect to each other. But the Hamitonian in general doesn't possess this invariance: Galilei boosts obviously change the values of the momenta, and the Hamiltonian in general depends on them. The set of operators which commute with all transformations that leave the action invariant defines the quantum numbers carried by elementary excitations of the system (elementary particles).

A nonrelativistic theory must have an action which is invaraint upon the Galilei group. The Galilei transformations, along with the associate quantum-mechanical operators are [2]:

- space translations: $x_i \to x_i' = x_i + a_i$; $\quad P_i = -i\partial_i$
- time translation: $t \to t' = t + a$; $\quad H = i\frac{d}{dt}$
- Galilei boosts: $x_i \to x_i' = x_i + v_i t$; $p_i \to p_i + m v_i$; $\quad K_i = -it\partial_i - m x_i$
- rotations: $x_i \to x_i' = R_{ij} x_j$; $\quad J_i = \epsilon_{ijk} x^j \partial_k + \sigma_i$

The generator of rotations is the sum of orbital angular momentum and *spin*.

The generators of the Galilei group form the Galilei algebra:

$$\begin{aligned} &[J_i, J_j] = \epsilon_{ijk} J_k; \quad [P_i, P_j] = 0; \quad [K_i, K_j] = 0; \quad [J_i, H] = [K_i, H] = 0; \\ &[k_i, H] = iP_i; \quad [J_i, P_j] = \epsilon_{ijk} P_k; \quad [J_i, K_j] = \epsilon_{ijk} K_k; \quad [K_i, P_j] = iM\delta_{ij} \; . \end{aligned} \tag{1}$$

In order to close the algebra it is necessary to introduce a (trivial) mass operator M which commutes with everything else:

$$[M, P_i] = [M, K_i] = [M, J_i] = [M, H] = 0. \tag{2}$$

The Casimir operators, which commute with all generators, are

$$C_1 = M; \quad C_2 = 2MP_0 - P_i P_i; \quad C_3 = (MJ_i - \epsilon_{ijk} P_j K_k)\,(MJ_i - \epsilon_{ilm} P_l K_m)\,. \tag{3}$$

In terms of quantum-mechanical operators the Casimirs correspond to

- $C_1 = m$ (mass);
- $\frac{1}{2m}C_2 = -i\frac{d}{dt} - \frac{p^2}{2m}$ (internal energy);
- $C_3 = \sigma_i\sigma_i$ (**spin**).

We see that spin is one of the three numbers which classify nonrelativistic elementary excitations, along with mass and internal energy.

In the relativistic case, the action is invariant under the Poincaré group. The transformations and associate operators are now:

- translations: $x_\mu \to x'_\mu = x_\mu + a_\mu$; $\qquad P_\mu = -i\partial_\mu$
- Lorentz transf.:
  $x^\mu \to x'^\mu = \Lambda^\mu{}_\nu x^\nu$; $p^\mu \to p'^\mu = \Lambda^\mu{}_\nu p^\nu$;
  $$J^{\mu\nu} = x^\mu P^\nu - x^\nu P^\mu + \Sigma^{\mu\nu}$$
  $$\text{rotations: } J_i = \tfrac{1}{2}\epsilon_{ijk}J^k = \epsilon_{ijk}x_j P_k + \sigma^i$$
  $$\text{boosts: } K_i = J^{i0}.$$

The Poincaré generators form the algebra

$$\begin{aligned}
&[J^{\mu\nu}, J^{\rho\sigma}] = i\left(g^{\mu\rho}J^{\nu\sigma} - g^{\nu\rho}J^{\mu\sigma} + g^{\mu\sigma}J^{\nu\rho} + g^{\nu\sigma}J^{\mu\rho}\right);\\
&[P^\mu, J^{\rho\sigma}] = -i\left(g^{\mu\rho}P^\sigma - g^{\mu\sigma}P^\rho\right); \qquad [K^\mu, P^\nu] = 0
\end{aligned} \tag{4}$$

Explicitly, the algebra of boosts and rotations is

$$\begin{aligned}
&[J_i, J_j] = \epsilon_{ijk}J_k; \quad [J_i, K_j] = \epsilon_{ijk}K_k; \quad [K_i, K_j] = -i\epsilon_{ijk}K_k\\
&[J_i, P_j] = \epsilon_{ijk}P_k; \quad [K_i, H] = iP_i; \quad [K_i, P_j] = iH\delta_{ij}\ .
\end{aligned} \tag{5}$$

The Casimir operators are now just two:

$$C_1 = P_\mu P^\mu; \quad C_2 = W_\mu W^\mu, \tag{6}$$

in terms of the momentum generator and the *Pauli-Lubanski* operator

$$W^\mu = \epsilon^{\mu\nu\rho\sigma}P_\nu J_{\rho\sigma}. \tag{7}$$

The corresponding quantum-mechanical operators are

- $C_1 = P^2$; eigenvalue M^2 (mass);
- $C_2 = W^2 = m\sigma^2$; eigenvalue $M^2 s(s+1)$ (mass×**spin**),

where the latter identification is clear if one chooses the rest frame, as we shall discuss in greater detail in Sect. 4.3.

Galilei transformations can be obtained from Poincaré transformations in the limit $v \ll 1$ by assuming the scaling laws $M \sim 1$, $J \sim 1$, $P \sim v$, $H \sim v^2$ $K \sim 1/v$.

Summarizing, both in nonrelativistic and relativistic theories spin is one of the quantum numbers that classify elementary excitations. In quantum mechanics, the state vectors of physical systems are expanded on a basis

of irreducible representations of the rotation group (in the nonrelativistic case) or the Lorentz group (in the relativistic case). In quantum field theory, one-particle states are, respectively, Galilei or Poincaré irreducible representations. In the relativistic case, rotations are implicitly defined by the Pauli-Lubanski vector Eq. (7) as the subgroup of the Lorentz group which leaves the four-momentum invariant.

In more than two spatial dimensions the rotation group $\mathrm{O}(d)$ is doubly connected (i.e., $\pi_1[\mathrm{O}(d)] = \mathbb{Z}_2$); its universal cover is the group $\mathrm{Spin}(d)$, which, in the usual $d = 3$ case, is isomorphic to SU(2). When $d = 2$ (planar systems) the rotation group is O(2), which, being isomorphic to the circle S^1 is infinitely connected ($\pi_1[O(2)] = \mathbb{Z}$); its universal cover is the real line $\mathbb{R}$. It follows that in more than two dimension the wave function can carry either a simple-valued (Bosons) or a double valued (Fermions) representation of the rotation group, and in two dimensions it may carry an arbitrarily multivalued one (anyons [4]).

The multivaluedness of the representation of rotations is classified by the value of the phase which the wave function acquires upon rotation by 2π about an arbitrary axis (the z axis, say):

$$R_z^{2\pi}\psi(q_1,\ldots,q_n) = e^{2\pi i J_z}\psi(q_1,\ldots,q_n) = e^{2\pi i\sigma}\psi(q_1,\ldots,q_n)\ , \tag{8}$$

where $J_z = L_z + \sigma$, and in the last step we have used the fact that the spectrum of orbital angular momentum is given by the integers, so upon 2π rotation it is only spin that contributes to the phase.

2.2 Statistics and Topology

The wave function for a system of n identical particles must be invariant in modulus, and thus acquire a phase, upon interchange of the full set of quantum numbers q_i of the i-th and j-th particle:

$$\psi(q_1,\ldots,q_i,\ldots,q_j,\ldots,q_n) = e^{2\pi i\sigma}\psi(q_1,\ldots,q_j,\ldots,q_i,\ldots,q_n)\ . \tag{9}$$

The parameter σ, which is only defined modulo integers, is the statistics of particles i, j. We now prove the spin-statistics theorem, which states that the statistics is a universal property of particles i, j, and it is equal to their spin (also in $d = 2$, where the spin as we have seen can be generic).

The proof is based on an analysis of the quantisation of systems defined on topologically nontrivial configuration spaces. Indeed, if $\mathcal{C}_d$ is the configuration space for a single particle in d dimensions, the configuration space for a system of n particles in d dimensions is

$$\bar{\mathcal{C}_d}^n = {C_d}^n - \mathcal{D}\ , \tag{10}$$

where $\mathcal{D}$ is the set of points where the full set of quantum numbers of two or more particles coincide. These points must be excised from space, otherwise Eq. (9) with $x_i = x_j$ implies that necessarily $\sigma = 0$.

If the particles are identical, points which differ by their interchange must be identified. The configuration space then becomes the coset space

$$\mathcal{C}_d{}^n = \frac{\bar{\mathcal{C}}_d{}^n}{S_n} \,, \tag{11}$$

where S_n is the group of permutations of n objects. The topological structure of the configuration space changes when going from two to more than two dimensions, just like the topological structure of the rotation group discussed in Sect. 2.2. Indeed, if $d = 2$ the space Eq. (10), i.e. before dividing out permutations, is multiply connected: a closed path traversed by the i-th particle in which particle j is inside the loop formed by particle i cannot be deformed into a path in which particle j is outside the loop. The configuration space $\mathcal{C}_2^n$ is then also multiply connected, and its fundamental group is the braid group $\pi_1(\mathcal{C}_2^n) = B_n$, as we shall discuss explicitly below.

In more than two dimensions, the space $\bar{\mathcal{C}}_d^n$ is simply connected: all closed path traversed by a particle can be continuously deformed into each other, because in more than two dimensions one cannot distinguish the inside of a one-dimensional curve from its outside. However, the configuration space $\mathcal{C}_d^n$ is multiply connected. This implies that a topologically nontrivial closed path in $\mathcal{C}_d^n$ must correspond to an open path in $\bar{\mathcal{C}}_d^n$, because all closed paths in $\bar{\mathcal{C}}_d^n$ can be deformed into each other. Furthermore, points in $\mathcal{C}_d^n$ that correspond to same point in $\bar{\mathcal{C}}_d^n$ are in one-to-one correspondence with elements of S_n, because S_n acts effectively, i.e. only the identity of S_n maps all points of $\bar{\mathcal{C}}_d^n$ onto thenselves. It follows that equivalence classes of paths in $\bar{\mathcal{C}}_d^n$ are in one-to-one correspondence with elements of the permutation group:

$$\pi_1(\mathcal{C}_d^n) = S_n \,. \tag{12}$$

Hence, the multiply connected nature of the configuration space is directly linked with the presence of identical particle, and specifically to the response of the system upon permutations, i.e. to statistics.

Therefore, let us consider quantization on a multiply connected space. The way nontrivial statistics is obtained can be understood by studying this problem in a path-integral approach [3], where transition amplitudes are written in terms of the propagator $K(q', q)$

$$S_{fi} \equiv \langle \psi_f | \psi_i \rangle = \langle \psi_f | q't' \rangle \langle q't' | qt \rangle \langle qt | \psi_i \rangle = \int dq\, dq'\, \psi_f^*(q') K(q', q) \psi_i(q) \,, \tag{13}$$

which in turn can be written as a sum over paths

$$K(q', t'; q, t) = \int_{q(t)=q;\, q(t')=q'} Dq(t_0)\, e^{i \int_t^{t'} dt_0\, L[q(t_0)]} \,. \tag{14}$$

Closed paths on a multiply connected space fall into homotopy classes. Moreover, open paths can also be classified in homotopy classes by a choice

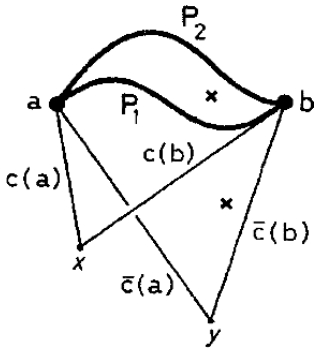

Fig. 1: Paths P_i are assigned to homotopy classes by connecting them to a base point through a mesh. Changing the base point from x to y can change the absolute class assigment of a path, but not the relative assignment of a pair of paths

of *mesh* (Fig. 1). Namely, one chooses an arbitrary reference point x (base point) and then one assign to each point in space a path connecting it to the base point. The homotopy class of an open path can then be defined as the homotopy class of the closed path formed by the given open path and the mesh that connects it to the base point. Once all paths (closed and open) are grouped into equivalence classes, the path integral is in general defined as follows

$$K(q',t';q,t) = \sum_{\alpha} \chi(\alpha) K^{\alpha}(q',t';q,t) \; , \tag{15}$$

where $K^{\alpha}(q',t';q,t)$ is computed including in the sum over paths only paths in the α-th homotopy class, and $\chi(\alpha)$ are weights which depend only on the equivalence class (homotopy class) of a given path.

The weighted sum (15) must satisfy the following physical requirements:

- (a) physical result must be independent of the choice of mesh;
- (b) amplitudes must satisfy the superposition principle, which in turn implies the convolutive property

$$K(q'',t'';q,t) = \int dq' \langle q''t''|q't'\rangle\langle q't'|qt\rangle = \int dq' K(q'',t''q't';)K(q't';q,t) \; . \tag{16}$$

The necessary and sufficient condition for these requirements to be satisfied is that the weights $\chi(\alpha)$ satisfy

$$|\chi(\alpha)| = 1 \tag{17}$$

$$\chi(\alpha \circ \beta) = \chi(\alpha)\chi(\beta) \; , \tag{18}$$

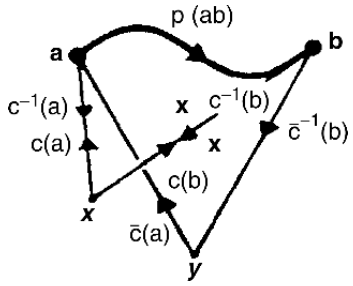

Fig. 2: Graphical representation of Eq. (20)

where in Eq. (18) α and β are the homotopy classes of paths with a common endpoint, and $\alpha \circ \beta$ is the homotopy class of the path obtained by joining them.

The proof that Eq. (18) implies property (b) is immediate:

$$\sum_{\gamma} \chi(\gamma) K^{\gamma}(q', t'; q, t) = \sum_{\alpha,\beta} \chi(\alpha)\chi(\beta) \int dq' K^{\alpha}(q'', t'' q' t';) K^{\beta}(q' t'; q, t) \; . \tag{19}$$

The proof that Eq. (17) implies property (a) is also easy: let P be the closed path obtained composing the open path p which connectes points a and b with a mesh C (Fig. 2)). Upon changing the mesh to $\bar{C}$, the path P becomes the path $\bar{P}$, which in turn can be obtained by composing P with the closed paths $\lambda \equiv \bar{C}(a)C^{-1}(a)$ and $\mu = C(b)\bar{C}^{-1}(b)$:

$$\begin{aligned} \bar{P}(ab) &= \bar{C}(a) p(ab) \bar{C}^{-1}(b) \\ &= \bar{C}(a) C^{-1}(a) C(a) p(ab) C^{-1}(b) C(b) \bar{C}^{-1}(b) \\ &= \lambda P(ab) \mu \; . \end{aligned} \tag{20}$$

Because μ and λ do not depend on the original path, but only on the two meshes, the factor $\chi(\lambda\mu)$ which relates the two class assignments

$$\bar{\chi}(\alpha) = \chi(\lambda\mu)\chi(\alpha) \tag{21}$$

is universal. It follows that

$$\sum_{\gamma} \bar{\chi}(\gamma) K^{\alpha}(q', t'; q, t) = \chi(\lambda\mu) \sum_{\gamma} \chi(\gamma) K^{\alpha}(q', t'; q, t) \; , \tag{22}$$

so if χ are phases the transition probability is unchanged.

This proves that conditions (17–18) are sufficient for requirements (a-b) to be satisfied, the proof that they are also necessary is somewhat more technical and we shall omit it. Conditions Eqs. (17–18), taken jointly, mean that phases χ provide one-dimensional unitary representation of $\pi_1(\mathcal{C}_d^n)$, i.e. the permutation group S_n (in more than two dimensions) or the braid group (in two dimensions).

2.3 Bosons, Fermions and Anyons

The relation between spin and statistics now follows from the structure of the path integral. First, we observe that there are only two unitary one-dimensional irreducible representations of the permutation group S_n: the trivial one (where $\chi = 1$ for all permutations), and the alternating one, where $\chi = 1$ if the permutation is even and $\chi = -1$ if it is odd (i.e., if they may be performed by an even or odd number of interchanges, respectively). Now, note that the wave function at time t is given by the path integral in terms of some boundary condition at time t_0:

$$\langle q, t|\psi\rangle = \int dq_0 K(q, t; q_0, t_0)\psi_0(q_0, t_0) \; . \tag{23}$$

Two evolutions that lead to final states which only differ by the interchange of the coordinates q_i, q_j in configuration space differ by the factor χ: hence, $\chi = -1$ correspond to $\sigma = \frac{1}{2}$ ($\sigma = 0$). However, an interchange of coordinates q_i, q_j can also be realized by a rotation by π of the system about any axis through the center of mass of the two particles (or a rotation about any axis followed by a translation), which in turn is generated by the corresponding angular momentum operator

$$|q_j q_i\rangle = e^{i\pi J_z^{ij}}|q_i q_j\rangle \; , \tag{24}$$

where J_z^{ij} is the component along the (arbitrarily defined) z axis of the angular momentum of particles i, j. The constraint that σ can only be either integer or half-integer is understood as a consequence of the trivial fact that two interchanges, or a rotation by 2π, must bring back to the starting configuration.

It follows that if $\chi = -1$, so $\sigma = \frac{1}{2}$, and the spectrum of J_z^{ij} is given by the odd integers. We can then view the contribution of χ to the path integral as the result of having added an extra internal effective interaction, which shifts the angular momentum of the pair of particles i, j by an integer, i.e. the angular momentum of each particle by a half-integer. This establishes the spin-statistics theorem in a nonrelativistic theory. The results is a consequence of the fact that fermionic statistics, which is usually given as a property of wave functions, has been lifted through the path-integral formalism to a property of particle paths, and attributed to a weight given to paths. The fact that either trivial or alternating representations of permutations are possible is then directly related to the existence of either single-valued or double-valued representations of rotations.

In order to understand this better, let us now consider the case of planar systems [4], both because we can then generalize this spin-statistics connection to arbitrary spin and statistics (anyons), and also because we can then work out an explicit representation for the effective interaction associated to the χ weigths, which will lead us to the spin action which we shall then discuss in the next section.

In $d = 2$, χ^α provide an abelian irreducible representation of the braid group. Indeed, each particle trajectory on a multiply-connected space defines an inequivalent braid (Fig. 3). Each braid, in turn, is uniquely defined as a sequence of interchanges of pairs of neighbouring particles. This can be represented algebraically by introducing the operator σ_i which exchanges particles i and $i + 1$ (Fig. 4). Two braids are equivalent if they can be deformed into each other. For instance (Fig. 5)

$$\sigma_i \sigma_{i+1} \sigma_i = \sigma_{i+1} \sigma_i \sigma_{i+1} \tag{25}$$

$$\sigma_i \sigma_{i+1} \sigma_i \neq \sigma_i \sigma_{i+1} \sigma_i^{-1} \, . \tag{26}$$

In fact, all independent relations between braids are Eq. (25) and

$$\sigma_i \sigma_j = \sigma_j \sigma_i \quad \text{if } |i - j| > 1 \, . \tag{27}$$

In terms of χ Eq. (27) implies

$$\chi(\sigma_i \sigma_j) = \chi(\sigma_i)\chi(\sigma_j) \quad \text{if } |i - j| > 1 \, , \tag{28}$$

while Eq. (25) implies

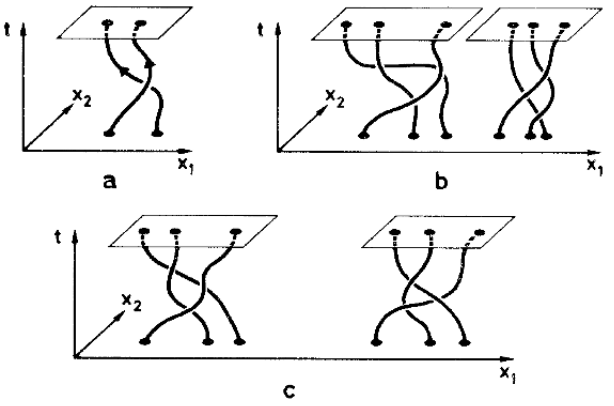

Fig. 3: Braids defined by particles' trajectories

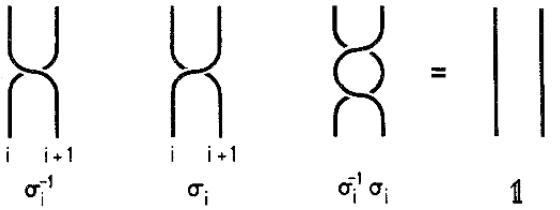

Fig. 4: The exchange operator σ_i and its inverse

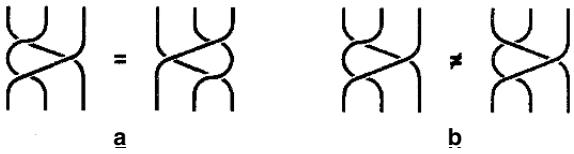

Fig. 5: Graphical representation of Eq. (25) (**a**) and Eq. (26) (**b**)

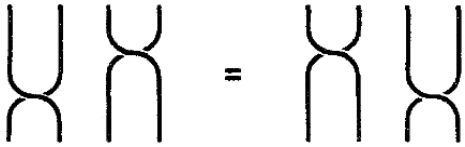

Fig. 6: Graphical representation of Eq. (27)

$$\chi(\sigma_i) = \chi(\sigma_j) \quad \text{for all } i, j \; . \tag{29}$$

Equations (28,29) in turn imply that the weight for a generic path (braid) is

$$\chi(\sigma_{i_1} \dots \sigma_{i_n}) = \chi(\sigma_{i_1}) \dots \chi(\sigma_{i_n}) = \exp\left(2i\sigma \sum_{k=1}^{n} \epsilon_k\right) \; , \tag{30}$$

where $\epsilon = +1$ for a direct exchange and $\epsilon = -1$ for its inverse σ_i^{-1}, and σ coincides with the statistics parameter Eq. (9). The cases of bosons and fermions are recovered when $\sigma = 0$ or $\sigma = \frac{1}{2}$, respectively, but now σ can take any real value (anyons). Indeed, in two dimensions two subsequent interchanges do not necessarily take back to the starting point, because a path where particle i traverses a loop encircling particle j cannot be shrunk to a point (identity). Hence, two interchanges do not necessarily bring back to the starting configuration, and the constraint that $2\sigma = 1$ no longer applies. Accordingly, as already mentioned, in two dimension the rotation group admits arbitrarily multivalued representations.

The χ weights can be represented explicitly in in terms of the variation of relative polar angle $\Theta(\boldsymbol{x}) \equiv \tan^{-1}\left(\frac{x^2}{x^1}\right)$ of particles i and j along the particles' paths:

$$\chi = \exp\left(-2i\sigma \sum_{i<j} \Delta\Theta_{ij}\right) = \exp\left(-i\sigma \sum_{i\neq j} \int dt \frac{d}{dt} \Theta\left(\boldsymbol{x}_i(t) - \boldsymbol{x}_j(t)\right)\right) \; . \tag{31}$$

Using this representation, the weighted path integral Eq. (15) becomes

$$K(q', t'; q, t) = \int_{q(t)=q;\, q(t')=q'} Dq(t_0)\, e^{i\int_t^{t'} dt_0 \left(L[q(t_0)] - \sigma \sum_{i\neq j} \frac{d}{dt_0} \Theta[\boldsymbol{x}_i(t_0) - \boldsymbol{x}_j(t_0)]\right)} \tag{32}$$

$$= \sum_{n_{ij,\,(i\neq j)}=-\infty}^{\infty} e^{-i\sigma\left(\sum_{i\neq j}\hat{\Theta}\left(\boldsymbol{x}_i(t')-\boldsymbol{x}_j(t')\right)+2\pi n_{ij}\right)} K_0^{(n)}(q',t';q,t) e^{i\sigma\sum_{i\neq j}\hat{\Theta}(\boldsymbol{x}_i(t)-\boldsymbol{x}_j(t))} .$$

Hence, the weights χ can be viewed as the consequence of having added to the Lagrangian L the effective interaction term

$$L_{\text{eff}}[q(t)] = -\sigma \sum_{i\neq j} \frac{d}{dt}\Theta[\boldsymbol{x}_i(t) - \boldsymbol{x}_j(t)] . \tag{33}$$

If the starting Lagrangian L described bosonic exitations, the interaction Eq. (33) endows these excitations with statistics σ.

Equation (32) shows that the effect of the statistics-changing interaction can be absorbed in a redefinition of the wave function by a phase:

$$\psi_0(q,t) = e^{i\sigma\sum_{i\neq j}\Theta_{ij}(t)}\psi(q,t) \ : \tag{34}$$

the wave function ψ_0 is propagated by the path-integral defined in terms of the bosonic Lagrangian L. However, it is defined on a space of paths rather than a space of coordinates, and it satisfies "twisted" boundary conditions: upon rotation by 2π it acquires a phase

$$R^{2\pi}\psi_0(q,t) = e^{i2\pi\sigma n(n-1)}\psi_0(q,t) , \tag{35}$$

and correspondingly the spectrum of eigenvalue of the angular momentum operator (which generates rotations) is

$$j = j_0 + \sigma n(n-1) , \tag{36}$$

where j_0 is the spectrum of angular momentum for the original Lagrangian. We see explicitly that for a system of of particles the angular momentum of the pair is shifted by 2σ i.e. each particle has acquired spin σ.

The effective statistics-changing Lagrangian L_{eff} Eq. (33) looks intrinsically nonrelativistic, in that it depends on the polar angle as a function of time. However, it also admits a covariant formulation, which will turn out to be closely related to the formulation of a path integral for spin. To see this, define a covariant particle current

$$j^\mu = \sum_{i=1}^{n}\left(1, \frac{d\boldsymbol{x}_i}{dt}\right)\delta^{(2)}(\boldsymbol{x}-\boldsymbol{x}_i) = \sum_{i=1}^{n}\int ds\, \delta^{(3)}(x-x_i)\,\frac{dx^\mu}{ds} , \tag{37}$$

where s is any covariant parametrization along the particle path (e.g. the path-length). Furthermore, add to the action $I_0 = \int dt L(t)$ a covariant coupling of the current to a gauge potential A_μ:

$$I = I_0 + I_c + I_f \tag{38}$$

$$I_c = \int d^3x\, j^\mu(x) A_\mu(x) \tag{39}$$

$$I_f = -\frac{1}{2\sigma}\int d^3x\, \epsilon^{\mu\nu\rho} A_\mu(x)\partial_\nu A_\rho(x) . \tag{40}$$

The action I_c for the gauge potential A_μ is quadratic and can be integrated out:

$$I_{\rm eff}[j] \equiv -i \ln \int \mathcal{D}A^\mu \, e^{i(I_c+I_f)} = \pi\sigma \int d^3x \, d^3y \, j^\mu(x) K_{\mu\nu}(x,y) j^\nu(y) \; , \quad (41)$$

in terms of the Green function $K_{\mu\nu}(x,y)$ for the operator $\epsilon_{\mu\rho\nu}\partial_\nu$:

$$K_{\mu\nu}(x,y) = -\frac{1}{2\pi}\epsilon_{\mu\rho\nu}\frac{(x-y)^\rho}{|x-y|^3} \quad (42)$$

$$\epsilon_{\mu\nu\rho}\partial_\nu K^{\rho\sigma}(x,y) = \delta_\mu{}^\sigma \delta^{(3)}(x-y) \; . \quad (43)$$

The effective current-current interaction

$$I_{\rm eff} = \sigma \sum_{i,j} I_{ij}, \quad I_{ij} = -\frac{1}{2}\int dx_i^\mu \, dx_j^\nu \epsilon_{\mu\rho\nu}\frac{(x_i-x_j)^\rho}{|x_i-x_j|^3} \quad (44)$$

is formally identical to the interaction of the current j^μ with a Dirac magnetic monopole potential $\tilde{A}_\mu$:

$$\frac{x^\mu}{|x|^3} = \epsilon^{\mu\alpha\beta}\partial_\alpha \tilde{A}_\beta(x) \; . \quad (45)$$

It is now easy to recover the form Eq. (33) of the spin-statistics changing interaction. To this purpose, we choose an explicit "Coulomb gauge" representation for the potential $\tilde{A}_\beta(x)$:

$$\tilde{A}_\mu(t,\boldsymbol{x}) = \left(0, -\frac{\epsilon_{ab}x^b}{r(t-r)}\right), \quad r^2 = |x|^2 = t^2 - x_1^2 - x_2^2 \; , \quad (46)$$

and we parametrize paths with time, $s = t$.

We get

$$\begin{aligned} I_{ij} &= -\frac{1}{2}\int_0^T dt \int_0^T dt' \frac{dx_i^\mu(t)}{dt}\left(\partial_\mu \tilde{A}_\nu(x_i-x_j) - \partial_\nu \tilde{A}_\mu(x_i-x_j)\right)\frac{dx_j^\nu(t')}{dt'} \\ &= \int_0^T dt \, \epsilon^{ab}\left(\frac{dx_i^a}{dt} - \frac{dx_j^a}{dt}\right)\frac{(x_i(t)-x_j(t))^b}{|x_i(t)-x_j(t)|^2} + I_g \; , \end{aligned} \quad (47)$$

where I_g is a rotationally invariant surface term which has no effect on spin and statistics. Now, terms with $i = j$ in Eq. (47) vanish by antisymmetry, while terms with $i \neq j$ can be rewritten using the identity

$$\partial_a \Theta(\boldsymbol{x}) = -\epsilon^{ab}\frac{x^b}{|x|^2} \; , \quad (48)$$

which immediately implies that

$$I_{ij} = -\int dt \, \frac{d}{dt}\Theta(\boldsymbol{x}_i - \boldsymbol{x}_j) + I_g \; , \quad (49)$$

i.e., up to the irrelevant I_g, the same as the action obtained from the effective Lagrangian Eq. (33).

Summarizing, we have found that nontrivial statistics is enforced by weighing topologically inequivalent paths in the path integral, that inequivalent paths are those which correspond to interchanging the coordinates of two or more particles, and that these weights can be obtained as the result of adding to the Lagrangian an effective interaction term, which shifts the spectrum of the total angular momentum by a half-integer contribution per particle. Furthermore, in two dimensions we have obtained an explicit local representation of this effective interaction term, which is formally equivalent to the interaction of the particle current with a Dirac magnetic monopole localized on each other particle.

3 A Path Integral for Spin

Spin is usually quantized by introducing degrees of freedom which live in an internal space. In particular, the quantization of Fermions is usually performed by introducing anticommuting variables. However, in the previous section we have seen that it is possible to represent the effect of fermionic statistics in terms of an interaction defined in configuration space, and then path-integrating over this space. In this section we shall see that it is also possible to obtain the path-integral quantization of a spin degree of freedom by constructing the configuration space for a classical spin, and then path-integrating over evolutions in this configuration spaces with a suitable weight.

3.1 The Spin Action

It is well-known that the classical action for a free (relativistic) particle coincides with the arc-length L of the path $x^\mu(s)$ traversed by it, and in fact its quantization [10] can be obtained by by summing over paths with a weight given by an action which coincides with the arc-length L:

$$I = m \int ds \sqrt{\left(\frac{dx^\mu}{ds}\right)^2} = mL \; . \tag{50}$$

Hence, the quantization of the spinning particle is obtained by first defining the space of paths, and then introducing as a weight over it the simplest geometric invariant of the paths.

The path-integral quantization of spin can be obtained in a similar way. The configuration space for spin is the set of points spanned by a vector

$$\boldsymbol{s} = \sigma \boldsymbol{e} \tag{51}$$

with fixed modulus $|\boldsymbol{s}| = \sigma$, namely the two-sphere S^2. This can be viewed as the result of the action of the rotation group on a reference vector, namely, the coset of the rotation group over the subgroup of rotations that leave the reference vector invariant (little group): $S^2 = SO(3)/SO(2)$. The simplest invariant over this space is the solid angle subtended by a closed path. Therefore, parametrizing the vector $\boldsymbol{e}$ in spherical coordinates

$$\boldsymbol{e} = \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix} \tag{52}$$

we define the spin action as

$$I_s = \int dt\, \mathcal{L}(\theta,\phi) = s \int dt\, \cos\theta\dot{\phi}\,. \tag{53}$$

Equation (53) provides us with a spin action in the sense that the time-evolution (transition amplitude) for spin wave functions

$$|\phi\rangle = |m\rangle\langle m|\phi\rangle; \quad \langle m|\phi\rangle = \frac{e^{-im\phi}}{\sqrt{2\pi}} \tag{54}$$

is given by

$$\langle f|i\rangle = \langle\phi_f|e^{i\int H(t)\,dt}|\phi_i\rangle, = \int_{\boldsymbol{e}(t_f)=\boldsymbol{e}(\phi_f);\,\boldsymbol{e}(t_i)=\boldsymbol{e}(\phi_i)} D\boldsymbol{e}\quad e^{i\int dt\,\mathcal{L}_s - V(\boldsymbol{J})}\,, \tag{55}$$

where $H(t)$ is a Hamiltonian which describes the spin dynamics (e.g. the coupling with an external magnetic field, $H = \boldsymbol{s}\cdot\boldsymbol{B}$) and the boundary conditions are given in terms of ϕ only (which is equivalent to specifying an eigenvalue m of the third component of angular momentum). This result can be proven by direct computation [5, 6]. We shall instead first show that the action I_s Eq. (53) leads to the correct classical dynamics of spin, then quantize it using the general principles of geometric quantization.

Let us first take a closer look at the spin action. Its geometric interpretation becomes apparent by rewriting it as

$$I_s = \sigma\int_C \cos\theta\dot{\phi}dt = \sigma\int_C \cos\theta d\phi \tag{56}$$

$$= \sigma\int_S d\cos\theta d\phi = \sigma\int_S d\boldsymbol{S}\cdot\boldsymbol{e} = \sigma\int_S \left(\frac{\partial\boldsymbol{e}}{\partial s}\times\frac{\partial\boldsymbol{e}}{\partial t}\right)\cdot\boldsymbol{e}\,, \tag{57}$$

where C is the path travsersed by the vector $\boldsymbol{e}$ Eq. (52), and the second step Eq. (57) holds when the path is closed, in which case S is th surface bound by C. In such case, Eq. (57) shows explicitly that the spin action coincides with the solid angle subtended by C.

Equation (57) shows manifestly that there is a 4π ambiguity in the definition of the spin action, which corresponds to the possibility of choosing the

upper or lower solid angle subtended by C on the sphere. In order for this ambiguity to be irrelevant, the parameter σ, which as we shall see corresponds to the value of spin, must be quantized in half-integer units. The connection between the spin action and the effective two-dimensional statistics action Eq. (44) becomes clear by rewriting it as

$$I_s = \sigma \int_S d\boldsymbol{S} \cdot \boldsymbol{\nabla} \times \tilde{\boldsymbol{A}}[\boldsymbol{e}] \,, \tag{58}$$

where $\tilde{A}$ is the Dirac monopole potential Eq. (45), in the space of spin vectors:

$$\boldsymbol{e} = \boldsymbol{\nabla} \times \tilde{\boldsymbol{A}}[\boldsymbol{e}] \tag{59}$$

3.2 Classical Dynamics

In order to verify that the spin action defines the action for a classical spin degree of freedom, we first check that it leads to the Poisson bracket

$$\{s^i, s^j\} = \epsilon^{ijk} s^k \,. \tag{60}$$

This is easily done using the Faddeev-Jackiw formalism [7] for the Hamiltonian treatment of systems defined by first-order Lagrangians, i.e. by a Lagrangian of the form

$$L = f_i(x) \frac{dx_i}{dt} - V(x) \,. \tag{61}$$

Namely, it easy to see that the Euler-Langrange equations implied by the Langrangian Eq. (61) have the form

$$f_{ij} \frac{dx_j}{dt} = \frac{\partial V}{\partial x_i} \tag{62}$$

$$f_{ij} \equiv \frac{\partial f_j}{\partial x_i} - \frac{\partial f_i}{\partial x_j} \,. \tag{63}$$

This coincides with the canonical Hamiltonian form

$$\frac{dx^i}{dt} = \{x^j, x^i\} \frac{\partial V}{\partial x^j} = \{V, x^i\} \tag{64}$$

if the Poisson brackets are given by

$$\{x^i, x^j\} = (f^{-1})^{ji} \tag{65}$$

It can be shown that the same result is found in the more standard approach, where the Lagrangian Eq. (61) is viewed as defining a constrained dynamics, which is then treated defining suitable Dirac brackets.

Specializing this formalism to the spin action we see that its Dirac monompole form Eq. (58) has the form of Eq. (61) with

$$f_i = \sigma \tilde{A}_i[\boldsymbol{e}] \ . \tag{66}$$

Using the definition Eq. (63) this leads to

$$f_{ij} = \sigma \left(\partial_i \tilde{A}_j - \partial_j \tilde{A}_i \right) = \sigma \epsilon^{ijk} e^k \ . \tag{67}$$

Because

$$f_{ij}^{-1} = \frac{1}{\sigma^2} f_{ij} \tag{68}$$

the Poisson Brackets are

$$\{e^i, e^j\} = \frac{1}{\sigma} \epsilon^{ijk} e^k \ , \tag{69}$$

which, identifying the spin vectore $\boldsymbol{s}$ with

$$\boldsymbol{s} = \sigma \boldsymbol{e} \ , \tag{70}$$

immediately lead to the spin Poisson brackets Eq. (60). This also shows that the parameter σ gives the value of spin.

3.3 Geometric Quantization

The spin action can be quantized using the formalism of geometric or "coadjoint orbit" quantization [8, 9]. Namely, we view the time evolution of the (unit) spin vector $\boldsymbol{e}(t)$ as the result of the action of a rotation matrix $\Lambda(t)$ on a reference vector $\boldsymbol{e}_0$:

$$\boldsymbol{e}(t) = \Lambda(t) \boldsymbol{e}_0 \ . \tag{71}$$

This defines a path (orbit) in $S^2 = SO(3)/SO(2)$, where $SO(2)$ is the little group of $\boldsymbol{e}_0$ (the set of Λ matrices which leaves $\boldsymbol{e}_0$ invariant).

The path in S^2 can be lifted to a path in $SO(3)$ by assigning a frame, e.g. by defining the vector

$$\boldsymbol{n}(t) \equiv \frac{\dot{\boldsymbol{e}}(t)}{|\dot{\boldsymbol{e}}(t)|} \tag{72}$$

which satisfies

$$\boldsymbol{n} \cdot \boldsymbol{e} = 0 \ . \tag{73}$$

The triple $\boldsymbol{e}$, $\boldsymbol{n}$, and

$$\boldsymbol{b}(t) \equiv \boldsymbol{e}(t) \times \boldsymbol{n}(t) \ . \tag{74}$$

defines a time-dependent frame, which coincides with the standard Frenet frame if $\boldsymbol{e}(t)$ is viewed as the tangent vector to some path, in which case $\boldsymbol{n}$ and $\boldsymbol{b}$ are the unit normal and binormal, respectively. The matrix Λ is then fully specified by Eq. (71) and

$$\boldsymbol{n}(t) = \Lambda(t) \boldsymbol{n}_0 \ . \tag{75}$$

It is convenient in particular to choose the set of reference vectors

$$\begin{pmatrix} \boldsymbol{v}^{(3)_0} = \boldsymbol{e}_0 \\ \boldsymbol{v}_0^{(1)} = \boldsymbol{n}_0 \\ \boldsymbol{v}_0^{(2)} = \boldsymbol{b}_0 \end{pmatrix} \tag{76}$$

as

$$v_0^{(a)}{}_i = \delta_i^a \ . \tag{77}$$

It is then easy to see that the quantity

$$\left(\Lambda^{-1}\dot{\Lambda}\right)^{ij} = \boldsymbol{v}^{(i)} \cdot \dot{\boldsymbol{v}}^{(j)} \tag{78}$$

is an element of the $SO(3)$ algebra, the so–called Maurer-Cartan form, given by

$$\left(\Lambda^{-1}\dot{\Lambda}\right)_{ij} = \sum_{ab} C_{ab}(M^{ab})_{ij}; \quad (M^{ab})_{ij} = \left(\delta_i^a\delta_j^b - \delta_j^a\delta_i^b\right) \ . \tag{79}$$

The coefficients C_{ij} can be extracted by exploiting the fact that the generators are orthogonal under tracing:

$$C_{ij} = \frac{1}{4}\,\mathrm{tr}\,\left(M_{ij}\Lambda^{-1}\dot{\Lambda}\right) = \frac{1}{2}\boldsymbol{v}^{(i)} \cdot \dot{\boldsymbol{v}}^{(j)} \ . \tag{80}$$

We can now use this gometric formalism to rewrite yet again the spin action Eq. (53) as

$$\begin{aligned} I_s &= \sigma \int_S \left(\frac{\partial \boldsymbol{e}}{\partial s} \times \frac{\partial \boldsymbol{e}}{\partial t}\right) \cdot \boldsymbol{e} = \sigma \int dt\ \dot{\boldsymbol{b}} \cdot \boldsymbol{n} + \text{integers} \\ &= \sigma \left(\mathrm{tr} \int dt\, \frac{1}{2}\left(\Lambda^{-1}\dot{\Lambda} M_{12}\right) + \text{integers}\right) \ . \end{aligned} \tag{81}$$

Note that any spin-dependent potential $V(\boldsymbol{\sigma})$ can be re-written in terms of Λ by exploiting Eq. (80) to express the spin vector $\boldsymbol{e}$ in terms of Λ:

$$e^i = \sigma\epsilon^{ijk}\left(\Lambda^{-1}\frac{M_{12}}{2}\Lambda\right)_{jk} \ . \tag{82}$$

This new form Eq. (81) of the spin action has a twofold advantage: first, it does not depend on the representation, and second, it is amenable to geometric quantization. To demonstrate its representation-independence, let us show how the spinor representation is recovered from it. For spin $\frac{1}{2}$, the generators are

$$M_{ij} = -i\epsilon^{ijk}\sigma_k \ , \tag{83}$$

where σ_i are the usual Pauli matrices. We then have

$$\mathrm{tr}\,\frac{1}{2}\left(\Lambda^{-1}\dot{\Lambda}M_{12}\right) = \mathrm{tr}\,\left(\Lambda^{-1}\dot{\Lambda}\frac{\sigma_3}{2i}\right) = \mathrm{tr}\,\left(\Lambda^{-1}\dot{\Lambda}\left(\frac{\mathbb{1} + \sigma_3}{2i}\right)\right) \ . \tag{84}$$

The connection to (Pauli) spinors is found by introducing the reference two-component spinor

$$\psi_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} , \tag{85}$$

upon which the matrix Λ is taken to act in the spinor representation, namely

$$\psi(t) = T[\Lambda(t)]\psi_0 , \tag{86}$$

where $T[\Lambda(t)]$ is the spinor representation of the rotation Λ. The relation between the spinor and vector representation is provided by constructing spinor bilinears

$$\psi^* \sigma^i \psi = \Lambda_{ij} \psi_0^* \sigma^i \psi_0 = \Lambda_{ij} e_0^j . \tag{87}$$

Using this relation, and noting that

$$|\psi_0\rangle\langle\psi_0| = \left(\frac{\mathbb{1} + \sigma_3}{2} \right) \tag{88}$$

it is easy to rewrite the spin action I_s Eq. (81) as

$$I_s = \frac{1}{2} \int \frac{dt}{i} \, \psi^*(t) \frac{d}{dt} \psi(t) , \tag{89}$$

which has the form of the kinetic term for a Pauli spinor $\psi(t)$. A generic spin-dependent potential $V(\boldsymbol{\sigma})$ can be written in terms of ψ by using the relation

$$\boldsymbol{e} = \psi^*(t) \boldsymbol{\sigma} \psi(t) . \tag{90}$$

The quantization of spin is now reduced to the general problem [9] of quantizing a system whose confiuration space is the space of states $|\psi\rangle$ which are orbits of a group G:

$$|\phi\rangle = T(g)|\phi_0\rangle , \tag{91}$$

where $g \in G$ is an element of the group of which $T(g)$ provides a unitary representation. The axioms of quantum mechanics imply that transition amplitudes for this system are given by the path integral

$$\langle f|i\rangle = \int Dg \, e^{iI_w[g]} \tag{92}$$

with the action

$$I_w[g] = \int dt \, \langle\phi_0| \left[T(g^{-1}(t)) \frac{d}{idt} T(g(t)) - H(g(t)) \right] |\phi_0\rangle , \tag{93}$$

where H is a generic spin-dependent potential (or Hamiltonian, which coincides with the potential for a first-order Lagrangian).

The spin action Eq. (81) is seen to coincide with the kinetic term of the geometric action Eq. (93) if one identifies the representation matrix $T(g)$

with Λ Eq. (71), and one observes that the projector on the state $|\phi_0\rangle$ can be expressed in terms of the generator $C^0_{ij}M^{ij}$ of the little group of $|\phi_0\rangle$:

$$|\phi_0\rangle\langle\phi_0| = C^0_{ij}M^{ij} \ . \tag{94}$$

Indeed, we get

$$\int dt\, \langle\phi_0| \left[T(g^{-1}(t))\frac{d}{idt}T(g(t))\right] |\phi_0\rangle = \int dt\ \mathrm{tr}\, C^0_{ij}M^{ij}\Lambda^{-1}(t)\dot{\Lambda}(t) \tag{95}$$

which coincides with the spin action if we choose $C^0_{ij}M^{ij} = M_{12}$. Hence, the spin path integral Eq. (55) with the spin action Eq. (53) follows from geometric quantization of the space of $SO(3)$ orbits.

The relation of this result to the usual sum over paths à la Feynman is apparent if we specialize again to the case of spin $\frac{1}{2}$. The sum over paths is performed by dividing the time evolution from t_i to t_f into discrete time steps $\Delta t = \frac{t_f - t_i}{N}$ so that $t_j = t_i + (j-1)\Delta t$, and then letting $N \to \infty$. For a spin system we get

$$\langle f|i\rangle = \langle\psi_f|e^{-i\int_{t_i}^{t_f} H\, dt}|\psi_i\rangle = \prod_{j=1}^{N}\int d\Lambda_j \langle\psi_{j+1}|e^{-i\Delta t H(t_j)}|\psi_j\rangle \ . \tag{96}$$

The evolution along an infinitesimal time slice is then given by

$$\begin{aligned}\langle\psi_{j+1}|e^{-i\Delta t H(t_i)}|\psi_j\rangle &\approx \langle\psi_{j+1}|\left(1 - i\Delta t H(t_j)\right)|\psi_j\rangle \\ &= 1 + \frac{1}{2}\Delta t\psi^*\frac{d}{dt}\psi - i\Delta t H(t_j) \\ &\approx e^{i\left[\psi^*\frac{d}{i dt}\psi - \Delta t H(t_i)\right]} \ ,\end{aligned} \tag{97}$$

which coincides with the geometric quantization result Eq. (93). The first-order quantization of spin is a simple consequence of the fact that a spin Hamiltonian does not contain a quadratic kinetic term: the action is then entirely determined by the first-order parallel transport of the spin vector.

The meaning of these results is that first, the probabilty for the time evolution between two spin states is given by

$$\langle f|i\rangle = \int_{\boldsymbol{e}_f(t_f)=\boldsymbol{e}_f(\Lambda_f);\, \boldsymbol{e}_i(t_i)=\boldsymbol{e}_i(\Lambda_i)} D\boldsymbol{e}\quad e^{i[I_s[e]-\int dt\, H(t,\boldsymbol{e})]} \tag{98}$$

and furthermore, the matrix element of any spin-dependent operator $F(\boldsymbol{\sigma})$ can be determined as

$$\langle f|F(\boldsymbol{\sigma})|i\rangle = \int_{\boldsymbol{e}_f(t_f)=\boldsymbol{e}_f(\Lambda_f);\, \boldsymbol{e}_i(t_i)=\boldsymbol{e}_i(\Lambda_i)} D\boldsymbol{e}\quad e^{i[I_s[e]-\int dt\, H(t,\boldsymbol{e})]}F(\boldsymbol{e}) \ . \tag{99}$$

Summarizing, we have seen that the path-integral quantization of a "static" spin degree of freedom – as e.g. in the Heisenberg model – can be given in terms of a geometrically determined first-order spin action. The usual formalism in the spin-$\frac{1}{2}$ case is obtained by specializing to the spinor representation of spin vectors, but it does not require anticommuting variables or relativity.

4 Relativistic Spinning Particles

As discussed in the introduction, in a relativistic theory physical states are irreducile representations of the Poincaré group, i.e. they carry mass and spin: the one-particle state $|m,s\rangle$ satisfies

$$P^2|m,s\rangle = m^2|m,s\rangle; \quad W^2|m,s\rangle = m^2 s(s+1) \,, \tag{100}$$

where the Pauli-Lubanski operator W, defined in Eq. (7), generates Lorentz tranformations which leave the particle momentum invariant, because by construction $W_\mu P^\mu = 0$. In particluar, in the rest frame of the particle (for massive particles) $p = (m, \mathbf{0})$, so $W = (0, \boldsymbol{s})$. In a general frame, spin spans the three dimensional ($d-1$ dimensional) space orthogonal to momentum. This introduces a coupling between spin and momentum which determines the dynamics of a relativistic spinning particle, both at the classical and quantum level.

4.1 Path Integral for Spinless Particles

Before discussing the quantization of spinning particles, let us review the path-integral quantization of a massive spinless particle [10]. As we mentioned already, the action Eq. (50) of a spinless free particle, or the kinetic term in the action for an interacting spinless particle, coincides with the arc-length of the path traversed by the particle. This can be written in various equivalent ways: the simple integral of the arc-length element $ds = \sqrt{dx^\mu dx_\mu}$ Eq. (50) can be rewritten in terms of an induced metric $g(s)$ along the path

$$I_0 = \int ds \left[\frac{1}{\sqrt{g}^{\frac{1}{2}}} \left(\frac{dx^\mu}{ds} \right)^2 + \frac{m^2}{2} \sqrt{g} \right] \,. \tag{101}$$

Both at the classical and at the quantum level, the equation of motion for g is the constraint

$$g = \frac{\dot{x}^2}{m^2} \,, \tag{102}$$

which shows that indeed $g(s)$ is the induced metric

$$dx^2 = g(s) ds^2 \,, \tag{103}$$

and leads back to the original form Eq. (50) of the action when substituted in Eq. (101).

The action Eq. (101) can in turn be rewritten in first-order form

$$I_0 = \int ds \left[p_\mu \frac{dx^\mu}{dt} - \frac{\sqrt{g}}{2} \left(p^2 - m^2 \right) \right] \,, \tag{104}$$

where the momentum (tangent vector) p^μ is also fixed by a constraint

$$p^\mu = \frac{1}{\sqrt{g}}\dot{x}^\mu \tag{105}$$

which again leads back to the original form Eq. (101) when subsitituted in the action Eq. (104). This first-order form of the action is the most suitable for geometric quantization, i.e. for describing the dynamics of the spinning particle similarly to the way we have described the dynamics of spin in Sect. 3.3. The classical equations of motion can be obtained from any of these equivalent forms of the action, and express energy-momentum conservation. For instance, using the first-order form Eq. (104) we get immediately the Euler-Lagrange equations

$$\frac{d}{dt}p^\mu = 0, \quad p^2 = m^2 \ . \tag{106}$$

Path-integral quantization [10] can be perfomed by exploiting the "gauge invariance", i.e. the reparametrization invariance of the system [8]. The (Euclidean) path integral

$$\langle x'|x\rangle = \mathcal{N}\int_{x(0)=x;\, x(1)=x'} Dx(s)\ e^{-m\int_0^1 ds\,\sqrt{\dot{x}^2}} \tag{107}$$

can be rewritten introducing the induced metric $g(s)$ Eq. (102) as

$$\langle x'|x\rangle = \mathcal{N}\int_{x(0)=x;\, x(1)=x'} Dx(s)Dg(s)\ \ \delta^{(\infty)}\big(\dot{x}^2 - g\big)\, e^{-m\int_0^1 ds\,\sqrt{g}} \ . \tag{108}$$

Reparametrization invariance is now manifest, because upon a general reparametrization $s \to f(s)$, the metric $g(s)$ transforms as $g(s) \to g(f(s))[\dot{f}(s)]^2$. We can now perform the path integral by fixing the gauge, e.g. by imposing the condition

$$\dot{g}(s) = 0 \ . \tag{109}$$

Because the path-length is

$$L = \int_0^1 ds\,\sqrt{\dot{x}^2} = \int_0^1 ds\,\sqrt{g(s)} \tag{110}$$

the gauge conditionEq. (109) implies

$$g(s) = L^2 \ . \tag{111}$$

We can thus write the gauge-fixed path-integral as

$$\begin{aligned}\langle x'|x\rangle &= \mathcal{N}\int_0^\infty dL \int_{x(0)=x;\, x(1)=x'} Dx(s)Dg(s)\ \ \delta^{(\infty)}\big(\dot{x}^2 - g\big)\,\delta\big(g - L^2\big)\, e^{-mL} \\ &= \mathcal{N}\int_0^\infty dL \int_{x(0)=x;\, x(1)=x'} Dx(s)\ \delta^{(\infty)}\left(\dot{x}^2 - L^2\right) e^{-mL} \ .\end{aligned} \tag{112}$$

After gauge-fixing, a residual integration over path lengths L remains.

The path-integral can be re-written in terms of geometric variables along the path: this leads to geometric quantization again. We introduce a tangent vector along the path, which for classical paths (those which satisfy the Euler-Lagrange equations) coincides with the particle four-momentum:

$$e^\mu = \frac{\dot{x}^\mu}{|\dot{x}|} = \frac{\dot{x}^\mu}{L} \ . \tag{113}$$

We can replace the path-integration over trajectories by a path-integration over the tangent vectors e^μ. However, the boundary conditions now become a non-local constraint:

$$x^{\mu\prime} - x^\mu = \int_0^L ds\, e^\mu(s) \ . \tag{114}$$

We thus get finally

$$\begin{aligned}\langle x'|x\rangle &= \mathcal{N}\int_0^\infty dL \int De(s)\, e^{-mL}\delta^{(\infty)}\big(e^2-1\big)\,\delta^{(3)}(x^{\mu\prime} - x^\mu - \int_0^L ds\, e^\mu(s)) \\ &= \mathcal{N}\int dL\, d\boldsymbol{p} \int De(s)\, e^{-mL}\delta^{(\infty)}\big(e^2-1\big)\, e^{ip\cdot\left(x'-x-\int_0^L ds\, e(s)\right)} \ . \end{aligned} \tag{115}$$

The usual expression of the bosonic (Klein-Gordon) propagator is obtained by regularizing the formal expression Eq. (115). To this purpose, we cut off paths which are coarse on a scale $\sim\epsilon$ (where, of course ϵ has the dimensions of [length]). We then take the continuum limit with a mass renormalization condition, expressed by defining a renormalized mass $M_{\rm phys}$ such that

$$m \propto \varepsilon M_{\rm phys}^2 \ . \tag{116}$$

The propagator $K(p)$ is obtained as the Fourier transform of the renormalized position-space amplitude:

$$\begin{aligned} K(p) &= \lim_{\varepsilon\to 0}\mathcal{N}\int dL\, e^{-mL}\int De(s) e^{-\frac{\varepsilon}{2}\int_0^L ds\,\dot{e}^2} e^{-ip\cdot\int_0^L ds\, e(s)}\delta^{(\infty)}\big(e^2-1\big) \\ &= \mathcal{N}\int dL\, e^{-L\varepsilon M_{\rm phys}^2} e^{-L\varepsilon p^2} = \mathcal{N}\frac{1}{p^2+M_{\rm phys}^2} \ . \end{aligned} \tag{117}$$

Up to the irrelevant albeit infinite normalization constant $\mathcal{N}$, we have thus recovered the standard form of the Klein-Gordon propagator.

4.2 The Classical Spinning Particle

The spinning particle is now obtained by coupling a spin degree of freedom to the spinless particle of Sect. 4.1, with dynamics governed by the action discussed in Sect. 3.1. This can be done in an elegant geometric way by combining the translational and spin configuration spaces. To this purpose,

in one time and $d-1$ space dimensions, we define a set of $d-1$ orthonormal vectors e^μ, $n^\mu_{(1)},\ldots,n^\mu_{(d-2)}$, which can in turn be obtained by action of a Lorentz transformation matrix Λ on a set of reference vectors

$$\begin{cases} e^\mu & = & \Lambda^\mu{}_\nu \hat{t}^\nu \\ n_{(i)}{}^\mu & = & \Lambda^\mu{}_\nu \hat{n}_{0\,(i)}{}^\mu \end{cases} . \tag{118}$$

The reference vectors

$$\hat{t}^\mu = \begin{pmatrix} 1 \\ \mathbf{0} \end{pmatrix}, \qquad \hat{n}_0^{(i)\mu} = \delta_i{}^\mu \tag{119}$$

define a basis in one time and $d-1$ space dimensions. The set of vectors e^μ, $n^\mu_{(i)}$ completely specifies the matrix Λ: indeed, the first vector has $d-1$ independent components (being unimodular), the second, orthogonal to it, has $d-2$ independent components and so on, so that overall they have $\sum_{i=0}^{d-2}(d-i-1) = \frac{1}{2}d(d-1)$ independent components, like the $O(d-1,1)$ matrix Λ.

In the four-dimensional case we are interested in, the matrix Λ has six independent components. We take the vector e^μ as the unit tangent to the particle trajectory, so that classically is is identified with momentum up to an overall factor of m:

$$e^\mu = \frac{\dot{x}^\mu}{|\dot{x}|}; \quad p^\mu = m e^\mu \tag{120}$$

and at the quantum level it is the variable one path-integrates over (compare Eq. (115)). The vector $n^\mu_{(1)}$ is then identified with the spin vector discussed in the previous section, it has two independent components and lives in the S^2 orthogonal to e^μ:

$$e_\mu n^\mu_{(1)} = 0; \quad s^\mu = \sigma n^\mu_{(1)} \tag{121}$$

At the quantum level, the two independent vectors p^μ and s^μ entirely specify the configuration of the system, whereas at the classical level the canonical coordinate x^μ must also be given.

The action for the spinning particle is now simply obtained by combining the action for the spinless particle Eq. (104) with the spin action Eq. (81): by writing both in terms of Λ, the momentum-spin orthogonality constraint is automatically enforced. We get

$$I = \int ds \left[p_\mu \frac{dx^\mu}{dt} - \frac{\sqrt{g}}{2}\left(p^2 - m^2\right)\right] + \sigma \,\mathrm{tr}\left(\Lambda^{-1}\dot{\Lambda} M_{12}\right) . \tag{122}$$

It is straightforward to check that, at the classical level, the correct dynamics is obtained: the Euler-Lagrange equations are found by varying the action upon the most general Poincaré transformation, namely a translation of x^μ, and a Lorentz transformation of Λ. The variation upon translations gives trivially the spinless equation of motion Eq. (106) (energy-momentum conservation). The most general Lorentz variation is

$$\delta\Lambda = i\omega^{\mu\nu} M_{\mu\nu}\Lambda \,, \tag{123}$$

upon which the action transforms as

$$\delta I = -i \,\mathrm{tr}\, (\omega^{\mu\nu} M_{\mu\nu} K) + i\sigma \,\mathrm{tr}\, \left(S \frac{d}{dt} \omega^{\mu\nu} M_{\mu\nu} \right) \tag{124}$$

$$K_{\mu\nu} \equiv (\dot{x}_\mu p_\nu - x_\nu \dot{p}^\nu) \tag{125}$$

$$S_{\mu\nu} = \sigma \left(\Lambda^{-1} M_{12} \Lambda \right)_{\mu\nu} \,. \tag{126}$$

Demanding that the action be stationary leads to the Euler-Lagrange equations

$$\frac{d}{dt} \left(x^\mu p^\nu - x^\nu p^\mu + S^{\mu\nu} \right) = 0 \,. \tag{127}$$

Equation (127) expresses the set of conservation laws of a Lorentz invariant Lagrangian: in particular, the (i, j) components give the conservation of (total) angular momentum, while the $(0, i)$ components give the equation $\boldsymbol{p} = \frac{d}{dt} (\boldsymbol{x} E)$ which relates momentum to velocity in the usual way.

4.3 Quantum Spinning Particles and Fermions

The dynamics of the spinning particle, described by the action Eq. (122), is given on the space of Lorentz orbits $\Lambda(t)$ which evolve according to Eq. (118) the pair of vectors p^μ Eq. (120), s^μ Eq. (121). The path integral then follows from geometric quantization [9] Eqs. (92,93):

$$\langle x', \boldsymbol{s}' | x, \boldsymbol{s} \rangle = \int d\boldsymbol{p}\, e^{ip\cdot(x'-x)} \int dL\, e^{-mL} \int D\Lambda(s) e^{-i\int_0^L ds \left[p\cdot\Lambda\hat{t} - \sigma \,\mathrm{tr}\, \left(\Lambda^{-1}\dot{\Lambda} M_{12} \right) \right]} \,. \tag{128}$$

In practice, the path integral is found by combining the spin path integral Eq. (98) and the spinless particle path integral Eq. (115).

Let us now discuss in particular the spin-$\frac{1}{2}$ case in the spinor formulation, and show how the Dirac equation is recovered. We can do this promoting to the Lorentz group the connection between spinor and vector representations of the rotation group Eq. (90). This is based on the transformation law of Dirac matrices, which connect the four-vector representation Λ of the Lorentz group with the corresponding spinor representation $T(\Lambda)$:

$$T(\Lambda^{-1}) \gamma^\mu T(\Lambda) = \Lambda^\mu{}_\nu \gamma^\nu \,. \tag{129}$$

Now, it is easy to show that given an unimodular vector v^μ, the spinor ψ such that

$$\psi^* \gamma^\mu \psi = v^\mu \tag{130}$$

satisfies the condition

$$v_\mu \gamma^\mu \psi = \psi \tag{131}$$

(in Euclidean space, in Minkowski space the spinor ψ^* must be replaced by $\bar{\psi} \equiv \psi^* \gamma^0$).

In our case, we associate to the one-particle state with normalized momentum e^μ the spinor $\psi[e^\mu]$ which satisfies the condition

$$p_\mu \gamma^\mu \psi = m\psi \ , \tag{132}$$

i.e. the Dirac equation. In practice, we can determine $\psi[e^\mu]$ by acting with the spinor representation $T(\Lambda)$ of the transformation Λ Eq. (118)

$$\psi = T(\Lambda)\psi_0 \tag{133}$$

on the reference spinor ψ_0 such that $\psi_0^* \gamma^\mu \psi_0 = \hat{t}^\mu$, i.e. (using Eq. (131) such that

$$\gamma^0 \psi_0 = \psi_0 \ . \tag{134}$$

If one uses the so-called Dirac representation for the γ matrices, $\gamma^0 = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}$ (where each entry is a 2×2 block), so

$$\psi = \begin{pmatrix} \phi \\ 0 \end{pmatrix} \ , \tag{135}$$

where ϕ is any two-component spinor.

The condition that the spin vector be given by s^μ fixes entirely the spinor (up to an overall U(1) phase): if Λ is such that $\Lambda^\mu_\nu s_0^\nu = s^\mu$, then, choosing according to Eq. (119) $s_0^\nu = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$, the spinor ψ is given by

$$\psi = T(\Lambda)\psi_0; \qquad \phi_0 \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix} \ . \tag{136}$$

It is easy to see that the spinor constructed in this way is an eigenstate of the projection of the Pauli-Lubanski operator along the spin vector $s^\mu = \frac{1}{2} n^\mu$:

$$W^\mu n_\mu \psi = \pm \frac{m}{2} \psi(p, s) \ . \tag{137}$$

This is obvious in the rest frame, because then $W^\mu s_\mu = m s_i \epsilon^{ijk} \sigma_{jk} = m \boldsymbol{s} \cdot \boldsymbol{\sigma}$, where $\boldsymbol{\sigma}$ are Pauli matrices, and Eq. (136) together with the relation between the spin vector and Pauli matrices Eq. (90) implies that

$$\boldsymbol{s} \cdot \boldsymbol{\sigma} \phi = \pm \tfrac{1}{2} \phi \ . \tag{138}$$

In other words, in the rest frame $W^\mu s_\mu$ is just the standard spin operator and ϕ is the two-component spinor Eq. (86) discussed in Sect. 3.3. But since in the rest frame the low components of ψ Eq. (135) vanish, this implies

$$W^\mu s_\mu \psi = \pm \tfrac{1}{2} m \psi \ . \tag{139}$$

In a generic frame, the four-vector s^μ is boosted by Λ, so

$$W^\mu n_\mu = W^\mu {\Lambda^{-1}}^\nu{}_\mu n_\mu = T(\Lambda) W^\mu T^{-1}(\Lambda) n_\mu \ , \tag{140}$$

but so is the spinor ψ in such a way that [Eq. (133)] the eigenvector condition still holds:

$$T(\Lambda) W^\mu T^{-1}(\Lambda) n_\mu T(\Lambda) \psi 0 = \pm \frac{m}{2} T(\Lambda) \psi_0 \tag{141}$$

Let us now consider the propagator $K(p)$, i.e. momentum-space path integral, related by Fourier transformation to the path-integral Eq. (128). We have found that in the spin-$\frac{1}{2}$ case, if the spinor representation is adopted, states along the path are instantaneous eigenstates of $e_\mu \gamma^\mu$, according to Eq. (132). In follows that momentum eigenstates, which are the boundary conditions to the momentum-space path-integral (i.e. states of definite e^μ) automatically satisfy the Dirac equation. Furthermore, the spinor states satisfy

$$\psi^* \gamma^\mu \psi = e^\mu \ , \tag{142}$$

i.e., e^μ is obtained by acting with γ^μ on the instantaneous spinor states along the path. But in Sect. 3.3 we have proven [Eq. (99)] that the expectation value of any function $F(\sigma)$ can be obtained by path-integration of the function $F(\Lambda)$ with a weight given by the spin action itself. Applying this in reverse, we see that averaging with the spin action produces the same result as taking matrix element of instantaneous (path-otrdered) functions of γ^μ, where γ^μ is identified with e^μ thanks to Eq. (142).

The propagator is therefore given by

$$\begin{aligned} K(p) &= \int dL\, e^{-mL} \int D\Lambda(s) e^{-i \int_0^L ds \left[p_\mu e^\mu - \sigma \,\mathrm{tr}\left(\Lambda \dot{\Lambda} M_{12} \right) \right]} \\ &= \int dL\, e^{-mL} e^{-iL p_\mu \gamma^\mu} \\ &= \frac{1}{\not{p}_\mu + m} \ , \end{aligned} \tag{143}$$

i.e. the usual Dirac form.

The link with Fermi statistics is understood by observing that the spin factor upon 2π rotation transforms as

$$\mathrm{tr}\left(\Lambda^{-1} \dot{\Lambda} R M_{12} R^{-1} \right) = R^i_j \hat{z}^j \epsilon^{ijk} \,\mathrm{tr}\left(\Lambda^{-1} \dot{\Lambda} M_{jk} \right) \tag{144}$$

so if $\sigma = \frac{1}{2}$ the path-integral Eq. (128) acquires a phase $e^{i\pi} = -1$. In the more conventional approach, this follows from the anticommuting properties of the γ matrices, and it requires anticommuting (Grassmann) variables. In the geometric approach which we have followed this is not necessary, because the

anticommutation properties follows automatically from the fact that physical states are localized on paths (so ordering along the path is enforced), and paths are given weights that transform nontrivially upon rotations. This provides an explicit realization of the general spin-statistics relation derived in Sect. 2: once spin is obtained as a consequence of an interaction defined in configuration space, the link with statistics follows from the fact that particle interchange can be perfomed by 2π rotation.

Finally, it is interesting to observe that the dynamical coupling of spin and momentum which follows from the geometric interpretation of spin as a vector in the space which is orthogonal to momentum actually changes the nature of the sum over paths: the Hausdorff dimension of paths d_h that contribute to the regularized and renormalized Euclidean path integral in the continuum limit is not the same for Bose and Fermi particles [11]. The Hausdorff dimension relates the typical length scale L of paths which dominate the propagator in the continuum limit to the momentum p which is propagated:

$$L \sim p^{d_H} \tag{145}$$

It can be proven that $d_H = 2$ for Bosons while $d_H = 1$ for Fermions [11]. A rough and ready way to see this is to compare the bosonic propagator Eq. (117) and the fermionic propagator Eq. (143): it appears that the scaling limit requires taking $Lm^a \sim$ constant with $a = 2$ for Bosons and $a = 1$ for Fermions. This means that Bosonic paths are coarser then Femionc paths: Bosonic propagation is an ordinary random walk (like Brownian motion), whereas Fermionic propagation is a directed random walk, essentially because the spin interaction quenches fluctuations of the tangent vector to the path.

5 Conclusion

The discussion of spin presented in these lectures was rooted in quantum mechanics, and has used few field-theoretic concepts. Yet, we have been able to derive many results which usually require the full framework of relativistic quantum field theory: the spin-statistics connection, multivalued spin wave functions, the spin propagator, the Dirac equation. In fact, we have shown that the quantization of spin both in a nonrelativistic and a relativistic setting follows from general properties of the configuration space for orbits of the rotation group, viewed as a subgroup of the Galilei or Poincaré group, respctively. It thus appears that the standard field-theoretic approach is is merely a convenient way of achieving the quantization of systems of elementary excitations which provide irreducible representations of the Galilei or Poincaré group, because field theory automatically combines quantum mechanics with the relevant symmetry group in a local, unitary way. Of course, the standard field-theoretic approach, with anticommuting variables and spinors, is by far more convenient for the sake of practical computations. However, we have

attempted to show that the origin of the quantum field theoretic features of spin in the way symmetry is realized in quantum mechanics.

Acknowledgement

I thank Walter Pötz for giving me the opportunity to present these lectures in a stimulating interdisciplinary environment, and all those who attended the lectures for their interest and enthousiasm.

References

1. See e.g. R. Jackiw: Topological Investigations of Quantized Gauge Theories, in: S. Trieman, R. Jackiw, B. Zumino and E. Witten, *Current Algebra and Anomalies*, (World Scientific, Singapore 1985) pp. 211–361
2. See e.g. W.I. Fushchich, A.G. Nikitin: *Symmetries of Equations of Quantum Mechanics* (Allerton, New York, 1994)
3. M. G. Laidlaw and C. M. DeWitt, Phys. Rev. D **3**, 1375 (1971)
4. See e.g. S. Forte, Rev. Mod. Phys. **64**, 193 (1992)
5. H. B. Nielsen and D. Rohrlich, Nucl. Phys. B **2**99, 471 (1988)
6. K. Johnson, Annals Phys. **192**, 104 (1989)
7. L. D. Faddeev and R. Jackiw, Phys. Rev. Lett. **60**, 1692 (1988)
8. A. M. Polyakov, *Gauge Field and Strings*, (Harwood, 1987); Two-Dimensional Quantum Gravity. Superconductivity At High T_c, in: *Fields, Strings and Critical Phenomena*, E. Brézin and J. Zinn-Justin, eds. (North-Holland, 1990) pp. 305–368
9. P. B. Wiegmann, Nucl. Phys. B **323**, 311 (1989)
10. R. P. Feynman and A. R Hibbs, *Quantum Mechanics and Path Integrals*, (McGraw-Hill, 1965)
11. T. Jaroszewicz and P. S. Kurzepa, Annals Phys. **210**, 255 (1991)

Nucleon Spin

E. Leader

Imperial College London, Prince Consort Road, London SW7 2BW
e.leader@imperial.ac.uk

Abstract. Spin in a relativistic context is a subtle concept!

This was brought out dramatically in 1988, when the results of the European Muon Collaboration (EMC) experiment at CERN on polarized deep inelastic lepton-hadron scattering (usually abbreviated as "polarized DIS") seemed to imply that the contribution to the proton's spin coming from the spin of its quarks was compatible with zero. The EMC paper was the most cited experimental paper for several years, and led to a major theoretical effort which eventually showed that the 20 year old, and long accepted, theoretical treatment was incomplete.

More recently, a classic and much cited paper on angular momentum sum rules, i.e. theoretical relations between the nucleon spin and the angular momentum of its constituents, was also shown to be incorrect.

Both these matters will be covered in the lectures.

1 Introduction

Deep inelastic lepton-hadron scattering has played a seminal role in the development of our present understanding of the sub-structure of elementary particles. The discovery of Bjorken scaling in the late nineteen-sixties provided the critical impetus for the idea that elementary particles contain almost pointlike constituents and for the subsequent invention of the Parton Model. DIS continued to play an essential role in the long period of consolidation that followed, in the gradual linking of partons and quarks, in the discovery of the existence of missing constituents, later identified as gluons, and in the wonderful confluence of all the different parts of the picture into a coherent dynamical theory of quarks and gluons – Quantum Chromodynamics (QCD).

Polarized DIS, involving the collision of a longitudinally polarized lepton beam on a polarized target (either longitudinally or transversely polarized) provides a different, complementary and equally important insight into the structure of the nucleon. Whereas ordinary DIS probes simply the number density of partons with a fraction x of the momentum of the parent hadron, polarized DIS can partly answer the more sophisticated question as to the number density of partons with given x and given spin polarization, in a hadron of definite polarization, either along or transverse to the motion of the hadron.

E. Leader: *Nucleon Spin*, Lect. Notes Phys. **712**, 95–128 (2007)
DOI 10.1007/3-540-38592-4_4

But what is quite extraordinary and unexpected ab initio is the richness and subtlety of the dynamical effects associated with the polarized case. Whereas the unpolarized scaling functions $F_{1,2}(x)$ have a simple interpretation in the Naive Parton Model (where the nucleon is considered as an ensemble of essentially free massless partons) and a straightforward generalisation in the framework of perturbative QCD, the spin dependent scaling functions $g_{1,2}(x)$ are much more subtle, each fascinating in its own way. The function $g_1(x)$ which, at first sight, seems trivial to deal with in the Naive Parton Model, turns out, within perturbative QCD, to have an anomalous gluon contribution associated with it. And $g_2(x)$ turns out not to have any interpretation at all in purely partonic language.

What is also fascinating is the extraordinary interplay of theory and experiment in the study of $g_1(x)$. For a long time the theory of $g_1(x)$ remained comfortably at the level of the Naive Parton Model. Then, in 1988, came the disturbing results of the European Muon Collaboration (EMC) [1], which differed significantly from the naive theoretical predictions. These results could be argued to imply that the sum of the spins carried by the quarks in a proton ($S_z^{\rm quarks}$) was consistent with zero, rather than with 1/2 as given in the quark model, suggesting a "spin crisis in the parton model" [2]. This led to an intense scrutiny of the basis of the theoretical calculation of $g_1(x)$ and the discovery of the anomalous gluon contribution [3]. So surprising was this discovery that the calculation was immediately checked by three different groups [4, 5, 6] who all arrived at the same result. (Somewhat fortuitously, as it turns out, as was demonstrated in [5]).

The above argument, leading to the spin crisis, is based upon a sum rule relating the projection along its direction of motion of the angular momentum of the nucleon to the spin and orbital angular momentum of its constituents. Although the sum rule is very intuitive in appearance, it turns out that the derivation of such sum rules is quite tricky.

A key element in deriving such sum rules is a precise knowledge of the tensorial structure of the expectation values of the angular momentum operators J_i in a state $|p\,,\sigma\rangle$ of the nucleon, labeled by its momentum p, and with some kind of specification of its spin state, denoted here non-commitally by σ.

In a much cited paper [7], Jaffe and Manohar stressed the subtleties involved in deriving general angular momentum sum rules. As they point out, too naive an approach leads immediately to highly ambiguous divergent integrals, and a careful limiting procedure has to be introduced in order to obtain physically meaningful results. In this it is essential to work with non-diagonal matrix elements $\langle p'\,,\sigma|\boldsymbol{J}|p\,,\sigma\rangle$ and this can have some unexpected consequences. Jaffe and Manohar comment that to justify rigorously the steps in such a procedure requires the use of normalizable wave packets, though they do *not* do this explicitly in their paper.

It turns out that despite all the care and attention to subtleties, there are flaws in the analysis in [7] and the results presented there are not entirely general [10]. Indeed there are cases where the results of [7] are incorrect. We shall present the correct results as well as a new angular momentum sum rule for the case of a fast moving nucleon polarized perpendicular to its direction of motion.

These lectures are based mainly on the Physics Report *The theory and phenomenology of polarized deep inelastic scattering* by Anselmino, Efremov and Leader [8], my book *Spin in particle physics* [9], and the article *Critique of the angular momentum sum rules and a new angular momentum sum rule* by Bakker, Leader and Trueman [10]. For a historical and fundamentally pedagogical introduction see *An introduction to gauge theories and modern particle physics* by Leader and Predazzi [11].

2 Polarized Lepton-Nucleon Deep Inelastic Scattering

Consider the inelastic scattering of polarized leptons on polarized nucleons:

$$l + N \to l' + X$$

in which only the final lepton is detected, so that X stands for a sum over all possible states compatible with the usual strong interaction conservation laws. We are interested in this reaction at high energy and large momentum transfer, where the parton model is applicable and where perturbative QCD calculations should be reliable.

We denote by m the lepton mass, k (k') the initial (final) lepton four-momentum and s (s') its covariant spin four-vector, such that $s \cdot k = 0$ ($s' \cdot k' = 0$) and $s \cdot s = -1$ ($s' \cdot s' = -1$); the nucleon mass is M and the nucleon four-momentum and spin four-vector are, respectively, P and S. Assuming one photon exchange, see Fig. 1, the differential cross-section for detecting the final polarized lepton in the solid angle $d\Omega$ and in the final energy range (E', $E' + dE'$) in the laboratory frame, $P = (M, \mathbf{0})$, $k = (E, \boldsymbol{k})$, $k' = (E', \boldsymbol{k}')$, can be written as (for a pedagogical and historical introduction see, for example [11])

$$\frac{d^2\sigma}{d\Omega\ dE'} = \frac{\alpha^2}{2Mq^4}\ \frac{E'}{E}\ L_{\mu\nu}\ W^{\mu\nu}\ , \tag{1}$$

where $q = k - k'$ and α is the fine structure constant.

In Eq. (1) the leptonic tensor $L_{\mu\nu}$ is known exactly and is given in terms of Dirac spinors by

$$L_{\mu\nu}(k, s; k', s') = [\bar{u}(k', s')\ \gamma_\mu\ u(k, s)]^*\ [\bar{u}(k', s')\ \gamma_\nu\ u(k, s)] \tag{2}$$

The unknown hadronic tensor $W_{\mu\nu}$ describes the interaction between the virtual photon and the nucleon and depends upon four scalar structure functions (like inelastic form factors) the unpolarized structure functions $F_{1,2}$ and

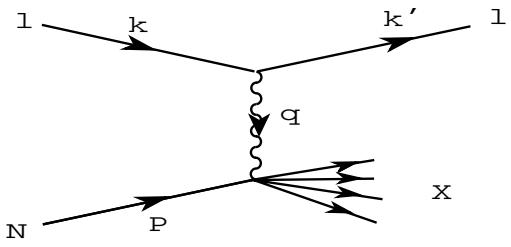

Fig. 1: Feynman diagram for inelastic lepton-hadron scattering

the spin-dependent sructure functions $g_{1,2}$, which have to be measured and can then be studied in theoretical models, in our case in the QCD-modified parton model. These can only be functions of the scalars q^2 and $q \cdot P$. It turns out to be more convenient to work with

$$Q^2 \equiv -q^2 \qquad \text{and} \qquad x_{Bj} \equiv Q^2/2q \cdot P = Q^2/2M\nu \tag{3}$$

where $\nu = E - E'$ is the energy of the virtual photon in the Lab frame. x_{Bj} is known as "x-Bjorken", and we shall simply write it as x.

The cross-section for unpolarized scattering is given by

$$\frac{d^2\sigma}{dx\,dy} = \frac{4\pi\alpha^2 s}{Q^4}[xy^2F_1 + (1-y)F_2] \tag{4}$$

where $y = P \cdot q/P \cdot k = \nu/E$ and $s = (P+k)^2$. (In this and all the following we neglect lepton masses and assume $s \gg M^2$ and $Q^2 \gg M^2$).

It was thought that analogous to the nucleon elastic form factors, $F_{1,2}$ would decrease rapidly as Q^2 increased. The discovery at SLAC in 1969 that there is almost no dependence on Q^2 at fixed x was referred to as "Bjorken scaling". By analogy with the elastic scattering of *point-like* particles, where the form factors are constants, the SLAC results were interpreted as the scattering of the virtual photon on some point-like constituents inside the nucleon, the *partons*. After many years of experimentation and analysis it became clear that the partons had the same internal quantum numbers as the *constituent quarks* which had been used to construct low energy models of the hadrons, and which had been assigned masses of order of 1/3 of the nucleon mass. The *partonic quarks* are usually taken to be massless in the kinematic regime where the parton model is applicable, and the precise relation between the constituent and partonic quarks (just *quarks* from now on) is not known.

In this picture the nucleon is viewed as simply a beam of parallel moving quarks and the reaction is visualized as in Fig. 2.

The input for this is the unpolarized quark number density $q(x')$ i.e. the number density of quarks with a fraction x' of the parent nucleon's momentum as shown in Fig. 3. It is easy to show that in DIS the photon interacts only with quarks which have $x' = x_{Bj} \equiv x$.

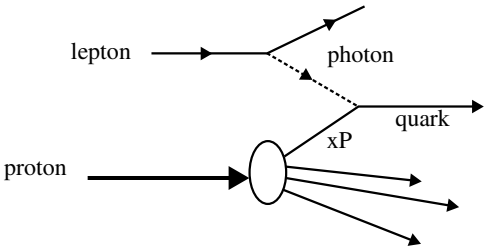

Fig. 2: Parton model interpretation of inelastic lepton-hadron scattering

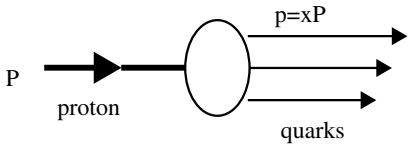

Fig. 3: The unpolarized quark density $q(x)$

Of course there is a $q(x)$ for each flavour of quark: $u(x), d(x), s(x)$ and their anti-quarks.

In this simple parton model F_1 and F_2 no longer depend on Q^2 are given by:

$$F_1(x) = \frac{1}{2}\left[\frac{4}{9}u(x) + \frac{1}{9}d(x) + \frac{1}{9}s(x) + antiquarks\right] \tag{5}$$

with $F_2 = 2xF_1$. In Eq. (5) the numerical factors are the squares of the quark electric charges. Note the crucial fact that the value of x appearing in the parton densities is the same as the value of the experimental quantity Bjorken-x at which F_1 is measured. This follows from the assumption that the quarks all move parallel to their parent hadron, the so-called *collinear* approximation i.e. with $\boldsymbol{p} = x\boldsymbol{P}$ and no transverse momentum $\boldsymbol{p_T}$.

Once the $q(x)$ were reasonably well determined from experiment it became clear that something was missing. Given the physical meaning of the $q(x)$ it is manifest that the "second moment"

$$\int_0^1 dx\, x q_f(x)$$

represents the fraction of the nucleon's momentum carried by the quarks of flavour f. Summing this over all the quark flavours should give 1, but was found to yield $\simeq 1/2$.

Fortunately the invention of QCD provided the (electrically neutral) gluons as the missing partons needed to carry the rest of the nucleon's momentum.

QCD also induced dynamical corrections to the simple parton model results, the most important for the unpolarized case being a gentle breaking of scaling, so that the parton densities become functions of (x, Q^2), with a slow calculable logarithmic dependence on Q^2. It may seem strange that the parton densities depend on Q^2. After all, Q^2 is a kinematical variable in a particular physical reaction, whereas the parton densities are a property of an isolated nucleon. What is hidden is the fact that in a field theoretic context the parton densities are defined as hadronic matrix elements of certain operators and these have to be renormalized. The *renormalization scale* or *renormalization point* μ_R^2 is supposed to be arbitrary, and it turns out to be most efficacious to choose $\mu_R^2 = Q^2$ in the theory of DIS. In addition infinities appear linked to the masslessness of the partons.These are rendered harmless by *factorizing* i.e. by splitting the QCD contribution into a soft and a hard part. The splitting is at some arbitrary *factorization* scale μ_F^2, and the infinite soft part is absorbed into the definition of the parton density. Strictly speaking then, the parton densities are functions of three variables: x, μ_R^2 and μ_F^2! Again, it is simplest to choose $\mu_F^2 = Q^2$. Thus the Q^2 in the structure functions refers to the kinematics of the experiment, whereas the Q^2 in the parton densities refers to the renormalization scale and the factorization scale at which they are defined. With this convention there is a very nice picture of what is happening physically. Namely, at fixed x what the virtual photon "sees" in the nucleon depends upon its virtuality Q^2 or frequency. As Q^2 increases the "resolution" increases.

Consider now the case where the lepton and the target nucleon are polarized longitudinally, i.e. along or opposite to the direction of the lepton beam, usually taken as along the positive OZ axis. Under reversal of the nucleon's spin direction the cross-section difference is given by

$$\frac{d^2\overrightarrow{\sigma}^{\Leftarrow}}{dx\,dy} - \frac{d^2\overrightarrow{\sigma}^{\Rightarrow}}{dx\,dy} = \frac{16\pi\alpha^2}{Q^2}\left[(1-\frac{y}{2})\,g_1 - \frac{2M^2xy}{Q^2}g_2\right]. \tag{6}$$

In the kinematic range of interest the second term should be negligible so that Eq. (6) provides a method of measuring the spin-dependent structure function $g_1(x)$.

If the nucleons are polarized transversely in the scattering plane, that is the nucleon spin is perpendicular to the direction of the incoming lepton, then one finds

$$\frac{d^2\sigma^{\rightarrow\Uparrow}}{dx\,dy} - \frac{d^2\sigma^{\rightarrow\Downarrow}}{dx\,dy} = -\frac{16\alpha^2}{Q^2}\left(\frac{2Mx}{Q}\right)\sqrt{1-y}\,\left[\frac{y}{2}g_1 + g_2\right]. \tag{7}$$

In principle Eqs. (6) and (7) allow measurement of both g_1 and g_2, but the transverse asymmetry is much smaller and therefore much more difficult to measure. Only in the past few years has it been possible to gather information on g_2 which turns out to be smaller than g_1.

Let us now turn to the interpretation of g_1 in the parton model. Note that there is no parton model for g_2 – its calculation is totally ambiguous.

3 The Spin Crisis in the Parton Model

For some twenty years prior to the EMC experiment the accepted expression for g_1 was completely analogous to Eq. (5) with the unpolarized quark density replaced by the (longitudinal) polarized density $\Delta q(x)$ shown in Fig. 4:

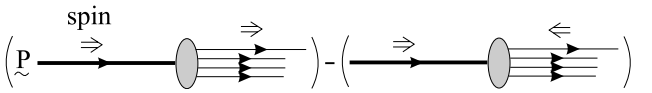

Fig. 4: The longitudinal polarized quark density $\Delta q(x)$

Let $q_{+/-}(x)dx$ be the number density of quarks whose helicities are along/opposite to the helicity of the nucleon. Then

$$\Delta q(x) \equiv q_+(x) - q_-(x) \tag{8}$$

Of course

$$q(x) = q_+(x) + q_-(x) \tag{9}$$

(Note that the third independent kind of quark density, the transverse polarized parton density (sometimes called transversality) $\Delta_T q(x)$ shown in Fig. 5 in which the nucleon is transversely polarized and $q_\uparrow$ and $q_\downarrow$ are the number density of quarks whose transverse spin is along or opposite to the spin of the nucleon, and

$$\Delta_T q(x) \equiv q_\uparrow(x) - q_\downarrow(x) \tag{10}$$

does not play a role in DIS, but will be of interest later on.)

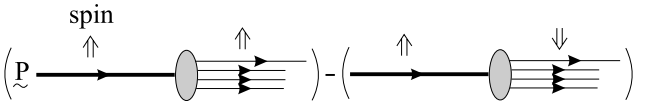

Fig. 5: The transversely polarized quark density $\Delta_T q(x)$

The expression for g_1 is then:

$$g_1(x) = \frac{1}{2}\left\{\frac{4}{9}\Delta u(x) + \frac{1}{9}\Delta d(x) + \frac{1}{9}\Delta s(x) + antiquarks\right\} \tag{11}$$

It is useful to rearrange this by defining combinations of quark densities which have specific transformation properties under the group of flavour transformations $SU(3)_F$. Define:

$$\Delta q_3 = (\Delta u + \Delta\overline{u}) - (\Delta d + \Delta\overline{d}) \tag{12}$$

$$\Delta q_8 = (\Delta u + \Delta\overline{u}) + (\Delta d + \Delta\overline{d}) - 2(\Delta s + \Delta\overline{s}) \tag{13}$$

$$\Delta\Sigma = (\Delta u + \Delta\overline{u}) + (\Delta d + \Delta\overline{d}) + (\Delta s + \Delta\overline{s}) \tag{14}$$

which transform respectively as the third component of an isotopic spin triplet, the eighth component of an $SU(3)_F$ octet and a flavour singlet. Eq. (11) then becomes:

$$g_1(x) = \frac{1}{9}\left[\frac{3}{4}\Delta q_3(x) + \frac{1}{4}\Delta q_8(x) + \Delta\Sigma\right] \tag{15}$$

Taking the first moment of this yields

$$\Gamma_1 \equiv \int_0^1 g_1(x)dx = \frac{1}{12}\left[a_3 + \frac{1}{\sqrt{3}}a_8 + \frac{4}{3}a_0\right] \tag{16}$$

where

$$\begin{aligned} a_3 &= \int_0^1 dx\ \Delta q_3(x) \\ a_8 &= \frac{1}{\sqrt{3}}\int_0^1 dx\ \Delta q_8(x) \\ a_0 &= \Delta\Sigma \equiv \int_0^1 dx\ \Delta\Sigma(x) \end{aligned} \tag{17}$$

The significance of these first moments lies in the fact that they can be related to hadronic matrix elements of currents which are measurable in *other* processes. To see this we need to briefly mention the Operator Product Expansion.

The fundamental understanding of the Q^2-behaviour of the moments in unpolarized DIS came originally from a study of the Operator Product Expansion. Later it was discovered that the same results could be obtained in the QCD Improved Parton Model.

It can be shown that the hadronic tensor $W^{\mu\nu}$ involved in the expression for the deep inelastic cross-section [see Eq. (1)] is given by the Fourier transform of the nucleon matrix elements of the commutator of electromagnetic currents $J_\mu(x)$:

$$W_{\mu\nu}(q;P,S) = \frac{1}{2\pi}\int d^4x\ e^{iq\cdot x}\langle P,S|[J_\mu(x),J_\nu(0)]|P,S\rangle \tag{18}$$

where S^μ is the covariant spin vector specifying the nucleon state of momentum P^μ. It is convenient to introduce an amplitude $T_{\mu\nu}$ which is closely

related to the forward T-matrix element for Compton scattering of a virtual photon of 4-momentum q and helicities λ and λ'. In the convention for the $\hat{T}$-operator

$$\hat{S} = \hat{I} + i(2\pi)^4 \, \delta^4(P_f - P_i) \, \hat{T}$$

one has

$$\langle P, S; q, \lambda' | \hat{T} | P, S; q, \lambda \rangle = 4\pi\alpha \, \varepsilon^*_\mu(\lambda') \, T^{\mu\nu} \, \varepsilon_\nu(\lambda) \tag{19}$$

where

$$T_{\mu\nu}(q; P, S) = i \int d^4x \; e^{iq\cdot x} \langle P, S | T(J_\mu(x) J_\nu(0)) | P, S \rangle \tag{20}$$

is given in terms of the time ordered product of the currents. It can be shown that

$$W_{\mu\nu} = \frac{1}{\pi} \operatorname{Im} T_{\mu\nu} \; . \tag{21}$$

The behaviour of $T_{\mu\nu}$ (and therefore $W_{\mu\nu}$) in the deep inelastic limit is controlled by the behaviour of the product of currents near the light cone $x^2 = 0$ and can be derived from Wilson's Operator Product Expansion.

It is important to note that the expressions Eqs. (18) and (20), for which the operator product approach can be utilized, only arise because of the fully inclusive nature of the deep inelastic reaction being considered. Indeed the starting point from which Eq. (18) can be derived is the expression

$$W_{\mu\nu} \propto \sum_X \langle P, S | J_\mu | X \rangle \langle X | J_\nu | P, S \rangle$$

which appears in the formula for the cross-section. Only if the sum is over *all* final states $|X\rangle$ does this reduce to Eq. (18).

The Operator Product Expansion gives results for the moments of $g_{1,2}$ in terms of hadronic matrix elements of certain operators multiplied by perturbatively calculable coefficient functions. It turns out that the the a_i in Eq. (16) are hadronic matrix elements of the octet of quark $SU(3)_F$ axial-vector currents $J^j_{5\mu}$ $(j = 1, \ldots, 8)$ and the flavour singlet axial current $J^0_{5\mu}$ taken between proton states of definite momentum and spin direction.

The octet currents are

$$J^j_{5\mu} = \bar{\psi} \gamma_\mu \gamma_5 \left(\frac{\lambda_j}{2} \right) \psi \qquad (j = 1, 2, \ldots, 8) \tag{22}$$

where the λ_j are the usual Gell-Mann matrices and ψ is a column vector in flavour space

$$\psi = \begin{pmatrix} \psi_u \\ \psi_d \\ \psi_s \end{pmatrix} \; , \tag{23}$$

and the flavour singlet current is

$$J^0_{5\mu} = \bar{\boldsymbol{\psi}} \gamma_\mu \gamma_5 \boldsymbol{\psi} \, . \tag{24}$$

The forward matrix elements of the $J^j_{5\mu}$ can only be proportional to S^μ and the a_j are defined by

$$\langle P,S|J^j_{5\mu}|P,S\rangle = Ma_jS_\mu$$
$$\langle P,S|J^0_{5\mu}|P,S\rangle = 2Ma_0S_\mu \ . \tag{25}$$

Analogous to Eq. (22) one introduces an octet of vector currents

$$J^j_\mu = \bar\psi\gamma_\mu\left(\frac{\lambda_{\mathbf{j}}}{\mathbf{2}}\right)\psi \qquad (\mathbf{j}=\mathbf{1},\ldots,\mathbf{8}) \tag{26}$$

which are *conserved currents* to the extent that $SU(3)_F$ is a symmetry of the strong interactions.

Now it is precisely the above octets of currents that control the β-decays of the neutron and of the octet of hyperons. This means that the values of a_3 and a_8 are known from *other* measurements. Consequently the measurement of Γ_1 can, via Eq. (16), be interpreted as a measurement of the flavour singlet a_0.

The result of the EMC experiment was that $a_0 \approx 0$, which, as we shall see, is somewhat startling.

Let us consider now the physical significance of $\Delta\Sigma(x)$. Since, via Eq. (8), $q_\pm(x)$ count the number of quarks of momentum fraction x with spin component $\pm\frac{1}{2}$ along the direction of motion of the proton (let us call this the z-direction), the total contribution to J_z coming from a given flavour quark is

$$S_z = \int_0^1 dx\left\{\left(\frac{1}{2}\right)q^+(x) + \left(-\frac{1}{2}\right)q^-(x)\right\}$$
$$= \frac{1}{2}\int_0^1 dx\ \Delta q(x)\ . \tag{27}$$

It follows that

$$a_0 = 2S_z^{quarks} \tag{28}$$

where S_z^{quarks} is the contribution to J_z from the spin of all quarks and anti-quarks.

We see that a_0 is playing two roles. Here we have argued that it measures the z component of the spin carried by the quarks; in the second of equations Eq. (25) $2a_0$ measures the expectation value of the flavour singlet axial-vector current. What is the connection? From Noether's Theorem the spin density operator for a spin 1/2 particle is $1/2\bar\psi(x)\gamma^\rho\sigma^{\nu\lambda}\psi(x)$. For $\rho=0,\nu=1,\lambda=2$ corresponding to spin in the z direction, this is just 1/2 times the axial vector current with $\mu=3$. Thus, if the nucleon state is written as a superposition of partonic states, the expectation value of the axial-vector current should indeed be twice the spin carried by the spin 1/2 quarks. But, as we shall see presently, the existence of the axial anomaly complicates this somewhat.

Now in the simple parton model $\mathbf{p}_\perp = \mathbf{0}$ and all quarks move parallel to the parent hadron, i.e. for a quark of momentum $\mathbf{p}$, $\mathbf{p} = \mathbf{xP}$. Hence any orbital angular momentum carried by the quarks is perpendicular to $\mathbf{P}$ and thus does not contribute to J_z. Hence, in the simple parton model, one expects for a proton of helicity $+1/2$:

$$S_z^{quarks} = J_z = 1/2 \; . \tag{29}$$

We stress that this ignores $\mathbf{p}_\perp$ effects and assumes only quark and antiquark constituents are polarized.

The EMC result for the value of a_0, on the contrary, implied that

$$\left(S_z^{quarks} \right)_{Exp} = 0.03 \pm 0.06 \pm 0.09 \; . \tag{30}$$

It was this highly unexpected result which was termed the "spin crisis in the parton model" [2].

4 Resolution of the Spin Crisis: The Axial Anomaly

Consider the axial current

$$J_{5\mu}^f = \bar{\psi}_f(x) \, \gamma_\mu \gamma_5 \, \psi_f(x) \tag{31}$$

made up of quark operators of definite flavour f. (An implicit colour sum is always implied). From the free Dirac equation of motion one finds that

$$\partial^\mu J_{5\mu}^f = 2im_q \bar{\psi}_f(x) \, \gamma_5 \, \psi_f(x) \tag{32}$$

where m_q is the mass of the quark of flavour f.

In the chiral limit $m_q \to 0$ Eq. (32) appears to imply that $J_{5\mu}^f$ is conserved. If this were really true there would be a symmetry between left and right-handed quarks, leading to a parity degeneracy of the hadron spectrum e.g. there would exist two protons, of opposite parity. However, the formal argument from the free equations of motion is not reliable and, as shown originally by Adler, and by Bell and Jackiw [12] (in the context of QED), there is an anomalous contribution arising from the triangle diagram shown in Fig. 6.

As a consequence the axial current is not conserved when $m_q = 0$. One has instead, for the QCD case

$$\partial^\mu J_{5\mu}^f = \frac{\alpha_s}{4\pi} \, G_{\mu\nu}^a \widetilde{G}_a^{\mu\nu} = \frac{\alpha_s}{2\pi} \, \mathrm{Tr} \left[\mathbf{G}_{\mu\nu} \; \widetilde{\mathbf{G}}^{\mu\nu} \right] \tag{33}$$

where α_s is the QCD analogue of the fine structure constant, and where $\widetilde{G}_{\mu\nu}^a$ is the dual field tensor

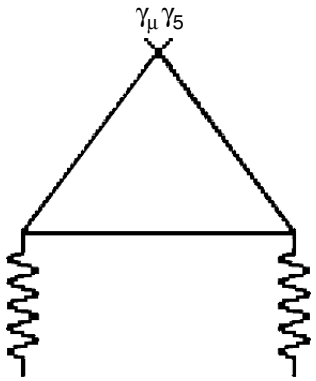

Fig. 6: The anomalous triangle diagram

$$\widetilde{G}^a_{\mu\nu} \equiv \frac{1}{2}\,\varepsilon_{\mu\nu\rho\sigma}\mathbf{G}_a^{\rho\sigma} \tag{34}$$

and where a field vector or tensor without a colour label stands for a matrix. In this case

$$\mathbf{G}_{\mu\nu} \equiv \left(\frac{\lambda_{\mathbf{a}}}{\mathbf{2}}\right)\mathbf{G}^{\mathbf{a}}_{\mu\nu}\ . \tag{35}$$

The result Eq. (33), which emerges from a calculation of the triangle diagram (Fig. 6) using $m_q = 0$ and the gluon virtuality $k^2 \neq 0$, is really a particular limit of a highly non-uniform function. If we take $m_q \neq\ 0,\ k^2 \neq 0$ the RHS of Eq. (33) is multiplied by

$$T(m_q^2/k^2) = 1 - \frac{2m_q^2/k^2}{\sqrt{1+4m_q^2/k^2}}\ \ln\left(\frac{\sqrt{1+4m_q^2/k^2}+1}{\sqrt{1+4m_q^2/k^2}-1}\right)\ . \tag{36}$$

We see that the anomaly corresponds to $T \to 1$ for $(m_q^2/k^2) \to 0$. On the other hand, for on-shell gluons, $k^2 = 0$, and $m_q \neq 0$, i.e. in the limit $(m_q^2/k^2) \to \infty$ the terms cancel, $T \to 0$, and there is no anomaly.For gluons bound inside a nucleon one sould utilize $k^2 \neq 0$ and the anomalous contribution is relevant.

The anomaly induces a *pointlike* interaction between $J^0_{5\mu}$ and gluons. That it is pointlike can be seen by taking different gluon momenta k_1 and k_2 in Fig. 6 and noting that the amplitude does not depend on the momentum transfer $k_1 - k_2$ when $m_q = 0$. Therefore, in computing the matrix element of $J^0_{5\mu}$ in a hadron state, we will get a contribution from the gluon components of the hadron as well as the more obvious contribution from quarks. From Adler's expression [ABJ 69] for the triangle diagram, modified to QCD, one finds for the forward gluonic matrix element of the flavour f current (our convention is $\varepsilon_{0123} = 1$)

$$\langle k,\lambda|J^f_{5\mu}|k,\lambda\rangle = \frac{i\alpha_s}{2\pi}\,\varepsilon_{\mu\nu\rho\sigma}\,k^\nu\varepsilon^{*\rho}(\lambda)\varepsilon^\sigma(\lambda)\,T(m_q^2/k^2)$$
$$= -\frac{\alpha_s}{2\pi}\,S^g_\mu(k,\lambda)\,T(m_q^2/k^2) \tag{37}$$

where λ is the gluon helicity and we may take

$$S^g_\mu(k,\lambda) \approx \lambda k_\mu \tag{38}$$

as the covariant spin vector for almost massless gluons.

We may then compute the gluonic contributions to the hadronic expectation value $\langle P,S|J^0_{5\mu}|P,S\rangle$. In this case the gluons being bound will be slightly off-shell i.e. $k^2 \neq 0$, but small. The full triangle contribution involves a sum over all quark flavours. We take m_u, m_d and m_s to be $\ll k^2$ whereas m_c, m_b and m_t are $\gg k^2$. The function $T(m_q^2/k^2)$ thus takes the values:

$$\begin{aligned} T &= 1 \qquad \text{for} \qquad u,\ d,\ s \\ T &= 0 \qquad \text{for} \qquad c,\ b,\ t \end{aligned} \tag{39}$$

and the gluon contribution is then given by [see Eq. (25)]

$$a_0^{gluons}(Q^2) = -3\,\frac{\alpha_s}{2\pi}\int_0^1 dx\ \Delta G(x,Q^2)$$
$$\equiv -3\,\frac{\alpha_s(Q^2)}{2\pi}\,\Delta G(Q^2) \tag{40}$$

or from Eq. (16)

$$\Gamma_{1p}^{gluons}(Q^2) = -\frac{1}{3}\,\frac{\alpha_s(Q^2)}{2\pi}\,\Delta G(Q^2)\,. \tag{41}$$

In the above we have implicitly taken into account QCD corrections which have the effect of replacing α_s by the *effective* or *running coupling* $\alpha_s(Q^2)$ which varies logarithmically with Q^2. $\Delta G(x,Q^2)$, the polarized gluon density, is the difference between the number density of gluons with the same helicity as the nucleon and those with opposite helicity; its integral $\Delta G(Q^2)$ is the total helicity carried by the gluons. In the above we have explicitly stressed that the gluon density is Q^2 dependent. This will be of interest presently.

The result Eq. (40) is of fundamental importance. It tells us that the simple parton model formulae Eq. (17) and (28) for a_0 (and hence for Γ_1^p in terms of the Δq_f) are incorrect. We now have, instead,

$$a_0 = \Delta\Sigma - 3\,\frac{\alpha_s}{2\pi}\,\Delta G \tag{42}$$

We see immediately that it has the fundamental implication *that the small measured value of a_0 does not necessarily imply that $\Delta\Sigma$ is small.*

There are several interesting and surprising aspects to the result Eq. (40). Firstly, we argued above that the axial-vector current is, effectively, the spin

density operator for *spin 1/2* particles. Yet in Eq. (37) its matrix element between gluon states is proportional to the spin vector of the gluons. The point is that the fields in the axial-vector current are not free fields (as signalled by the factor $\alpha(Q^2)$) and in a perturbative expansion would contain terms involving gluon fields. Secondly, the simple parton model is usually thought of as the limit when the QCD coupling is switched off. Moreover because QCD possess the property of *asymptotic freedom* i.e the effective coupling goes to zero logarithmically as $Q^2 \to \infty$, we would expect that as $Q^2 \to \infty$ the term $a_0^{gluons}(Q^2)$ should vanish, and Eq. (42) would reduce to the simple parton model result Eq. (17). However it can be shown that the first moment $\Delta G(Q^2)$ tends to infinity logarithmically as $Q^2 \to \infty$, thus exactly cancelling the decrease in $\alpha(Q^2)$ and the gluonic term survives!

The "spin crisis" was signalled via Eq. (29), which is an intuitive statement that the angular momentum of the nucleon should be made up of the angular momentum of its constituents. This is an example of an angular momentum sum rule, and it seems obviously true. However, as we shall now see, such relations in a relativistic theory are highly non-trivial. After studying these sum rules we shall return to reconsider Eq. (29) when allowing for the fact that partons have intrinsic perpendicular momentum and that the parton densities depend on the renormalization scale.

Angular momentum sum rules require explicit expressions for the matrix elements of the angular momentum operators. Obtaining these is non-trivial.

5 Matrix Elements of Angular Momentum Operators: The Problem

In the standard approach one relates the matrix elements of the angular momentum operators to those of the energy-momentum tensor.

Let $T^{\mu\nu}(x)$ be the total energy-momentum density which is conserved

$$\partial_\mu T^{\mu\nu}(x) = 0 \; . \tag{43}$$

Later we shall distinguish between the conserved *canonical* energy-momentum tensor $T_{\rm C}^{\mu\nu}$, which emerges from Noether's theorem, and which is, generally, not symmetric under $\mu \leftrightarrow \nu$, and the *symmetrised* Belinfante tensor $T^{\mu\nu}$, which for QCD is given by

$$T^{\mu\nu}(x) = \frac{1}{2}\left(T_{\rm C}^{\mu\nu}(x) + T_{\rm C}^{\nu\mu}(x)\right) \tag{44}$$

and which is also conserved. For the moment, however, this distinction is irrelevant.

Being a *local* operator, $T^{\mu\nu}(x)$ transforms under translations as follows

$$T^{\mu\nu}(x) = e^{ix\cdot P} T^{\mu\nu}(0) e^{-ix\cdot P} \; , \tag{45}$$

where the P^μ are the total momentum operators of the theory.

By contrast the various angular momentum density operators which are of interest, the orbital angular momentum densities

$$M^{\mu\nu\lambda}_{\rm orb}(x) \equiv x^\nu T^{\mu\lambda}_{\rm C}(x) - x^\lambda T^{\mu\nu}_{\rm C}(x) \tag{46}$$

or the version constructed using the symmetrised stress-energy tensor,

$$M^{\mu\nu\lambda}(x) \equiv x^\nu T^{\mu\lambda}(x) - x^\lambda T^{\mu\nu}(x) \tag{47}$$

are not local operators (we shall call them *compound*) and do not transform according to Eq. (45).

Note that, strictly speaking, the operators relevant to the angular momentum are the components M^{0ij} where i, j are spatial indices. However, for reasons of simplicity in utilising the Lorentz invariance of theory, it may be preferable to deal covariantly with the entire tensor $M^{\mu\nu\lambda}$. We shall loosely refer to them also as angular momentum densities.

The total angular momentum density is

$$J^{\mu\nu\lambda}(x) = M^{\mu\nu\lambda}_{\rm orb}(x) + M^{\mu\nu\lambda}_{\rm spin}(x) \ , \tag{48}$$

where the structure of $M^{\mu\nu\lambda}_{\rm spin}$ depends on the type of fields involved. From Noether's theorem $J^{\mu\nu\lambda}(x)$ is a set of conserved densities, i.e.,

$$\partial_\mu J^{\mu\nu\lambda}(x) = 0 \ . \tag{49}$$

As a consequence of the densities being conserved, it follows that the total momentum operators

$$P^\nu \equiv \int d^3x T^{0\nu}(x) \tag{50}$$

and the total angular momentum operators $\boldsymbol{J}$,

$$J_z = J^3 = J^{12} \ , \quad \text{(cyclical)} \tag{51}$$

with

$$J^{ij} \equiv \int d^3x J^{0ij}(x) \tag{52}$$

are conserved quantities, independent of time.

The relationship between the $M^{\mu\nu\lambda}(x)$, constructed using the symmetrical energy-momentum density and the $J^{\mu\nu\lambda}(x)$ constructed from the canonical energy-momentum tensor is extremely interesting. One can show that

$$M^{0ij}(x) = J^{0ij}(x) + [\text{E of M terms}] + [\text{divergence terms}] \ . \tag{53}$$

The [E of M terms] vanish if it is permissible to use the equations of motion of the theory. The [divergence terms] are of the form $\partial_\alpha F^{\alpha 0ij}(x)$.

We shall be primarily interested in the expectation values of the physical operators, i.e. in their forward matrix elements. If $F^{\alpha 0ij}(x)$ were a local operator, it would follow directly that the forward, momentum-space, matrix elements of the divergence terms in Eq. (53) vanish. But it is not a local operator. Nonetheless, a careful treatment using wave packets [13] demonstrates that the forward matrix elements do indeed vanish.

Dropping, as is customary, the [E of M terms], we shall thus assume the validity of

$$\langle p,\sigma| \int d^3x M^{0ij}(\boldsymbol{x},0)|p,\sigma\rangle = \langle p,\sigma| \int d^3x J^{0ij}(\boldsymbol{x},0)|p,\sigma\rangle \ . \tag{54}$$

Of primary interest are the matrix elements of the angular momentum operators J^k or, equivalently, the J^{ij}. Consider the forward matrix element, at $t=0$,

$$\begin{aligned}
\mathcal{M}^{0ij}(p,\boldsymbol{s}) &\equiv \langle p,\boldsymbol{s}| \int d^3x M^{0ij}(\boldsymbol{x},0)|p,\boldsymbol{s}\rangle \\
&= \int d^3x \langle p,\boldsymbol{s}|x^i T^{0j}(x) - x^j T^{0i}(x)|p,\boldsymbol{s}\rangle \\
&= \int d^3x\, x^i\, \langle p,\boldsymbol{s}|e^{iP\cdot x} T^{0j}(0) e^{-iP\cdot x}|p,\boldsymbol{s}\rangle - (i \leftrightarrow j) \\
&= \int d^3x\, x^i\, \langle p,\boldsymbol{s}|T^{0j}|p,\boldsymbol{s}\rangle - (i \leftrightarrow j) \ .
\end{aligned} \tag{55}$$

where $\boldsymbol{s}$ is the rest frame spin vector. The integral in Eq. (55), $\int d^3x\, x^i$, is totally ambiguous, being either infinite or, by symmetry, zero.

The essential problem is to obtain a sensible physical expression, in terms of p and $\boldsymbol{s}$, for the above matrix element. The fundamental idea is to work with a non-forward matrix element and then to try to approach the forward limit. This is similar to what is usually done when dealing with non-normalizable plane wave states and it requires the use of wave packets for a rigorous justification.

It will turn out that the results are sensitive to the type of relativistic spin state employed, so in the next Section we present a brief resumé of the properties of relativistic spin states.

We mentioned in the Introduction that one of the classic papers on this subject is incorrect. The most crucial error in that treatment is the mishandling of the matrix elements of a covariant tensor operator. If $T^{\mu\lambda}$ transforms as a second-rank tensor its *non-forward* matrix elements do *not* transform covariantly. This was the motivation, decades ago, for Stapp to introduce M-functions [14]. Namely, even if the spin state is specified by the covariant spin vector S, the covariance is spoilt, for canonical spin states by the Wigner rotation, and, for helicity states by the analogous Wick helicity rotation [15]. Only by first factoring out the wave-functions (in our case Dirac spinors) i.e. by writing

$$\langle p', S'|T^{\mu\lambda}|p, S\rangle = \bar{u}(p', S')\mathcal{T}^{\mu\nu}(p', p)u(p, S) \ . \tag{56}$$

does the remaining M-function, in this case $\mathcal{T}^{\mu\nu}(p', p)$, transform covariantly. For local operators the transformations of the spinors u and $\bar{u}$ cancel between themselves for forward matrix elements and so the result does have the naively expected tensor expansion. This is not true in general for compound operators, in particular the angular momentum and boost operators.

6 Relativistic Spin States

The definition of a spin state for a particle in motion, in a relativistic theory, is non-trivial, and is convention dependent. Namely, starting with the states of a particle at rest, which we shall denote by $|0, m\rangle$, where m is the spin projection in the z-direction, one defines states $|p, \sigma\rangle$ for a particle with four-momentum p by acting on the rest frame states with various boosts and rotations, and the choice of these is convention-dependent. Note that in this section and the following ones we are using $\boldsymbol{p}$ to denote the nucleon momentum. The states are on-shell so $p^2 = M^2$

There are three conventions in general use [9]

(a) Canonical or boost states as used e.g. in Bjorken and Drell [16] or Peskin and Schroeder [17]

$$|p, m\rangle = B(\boldsymbol{v})|0, m\rangle \tag{57}$$

where $B(\boldsymbol{v})$ is a pure boost along $\boldsymbol{v} = \boldsymbol{p}/p_0$, and $\boldsymbol{p} = (p, \theta, \phi)$ denotes the three-vector part of p^μ

(b) Jacob-Wick helicity states [18]

$$|p, \lambda\rangle_{JW} = R_z(\phi)\, R_y(\theta)\, R_z(-\phi)\, B_z(v)|0, m = \lambda\rangle \tag{58}$$

where B_z is a boost along OZ, and the later introduced, somewhat simpler

(c) Wick helicity states [15]

$$|p, \lambda\rangle = R_z(\phi)\, R_y(\theta)\, B_z(v)|0, m = \lambda\rangle \tag{59}$$

Recall that the physical meaning of helicity is the projection of $\boldsymbol{J}$ along the direction of motion of the particle i.e. along $\hat{\boldsymbol{P}}$.

From the canonical states of spin-$\frac{1}{2}$ one can construct the states

$$|p, \boldsymbol{s}\rangle = B(\boldsymbol{v})|0, \boldsymbol{s}\rangle = B(\boldsymbol{v})\mathcal{D}^{1/2}_{m1/2}(R(\boldsymbol{s}))|0, m\rangle \tag{60}$$

which, in the rest frame, are eigenstates of spin with spin eigenvector along the unit vector $\boldsymbol{s}$. The rotation $R(\boldsymbol{s})$ rotates a unit vector in the z-direction into $\boldsymbol{s}$ by first a rotation about y and then a rotation about z.

The canonical states, with their reference to a rest frame, are clearly not suitable for massless particles like gluons. Helicity states, on the other hand, can be used for both massive and massless particles. However, it turns out that the results for the matrix elements for the canonical states are much more intuitive, so we will generally use them for $M \neq 0$.

The reason we are emphasizing this distinction between canonical and helicity states is that the matrix elements of the angular momentum operators between helicity states are quite bizarre! Since, for arbitrary p, helicity states are just linear superpositions of canonical states, one may wonder why this is so. It results from the facts (i) that the coefficients in the linear superposition are p-dependent, i.e. depend upon the polar angles of $\boldsymbol{p}$ and (ii) that the matrix elements of the angular momentum operators contain derivatives of δ-functions, and these, as usual, must be interpreted in the sense of partial integration, i.e.

$$f(p,p')\frac{\partial}{\partial p_i}\delta^3(p-p') = -\delta^3(p-p')\frac{\partial}{\partial p_i}f(p,p') \tag{61}$$

In almost all studies of hard processes, where a mixture of perturbative and non-perturbative QCD occurs, nucleons are taken to be in helicity states moving with high energy along the z-axis, and typically one is utilizing matrix elements of local products of quark or gluon field operators between these states. For these operators there is no problem in dealing with diagonal matrix elements. But when it comes to an angular momentum sum rule for the nucleon, care must be taken to decide whether one is dealing with helicity states $|p_z,\lambda\rangle$, where $p_z = (E,0,0,p)$ or with canonical states $|p_z,\boldsymbol{s}_z\rangle$, where $\boldsymbol{s}_z = (0,0,2\lambda)$. The point is that even though the initial states are the same,

$$|p_z,\lambda\rangle = |p_z,\boldsymbol{s}_z\rangle \tag{62}$$

the singular nature of J_i forces one to deal with non-diagonal matrix elements i.e. to utilize $\langle p',\sigma|$ where $\boldsymbol{p'}$ is not along the z-axis, and for these

$$\langle p',\lambda| \neq \langle p',\boldsymbol{s}_z| \tag{63}$$

In [10] it is shown that it is possible to give a rigorous derivation of the structure of the expectation values for canonical states

$$\langle p,\boldsymbol{s}|J_i|p,\boldsymbol{s}\rangle \equiv \mathcal{L}_{p'\to p}\langle p',\boldsymbol{s}|J_i|p,\boldsymbol{s}\rangle \tag{64}$$

where $\boldsymbol{s}$ is a unit vector along the rest frame spin eigenvector, and for helicity states

$$\langle p,\lambda|J_i|p,\lambda\rangle \equiv \mathcal{L}_{p'\to p}\langle p',\lambda|J_i|p,\lambda\rangle. \tag{65}$$

In general, for the arbitrary i^{th} component of $\boldsymbol{J}$, for spin-$\frac{1}{2}$

$$\langle p,\lambda|J_i|p,\lambda\rangle \neq \langle p,\boldsymbol{s}=2\lambda\hat{\boldsymbol{p}}|J_i|p,\boldsymbol{s}=2\lambda\hat{\boldsymbol{p}}\rangle \tag{66}$$

even though $\boldsymbol{s}$ lies along the direction of $\hat{p}$ in both cases, and even if $\boldsymbol{p}$ is along OZ where Eq. (62) holds. Only for the specific component of $\boldsymbol{J}$ along $\hat{\boldsymbol{p}}$ do the matrix elements agree, i.e. for arbitrary $\boldsymbol{p}$,

$$\langle p, \lambda | \boldsymbol{J} \cdot \boldsymbol{p} | p, \lambda \rangle = \langle p, \boldsymbol{s} = 2\lambda \hat{\boldsymbol{p}} | \boldsymbol{J} \cdot \boldsymbol{p} | p, \boldsymbol{s} = 2\lambda \hat{\boldsymbol{p}} \rangle \ . \tag{67}$$

In using the sum rules based on Eq. (64) or Eq. (65) for arbitrary i to test any model of the nucleon in terms of its constituents, it is essential to construct wave-functions appropriate to the type of spin state being used for the nucleon. The equations Eq. (64) and (65) contain delta functions and *derivatives* of delta-functions and this is the reason for the special care required. In the following we use canonical states wherever possible, except when discussing gluons where we are forced to use helicity states.

7 Matrix Elements of Angular Momentum Operators: The Results

Since it is claimed that well known results in the literature are incorrect, in [10] pains were taken to derive the correct expressions in three different ways, two involving explicit physical wave packets and the third, totally independent, based upon the rotational properties of the state vectors. Surprisingly it turns out that the results are very sensitive to the type of relativistic spin state used to describe the motion of the particle i.e. whether a canonical (i.e. boost) state or a helicity state is utilized. We shall present results for the matrix elements of the angular momentum operators, valid in an arbitrary Lorentz frame, both for helicity states and canonical states.

The bulk of the analysis in [10] is based on a straightforward wave-packet approach. However, this is rather subtle for particles with non-zero spin. The key points requiring care are:

(1) The wave packets should be strictly physical, i.e. a superposition of physical plane-wave states. This requirement turns out to be incompatible with some of the steps in [7].
(2) Care is needed in the treatment of the Lorentz covariance properties of the matrix elements involved in the subsidiary steps of the analysis, as mentioned in the discussion of Eq. (56). This leads to tensorial structures which differ in some cases from those in [7].
(3) Because the results differ from [7] the matrix elements have been studied in a totally different and independent way. This does not use wave packets and is based upon the transformation properties of momentum states under rotations. This very direct approach holds for arbitrary spin, whereas in the wave packet treatment it was only possible to deal with spin-1/2 particles. It is much simpler than the wave packet approach and it avoids all the issues of ambiguous integrals, which in the old fashioned

treatment force one to utilize wave packets. It also brings to light some peculiar and unintuitive properties of helicity states, and this must be taken into account when deriving spin sum rules. This is important since we have to deal with gluons in our sum rules.

The approach via wave packets is long and complicated, so we shall only utilize the latter method here.

7.1 Canonical Spin State Matrix Elements

In order to utilize the rotational properties of the canonical or boost spin states we need to display explicitly the Wigner boost operators used in defining the states of a moving particle in terms of the rest frame spin states quantized in the z-direction $|0, m\rangle$. Then the transformation properties of the states becomes explicit.

The state for a spin-s particle at rest in an eigenstate of spin, with spin pointing along the $\boldsymbol{s}$ direction, is given by

$$|0, \boldsymbol{s}\rangle = \mathcal{D}^s_{ms}(R(\boldsymbol{s}))|0, m\rangle \tag{68}$$

where $R(\boldsymbol{s})$ rotates a unit vector in the z-direction into $\boldsymbol{s}$ by first a rotation about y and then a rotation about z, a common convention. Note that we have constructed an eigenstate with spin eigenvector along $\boldsymbol{s}$ and the state is thus completely specified no matter what the spin of the particle is.

We will use the definitions of the Lorentz group generators given in Weinberg [19]. In this section we will write the three vector rotation and boost operators in terms of the integrated tensor M^{ij}

$$\begin{aligned} J_i &= \frac{1}{2}\epsilon_{ijk}M^{jk} \\ K_i &= M^{i0} \ . \end{aligned} \tag{69}$$

where on the left hand side $i = x, y, z$. For a particle of mass M in motion the boost (or canonical) state is defined by

$$|p, m\rangle = B(\boldsymbol{v})|0, m\rangle = \exp(i\zeta\, \hat{\boldsymbol{p}} \cdot \boldsymbol{K})|0, m\rangle \ , \tag{70}$$

where $\boldsymbol{v} = \boldsymbol{p}/p_0$, $\cosh\zeta = p_0/M$, and $\hat{\boldsymbol{p}}$ is the unit vector along $\boldsymbol{p}$.

Now consider a rotation about axis-i through an angle β. The unitary operator which effects this in Hilbert space is given in terms of the angular momentum operator J_i:

$$U[R_i(\beta)] = \exp(-i\beta J_i) \tag{71}$$

and for a particle of arbitrary spin-s

$$U[R_i(\beta))]|p, m\rangle = |R_i(\beta)p, n\rangle \mathcal{D}^s_{nm}(R_W(p, \beta)) \ . \tag{72}$$

where $R_W(p,\beta)$ is the Wigner rotation (see for example [9]). For a pure rotation the Wigner rotation R_W is very simple

$$R_W(p,\beta) = R_i(\beta);$$

independent of p. Therefore,

$$\begin{aligned}\langle p', m'|U[R_i(\beta)]|p,m\rangle &= \langle p', m'|R_i(\beta)p, n\rangle \mathcal{D}^s_{nm}(R_i(\beta)) \\ &= 2p_0(2\pi)^3\delta^{(3)}(\boldsymbol{p'} - R_i(\beta)\boldsymbol{p})\mathcal{D}^s_{m'm}(R_i(\beta)) \ ,\end{aligned}$$

using the conventional normalization

$$\langle p', m'|p,m\rangle = 2p_0(2\pi)^3\delta^{(3)}(\boldsymbol{p'} - \boldsymbol{p})\delta_{m'm} \ . \tag{73}$$

Thus

$$\begin{aligned}\langle p', m'|J_i|p,m\rangle &= i\,\frac{\partial}{\partial\beta}\langle p', m'|U[R_i(\beta)]|p,m\rangle|_{\beta=0} \\ &= 2p_0(2\pi)^3\left[i\epsilon_{ijk}p_j\frac{\partial}{\partial p_k}\delta_{m'm}\right. \\ &\quad \left. + i\frac{\partial}{\partial\beta}\mathcal{D}^s_{m'm}(R_i(\beta))\big|_{\beta=0}\right]\delta^{(3)}(\boldsymbol{p'} - \boldsymbol{p}) \ .\end{aligned}$$

Now [22]

$$i\frac{\partial}{\partial\beta}\mathcal{D}^s_{m'm}(R_i(\beta))\big|_{\beta=0} = (\mathrm{S}_i)_{m'm}$$

where the three $(2s+1)$ dimensional matrices S_i are the spin matrices for spin-s which satisfy

$$[\mathrm{S}_j\,, \mathrm{S}_k] = i\epsilon_{jkl}\,\mathrm{S}_l$$

Thus, our final result for the matrix elements of the angular momentum operators for arbitrary spin, from Eq. (74), becomes

$$\langle p', m'|J_i|p,m\rangle = 2p_0(2\pi)^3\left[\mathrm{S}_i + i\epsilon_{ijk}p_j\frac{\partial}{\partial p_k}\right]_{m'm}\delta^{(3)}(\boldsymbol{p'} - \boldsymbol{p}) \ . \tag{74}$$

For spin-$\frac{1}{2}$, of course, the S_i are just $\frac{1}{2}$ times the Pauli matrices σ_i. For arbitrary spin they are still very simple:

$$\begin{aligned}(\mathrm{S}_z)_{m'm} &= m\,\delta_{m'm} \\ (\mathrm{S}_x)_{m'm} &= \frac{1}{2}\left[C(s,m)\,\delta_{m',m+1} + C(s,-m)\,\delta_{m',m-1}\right] \\ (\mathrm{S}_y)_{m'm} &= \frac{-i}{2}\left[C(s,m)\,\delta_{m',m+1} - C(s,-m)\,\delta_{m',m-1}\right]\end{aligned}$$

where

$$C(s,m) = \sqrt{(s-m)(s+m+1)}$$

For the case of spin-$\frac{1}{2}$ Eq. (74) is exactly equivalent to the result one obtains after much labor using the wave packet approach. It is completely general. The second term will vanish if integrated over symmetric wave packets, so does not appear in the wave packet treatment. However it must be kept for analyzing the Lorentz transformation properties, as we will see, and must, as usual, always be interpreted in the sense of partial integration. It is very easy to verify that the form Eq. (74) satisfies the usual commutation relation relations and so is consistent with rotational invariance.

Combining the result Eq. (74) for the case of spin-$\frac{1}{2}$ with Eq. (68) leads directly to

$$\langle p', \boldsymbol{s}|J_i|p, \boldsymbol{s}\rangle = 2p^0(2\pi)^3 \left[\frac{1}{2}\, s_i + i(\boldsymbol{p} \times \boldsymbol{\nabla}_{\boldsymbol{p}})_i\right] \delta^3(\boldsymbol{p'} - \boldsymbol{p}) \,. \qquad (75)$$

We shall compare this result with that of [7] in Sect. 8.4.

Note that the term involving the derivative of a delta-function is a particular manifestation of the ambiguous integral in Eq. (55) and in a wave packet treatment corresponds to the orbital angular momentum about the origin of the packet as a whole, and vanishes for a symmetric packet. However, when dealing with the matrix elements between particles of definite momentum, as in Eq. (75), it is essential to keep the delta-function term and, moreover, to interpret it as explained in Eq. (61). In fact it will play a crucial role in the comparison between matrix elements involving canonical states and helicity states. We shall refer to this term as the *orbital angular momentum*.

For the purpose of deriving sum rules our result for the matrix elements *non-diagonal* in the spin label is actually more useful, namely, for a spin-1/2 particle

$$\langle p', m'|J_i|p, m\rangle = 2p_0(2\pi)^3 \left[\frac{1}{2}\sigma_i + i\epsilon_{ijk}p_j\frac{\partial}{\partial p_k}\right]_{m'm} \delta^{(3)}(\boldsymbol{p'} - \boldsymbol{p}) \,. \qquad (76)$$

The proof that these results are consistent with Lorentz invariance is rather complicated. It can be shown [10] that the matrix elements of the rotation and boost operators are consistent with the commmutation relations of these operators.

7.2 Helicity State Matrix Elements

We now turn to the case of helicity states which have some rather surprising properties. One can proceed just as in the canonical case; the main difference is that the Wigner rotation becomes a Wick helicity rotation (see for example [9]), always about the z-axis. This simplifies things somewhat; all the complication is in calculating the rotation angle. The result is also convention

dependent, depending on whether one uses the original Jacob and Wick definition [18] or the later one due to Wick [15] [see Eqs. (58) and (59)]. We give here the result for the first case. The result of this messy calculation is that, for $\boldsymbol{p} = (p, \theta, \phi)$

$$\langle p', \lambda' | J_i | p, \lambda \rangle_{JW} = (2\pi)^3 2p_0 \left[\lambda \eta_i + i(\boldsymbol{p} \times \boldsymbol{\nabla})_i\right] \delta^{(3)}(\boldsymbol{p}' - \boldsymbol{p}) \delta_{\lambda'\lambda} \tag{77}$$

where

$$\eta_x = \cos(\phi)\tan(\theta/2), \qquad \eta_y = \sin(\phi)\tan(\theta/2), \qquad \eta_z = 1. \tag{78}$$

Although these components look a little odd – the singularity at $\theta = \pi$ results from the ambiguity of Jacob and Wick helicity states at that point – it is easy to verify some important properties: they are manifestly diagonal in λ, which is required since rotations preserve the helicity, and they satisfy the requirement that the projection of $\boldsymbol{J}$ along the direction of motion i.e. along $\hat{\boldsymbol{p}}$ gives the helicity. Namely one finds

$$\langle p', \lambda' | \hat{\boldsymbol{p}} \cdot \boldsymbol{J} | p, \lambda \rangle_{JW} = \lambda \; 2p_0 \, (2\pi)^3 \delta^{(3)}(\boldsymbol{p}' - \boldsymbol{p}) \; \delta_{\lambda'\lambda}$$

and no orbital angular momentum piece survives as expected.

It is enlightening to consider these amplitudes from a different direction: comparing the definitions of canonical (boost) states to helicity states we have for the case of spin-1/2

$$\begin{aligned} |p, \boldsymbol{s}\rangle &= |p, m\rangle \mathcal{D}^{1/2}_{m1/2}(R(\boldsymbol{s})) \\ &= |p, \lambda\rangle_{JW} \mathcal{D}^{1/2}_{\lambda m}(R^{-1}(\boldsymbol{p})) \mathcal{D}^{1/2}_{m1/2}(R(\boldsymbol{s})) \\ &= |p, \lambda\rangle_{JW} \mathcal{D}^{1/2}_{\lambda 1/2}(R^{-1}(\boldsymbol{p}) R(\boldsymbol{s})) \; . \end{aligned}$$

This has the appearance of an ordinary unitary change of basis, but because of the compound nature of J_i when we apply this to the canonical form, using the spin-1/2 version of Eq. (75), we get

$$\begin{aligned} \langle p', \lambda' | J_i | p, \lambda \rangle_{JW} = (2\pi)^3 \, 2p_0 \mathcal{D}^{1/2}_{m\lambda}(R(\boldsymbol{p})) \mathcal{D}^{1/2}_{m'\lambda'}(R(\boldsymbol{p}'))^* \\ \times \left[i\epsilon_{ijk} p_j \partial_k + \frac{1}{2} \sigma_i \right]_{m'm} \delta^{(3)}(\boldsymbol{p}' - \boldsymbol{p}) \; . \end{aligned}$$

We *cannot* use the unitarity of the $\mathcal{D}$'s because $\boldsymbol{p} \neq \boldsymbol{p}'$, and we must first pass the first $\mathcal{D}^{1/2}(R(\boldsymbol{p}))$ through the derivative before setting them equal. This produces an extra term

$$-(2\pi)^3 \, 2p_0 \, i\epsilon_{ijk} \mathcal{D}^{1/2}_{m'\lambda'}(R(\boldsymbol{p}'))^* p_j \partial_k \mathcal{D}^{1/2}_{m\lambda}(R(\boldsymbol{p})) \delta^{(3)}(\boldsymbol{p}' - \boldsymbol{p}) \tag{79}$$

which is tedious to evaluate in the general case. The result of this labour is identical to Eqs. (77, 78).

8 Angular Momentum Sum Rules

Now that we have an expression for the matrix elements of the angular momentum operators we can equate the results using on the one hand the nucleon state itself, on the other an expression for the nucleon state in terms of its constituents i.e. a Fock space expansion of the nucleon state.

8.1 General Structure of Sum Rules: Parton Transverse Momentum

Consider a nucleon with momentum along OZ, $\boldsymbol{p} = (0, 0, p)$, in a canonical spin state with rest-frame spin eigenvector along $\boldsymbol{s}$, where $\boldsymbol{s}$ could be longitudinal $\boldsymbol{s}_L$ or transverse $\boldsymbol{s}_T$. Sum rules can be constructed by equating the expression Eq. (76) for the nucleon matrix elements $\langle \boldsymbol{p}', m' | J_i | \boldsymbol{p}, m \rangle$ with the expression obtained when the nucleon state is expressed in terms of the wave functions of its constituents (partons; quarks and gluons). Recall, however, that the parton picture of the nucleon is only supposed to be valid when observing a very fast moving nucleon, so we will take limit $p \to \infty$ at the end in order to obtain the parton model sum rules.

There is great interest in such sum rules especially if the partonic quantities can be related to other physically measurable quantities. A classic example was Eq. (29), which, as discussed earlier, gave rise to the "spin crisis". We will now investigate more carefully the origin and generalization of Eq. (29) and look at other similar possibilities, using Eq. (76) as the relevant starting point.

We have mentioned the importance of a wave-packet approach in order to deal with the derivative of the delta-function in the equations above. As it happens, however, when constructing sum rules, the expression in terms of constituents automatically produces a term which cancels the delta-function, irrespective of the actual model wave-functions used, so we need only concern ourselves with the spin term in Eq. (76) for the *nucleon state* matrix element. However, this term must be retained in the partonic state matrix elements.

The nucleon state is expanded as a superposition of n-parton Fock states. It is clear from such a basic concept as the Uncertainty Relations that if the partons are confined inside the nucleon they must possess both longitudinal AND transverse momentum. In the high energy, high momentum transfer reactions we have been considering it was usually assumed that the *intrinsic* parton transverse momentum $\boldsymbol{k}_T$ could be ignored (the collinear approximation discussed previously), but recently it has been appreciated that certain physical effects which are observed suggest that the transverse momentum is in some circumstances non-negligible. Moreover, it turns out that some partonic effects of transverse momentum are surprisingly large [20, 21] and can generate phenomena which would be impossible to reproduce in the collinear treatment:

– the presence of an intrinsic $\boldsymbol{k}_T$ alters the relationship between the light-cone momentum fraction x of the parton and the Bjorken-x_{Bj}, so that $x \neq x_{Bj}$. Although the shift is small and proportional to $\boldsymbol{k}_T^2/(x\sqrt{s})^2$, it can have a substantial effect in the region of x where the parton densities are varying rapidly. This is a kind of *enhanced* effect and can lead to up to an order of magnitude change in a cross section.
– In the presence of transverse momentum, certain spin-dependent effects can be generated by soft mechanisms and can be used to understand the large transverse single spin asymmetries (SSA) found in many reactions like $A^{\uparrow} + B \to C + X$ and the large hyperon polarizations in processes like $A + B \to H^{\uparrow} + X$.

Thus in the partonic Fock states and their associated wave functions the state of the i^{th} parton must be specified by a full three-dimensional momentum $\boldsymbol{k}_i$.

For a nucleon moving along OZ with momentum P, the parton densities $q(\boldsymbol{k}) = q(k_z, \boldsymbol{k}_T)$ will now specify the number density of partons with momentum in $d^3\boldsymbol{k} = dk_z d^2\boldsymbol{k}_T = P dx d^2\boldsymbol{k}_T$, with

$$q(x) = \int d^2\boldsymbol{k}_T q(x, \boldsymbol{k}_T) \tag{80}$$

where $q(x, \boldsymbol{k}_T) = P\, q(k_z, \boldsymbol{k}_T)$.

The calculation of the angular momentum matrix elements is long and notationally complicated, so we shall be rather schematic, and will not display flavour and colour labels. We write, for the nucleon state,

$$|\boldsymbol{p}, m\rangle = [(2\pi)^3 2p_0]^{1/2} \sum_n \sum_{\{\sigma\}} \int \frac{d^3\boldsymbol{k}_1}{\sqrt{(2\pi)^3 2k_1^0}} \cdots \frac{d^3\boldsymbol{k}_n}{\sqrt{(2\pi)^3 2k_n^0}} \times \psi_{\boldsymbol{p},m}(\boldsymbol{k}_1, \sigma_1, \cdots \boldsymbol{k}_n, \sigma_n) \delta^{(3)}(\boldsymbol{p} - \boldsymbol{k}_1 \cdots - \boldsymbol{k}_n) |\boldsymbol{k}_1, \sigma_1, \cdots \boldsymbol{k}_n, \sigma_n\rangle \ . \tag{81}$$

where σ_i denotes either the spin projection on the z-axis or the helicity, as appropriate. $\psi_{\boldsymbol{p},m}$ is the partonic wave function of the nucleon normalized so that

$$\sum_{\{\sigma\}} \int d^3\boldsymbol{k}_1 \cdots d^3\boldsymbol{k}_n |\psi_{\boldsymbol{p},m}(\boldsymbol{k}_1, \sigma_1, \cdots \boldsymbol{k}_n, \sigma_n)|^2 \delta^{(3)}(\boldsymbol{p} - \boldsymbol{k}_1 - \cdots - \boldsymbol{k}_n) = \mathcal{P}_n \ . \tag{82}$$

with $\mathcal{P}_n$ denoting the probability of the n-parton state. We substitute this expansion for the nucleon state in the matrix element of the angular momentum operators and we take for the Fock-state matrix elements

$$\langle \boldsymbol{k}_1', \sigma_1', \ldots, \boldsymbol{k}_n', \sigma_n' | J_i | \boldsymbol{k}_1, \sigma_1, \ldots, \boldsymbol{k}_n, \sigma_n \rangle = \sum_r \langle \boldsymbol{k}_r', \sigma_r' | J_i | \boldsymbol{k}_r, \sigma_r \rangle \prod_{l \neq r} (2\pi)^3 \times 2k_l^0 \delta^{(3)}(\boldsymbol{k}_l' - \boldsymbol{k}_l) \delta_{\sigma_l' \sigma_l} \ . \tag{83}$$

After some manipulation the nucleon matrix element can be written as:

$$\langle \boldsymbol{p}', m'|J_i|\boldsymbol{p}, m\rangle = (2\pi)^3\, 2p_0 \sum_n \sum_{\sigma,\sigma'} \int d^3k\, d^3k'\, \delta^{(3)}(\boldsymbol{p}' - \boldsymbol{p} + \boldsymbol{k} - \boldsymbol{k}') \times \rho^{m'\,m}_{\sigma'\,\sigma}(\boldsymbol{k}', \boldsymbol{k})^a \frac{1}{\sqrt{(2\pi)^3 2k'_0}} \langle \boldsymbol{k}', \sigma'|J_i|\boldsymbol{k}, \sigma\rangle \frac{1}{\sqrt{(2\pi)^3 2k_0}} \quad (84)$$

where we have introduced a density matrix for the internal motion of type "a" partons in a proton of momentum $\boldsymbol{p}$:

$$\begin{aligned}\rho^{m'\,m}_{\sigma'\,\sigma}(\boldsymbol{k}', \boldsymbol{k})^a \equiv & \sum_{n,r(a)} \sum_{\sigma_i} \sum_{\sigma'_r} \delta_{\sigma\,\sigma_r} \delta_{\sigma'\,\sigma'_r} \\ & \times \int d^3k'_r\, d^3k_1 \cdots d^3k_r \cdots d^3k_n\, \delta^{(3)}(\boldsymbol{k} - \boldsymbol{k_r})\, \delta^{(3)}(\boldsymbol{k}' - \boldsymbol{k}'_{\boldsymbol{r}}) \\ & \times \psi^*_{p'm'}(\boldsymbol{k_1}, \sigma_1, \cdots \boldsymbol{k}'_{\boldsymbol{r}}, \sigma'_r, \cdots \boldsymbol{k_n}, \sigma_n)\, \psi_{pm}(\boldsymbol{k_1}, \sigma_1, \cdots \boldsymbol{k_r}, \sigma_r, \cdots \boldsymbol{k_n}, \sigma_n) \\ & \times \delta^{(3)}(\boldsymbol{p} - \boldsymbol{k_1}\ \cdots - \boldsymbol{k_r}\ \cdots - \boldsymbol{k_n}) \end{aligned} \quad (85)$$

Here a, which we will frequently suppress, denotes the type of parton: quark, anti-quark or gluon. The sum goes over all Fock states and, within these states, over the spin and momentum labels r corresponding to the parton type a. Equations. (84) and (85) are the basis for the angular momentum sum rules.

The two terms in Eq. (76) applied to the parton matrix elements in Eq. (84) suggest a spin part and an orbital part for quarks and gluons. First consider the spin part of the matrix element when k is the momentum carried by a quark.

$$\langle \boldsymbol{p}', m'|J_i|\boldsymbol{p}, m\rangle^{\text{quarkspin}} = (2\pi)^3\, 2p_0\, \delta^{(3)}(\boldsymbol{p}' - \boldsymbol{p}) \int d^3k d^3k'\ \delta^{(3)}(\boldsymbol{k} - \boldsymbol{k}') \times \sum_{\sigma,\sigma'} \frac{1}{2} (\boldsymbol{\sigma}_i)_{\sigma'\,\sigma}\ \rho^{m'\,m}_{\sigma'\,\sigma}(\boldsymbol{k}', \boldsymbol{k})^q\ , \quad (86)$$

where here $\boldsymbol{\sigma}_i$ denotes the Pauli spin matrix of Eq. (76).

The spin part for the gluons is completely analogous, but now σ and σ' in Eqs. 84 and (85) refer to the gluon helicity λ. From Eq. (77), which is diagonal in helicity, we obtain

$$\langle \boldsymbol{p}', m'|J_i|\boldsymbol{p}, m\rangle^{\text{gluonspin}} = (2\pi)^3\, 2p_0\, \delta^{(3)}(\boldsymbol{p}' - \boldsymbol{p}) \int d^3k d^3k'\ \delta^{(3)}(\boldsymbol{k} - \boldsymbol{k}') \times \eta_i\, \lambda\, \rho^{m'\,m}_{\lambda\,\lambda}(\boldsymbol{k}', \boldsymbol{k})^G\ . \quad (87)$$

The orbital part is somewhat different because of the derivative of the δ-function that enters. We have mentioned the need for a proper wave packet treatment when dealing with states of definite momentum, but here the partons are not in plane wave states and the partonic wave function ψ plays the

role of a wave packet. Thus we may proceed directly by inserting the orbital piece of Eqs. (76,77) as was done for the spin part.

After some manipulation and integration by parts, the orbital piece produces two terms. One involves a derivative of a delta-function containing the nucleon momenta and eventually yields

$$2p_0(2\pi)^3 i\epsilon_{ijk}p_j\frac{\partial}{\partial p_k}\delta^{(3)}(\boldsymbol{p'}-\boldsymbol{p})\,\delta_{mm'} \tag{88}$$

which, as mentioned earlier, will just cancel the derivative of the delta-function in Eq. (76) arising from the matrix element between nucleon states.

The other term yields

$$2p_0(2\pi)^3\delta^{(3)}(\boldsymbol{p'}-\boldsymbol{p})\langle L_i\rangle^a_{m'\,m} \tag{89}$$

where $\langle L_i\rangle^a_{m'\,m}$ is the contribution from the *internal* angular momentum arising from partons of type a, given by

$$\begin{aligned}\langle L_i\rangle^a_{m'\,m} = &\sum_n\sum_{\{\sigma\}}\int d^3k_1\cdot\cdots\cdot d^3k_n\psi^*_{\boldsymbol{p},m'}(\boldsymbol{k}_1,\sigma_1,\ldots,\boldsymbol{k}_n,\sigma_n)\\ &\sum_{r(a)}\{[-i(\boldsymbol{k}_r\times\boldsymbol{\nabla}_{k_r})_i]\psi_{\boldsymbol{p},m}(\boldsymbol{k}_1,\sigma_1,\cdots\boldsymbol{k}_r,\sigma_r,\cdots\boldsymbol{k}_n,\sigma_n)\}\\ &\delta^{(3)}(\boldsymbol{p}-\boldsymbol{k}_1-\cdots-\boldsymbol{k}_n)\end{aligned} \tag{90}$$

where the sum over $r(a)$ means a sum over those r-values corresponding to partons of type a. Note that a can refer to both quarks and gluons; the structure of Eq. (90) is the same for both.

Putting Eqs. (89), (86) and (87) into Eq. (84), utilizing Eq. (76) for its LHS, and cancelling the factors $2p_0(2\pi)^3\delta(\boldsymbol{p'}-\boldsymbol{p})$, we end up with the most general sum rule for a spin-1/2 nucleon:

$$\begin{aligned}\frac{1}{2}(\boldsymbol{\sigma}_i)_{m'\,m} = \int d^3\boldsymbol{k}\ &\left[\frac{1}{2}(\boldsymbol{\sigma}_i)_{\sigma'\,\sigma}\ \rho^{m'\,m}_{\sigma'\,\sigma}(\boldsymbol{k},\boldsymbol{k})^{q+\bar{q}}+\lambda\ \eta_i(\boldsymbol{k})\ \rho^{m'\,m}_{\lambda\,\lambda}(\boldsymbol{k},\boldsymbol{k})^G\right]\\ &+\langle L_i\rangle^{q+\bar{q}}_{m'\,m}+\langle L_i\rangle^G_{m'\,m}\ .\end{aligned} \tag{91}$$

In fact Eq. (91) contains only two independent sum rules, a longitudinal one and a transverse one. Recall that the nucleon is moving along OZ. Then for $m=m'$ the terms in Eq. (91) are only non-vanishing when $i=z$ i.e. for J_z and parity ensures that the two cases $m=\pm 1/2$ give identical results – this leads to the longitudinal sum rule. For $m=-m'$ the terms are only non-vanishing for $i=x$ or $i=y$ i.e. for $J_{x,y}$. Again all four possible cases $m=\pm m'$; $i=x,y$ lead to the same result – this is the transverse sum rule.

The density matrix appearing in Eq. (91) is defined in terms of parton wave-functions in Eq. (85). We shall now show how, in the sum rules, this can be related to the parton densities utilized in DIS and other hard processes. We shall suppress all irrelevant labels in the following.

8.2 The Longitudinal Sum Rule

Consider the case of longitudinal polarization with $m = m' = 1/2$. The LHS of Eq. (91) is equal to 1/2. On the RHS, for the quark spin term in Eq. (91) the integrand will contain

$$\frac{1}{2}\rho^{++}_{++} - \frac{1}{2}\rho^{++}_{--} \tag{92}$$

where we have used $\pm$ to indicate $\pm 1/2$.

Now, schematically,

$$\rho^{m'\,m}_{\sigma'\,\sigma} = \sum_{X=all} \psi^*_{m'}(\sigma', X)\,\psi_m(\sigma, X)\ . \tag{93}$$

But the number density of quarks with spin along or opposite to OZ for a proton with spin along OZ can be expressed in terms of the quark wave-functions, namely

$$q_\pm(\boldsymbol{k}) = \sum_{X=all} |\psi_+(\pm, X)|^2\ . \tag{94}$$

We see therefore that the expression in Eq. (92) is just

$$\frac{1}{2}[q_+(\boldsymbol{k}) - q_-(\boldsymbol{k})]\ . \tag{95}$$

Substituting this expression into Eq. (91) and carrying out the integration over $d^2\boldsymbol{k}_T$ and using Eq. (80) and then Eq. (8) yields $\frac{1}{2}\int dx \Delta q(x)$. Summing over all quarks and antiquarks thus gives $\frac{1}{2}\Delta\Sigma$.

A similar analysis for the gluons yields $\Delta G = \int dx \Delta G(x)$.

Including the orbital terms in Eq. (91) we finally have the longitudinal sum rule

$$\frac{1}{2} = \frac{1}{2}\Delta\Sigma + \Delta G + \langle L^q_z\rangle + \langle L^G_z\rangle \tag{96}$$

to be compared to the simple parton model result Eq. (29)

$$S^{quarks}_z = \frac{1}{2}\Delta\Sigma = \frac{1}{2}\ . \tag{97}$$

Equation (96) is not new. It agrees with the result derived in [7], because the expression for the matrix elements of the angular momentum operators given in [7] is correct precisiely for the case where the spin of the nucleon is along its momentum. This is very fortunate, since the sum rule has such an intuitive structure that it had been in use for years, long before the formal proof was given in [7].

We will return to discuss Eq. (96) later, in particular to comment on the question of the Q^2 dependence which we have so far suppressed in Eq. (96).

8.3 The Transverse Case: The New Sum Rulexs

Let us turn now to the transverse sum rule. Here the nucleon is moving along OZ and is polarized along OX. The LHS of Eq. (91) is then equal to 1/2. The quark spin contribution to the RHS is

$$\frac{1}{2}\int d^3\boldsymbol{k}\frac{1}{2}\left[\rho^{+\,-}_{+\,-}+\rho^{+\,-}_{-\,+}+\rho^{-\,+}_{-\,+}+\rho^{-\,+}_{+\,-}\right] \tag{98}$$

where + and - refer to $\pm 1/2$. The relation of this expression to quark densities is here rather complicated. By rotating the system through π about the z-axis, it is easy to see that elements of $\rho^{m',m}_{\sigma',\sigma}$ with $(-1)^{m-m'-\sigma+\sigma'}=-1$ are odd under this rotation and so will integrate to zero when integrated over $\boldsymbol{k_T}$. This enables us to rewrite the expression Eq. (98), the quark contribution, in a way that has a nice interpretation, viz.

$$\frac{1}{2}\int d^3\boldsymbol{k}\,\frac{1}{2}\left[\rho^{+\,+}_{+\,-}+\rho^{+\,+}_{-\,+}+\rho^{-\,-}_{+\,-}+\rho^{-\,-}_{-\,+}+\rho^{+\,-}_{+\,-}+\rho^{+\,-}_{-\,+}+\rho^{-\,+}_{-\,+}+\rho^{-\,+}_{+\,-}\right]\ . \tag{99}$$

Consider the proton state with spin oriented along OX, perpendicular to the direction of motion

$$|\boldsymbol{p},\boldsymbol{s}_x\rangle=\frac{1}{\sqrt{2}}\{|\boldsymbol{p},m=1/2\rangle+|\boldsymbol{p},m=-1/2\rangle\} \tag{100}$$

To understand the content of expression Eq. (99) write schematically

$$\rho^{m'\,m}_{\sigma'\,\sigma}=\sum_{X=all}\psi^*_{m'}(\sigma',X)\,\psi_m(\sigma,X) \tag{101}$$

Now the number density of quarks with spin along or opposite to OX, denoted by $\pm\hat{\boldsymbol{s}}_x$ in a proton spinning along OX is

$$q_{\pm\hat{\boldsymbol{s}}_x/\boldsymbol{s}_x}(\boldsymbol{k})=\sum_{X=all}\left|\psi_{\boldsymbol{s}_x}(\pm\hat{\boldsymbol{s}}_x,X)\right|^2 \tag{102}$$

where, via Eq. (100),

$$\psi_{\boldsymbol{s}_x}(\pm\hat{\boldsymbol{s}}_x)=\frac{1}{2}[\psi_+(+)\pm\psi_+(-)+\psi_-(+)\pm\psi_-(-)] \tag{103}$$

so that (suppressing the $\sum_{X=all}$)

$$q_{\hat{\boldsymbol{s}_x}/\boldsymbol{s}_x}(\boldsymbol{k})-q_{-\hat{\boldsymbol{s}_x}/\boldsymbol{s}_x}(\boldsymbol{k})=Re\{\,[\,\psi_+(+)+\psi_-(+)\,]^*[\,\psi_-(-)+\psi_+(-)\,]\,\} \tag{104}$$

which, via Eq. (101), is exactly the integrand in Eq. (99). Thus the expression Eq. (98) is equal to

$$\frac{1}{2}\int d^3\boldsymbol{k}\,[\,q_{\hat{\boldsymbol{s}_x}/\boldsymbol{s}_x}(\boldsymbol{k})-q_{-\hat{s}_x/\boldsymbol{s}_x}(\boldsymbol{k})\,]=\frac{1}{2}\int dx\,d^2\boldsymbol{k}_T\,[\,q_{\hat{\boldsymbol{s}_x}/\boldsymbol{s}_x}(x,\boldsymbol{k}_T)-q_{-\hat{s}_x/\boldsymbol{s}_x}(x,\boldsymbol{k}_T)\,] \tag{105}$$

and there is an an analogous term for the antiquarks.

Now the structure of the integrand in Eq. (105) is known [23]. One has

$$q_{\hat{s}_x/s_x}(x,\boldsymbol{k}_T) - q_{-\hat{s}_x/s_x}(x,\boldsymbol{k}_T) = \Delta'_T(x,k_T^2) + \cos 2\phi \frac{k_T^2}{2M^2} h_{1T}^{\perp}(x,k_T^2) + \sin\phi \frac{k_T}{M} h_1^{\perp}(x,k_T^2) \tag{106}$$

where ϕ is the azimuthal angle of $\boldsymbol{k}_T$ and $\Delta'_T q^a(x,k_T^2)$ is related to the transverse density introduced in Eq. (10), namely,

$$\Delta_T q^a(x) = \int d^2\boldsymbol{k}_T \Delta'_T q^a(x,k_T^2) \; . \tag{107}$$

The unknown functions $h_{1T}^{\perp}(x,k_T^2)$ and $h_1^{\perp}(x,k_T^2)$ play no role in the sum rule, since their terms integrate out to zero.

Substituting Eq. (106) into Eq. (105) and integrating over the direction of $\boldsymbol{k}_T$, we end up with the quark spin contribution to the RHS of Eq. (91):

$$\frac{1}{2} \int dx \sum_{a,\bar{a}} \Delta_T q^a(x) \; . \tag{108}$$

We turn now to the gluon contribution to the RHS of Eq. (91), which is

$$\int d^3\boldsymbol{k}\, \eta_x(\boldsymbol{k}) \left(\frac{1}{2} \left[\rho_{1\,1}^{+\,+} - \rho_{-1\,-1}^{+\,+} + \rho_{1\,1}^{-\,-} - \rho_{-1\,-1}^{-\,-} + \rho_{1\,1}^{+\,-} - \rho_{-1\,-1}^{+\,-} + \rho_{1\,1}^{-\,+} - \rho_{-1\,-1}^{-\,+} \right] \right) , \tag{109}$$

where ± 1 refers to the gluon helicity. Once again we have added in terms which integrate to zero in order to get a nice interpretation in terms of densities. (Recall that in Eq. (78) η_x contains the factor $\cos\phi$, and the factor $\rho_{\lambda'\,\lambda}^{m',m}$ with $(-1)^{m'-m-\lambda'+\lambda} = +1$ can be shown to be even under $\phi \to \pi \pm \phi$.)

Now consider

$$\begin{aligned} \Delta G_{h/\boldsymbol{s}_x} &\equiv G_{1/\boldsymbol{s}_x} - G_{-1/\boldsymbol{s}_x} \\ &= \sum_{X=all} \{ |\psi_{\boldsymbol{s}_x}(1,X)|^2 - |\psi_{\boldsymbol{s}_x}(-1,X)|^2 \} \end{aligned} \tag{110}$$

Carrying out the analogue of Eq. (103) we find that the RHS of Eq. (110) is exactly equal to the terms in parenthesis in Eq. (109). Thus the gluon spin contribution to the RHS of Eq. (91) is

$$\int d^3\boldsymbol{k}\, \eta_x(\boldsymbol{k})\, \Delta G_{h/\boldsymbol{s}_x}(\boldsymbol{k}) = \int dx\, d^2\boldsymbol{k}_T\, \eta_x(\boldsymbol{k})\, \Delta G_{h/\boldsymbol{s}_x}(x,\, \boldsymbol{k}_T) \tag{111}$$

It is easy to see, geometrically, that $\Delta G_{h/\boldsymbol{s}_x}(x,\boldsymbol{k}_T)$ contains a factor k_x and we make this explicit by writing [24]

$$\Delta G_{h/\boldsymbol{s}_x}(x, \boldsymbol{k}_T) = \frac{k_x}{M} g_{1T}^G(x, k_T^2) \; . \tag{112}$$

Then the contribution of the gluon spin to the RHS of Eq. (91) i.e. to the proton whose spin is in the x-direction is

$$\begin{aligned} \Delta G_{h/\boldsymbol{s}_x} &= \int dx \, d^2\boldsymbol{k}_T \eta_x \frac{k_x}{M} g_{1T}^G(x, k_\perp^2) \\ &= \pi \int dx \, k_T dk_T \frac{\sqrt{x^2 p^2 + k_T^2} - x \; p}{M} g_{1T}^G(x, k_T^2)) \; . \end{aligned} \tag{113}$$

where we have used Eq. (78). As $p \to \infty$ this piece vanishes and so the gluon spin does not contribute to the transverse spin sum rule.

Finally, the internal orbital angular momentum terms $\langle L_x \rangle_{\boldsymbol{s}_x}^q$ and $\langle L_x \rangle_{\boldsymbol{s}_x}^G$ are obtained from Eq. (90) by the replacement

$$\psi_{\boldsymbol{p}, m} \to \psi_{\boldsymbol{p}, \boldsymbol{s}_x} = \frac{1}{\sqrt{2}} \left[\psi_{\boldsymbol{p}, 1/2} + \psi_{\boldsymbol{p}, -1/2} \right] \tag{114}$$

Putting together the various pieces of the RHS of Eq. (91) we obtain a new, transverse spin sum rule. Since the same result holds when considering J_y with the proton polarized along OY, we prefer to state the result in the more general form: for a proton in an eigenstate of transverse spin with eigenvector along $\boldsymbol{s}_T$

$$\frac{1}{2} = \frac{1}{2} \sum_{q, \bar{q}} \int dx \, \Delta_T q^a(x) + \sum_{q, \bar{q}, G} \langle L_{\boldsymbol{s}_T} \rangle^a \tag{115}$$

where $L_{\boldsymbol{s}_T}$ is the component of $\boldsymbol{L}$ along $\boldsymbol{s}_T$.

This has a very intuitive appearance, very similar to Eq. (96).

The function $\Delta_T q^a(x)$ does not play any role in DIS but could be measured in other processes, notably doubly polarized Drell-Yan reactions like

$$p(\boldsymbol{s}_T) + p(\boldsymbol{s}_T) \to leptonpair + X \; ,$$

in semi-inclusive hadronic reactions like

$$p + p(\boldsymbol{s}_T) \to H + X$$

where H is a detected hadron, typically a pion, [25], and in semi-inclusive DIS reactions with a transversely polarized target [26, 27] like

$$\ell + p(\boldsymbol{s}_T) \to \ell + H + X \; .$$

8.4 Comparison with Results in the Literature

The sum rules Eq. (96) and Eq. (115) are based upon our expression Eq. (91) for the matrix elements of the angular momentum operators, and as stressed

earlier, this expression is in disagreement with the results in the literature. It is interesting to compare results and the consequences for the sum rules.

If we rewrite the spin type term in the Jaffe-Manohar result [7] in terms of the independent vectors $\boldsymbol{p}$ and $\boldsymbol{s}$, we find, for the expectation value

$$\langle J_i \rangle_{JM} = \frac{1}{4Mp^0} \left\{ (3p_0^2 - M^2)s_i - \frac{3p_0 + M}{p_0 + M} (\boldsymbol{p} \cdot \boldsymbol{s}) p_i \right\} \tag{116}$$

to be compared to our result

$$\langle J_i \rangle = \frac{1}{2} s_i \tag{117}$$

arising from the first term in Eq. (75). In general these are different. However, one may easily check that if $\boldsymbol{s} = \hat{\boldsymbol{p}}$ the Jaffe-Manohar value agrees with Eq. (117), while if $\boldsymbol{s} \perp \hat{\boldsymbol{p}}$ they are not the same.

The agreement for $\boldsymbol{s} = \hat{\boldsymbol{p}}$ is consistent with the much used and intuitive longitudinal sum rule given in Eq. (96).

But for the transverse case no sum rule is possible with the Jaffe-Manohar formula because, as $p \to \infty$, Eq. (116) for $i = x, y$ diverges.

9 Interpretation of the Sum Rules

Once it is accepted that the partons possess intrinsic transverse momentum i.e. perpendicular to the motion along OZ of the nucleon, they can have orbital angular momentum $\boldsymbol{L}$ with a component along OZ, as exemplified in Eq. (96), which also allows for the possibility that the gluons are polarized and can have orbital angular momentum. As discussed earlier, the axial anomaly allows us to escape the conclusion that the EMC result implies a very small value of $\Delta\Sigma$, so it is not impossible that the RHS of Eq. (91) is dominated by $\Delta\Sigma$. On the other hand, for an object the size of a nucleon the r in an orbital angular momentum term like $\boldsymbol{r} \times \boldsymbol{p}$ can be of order 1 fermi and with p_T of order a few hundred MeV, it is easy to produce half a unit of orbital angular momentum along OZ.

In the simple parton model we visualize partons almost like real physical particles and the above comments are meaningful in such a framework. We assume implicitly that quantities like $\Delta\Sigma$, for example, have some objective physical significance. Unfortunately QCD teaches us the unpleasant fact that this is not so. Firstly the quantities $\Delta\Sigma$ and ΔG on the RHS of Eq. (96), as determined from studies of polarized DIS, depend on Q^2. Of course the Q^2 of a DIS experiment has no meaning in Eq. (96) which refers to an isolated nucleon, but, as discussed earlier, what Q^2 actually corresponds to, in a parton density, is the value of the *renormalization scale* and the *factorization sclae*, taken equal for simplicity, and taken equal to Q^2 in DIS. So measuring a given parton density in a reaction at some value of Q^2 really means measuring the parton density at renormalization scale $\mu_R^2 =$ factorization scale $\mu_F^2 = Q^2$.

But what makes the interpretation even more problematic is that at the level of what is called *next to leading order* or NLO perturbative QCD the behaviour of the quantities in Eq. (96) becomes also scheme dependent i.e. depends upon what type of renormalization scheme is being utilized, so they cannot be *physical quantities*. Moreover, there are, in principle, an infinite number of possible renormalization schemes!

Perhaps the most natural scheme is the so-called JET scheme (for a discussion of this and other schemes see e.g. [28]), which has the nice property that the first moment $\Delta\Sigma$ is independent of the renormalization scale, or, in the context of DIS, is independent of Q^2. Thus there would seem to be some physical sense in thinking of this as the spin carried by the quarks. In the most recent analysis of the world data on polarized DIS it was found that $\Delta\Sigma = 0.32 \pm 0.06$ in the JET scheme [29]. This is considerably larger than the result in the infamous EMC experiment, but still a long way from the naive value 1. However, in this scheme, the other three terms on the RHS of Eq. (96) do depend on μ^2, so it is not clear what physical significance they have. Since, as already mentioned, $\Delta G(Q^2)$ grows like lnQ^2 as Q^2 increases, it must be that the orbital angular momentum terms become large and negative as Q^2 increases.

All of this is highly unintuitive, and we are left with the unpleasant, but unavoidable, conclusion that in higher orders of QCD the partons more and more lose their nice simple particle-like properties, and the question, as to which constituents carry what fraction of the nucleon's spin, becomes more and more intractable and, maybe, meaningless.

References

1. European Muon Collaboration: J. Ashman et al: Phys. Lett. **B** 206, 364 (1988)
2. M. Anselmino and E. Leader: Z. Phys. **C 41**, 239 (1988)
3. A.V. Efremov and O.V. Teryaev: JINR, Report No. E2–88–287 (1988); unpublished
4. G. Altarelli and G.G. Ross: Phys. Lett. **B 212**, 391 (1988)
5. R.D. Carlitz, J.C. Collins and A.H. Mueller: Phys. Lett. **B 214**, 229 (1988)
6. E. Leader and M. Anselmino: Santa Barbara Preprint NSF–88–142, July 1988, unpublished
7. R. L. Jaffe and A. Manohar: Nucl. Phys. **B 337**, 509 (1990)
8. M. Anselmino, A. Efremov, and E. Leader: Phys. Rep. **261**, 4 (1995)
9. E. Leader: *Spin in Particle Physics*, Cambridge University Press (2001)
10. B. L. G. Bakker, E. Leader, and T. L. Trueman: Phys. Rev. **D 70**, 114001 (2004)
11. E. Leader and E. Predazzi: *An Introduction to Gauge Theories and the New Physics*, Cambridge University Press (1985)
12. S.I. Adler: Phys. Rev. **177** (1969) 2426; J.S. Bell and R. Jackiw, N.C. **51A** (1969) 47
13. G. M. Shore and B. E. White: Nucl. Phys. **B 581**, 409 (2000)

14. H. P. Stapp: Phys. Rev. **125**, 2139 (1962)
15. G. C. Wick: Ann. Phys. **18**, 65 (1962)
16. J. D. Bjorken and S. C. Drell: *Relativistic Quantum Mechanics*, Mcgraw-Hill (1964)
17. M. E. Peskin and D. V. Schroeder: *An Introduction to Quantum Field Theory*, Westview Press (1995)
18. M. Jacob and G. C. Wick: Ann. Phys. **7**, 404 (1959)
19. S. Weinberg: *Quantum Theory of Fields*, Cambridge University Press (1995)
20. M. Anselmino, M. Boglione, U. D'Alesio, E. Leader, and F. Murgia: Phys. Rev. **D 71**, 014002 (2005)
21. U. D'Alesio and F. Murgia: Phys. Rev. **D 70**, 074009 (2004)
22. M. Rose: *Elementary Theory of Angular Momentum*, Wiley (1957)
23. V. Barone, A. Drago, and P. G. Ratcliffe: Phys. Rep. **359**, 1 (2002)
24. P. J. Mulders and R. D. Tangerman: Nucl. Phys. B **461**, 197 (1996); erratum Nucl. Phys. **B 484**, 538 (1997)
25. For access to the literature see: M. Boglione and E. Leade:, Phys. Rev. **D 57**, 5780 (1998)
26. D. Boer, and P. J. Mulders: Phys. Rev. **D 57**, 5780 (1998)
27. A. M. Kotzinian and P. J. Mulders: Phys. Lett. **B 406**, 373 (1997)
28. E. Leader, A. V. Sidorov, and D. B. Stamenov: Phys. Rev. **D 58**, 114028 (1998)
29. E. Leader, A. V. Sidorov, and D. B. Stamenov: Part. Nucl. Phys. Lett. **V1, No5**, 7 (2004)

Lecture Notes in Physics

For information about earlier volumes
please contact your bookseller or Springer
LNP Online archive: springerlink.com

Vol.666: D. Britz, Digital Simulation in Electrochemistry

Vol.667: W. D. Heiss (Ed.), Quantum Dots: a Doorway to Nanoscale Physics

Vol.668: H. Ocampo, S. Paycha, A. Vargas (Eds.), Geometric and Topological Methods for Quantum Field Theory

Vol.669: G. Amelino-Camelia, J. Kowalski-Glikman (Eds.), Planck Scale Effects in Astrophysics and Cosmology

Vol.670: A. Dinklage, G. Marx, T. Klinger, L. Schweikhard (Eds.), Plasma Physics

Vol.671: J.-R. Chazottes, B. Fernandez (Eds.), Dynamics of Coupled Map Lattices and of Related Spatially Extended Systems

Vol.672: R. Kh. Zeytounian, Topics in Hyposonic Flow Theory

Vol.673: C. Bona, C. Palenzula-Luque, Elements of Numerical Relativity

Vol.674: A. G. Hunt, Percolation Theory for Flow in Porous Media

Vol.675: M. Kröger, Models for Polymeric and Anisotropic Liquids

Vol.676: I. Galanakis, P. H. Dederichs (Eds.), Half-metallic Alloys

Vol.677: A. Loiseau, P. Launois, P. Petit, S. Roche, J.-P. Salvetat (Eds.), Understanding Carbon Nanotubes

Vol.678: M. Donath, W. Nolting (Eds.), Local-Moment Ferromagnets

Vol.679: A. Das, B. K. Chakrabarti (Eds.), Quantum Annealing and Related Optimization Methods

Vol.680: G. Cuniberti, G. Fagas, K. Richter (Eds.), Introducing Molecular Electronics

Vol.681: A. Llor, Statistical Hydrodynamic Models for Developed Mixing Instability Flows

Vol.682: J. Souchay (Ed.), Dynamics of Extended Celestial Bodies and Rings

Vol.683: R. Dvorak, F. Freistetter, J. Kurths (Eds.), Chaos and Stability in Planetary Systems

Vol.684: J. Dolinšek, M. Vilfan, S. Žumer (Eds.), Novel NMR and EPR Techniques

Vol.685: C. Klein, O. Richter, Ernst Equation and Riemann Surfaces

Vol.686: A. D. Yaghjian, Relativistic Dynamics of a Charged Sphere

Vol.687: J. W. LaBelle, R. A. Treumann (Eds.), Geospace Electromagnetic Waves and Radiation

Vol.688: M. C. Miguel, J. M. Rubi (Eds.), Jamming, Yielding, and Irreversible Deformation in Condensed Matter

Vol.689: W. Pötz, J. Fabian, U. Hohenester (Eds.), Quantum Coherence

Vol.690: J. Asch, A. Joye (Eds.), Mathematical Physics of Quantum Mechanics

Vol.691: S. S. Abdullaev, Construction of Mappings for Hamiltonian Systems and Their Applications

Vol.692: J. Frauendiener, D. J. W. Giulini, V. Perlick (Eds.), Analytical and Numerical Approaches to Mathematical Relativity

Vol.693: D. Alloin, R. Johnson, P. Lira (Eds.), Physics of Active Galactic Nuclei at all Scales

Vol.694: H. Schwoerer, J. Magill, B. Beleites (Eds.), Lasers and Nuclei

Vol.695: J. Dereziński, H. Siedentop (Eds.), Large Coulomb Systems

Vol.696: K.-S. Choi, J. E. Kim, Quarks and Leptons From Orbifolded Superstring

Vol.697: E. Beaurepaire, H. Bulou, F. Scheurer, J.-P. Kappler (Eds.), Magnetism: A Synchrotron Radiation Approach

Vol.698: S. Bellucci (Ed.), Supersymmetric Mechanics – Vol. 1

Vol.699: J.-P. Rozelot (Ed.), Solar and Heliospheric Origins of Space Weather Phenomena

Vol.700: J. Al-Khalili, E. Roeckl (Eds.), The Euroschool Lectures on Physics with Exotic Beams, Vol. II

Vol.701: S. Bellucci, S. Ferrara, A. Marrani, Supersymmetric Mechanics – Vol. 2

Vol.702: J. Ehlers, C. Lämmerzahl, Special Relativity

Vol.703: M. Ferrario, G. Ciccotti, K. Binder (Eds.), Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology Volume 1

Vol.704: M. Ferrario, G. Ciccotti, K. Binder (Eds.), Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology Volume 2

Vol.705: P. Bhattacharyya, B.K. Chakrabarti (Eds.), Modelling Critical and Catastrophic Phenomena in Geoscience

Vol.706: M.A.L. Marques, C.A. Ullrich, F. Nogueira, A. Rubio, K. Burke, E.K.U. Gross (Eds.), Time-Dependent Density Functional Theory

Vol.707: A.V. Shchepetilov, Calculus and Mechanics on Two-Point Homogenous Riemannian Spaces

Vol.708: F. Iachello, Lie Algebras and Applications

Vol.709: H.-J. Borchers and R.N. Sen, Mathematical Implications of Einstein-Weyl Causality

Vol.710: K. Hutter, A.A.F. van de Ven, A. Ursescu, Electromagnetic Field Matter Interactions in Thermoelastic Solids and Viscous Fluids

Vol.711: H. Linke, Controlled Nanoscale Motion in Biological and Artifical Systems

Vol.712: W. Pötz, J. Fabian, U. Hohenester (Eds.), Modern Aspects of Spin Physics