THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par *l'Université Toulouse III - Paul Sabatier*
Discipline ou spécialité : *MATHEMATIQUES APPLIQUEES*

Présentée et soutenue par *Stefan Possanner*
Le *07 âout 2012*

Titre : *Modeling and Simulation of Spin-Polarized Transport at the Kinetic and Diffusive Level*

JURY
M. Ferdinand Schuerrer (Rapporteur)
M. Ansgar Juengel (Rapporteur)
M. Roland Wurschum (Examinateur)
Mme. Claudia Negulescu (Examinateur)

Ecole doctorale : *EDMITT – Ecole Doctorale Mathématiques, Informatique et Télécommunication de Toulouse*
Unité de recherche : *IMT – Institut de Mathématiques de Toulouse*
Directeur(s) de Thèse : *Pierre DEGOND, Ferdinand SCHUERRER*
Rapporteurs :

UNIVERSITÉ TOULOUSE III - PAUL SABATIER
GRAZ UNIVERSITY OF TECHNOLOGY

DOCTORAL THESIS

for obtaining the academic degree of

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE
from the Université Toulouse III - Paul Sabatier

and

DOKTOR DER TECHNISCHEN WISSENSCHAFTEN
from the Graz University of Technology

in the field of: **Applied Mathematics, Theoretical Physics**

written by

Stefan POSSANNER

Title:

Modeling and Simulation of Spin-Polarized Transport at the Kinetic and Diffusive Level

---

supervised by

Pierre Degond	CNRS France
Ferdinand Schürrer	Graz University of Technology

Institut de Mathématiques de Toulouse
Équipe Mathématiques pour l'Industrie et la Physique (MIP)

Institute of Theoretical and Computational Physics
Graz University of Technology

Deutsche Fassung:
Beschluss der Curricula-Kommission für Bachelor-, Master- und Diplomstudien vom 10.11.2008
Genehmigung des Senates am 1.12.2008

EIDESSTATTLICHE ERKLÄRUNG

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommene Stellen als solche kenntlich gemacht habe.

Graz, am …… ……..
(Unterschrift)

Englische Fassung:

STATUTORY DECLARATION

I declare that I have authored this thesis independently, that I have not used other than the declared sources / resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

…… ……..
date (signature)

Acknowledgements

I am especially indebted to the late Naoufel Ben Abdallah, who supported me at the beginning of this thesis in any possible way. Naoufel gave me the opportunity to start my PhD in his group at the Institut de Mathématiques de Toulouse, which is a wonderful place for a young researcher. His kind personality and his enthusiasm for mathematics are the things I will remember the most about him. This thesis would not exist if not for Naoufel's effort to bring together people from different scientific disciplines and with different nationalities.

I am very grateful to Ferdinand Schürrer, who supported me in starting my scientific career at the Graz University of Technology even before this PhD thesis during the master thesis. One of his numerous international collaborations led to the stay in Toulouse during the first half of my PhD. I thank professor Schürrer for his steady support and for many fruitful and joyful discussions, in which his passion for physics (and for winter sports) gave me a great deal of inspiration.

I am much obliged to Claudia Negulescu for resuming the supervision of my scientific work after the demise of N. Ben Abdallah. I consider myself very fortunate to work with her and I am very grateful for the many hours she put into the development of this thesis. I am looking forward to joining Claudia for many scientific projects in the future.

During this thesis I was fortunate to meet a lot of wonderful people who share my enthusiasm for physics and mathematics and who I thank for contributing to portions of this work: Raymond el Hajj, Benjamin Stickler, Ewald Schachinger, Luigi Barletti, Florian Méhats.

Special thanks also to Brigitte Schwarz, Agnès Requis, Zohra Kallel and Janani Chandran for supporting me with administrative issues which can be a difficult task in a co-tutelle agreement.

Summary

The aim of this thesis is to contribute to the understanding of spin-induced phenomena in electron motion. These phenomena arise when electrons move through a (partially) magnetic environment, in such a way that its magnetic moment (spin) may interact with the surroundings. The pure quantum nature of the spin requires transport models that deal with effects like quantum coherence, entanglement (correlation) and quantum dissipation. On the meso- and macroscopic level it is not yet clear under which circumstances these quantum effects may transpire. The purpose of this work is, on the one hand, to derive novel spin transport models from basic principles and, on the other hand, to develop numerical algorithms that allow for a solution of these new and other existing model equations.

The thesis consists of four parts. The first part has introductory character; it comprises an overview of fundamental spin-related concepts in electronic transport such as the giant-magneto-resistance (GMR) effect, the spin-transfer torque in metallic magnetic multilayers and the matrix-character of transport equations that take spin-coherent electron states into account. Special emphasis is placed on the modeling of the spin-transfer torque which represents the intersection of these concepts. In particular, we consider the diffusive Zhang-Levy-Fert (ZLF) model, an exchange-torque model that consists of the Landau-Lifshitz equation and a heuristic matrix spin-diffusion equation. A finite difference scheme based on Strang operator splitting is developed that enables a numerical, self-consistent solution of this non-linear system within multilayer structures. Finally, the model is tested by comparison of numerical results to recent experimental data.

Parts two and three are the thematic core of this thesis. In part two we propose a matrix-Boltzmann equation that allows for the description of spin-coherent electron transport on a kinetic level. The novelty here is a linear collision operator in which the transition rates from momentum k to momentum k' are modeled by a 2×2 Hermitian matrix; hence the mean-free paths of spin-up and spin-down electrons are represented by the eigenvalues of this scattering matrix. After a formal derivation of the matrix-Vlasov equation as the semi-classical limit of the one-electron Wigner equation, the ensuing kinetic equation is studied with regard to existence, uniqueness

and positive semi-definiteness of a solution. Furthermore, the new collision operator is investigated rigorously and the diffusion limit $\tau_c \to 0$ of the mean scattering time is performed. The obtained matrix drift-diffusion equations are an improvement over the heuristic spin-diffusive model treated in part one. The latter is obtained in the limit of identical eigenvalues of the scattering matrix.

Part three is dedicated to a first step towards the derivation of the matrix collision operator, introduced in part two, from first principles. For this, we augment the von Neumann equation of a composite quantum system by a dissipative term that relaxes the total state operator towards the Born approximation. Under the premise that the relaxation is the dominant process we obtain a hierarchy of non-Markovian master equations. The latter arises from an expansion of the total state operator in powers of the relaxation time τ_r. In the Born-Markov limit $\tau_r \to 0$ the Lindblad master equation is recovered. It has the same structure as the collision operator proposed in part two heuristically . However, the Lindblad equation is still a microscopic equation; thus the next step would be to carry out the semi-classical limit of the result obtained.

In part four we perform a numerical study of a quantum-diffusive, two-component spin model of the transport in a two-dimensional electron gas with Rashba spin-orbit coupling. This model assumes the electrons to be in a quantum equilibrium state in the form of a Maxwellian operator. We present two space-time discretizations of the model which also comprise the Poisson equation. In a first step pure time discretization is applied in order to prove the well-posedness of the two schemes, both of which are based on a functional formalism to treat the non-local relations between spin densities via the chemical potentials. We then use fully space-time discrete schemes to simulate the dynamics in a typical transistor geometry. Finite difference approximations applied in these schemes are first order in time and second order in space. The discrete functionals introduced are minimized with the help of a conjugate gradient-based algorithm in which the Newton method is applied to find the desired line minima.

Keywords. *Spintronics, spin-transfer torque, Landau-Lifshitz equation, spin-diffusion, spin-coherent transport, Strang operator splitting, Wigner transform, semi-classical limit, matrix-Boltzmann equation, spin-polarized electron conduction, matrix collision operator, maximum principle, diffusion limit, matrix drift-diffusion, composite quantum systems, Lindblad equation, non-Markovian quantum dynamics, Rashba spin-orbit coupling, maximum entropy principle, finite difference, conjugate gradient method.*

Résumé

L'objectif de cette thèse est de contribuer à la compréhension des phénomènes de mouvement de l'électron induits par le spin. Ces phénomènes aparaissent lorsqu'un électron se déplace à travers un environnement (partiellement) magnétique, de telle sorte que son moment magnétique (spin) peut interagir avec l'environnement. La nature quantique pure du spin nécessite des modèles de transport qui traitent des effets comme la cohérence quantique, l'intrication (corrélation) et la dissipation quantique. Sur le niveau méso- et macroscopique, il n'est pas encore clair dans quelles circonstances ces effets quantiques du spin peut transparaitre. Le but de ce travail est, d'une part, de dériver des nouveaux modèles de transport de spin à partir des principes de base et, d'autre part, de développer des algorithmes numériques qui permettent de trouver une solution de ces modèles.

Cette thèse se compose de quatre parties. La première partie introductive contient un aperçu des concepts fondamentaux liés au transport polarisé en spin, tels que la magnéto-résistance géante (GMR), le couple de transfert de spin dans les multi-couches magnétiques et le caractère matriciel des équations de transport qui prennent en compte la cohérence de spin. L'accent est mis sur la modélisation du couple de transfert de spin, qui représente l'intersection de ces concepts. En particulier, nous considérons pour sa description le modèle diffusif de Zhang-Levy-Fert (ZLF) qui se compose de l'équation de Landau-Lifshitz et d'une équation de diffusion matricielle pour le spin. Un schéma de différences finies est développé pour résoudre numériquement ce système non-linéaire dans des structures multi-couches. Le modèle est testé par comparaison des résultats obtenus aux données expérimentales récentes.

Les parties deux et trois forment le noyau thématique de cette thèse. Dans la deuxième partie nous proposons une équation de Boltzmann matricielle qui permet la description de la cohérence de spin sur le niveau cinétique. La nouveauté est un opérateur de collision dans lequel les taux de transition de la quantité de mouvement sont modélisés par une matrice 2×2 hermitienne; par conséquent, les libre parcours moyens des électrons spin-up et spin-down sont représentés par les valeurs propres de cette matrice de scattering. Après une dérivation formelle de l'équation de Vlasov

matricielle à partir de l'équation de Wigner, l'équation cinétique qui suit est étudiée en ce qui concerne l'existence, l'unicité et la positivé d'une solution. En outre, le nouveau opérateur de collision est étudié rigoureusement et la limite de diffusion $\tau_c \to 0$, correspondant à l'annulation de la moyenne de temps de scattering, est effectué. Les équations de drift-diffusion matricielle qui sont obtenues représentent une amélioration par rapport au modèle traité dans la première partie. Ce dernier est obtenu dans la limite ou la différence entre les deux valeurs propres de la matrice de scattering va disparaître.

La troisième partie est consacrée à l'obtention de l'opérateur de collision matricielle introduit auparavant, à partir des principes quantiques. Pour cela, nous augmentons l'équation de von Neumann d'un système composite par un terme dissipatif qui fait tendre l'opérateur de densité totale vers l'approximation de Born. En vertu de la prémisse que la relaxation est le processus dominant, on obtient une hiérarchie d'équations non-Markoviennes. Celles-ci découlent d'une expansion de l'opérateur de densité en termes de τ_r, le temps de relaxation. Dans la limite de Born-Markov, $\tau_r \to 0$, l'équation de Lindblad est récupérée. Elle a la même structure que l'opérateur de collision proposé dans la deuxième partie. Cependant, l'équation de Lindblad est encore une équation microscopique; donc la prochaine étape serait de procéder à la limite semi-classique du résultat obtenu.

Dans la quatrième partie nous procédons à une étude numérique d'un modèle quantique-diffusif de spin qui décrit le transport dans un gaz d'électrons bidimensionnel avec un couplage spin-orbite de Rashba. Ce modèle suppose que les électrons sont dans un état d'équilibre quantique sous la forme d'un opérateur de Maxwell. Nous présentons deux discrétisations espace-temps du modèle couplé par l'équation de Poisson. Dans une première étape on applique une discrétisation en temps et on montre que les systèmes sont bien définis. Ceux-ci sont basés sur un formalisme fonctionnel pour traiter les relations non-locales entre les densités de spin. Nous utilisons ensuite des discrétisations espace-temps pour simuler la dynamique dans une géométrie typique d'un transistor. Les approximations différences finies sont du premier ordre en temps et du second ordre en espace. Les fonctionnelles discrètes sont minimisée à l'aide d'un algorithme du gradient conjugué et la méthode de Newton est appliquée afin de trouver les minima dans la direction désirée.

Table of contents

Introductory Overview

In conventional nanoelectronics the charge and the magnetic moment (spin) of electrons are used for different purposes; logical operations and fast (volatile) memory elements are implemented via the control of charge transport by means of electric fields, for instance in diodes, transistors or capacitors, whereas the magnetic properties are used mainly for the purpose of non-volatile, long-term data storage in hard disk drives. The miniaturization of these devices is gradually reaching its physical boundaries and, thus, considerable effort is put into the search for novel device concepts. A promising approach is the combined usage of the electron's charge- and spin-degree of freedom in so-called *spintronics* applications [1, 2, 3]. In these devices one uses either an electric field to control magnetic properties or a magnetic field to control the properties of charge transport. This thesis is devoted to the modeling and the numerical simulation of some of the interesting phenomena arising in spintronics applications. The emphasis is on the meso- and the macroscopic level of description, which appears to be well suited for performing device-related numerical studies. The goal is, on the one hand, to improve some of the state-of-the-art spin-transport models and, on the other hand, to provide numerical data that enable a sophisticated interpretation of recent experimental findings. The thesis consists of four separate parts which are loosely related. These parts are sorted in chronological order of their making and they are concerned with the following topics:

***Part I*:** The first chapter of this part can be viewed as an introduction to the field of spintronics. We revisit the basic ideas of spin-polarized electron conduction in the transition-metal ferromagnets Fe, Ni and Co, of the Giant-Magneto-Resistance (GMR) effect and of the non-equilibrium spin accumulation at a non-magnetic/ferromagnetic interface that is traversed by an electronic current. In Chapter 2 we treat numerically the Zhang-Levy-Fert (ZLF) model [4] to describe the spin-transfer torque and the ensuing magnetization dynamics in magnetic multilayers under high current densities. For this, we solve self-consistently the following non-linear system of equations for the magnetization $\vec{m} : \mathbb{R} \times \mathbb{R}^+ \to S^2$ and the spin accumulation $\vec{s} : \mathbb{R} \times \mathbb{R}^+ \to \mathbb{R}^3$:

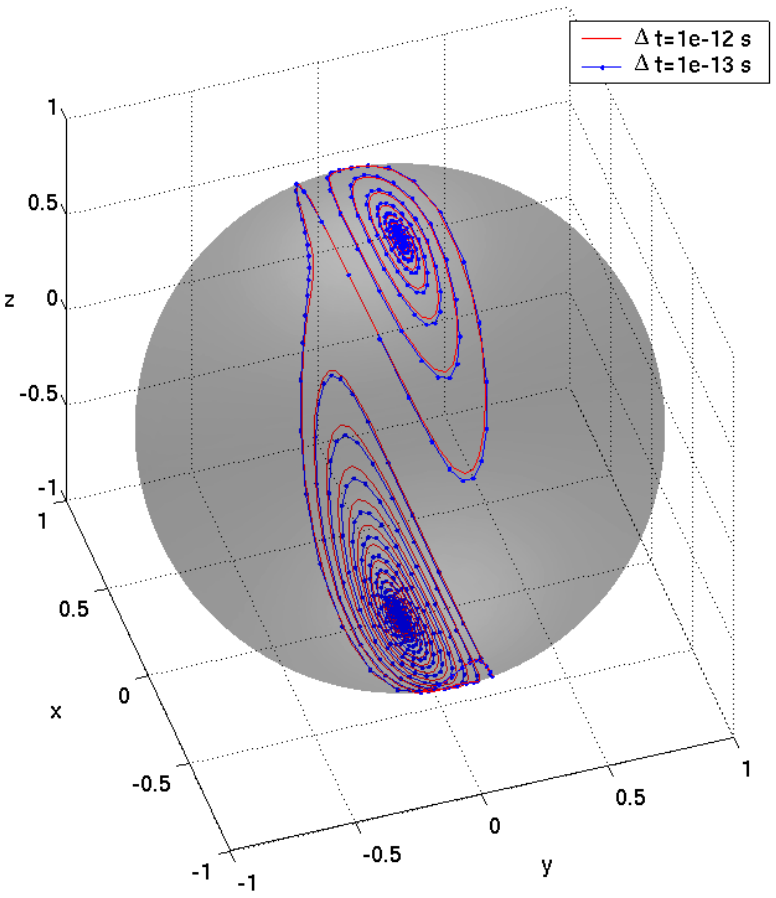


Figure 1: Simulated magnetization dynamics in the thin ferromagnetic layer of a spin-valve (trilayer system) with intial condition $\vec{m}(x,0) = (0,0,1)$ in the respective domain. Due to the spin-transfer torque exerted by the spin accumulation $\vec{s}$, the magnetization switches to the steady-state $\vec{m}(x,\infty) = (0,0,-1)$.

$$\partial_t \vec{m} = -\gamma\mu_0 \vec{m} \times \vec{H}_{eff} + \alpha \vec{m} \times \partial_t \vec{m}\,, \qquad \partial_t \vec{s} = D\Delta\vec{s} - \frac{\vec{s}}{\tau} - \frac{J_{ex}}{\hbar}\vec{s} \times \vec{m}\,,$$

where γ is the gyromagnetic ratio, μ_0 stands for the magnetic constant, α denotes the damping parameter, $\vec{H}_{eff}$ is the effective field, D the diffusion constant, Δ stands for the Laplacian, τ is the spin-flip scattering time and J_{ex} denotes the *sd*-exchange constant. The effective field $\vec{H}_{eff}$ depends on $\vec{m}$ (and spatial derivatives thereof) and $\vec{s}$ in a local manner. The equation for $\vec{m}$ is the Landau-Lifshitz-Gilbert equation (LLGE) and the equation for $\vec{s}$ is called the spin-diffusion equation (SDE). Initial and boundary conditions are chosen to model the spin-transfer torque, which arises due to the exchange coupling of strength J_{ex}, in a trilayer system, where a non-magnetic metallic spacer is sandwiched between a thick and a thin ferromagnet. The algorithm is based on Strang operator splitting to treat the different time scales and the strong non-linearities in the LLGE. A typical simulation result is depicted in Fig. 1. It shows the time evolution of the magnetization in the thin ferromagnetic

layer with the initial condition $\vec{m}(x,0) = (0,0,1)$ in the respective domain. It can can be seen that the magnetization switches in the course of time, ending its motion in the steady-state $\vec{m}(x,\infty) = (0,0,-1)$. The developed algorithm is used to reproduce and interpret recent experimental data by Krivorotov et al. [5]. It is shown that the ZLF model is appropriately suited for describing the spin-transfer torque in spin-valves. However, it is a simple model that has its limitations for higher current densities (i.e. if the modulus of $\vec{s}$ becomes large) and it does not describe each interface in the multilayer structure on an equal footing. Hence, it is the aim of the second part of this thesis to find a more sophisticated macroscopic model for describing the dynamics of the spin accumulation $\vec{s}$.

***Part II*:** In Chapter 3 we give an introduction to the Wigner-Weyl calculus for matrix valued operator symbols in the phase space [6]. In the semi-classical scaling the Weyl correspondence between the state operator ρ in an electron's Hilbert space and its symbol w in the phase space is given component-wise by

$$[\rho\phi]_i(x) = \varepsilon^d \sum_{j=1}^{n} \int \mathrm{d}y \int \mathrm{d}\xi \, w_{ij}\left(\frac{x+y}{2}, \varepsilon\xi\right) \phi_j(y) e^{i(x-y)\cdot\xi} \,,$$

which defines the mapping $\mathrm{Op} : w \mapsto \rho$ and its inverse $\mathrm{Op}^{-1}(\rho) = w$. Here, ϕ is a vector-valued one-electron wavefunction with n components, ε denotes the semi-classical parameter and ξ stands for the scaled momentum variable. Denoting the Moyal product [7] of a Hermitian operator A in the electron Hilbert space with the state operator ρ by $\mathrm{Op}_\varepsilon^{-1}(A\rho) = a \#_\varepsilon w^{(\varepsilon)}$, an expansion in orders of ε (formally) results in

$$a \#_\varepsilon w^{(\varepsilon)} = a w^{(\varepsilon)} + i\frac{\varepsilon}{2}\{a, w^{(\varepsilon)}\}_{(x,\xi)} + \mathcal{O}(\varepsilon^2) \,,$$

where $\{a, w^{(\varepsilon)}\}_{(x,\xi)}$ stands for the Poisson bracket in the phase space $\mathbb{R}^d_x \times \mathbb{R}^d_\xi$ and a is the symbol corresponding to A. The Moyal product enables us to perform the semi-classical limit in the von Neumann equation for a single electron with spin. Under the premise that the matrix-valued spin part h_s of the symbol h of the Hamilton operator scales with order ε, i.e.

$$h(x,\xi) = h_c(x,\xi)\mathbb{1}_n + \varepsilon h_s(x,\xi) \qquad \forall (x,\xi) \in \mathbb{R}^d_x \times \mathbb{R}^d_\xi \,,$$

where h_c denotes the scalar charge part of h, in the semi-classical limit $\varepsilon \to 0$ one obtains the matrix-Vlasov equation [8]. In Chapter 4 the latter appears as the left-hand-side of the proposed matrix-Boltzmann equation, which is a kinetic equation

for spin-coherent electron transport:

$$\partial_t F^{(\alpha)} + \frac{1}{\alpha}\left(\nabla_\xi h_c \cdot \nabla_x F^{(\alpha)} - \nabla_x h_c \cdot \nabla_\xi F^{(\alpha)}\right) + i[F^{(\alpha)}, \vec{h}\cdot\vec{\sigma}] = \frac{1}{\alpha^2} Q_{gl}(F^{(\alpha)}) + Q_{sf}(F^{(\alpha)})\,.$$

Here, $F^{(\alpha)} = w^{(0)}$ stands for the Hermitian 2×2 distribution matrix describing the spin-coherent electron(s) and we set $h_s = \vec{h}\cdot\vec{\sigma}$ for the spin part of h, where $\vec{\sigma}$ denotes the vector of the Pauli matrices. Moreover, we introduced the scaled mean scattering time $\alpha^2 \ll 1$ and the two collision operators Q_{gl} and Q_{sf}. The latter describes scattering events that flip the electron's spin and, thus, relax the matrix $F^{(\alpha)}$ towards a scalar distribution function (σ_0 stands for the 2×2 unit matrix),

$$Q_{sf}(F^{(\alpha)}) := \frac{1}{2}\mathrm{tr}\left(F^{(\alpha)}\right)\sigma_0 - F^{(\alpha)}\,.$$

The operator Q_{gl} accounts for spin-polarized momentum scattering; it is the true novelty in the present kinetic model. The subscripts g and l stand for "gain" and "loss", respectively,

$$Q_{gl}(F^{(\alpha)}) := Q_g^+(F^{(\alpha)}) - Q_l^-(F^{(\alpha)}) \qquad g, l \in \{1, 2\}\,,$$

where we propose two different structures for the gain and the loss term, respectively:

$$\begin{aligned}
Q_1^+(F^{(\alpha)}) &:= \int_{\mathbb{R}^d_{\xi'}} \left(\frac{1}{2}S' F^{(\alpha)}(x,\xi') + \frac{1}{2}F^{(\alpha)}(x,\xi')S'\right)\mathrm{d}\xi'\,,\\
Q_2^+(F^{(\alpha)}) &:= \int_{\mathbb{R}^d_{\xi'}} S'^{1/2} F^{(\alpha)}(x,\xi') S'^{1/2}\mathrm{d}\xi'\,,\\
Q_1^-(F^{(\alpha)}) &:= \frac{1}{2}\Lambda F^{(\alpha)}(x,\xi) + \frac{1}{2}F^{(\alpha)}(x,\xi)\Lambda\,,\\
Q_2^-(F^{(\alpha)}) &:= \int_{\mathbb{R}^d_{\xi'}} S^{1/2} F^{(\alpha)}(x,\xi) S^{1/2}\mathrm{d}\xi'\,.
\end{aligned}$$

Here, $\Lambda = \int S\mathrm{d}\xi'$ and the eigenvalues of the Hermitian 2×2 scattering matrix S account for the two different scattering probabilities (transition rates) of spin-up and spin-down electrons, thus modeling spin-polarized electron resistances. The matrix-Boltzmann equation is investigated regarding the existence, uniqueness and positivity of a solution $F^{(\alpha)}$ on the domain $[0,T] \times \mathbb{R}^d_x \times \mathbb{R}^d_\xi$. Furthermore, the Diffusion limit $\alpha \to 0$ is performed in case that $Q_{gl} = Q_{22}$ and for a scattering

matrix of the form

$$S(t,x,\xi,\xi') = s(t,x,\xi,\xi')P(t,x)\,, \qquad P(t,x) = \sigma_0 + p(t,x)\vec{\Omega}(t,x)\cdot\vec{\sigma}\,,$$

where $s \in \mathbb{R}^+$ denotes the mean scattering rate[1] from ξ to ξ' at (t,x) and P denotes the *polarization matrix* which does not depend on ξ and ξ'. The spin-polarization of the scattering rates is reflected by the polarization parameter $0 \leq p(t,x) < 1$. In the limit $\alpha \to 0$ we obtain a new matrix-drift-diffusion model for spin-polarized, spin-coherent electron transport. This model features a novel coupling between the charge- and the spin-degree of freedom of a single electron. Numerical simulations show that this coupling could indeed be exploited in future spintronics applications. The new model represents an improvement over the heuristic ZLF model treated in Part *I*, since it inherently describes the creation of non-equilibrium spin accumulation at each interface of a multilayer structure on an equal footing (this process is described in detail in the introductive Chapter 1).

***Part III*:** This part is concerned with the theory of open quantum systems [9]. It is the aim to derive, from first principles, the spin-coherent collision operators Q_{gl} proposed in Part *II*. For this, we consider a composite quantum system with the Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$, where A and B denote the two subsystems. The most general Hamiltonian of this system reads

$$H = H_A \otimes \mathbb{1}_B + \mathbb{1}_A \otimes H_B + H_I\,,$$

where the operator subscript $A(B)$ indicates an operator acting in $\mathcal{H}_A$ ($\mathcal{H}_B$), $\mathbb{1}_A$ ($\mathbb{1}_B$) denotes the respective identity and the operator H_I accounts for the interactions between A and B. The approach is to trace over the degrees of freedom B in the dynamics of the composite system AB. The interaction Hamiltonian H_I should then lead to dissipative dynamics of the reduced state operator $\rho_A := \mathrm{tr}_B(\rho)$. Here, ρ denotes the state operator of the composite system and $\mathrm{tr}_B(\cdot)$ stands for the trace over the degrees of freedom B. We start from a von Neumann equation that has been augmented by a relaxation operator,

$$\partial_t\rho = -i[H,\rho] + \frac{1}{\tau_r}\left(\mathrm{tr}_B(\rho)\otimes\chi_B - \rho\right)\,,$$

where χ_B is a predefined, time-independent state operator in $\mathcal{H}_B$, hence $\mathrm{tr}_B(\chi_B) = 1$. The dissipative term added to the von Neumann equation relaxes the state operator ρ of the composite system towards the Born approximation $\mathrm{tr}_B(\rho)\otimes\chi_B$ on a timescale τ_r. Our main result is that under the premise that this relaxation

[1] The mean value is taken with respect to spin-up and spin-down scattering rates.

is the dominant process and that, furthermore, the interaction between A and B is strong, in the limit $\tau_r \to 0$ one obtains the Lindblad master equation [10],

$$\partial_t \rho_A^{(0)} = -i[\widetilde{H}_A, \rho_A^{(0)}] - \mathrm{tr}_B \left([\widetilde{H}_I, [\widetilde{H}_I, \rho_A^{(0)} \otimes \chi_B]] \right) .$$

Here, we introduced the mean-field-corrected operators

$$\begin{aligned} \widetilde{H}_I &:= H_I - H_A^{mf} \otimes \mathbb{1}_B \,, \\ \widetilde{H}_A &:= H_A + H_A^{mf} \,, \end{aligned}$$

where the mean-field operator H_A^{mf} in $\mathcal{H}_A$ is defined by

$$H_A^{mf} := \mathrm{tr}_B \left(H_I \chi_B \right) .$$

The Lindblad equation is a Markovian (local in time) master equation. The second term on its right-hand side has the same matrix product structure as the collision operator Q_{21}. However, the Lindblad equation is still a microscopic equation. It is thus the purpose of a forthcoming work to regard its semi-classical limit. For this one could apply the scaled Wigner-Weyl calculus presented in Chapter 3. In this work, we derive non-Markovian corrections to the Lindblad master equation in the case that τ_r is small bot not zero. Denoting the scaled relaxation time by $\alpha \ll 1$, the state operator in subsystem A is written as $\rho_A = \rho_A^{(0)} + \alpha \rho_A^{(1)}$, where first-order correction satisfies

$$\partial_t \rho_A^{(1)} = -i[\widetilde{H}_A, \rho_A^{(1)}] - \mathrm{tr}_B \left([\widetilde{H}_I, [\widetilde{H}_I, \rho_A^{(1)} \otimes \chi_B]] \right) + \mathcal{S}_1 \,,$$

which features the local source term

$$\begin{aligned} \mathcal{S}_1 = &- \mathrm{tr}_B \left(i[\widetilde{H}_I, \rho_A^{(0)} \otimes [H_B, \chi_B] \right) \\ &+ \mathrm{tr}_B \left(i[\widetilde{H}_I, [\widetilde{H}_I, [\widetilde{H}_I, \rho_A^{(0)} \otimes \chi_B]]] \right) . \end{aligned}$$

Higher-order corrections are derived as-well, which permit the systematic treatment of non-Markovian quantum dynamics in a perturbative manner.

***Part IV*:** The last part of this thesis contains the numerical treatment of a quantum-diffusive spin model, which was developed in [11] on the basis of the maximum entropy principle [12, 13]. This model describes the time evolution of the spin densities n_1 and n_2 of a two-dimensional electron gas subjected to Rashba spin-orbit coupling in the case that the state operator ρ is a quantum Maxwellian,

$\rho = e^{-H}$, at all times. The basic equations to be solved for $n_1(t,x)$ and $n_2(t,x)$ read

$$\begin{aligned}
&\partial_t n_1 + \nabla\cdot(n_1\nabla(A_1 - V_s))+ \\
&\quad + \alpha(A_1 - A_2)\,\mathcal{R}\mathrm{e}(\mathcal{D}n_{21}) - 2\alpha\,\mathcal{R}\mathrm{e}(n_{21}\mathcal{D}(A_2 - V_s))- \\
&\quad - \frac{2\alpha}{\varepsilon}(A_1 - A_2)\,\mathcal{I}\mathrm{m}(J^x_{21} - iJ^y_{21}) = 0\,, \\
&\partial_t n_2 + \nabla\cdot(n_2\nabla(A_2 - V_s))+ \\
&\quad + \alpha(A_1 - A_2)\,\mathcal{R}\mathrm{e}(\mathcal{D}n_{21}) + 2\alpha\,\mathcal{R}\mathrm{e}(n_{21}\mathcal{D}(A_1 - V_s))+ \\
&\quad + \frac{2\alpha}{\varepsilon}(A_1 - A_2)\,\mathcal{I}\mathrm{m}(J^x_{21} - iJ^y_{21}) = 0\,,
\end{aligned}$$

where V_s denotes the self-consistent potential due to the electron-electron interaction, computet from the Poisson equation,

$$-\gamma^2\Delta V_s = n_1 + n_2\,,$$

and α, ε and γ are scaled natural constants. In the present model, a non-local coupling between n_1 and n_2 arises from the chemical potential $A = (A_1, A_2)$, which consists of the Lagrange multipliers A_1 and A_2 used in the entropy maximization. Hence the closure of the above system is ensured by the equilibrium expression for the density matrix ϱ and the current matrix J,

$$\begin{aligned}
\varrho &= \sum_l e^{-\lambda_l}\begin{pmatrix} |\psi^1_l|^2 & \psi^1_l\,\overline{\psi^2_l} \\ \psi^2_l\,\overline{\psi^1_l} & |\psi^2_l|^2 \end{pmatrix} = \begin{pmatrix} n_1 & \overline{n}_{21} \\ n_{21} & n_2 \end{pmatrix}, \\
J &= -\frac{i\varepsilon}{2}\sum_l e^{-\lambda_l}\begin{pmatrix} \overline{\psi^1_l}\nabla\psi^1_l - \psi^1_l\nabla\overline{\psi^1_l} & \overline{\psi^2_l}\nabla\psi^1_l - \psi^1_l\nabla\overline{\psi^2_l} \\ \overline{\psi^1_l}\nabla\psi^2_l - \psi^2_l\nabla\overline{\psi^1_l} & \overline{\psi^2_l}\nabla\psi^2_l - \psi^2_l\nabla\overline{\psi^2_l} \end{pmatrix} = \begin{pmatrix} J_1 & \overline{J}_{21} \\ J_{21} & J_2 \end{pmatrix}.
\end{aligned}$$

The wavefunctions $\psi^1_l, \psi^2_l \in H^2$ and the numbers $\lambda_l \in \mathbb{R}$ are the components of eigenfunctions $\psi_l \in (H^2)^2$ and the eigenvalues, respectively, of the system Hamiltonian $H(A)$ in $(H^2)^2$,

$$H(A)\psi_l(A) = \lambda_l(A)\psi_l(A)\,,$$

given by

$$H(A) = \begin{pmatrix} -\frac{\varepsilon^2}{2}\,\Delta + V_{ext,1} + A_1 & \varepsilon^2\alpha(\partial_x - i\partial_y) \\ -\varepsilon^2\alpha(\partial_x + i\partial_y) & -\frac{\varepsilon^2}{2}\,\Delta + V_{ext,2} + A_2 \end{pmatrix}.$$

The numerical solution of the present quantum-diffusive model is accomplished by means of a functional formalism similar to the one which is well-established in the spin-less case [14]. We investigate two different approaches, one that advances in time the chemical potentials and another that advances the spin densities. The

former leads to a more implicit scheme and, thus, to a better numerical stability. The used functionals depend on A_1 and A_2 (as well as on V_s in one case). They are shown to be convex and to have a unique minimum, solution of the above system. The developed finite difference algorithm is then applied to simulate the steady-state spin distribution in a typical transistor geometry.

References

[1] I. Zutic, J. Fabian, and Das Sarma S. Spintronics: Fundamentals and applications. *Rev. Mod. Phys.*, 76(2):323–410, 2004.

[2] G.A. Prinz. Magnetoelectronics applications. *J. Magn. Magn. Mater.*, 200:57–68, 1999.

[3] Chappert C., A. Fert, and F.N. Van Dau. The emergence of spin electronics in data storage. *Nature Mat.*, 6:813–823, 2007.

[4] S. Zhang, P.M. Levy, and A. Fert. Mechanisms of spin-polarized current-driven magnetization switching. *Phys. Rev. Lett.*, 88(23):236601–1, 2002.

[5] I.N. Krivorotov, N.C. Emley, J.C. Sankey, S.I. Kiselev, D.C. Ralph, and R.A. Buhrman. Time-domain measurements of nanomagnet dynamics driven by spin-transfer torques. *Science*, 307:228, 2005.

[6] L. Hörmander. The analysis of linear partial differential operators, vol. 3. pseudo-differential operators, volume 274 of grundlehren der mathematischen wissenschaften, 1985.

[7] C. Zachos, D. Fairlie, and T. Curtright. *Quantum mechanics in phase space: an overview with selected papers*, volume 34. World Scientific Pub Co Inc, 2005.

[8] R. El Hajj. *Etude mathématique et numérique de modèles de transport: application à la spintronique*. PhD thesis, Institut de Mathématiques de Toulouse (IMT), Université Paul Sabatier, 2008.

[9] H.P. Breuer and F. Petruccione. *The theory of open quantum systems*, volume 28. Oxford University Press Oxford, 2002.

[10] G. Lindblad. On the generators of quantum dynamical semigroups. *Communications in Mathematical Physics*, 48(2):119–130, 1976.

[11] L. Barletti and F. Méhats. Quantum drift-diffusion modeling of spin transport in nanostructures. *Journal of Mathematical Physics*, 51:053304, 2010.

[12] P. Degond and C. Ringhofer. Quantum moment hydrodynamics and the entropy principle. *Journal of statistical physics*, 112(3):587–628, 2003.

[13] P. Degond, F. Méhats, and C. Ringhofer. Quantum energy-transport and drift-diffusion models. *Journal of statistical physics*, 118(3):625–667, 2005.

[14] S. Gallego and F. Méhats. Entropic discretization of a quantum drift-diffusion model. *SIAM journal on numerical analysis*, pages 1828–1849, 2006.

Part I

Spin-Transfer Torques in Magnetic Multilayers

Chapter 1

Spin-induced Phenomena in Metallic Magnetic Multilayers under Current

---

Overview. We summarize the basic physical concepts of spin-polarized electron conduction in the transition-metal ferromagnets Fe, Ni and Co, of the giant magnetoresistance (GMR) effect and of the non-equilibrium spin accumulation at a nonmagnetic/ferromagnetic interface that is traversed by an electronic current. These very basics of the emerging field of spintronics will reappear throughout this thesis. Furthermore, we review the dissipative character of the Landau-Lifshitz equation and estimate the length scales occurring in the spin-diffusion equation in the transition metal ferromagnets. As a next step the notion of spin-coherence is discussed in detail on the quantum mechanical level. It is shown that only spin-coherent states can contribute to the spin-transfer torque in metallic magnetic multilayers. In the last chapter we present numerical, self-consistent solutions of the Landau-Lifshitz equation and the spin diffusion equation in multilayer structures. The simulated power spectra of the magnetic precessional states in thin ferromagnetic layers of spin-valves are compared to recent experiments and a satisfactory agreement is found.

1.1 Basic physical concepts

In 1935 Sir N.F. Mott predicted on the basis of quantum mechanical calculations that the conductivity in the transition metal ferromagnets Fe, Ni and Co is different for electrons whose magnetic moment is aligned with the magnetization (majority electrons)[1] and for electrons with anti-parallel magnetic moments (minority electrons) [1]. Experimental evidence of this *spin-polarized* electron conduction in the transition metal ferromagnets and in alloys thereof was found roughly 30 years later [2, 3]. As argued by Mott, the reason for this effect lies in the shifted band structure of majority and minority electrons, respectively, depicted in Fig. 1.1i for Co. The shift, or the spin-splitting, arises from the exchange interaction which favors parallel alignment of spins in partially filled atomic shells; an effect that may carry over to the band structure of crystals. In Fe, Ni and Co the narrow $3d$ band features a large exchange splitting and moreover crosses the Fermi energy only for the minority electrons [4, 5, 6]; it is thus responsible for the ferromagnetic character. On the other hand, the transport in these materials is dominated by the delocalized electrons in the much wider $4s$ band, which has a comparably small exchange splitting and crosses the Fermi energy for both spin species. Nevertheless, the conductivity is spin-polarized, because the main transition induced by scattering is $4s$ to $3d$. Since for minority electrons the density of states (DOS) of $3d$ electrons near the Fermi energy is much larger than for majority electrons, such a transition is much more likely for minority electrons (Fermi's Golden rule). Moreover, in the case that spin-flip scattering is negligible, the electron conduction can be viewed as a process that happens in two parallel channels with different resistances, one for majority and one for minority electrons, respectively [7].

It was not until the late 80's that the two-current conduction in transition metal ferromagnets would lead to a major advancement in the storage technology, when A. Fert and P. Grünberg independently discovered the giant magneto-resistance (GMR) effect in multilayer structures [8, 9]. Their underlying idea is sketched in Fig. 1.1ii. A non-magnetic spacer layer is sandwiched between two ferromagnetic layers with resistances $R_\uparrow$ for majority electrons and $R_\downarrow$ for minority electrons, respectively, and $R_\uparrow < R_\downarrow$. The magnetization of one of the ferromagnetic layers can be switched by means of an applied magnetic field. According to the two-current model, the total resistance in this structure is different for the parallel (P) and anti-parallel (AP)

[1] We remark that the spin of majority electrons is anti-parallel to the magnetization, since it is of opposite sign than the electron's magnetic moment.

configuration of the magnetizations:

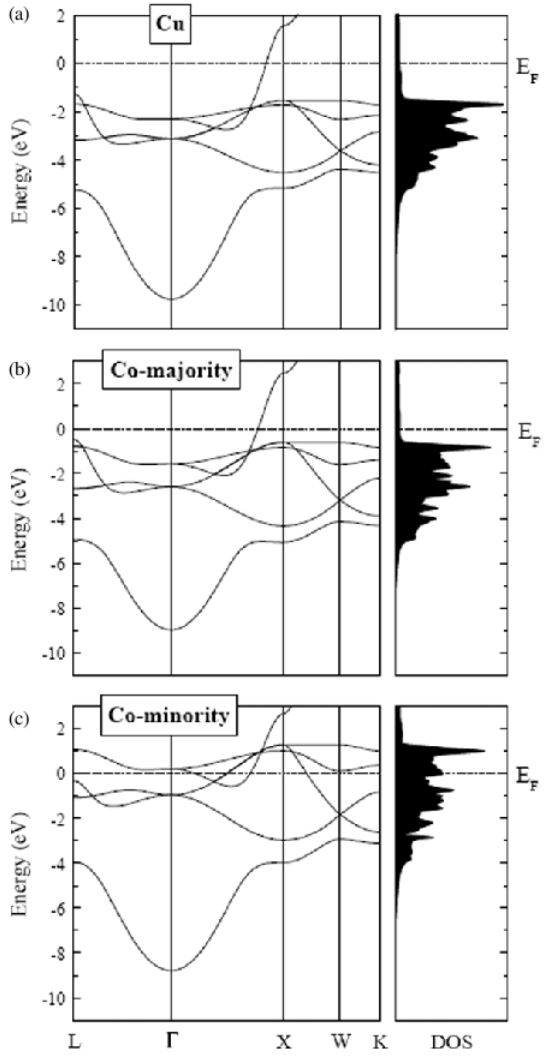
$$R_P = \frac{2R_\uparrow R_\downarrow}{R_\uparrow + R\downarrow}\,, \qquad R_{AP} = \frac{R_\uparrow + R_\downarrow}{2}\,. \tag{1.1.1}$$

Hence for the parallel alignment, majority electrons have a low resistance path through the multilayer, whereas in the AP configuration the role of the electrons changes in each layer and such a path is absent. From Eq. (1.1.1) we obtain

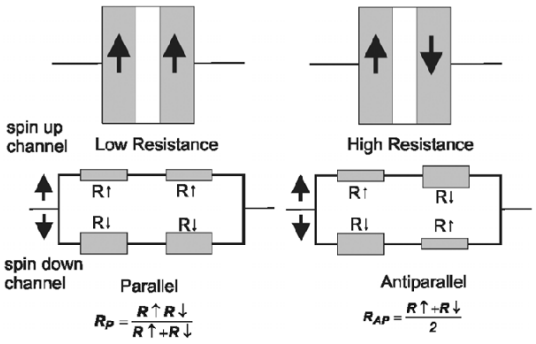
$$\frac{\Delta R}{R_P} = \frac{R_{AP} - R_P}{R_P} = \frac{(R_\uparrow - R_\downarrow)^2}{4R_\uparrow R_\downarrow} > 0 \tag{1.1.2}$$

for the relative change in resistance. The value of this relative change is usually around 5% for a trilayer, but can be enhanced up to around 100% or higher for multilayers consisting of a high number of periods, hence the name GMR. It is the basic concept for a memory unit ("bit") in state-of-the-art hard disk drives, which is why its discoverers were granted the Nobel prize for Physics in 2007 [10, 11]. The spin-transfer torques (STTs) considered here are an approach towards a further development of this technology with the aim to eliminate the magnetic field from the switching process, which may lead to a considerable gain in areal memory density because of the redundancy of the magnetic write head [12, 13].

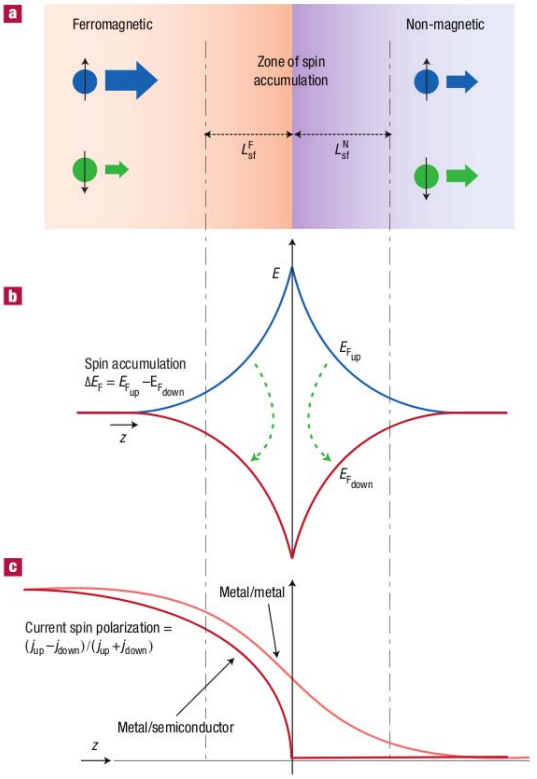
In order to understand the mechanism of STTs we need to elaborate on the effect of *spin accumulation* at the interface between a ferromagnet (F) and a non-magnet (NM) when a current flows through this interface. The situation is sketched in Fig. 1.1iii. Due to the spin-polarization of the resistivity in the F-layer the current densities j_{up} and j_{down} of majority and minority electrons are very different in the F-layer far away from the interface. In the NM layer, however, these current densities are equal away from the interface. Hence, there is an interfacial region where the spin-polarization of the current is gradually destroyed. We point out that this polarization arises solely from the spin-polarized scattering probabilities in the F-layer and not (!) from a difference in spin densities, since the $4s$ band responsible for the current has a minor spin splitting. Thus, in order to accommodate the spin polarization of the current in the NM-layer in the vicinity of the interface, the density of spin-up electrons must be augmented. This is pictured by a difference in the respective Fermi energy of spin-up and spin-down electrons, respectively. The length over which such a non-equilibrium polarization of the spin densities, called spin accumulation, is present around the interface is denoted as the spin diffusion length L_{sf}. It may be far larger than the usual mean free path, reaching more than 200 nm in some materials. We remark that the matrix-drift-diffusion model derived in Chapter 4 is well-suited for describing spin accumulation in non-collinear



(i) The electronic band structure and density of states (DOS) of a) non-magnetic Cu, b) the majority electrons in fcc Co, (c) the minority electrons in fcc Co. The large exchange splitting between Co-majority electrons and Co-minority electrons results in a considerable shift of the respective density of states, which has two main consequences: ferromagnetism and enhanced scattering probability for minority electrons near the Fermi level. (Picture reprinted from [11].)



(ii) Trilayer consisting of a non-magnetic spacer layer (white) sandwiched between two ferromagnetic layers with spin-polarized resistances $R_\uparrow$ and $R_\downarrow$ for the respective electron species. A magnetic field is used to switch between parallel (P) and anti-parallel (AP) alignment of the magnetizations. The P configuration offers a low resistance path for majority electrons. (Picture reprinted from [11].



(iii) a) The spin-polarization of the current density is destroyed in the zone of spin accumulation, whose width is determined by the spin diffusion lengths L_{sf}^{F} and L_{sf}^{NM}. b) Difference in Fermi energies for spin-up and spin-down electrons in the vicinity of the interface. c) Current polarization across the interface. (Picture reprinted from [5].)

Figure 1.1: Important properties of spin-polarized electron transport in transition metal ferromagnets: (i) band structures, (ii) the GMR-effect and (iii) spin accumulation at ferromagnetic/non-magnetic interfaces.

magnetic multilayers, see for instance the Figs. 4.1, 4.2 and 4.3.

We shall now introduce the concepts of the STT effect in spin-valves. A spin-valve basically consists of a thick F-layer (the "polarizing" or fixed layer), a thin non-magnetic metal spacer and a thin F-layer (free layer). Both F-layers are contacted by electrodes in order to pass a current through the structure which induces magnetization dynamics in the free magnetic layer. Two common experimental realizations of spin-valves, namely the point-contact geometry and the nanopillar geometry, are depicted in Fig. 1.2i. The magnetization dynamics in the free layer originate from a direct transfer of angular momentum from the spin-polarized current to the magnetic moments of the $3d$ electrons in the thin layer [14, 15]. In particular, electrons that enter the polarizing layer will quickly align their magnetic moments with the magnetization of that layer and subsequently travel to the free layer. The non-magnetic spacer layer is much thinner than the corresponding spin-diffusion length. Hence, a considerable spin accumulation in the direction of the polarizing layer will be present at the interface to the thin magnetic layer. In case that the magnetizations in the two magnetic layers are neither parallel nor anti-parallel, the perpendicular component of the angular momentum of the spin accumulation will exert a torque on the magnetic moments in the free layer, resulting in magnetic precession or even the switching of the magnetic direction. The mechanism along with the usual device dimensions and the applied fields is sketched in Fig. 1.2ii. Once the perpendicular component of the spin-angular momentum has decayed in the thin layer, the conservation of angular momentum implies a total transfer of this non-equilibrium spin to the thin layer, which is the origin of the magnetization dynamics. The corresponding length scale of the decay is 1-4 nm, hence the STT can be considered as an interface effect [16, 17, 4].

What makes the STT effect interesting is the possibility to switch between the P state (logical '1') and the AP (logical '0') state in a memory unit solely by passing a current through it, thereby eliminating the necessity of a magnetic write head. There have been many successful experiments confirming magnetic switching and magnetic precession in spin-valves under current [18, 19, 20, 21, 22, 23, 24]. However, the needed current densities are still too high and the device dimensions too large for industrial application. We shall briefly comment on the results of the insightful experiments conducted by Krivorotov et al. [22], who were the first to perform time-resolved measurements of the STT-induced magnetization dynamics. They used the nanopillar geometry with permalloy $Ni_{80}Fe_{20}$ as ferromagnets and Cu as the spacer layer. Figure 1.2iii shows the voltage drop over the device as a function of time for different magnitudes of applied currents I_p. This voltage drop is due to the GMR resistance, which varies as a function of the angle between the two magnetization

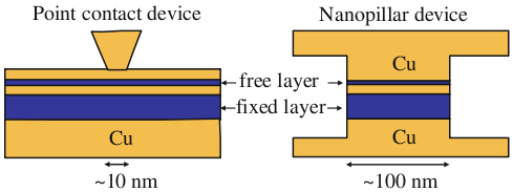
directions in the trilayer and has a maximum modulus for the AP configuration. In Fig. 1.2iii A)-C) it is transparent that the device switches from the AP alignment to the P alignment within a switching time τ_S that depends on I_p. Moreover, the switching is accompanied by oscillatory signals, which confirm a magnetic precession during the process. Figure 1.2iii D) shows numerical simulations of the Landau-Lifshitz equation (see the next section) which seem to fit quite well the experimental observations. The STT effect also plays a major role in the understanding of the current-induced domain wall motion, which is a growing topic of ongoing research [25, 26, 27, 28]. The matrix-drift-diffusion model developed in Chapter 4, as well as the algorithm presented in Chapter 2, enable a new route to investigate these phenomena from a numerical point of view.

1.2 A diffusive exchange-torque model for current-induced magnetization dynamics

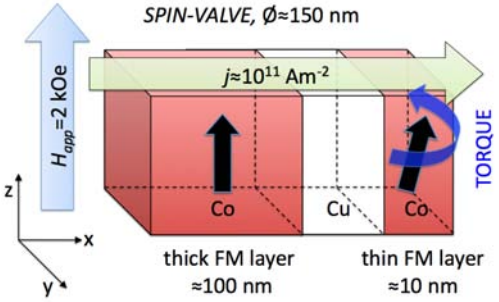
In general, the treatment of dynamical processes in transition metal ferromagnets is quite involved due to the *sp*-hybridization of the transport band and due to the itinerant nature of the semi-localized electrons in the $3d$ band. When a voltage is applied, a first simplifying assumption is often that the electric current originates solely from the $4s$ electrons, while the magnetization is constituted by the local moments of the $3d$ electrons. This approach is called the *sd*-model. In what follows we shall review a version of the *sd*-model proposed by Zhang et al. [29] which describes the spin dynamics in ferromagnets under current at the macroscopic (diffusive) level. Let $\Omega \subset \mathbb{R}^d_x$ stand for the regular domain representing the (multilayer) device, where d is the dimension of the position space, and let S^2 denote the unit sphere in $\mathbb{R}^3$. We search for the unknowns $\vec{m} : \Omega \times \mathbb{R}^+ \to S^2$, which is the local magnetization due to the $3d$ electrons, and $\vec{s} : \Omega \times \mathbb{R}^+ \to \mathbb{R}^3$, which is the spin accumulation due to the $4s$ electrons. These quantities are determined from the following system of equations:

$$\begin{cases} \partial_t \vec{m} = -\gamma_0 \vec{m} \times \vec{H}_{eff} + \alpha \vec{m} \times \partial_t \vec{m} \,, \\ \vec{m}(x,0) = \vec{m}_0(x) \,, \\ \\ \partial_t \vec{s} = D\Delta \vec{s} - \dfrac{\vec{s}}{\tau} - \dfrac{J_{ex}}{\hbar} \vec{s} \times \vec{m} \,, \\ \vec{s}(x,0) = \vec{s}_0(x) \,. \end{cases} \tag{1.2.1}$$

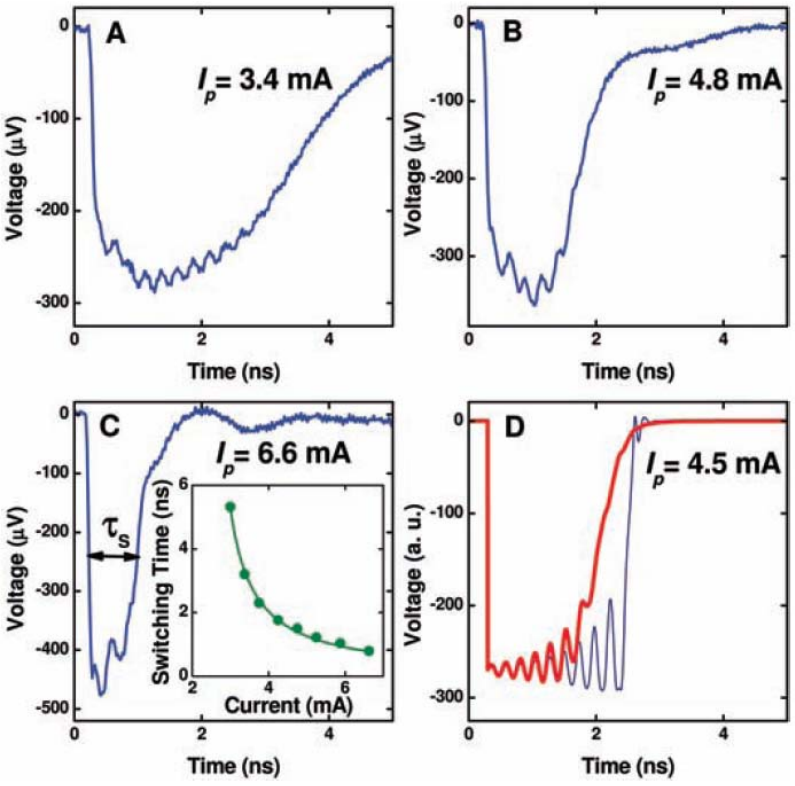
Here, $\gamma_0 = \gamma \mu_0$, where γ is the gyromagnetic ratio and μ_0 stands for the magnetic constant, $\alpha > 0$ denotes the damping parameter, D stands for the spin diffusion constant, τ is the spin-flip time, J_{ex} denotes the exchange constant, $\hbar$ is the Planck



(i) A thick magnetic layer (fixed layer) and a thin magnetic layer (free layer) are spaced by a non-magnetic metallic layer, which is much thinner than the spin-diffusion length. The area over which the current is injected differs considerably in the point-contact and in the nanopillar setup.



(ii) Common values for the current density j, the applied magnetic field H_{app} and the device dimensions in STT experiments in spin-valves of the nanopillar geometry. The direct transfer of angular momentum from the spin accumulation induced by the polarizing layer results in precession or even switching of the magnetization in the thin magnetic layer.



(iii) Time-resolved magnetic precession and switching in nanopillar spin-valves for different applied current densities I_p. The voltage drop over the device depends on the angle between the magnetizations of the two F-layers (GMR resistance), which allows to infer to the magnetic dynamics. A)-C) show experimental results while D) shows a signal simulated with a Landau-Lifshitz equation-based STT model. (Data reprinted from [22].)

Figure 1.2: Current-induced magnetization dynamics in spin-valves: (i) experimental device geometries, (ii) basic principle of the STT-effect and magnitudes of the applied fields, (iii) experimental evidence of magnetic switching.

constant divided by 2π and $\vec{m}_0$ and $\vec{s}_0$ denote the initial conditions. Furthermore, the effective field $\vec{H}_{eff} : \Omega \times \mathbb{R}^+ \to \mathbb{R}^3$ is given by

$$\vec{H}_{eff} := -\frac{1}{\mu_0 M_s} \delta\mathcal{E}(\vec{m})(\delta\vec{m}) \,, \tag{1.2.2}$$

where M_s stands for the saturation magnetization, $\mathcal{E}(\vec{m})$ denotes the energy functional corresponding to the magnetization and we define the functional derivative as

$$\delta\mathcal{E}(\vec{m})(\delta\vec{m}) = \lim_{\varepsilon \to 0} \frac{\mathcal{E}(\vec{m} + \varepsilon\delta\vec{m}) - \mathcal{E}(\vec{m})}{\varepsilon} \tag{1.2.3}$$

for some variation $\delta\vec{m}(x,t)$ which vanishes at the boundary $\partial\Omega$ of the domain Ω. Usually, the energy functional $\mathcal{E}$ consists of five contributions [17],

$$\mathcal{E} = \mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3 + \mathcal{E}_4 + \mathcal{E}_5 \,, \tag{1.2.4}$$

given by the energy due to an externally applied magnetic field,

$$\mathcal{E}_1 = -\mu_0 M_s \int_\Omega \vec{m} \cdot \vec{H}_{app} \, dx \,, \tag{1.2.5}$$

the magnetocrystalline and shape anisotropy energies,

$$\mathcal{E}_2 = -K_u \int_\Omega (\vec{u} \cdot \vec{m})^2 \, dx + \frac{\mu_0 M_s^2}{2} \int_\Omega (\vec{w} \cdot \vec{m})^2 \, dx \,, \tag{1.2.6}$$

where K_u is the anisotropy constant and $\vec{u}, \vec{w} \in S^2$ denote the magnetic easy axis and the normal to the magnetic easy plane[2], respectively; furthermore, the micromagnetic exchange energy,

$$\mathcal{E}_3 = A_{ex} \int_\Omega \sum_{i=1}^{d} (\partial_{x_i}\vec{m} \cdot \partial_{x_i}\vec{m}) \, dx \,, \tag{1.2.7}$$

where A_{ex} stands for the exchange constant; the magnetostatic dipole-dipole interaction energy (only for $d = 3$),

$$\mathcal{E}_4 = -\frac{\mu_0 M_s^2}{8\pi} \int_\Omega \int_\Omega \vec{m}(x) \cdot \frac{3(\vec{m}(x') \cdot x)x - \vec{m}(x')|x|^2}{|x|^5} \, dx' \, dx \,, \tag{1.2.8}$$

[2]The magnetic easy axis is a preferred axis for the magnetization arising from crystalline anisotropy; the magnetic easy plane is a preferred plane for the magnetization arising from the shape anisotropy, e.g. if the device has the shape of a thin disc.

and finally the exchange energy due to the interaction with the spin accumulation,

$$\mathcal{E}_5 = -J_{ex} \int_\Omega \vec{m} \cdot \vec{s}\, dx\,, \tag{1.2.9}$$

where J_{ex} stands for the characteristic exchange energy between $4s$ and $3d$ electrons. It is now straightforward to compute the effective field from (1.2.2),

$$\begin{aligned} \vec{H}_{eff} =& \vec{H}_{app} + \frac{2K_u}{\mu_0 M_s}(\vec{u}\cdot\vec{m})\vec{u} - M_s(\vec{w}\cdot\vec{m})\vec{w} + \frac{2A_{ex}}{\mu_0 M_s}\Delta\vec{m} \\ &+ \frac{M_s}{4\pi}\int_\Omega \frac{3(\vec{m}(x')\cdot x)x - \vec{m}(x')|x|^2}{|x|^5}\, dx' + \frac{J_{ex}}{\mu_0 M_s}\vec{s}. \end{aligned} \tag{1.2.10}$$

The values of the natural constants and the material parameters appearing in Eqs. (1.2.1) and (1.2.10) are given in Table 1.1 and Table 1.2, respectively. In (1.2.1) the first equation is the Landau-Lifshitz-Gilbert equation (LLGE) for $\vec{m}$ and the second equation is the spin-diffusion equation (SDE) for $\vec{s}$. These equations are coupled through the exchange terms whose strength is determined by J_{ex} (appearing in (1.2.10) in the LLGE). In Chapter 2 we develop a numerical algorithm that enables the self-consistent solution of Eqs. (1.2.1) in one space dimension ($d = 1$, thus neglecting the dipole-dipole interaction (1.2.8)). The strong non-linearity of the LLGE as well as the different timescales of the dynamics for $\vec{m}$ and $\vec{s}$ make the numerical treatment a challenging task.

The Landau-Lifshitz equation derived from the LLGE and damping

We shall show that the LLGE in (1.2.1) is dissipative due to the second (damping) term on the right-hand-side. Employing the short notation $\delta\mathcal{E} = \delta\mathcal{E}(\vec{m})(\delta\vec{m})$, we start from

$$\partial_t\vec{m} = \frac{\gamma_0}{\mu_0 M_s}\vec{m}\times\delta\mathcal{E} + \alpha\vec{m}\times\partial_t\vec{m} \tag{1.2.11}$$

and take the vector product with $\vec{m}$ which results in

$$\vec{m}\times\partial_t\vec{m} = \frac{\gamma_0}{\mu_0 M_s}\vec{m}(\vec{m}\cdot\delta\mathcal{E}) - \frac{\gamma_0}{\mu_0 M_s}\delta\mathcal{E}\,|\vec{m}|^2 + \alpha\vec{m}(\vec{m}\cdot\partial_t\vec{m}) - \alpha\partial_t\vec{m}\,|\vec{m}|^2\ . \tag{1.2.12}$$

We now perform the scalar multiplication of (1.2.12) with $\partial_t\vec{m}$ and subsequently we apply the relations $\vec{m}\cdot\partial_t\vec{m} = 0$ and $|\vec{m}|^2 = 1$ to obtain

$$\frac{1}{\mu_0 M_s}\delta\mathcal{E}\cdot\partial_t\vec{m} = -\frac{\alpha}{\gamma_0}\,|\partial_t\vec{m}|^2\ . \tag{1.2.13}$$

From Eqs. (1.2.4) to (1.2.10) we deduce that

$$\partial_t \mathcal{E} = \frac{1}{\mu_0 M_s} \int_\Omega \delta\mathcal{E} \cdot \partial_t \vec{m} \, dx \,. \tag{1.2.14}$$

Hence the integration over Ω in Eq. (1.2.13) makes the dissipation of energy in the course of time,

$$\partial_t \mathcal{E} = -\frac{\alpha}{\gamma_0} \int_\Omega |\partial_t \vec{m}|^2 \, dx \leq 0 \,, \tag{1.2.15}$$

transparent. We shall furthermore derive from (1.2.1) a form of the LLGE that is explicit in the time derivative $\partial_t \vec{m}$. Starting from

$$\partial_t \vec{m} = -\gamma_0 \vec{m} \times \vec{H}_{eff} + \alpha \vec{m} \times \partial_t \vec{m} \,, \tag{1.2.16}$$

and performing the vector product of this equation with $\vec{m}$ leads to

$$\begin{aligned} \vec{m} \times \partial_t \vec{m} &= -\gamma_0 \vec{m} \times (\vec{m} \times \vec{H}_{eff}) + \alpha \vec{m} \times (\vec{m} \times \partial_t \vec{m}) \\ &= -\gamma_0 \vec{m} \times (\vec{m} \times \vec{H}_{eff}) + \alpha \vec{m} (\vec{m} \cdot \partial_t \vec{m}) - \alpha \partial_t \vec{m} \, |\vec{m}|^2 \,. \end{aligned} \tag{1.2.17}$$

Applying the relations $\vec{m} \cdot \partial_t \vec{m} = 0$ and $|\vec{m}|^2 = 1$ and inserting Eq. (1.2.16) into the left-hand-side of (1.2.17) results in

$$\partial_t \vec{m} = -\frac{\gamma_0}{1+\alpha^2} \vec{m} \times \vec{H}_{eff} - \frac{\alpha \gamma_0}{1+\alpha^2} \vec{m} \times (\vec{m} \times \vec{H}_{eff}) \,, \tag{1.2.18}$$

which is known as the Landau-Lifshitz equation (LLE). Both the LLGE and the LLE describe the dynamics of the magnetization on an equal footing. However, it might be desirable from a numerical point of view to discretize Eq. (1.2.18) rather than Eq. (1.2.16) in order to obtain a linear system of equations that is norm-conserving.

Table 1.1: Natural constants.

Symbol	Value	Name
$\hbar$	$1.05457163 \times 10^{-34}$ J·s	reduced Planck constant
e	$1.60217649 \times 10^{-19}$ C	elementary charge
μ_0	$1.25663706 \times 10^{-6}$ N·A^{-2}	magnetic constant
μ_B	$9.27400915 \times 10^{-24}$ J·T^{-1}	Bohr magneton
g_e	-2.00231930	electron g-factor
γ	$1.76085978 \times 10^{11}$ rad·s^{-1}·T^{-1}	electron gyromagnetic ratio

Table 1.2: Typical values of the saturation magnetization, the anisotropy constant, the exchange constant and the damping parameter.

Ref.	Material	M_s [A·m^{-1}]	K_u [J·m^{-3}]	A_{ex} [J·m^{-1}]	α
[4]	$Ni_{80}Fe_{20}$	8.0×10^5	1.45×10^4	1.0×10^{-11}	0.01
[4]	Co/CoFe	1.5×10^6		1.3×10^{-11}	0.006
[28]	$Ni_{80}Fe_{20}$	8.0×10^5		1.0×10^{-11}	0.02
[30]	$Ni_{80}Fe_{20}$	8.0×10^5	5.0×10^2	1.3×10^{-11}	0.1

Spin-coherent states, the SDE and its characteristic length scales

Let us make transparent the notion of spin coherence. In what follows we shall use the notations from section 4.2. The spin of an electron is a quantum mechanical degree of freedom with two possible realizations with respect to an arbitrarily chosen *quantization axis*. For instance, at the point x in position space, let the local magnetization $\vec{m}(x)$ in a ferromagnet define this axis. Consequently, when the spin of a conduction electron is measured along $\vec{m}$ at point x, there are two possible results: the spin is found to be either parallel or anti-parallel with respect to $\vec{m}$. However, the specific outcome of the experiment is not predetermined in the general case; rather, there are probabilities related to each of the two possible results. Indeed, these probabilities are given by the two eigenvalues of the diagonal element $\varrho(x,x) \in \mathcal{H}_2^{0,+}(\mathbb{C})$ of the density matrix describing the electron. For simplicity, we consider a pure state, the generalization to mixed states is straightforward:

$$\varrho(x,x) = \begin{pmatrix} |\psi_\uparrow(x)|^2 & \psi_\uparrow(x)\overline{\psi}_\downarrow(x) \\ \psi_\downarrow(x)\overline{\psi}_\uparrow(x) & |\psi_\downarrow(x)|^2 \end{pmatrix} . \tag{1.2.19}$$

Here, $\psi_\uparrow$ and $\psi_\downarrow$ denote the spin-up and spin-down component, respectively, of the spinor wave function $\psi \in (L^2)^2$ associated to the spin-1/2 electron and written in the basis defined by the local magnetization $\vec{m}$:

$$\psi(x) = \begin{pmatrix} \psi_\uparrow(x) \\ \psi_\downarrow(x) \end{pmatrix} . \tag{1.2.20}$$

We want to remind of the normalization condition

$$\int dx \left(|\psi_\uparrow(x)|^2 + |\psi_\downarrow(x)|^2 \right) = 1 . \tag{1.2.21}$$

A state (1.2.19) or (1.2.20), respectively, is *spin-coherent with respect to* $\vec{m}$ at point x if both components $\psi_\uparrow(x)$ and $\psi_\downarrow(x)$ are non-zero and, thus, if the off-diagonal elements of $\varrho(x,x)$ are non-zero. This means that the electron is neither entirely

spin-up nor entirely spin-down; rather its state is a quantum superposition of these two. On the other hand, for an incoherent state the matrix (1.2.19) is diagonal and one of the two eigenvalues is zero[3]. Spin-coherent states play a crucial role in the mechanism of spin-transfer torques; in fact, it is solely due to electrons in spin-coherent states with respect to $\vec{m}(x)$ that a torque is exerted on this local magnetization. This can be deduced with the help of the relations (4.2.11) and (4.2.12). We write the density matrix (1.2.19) as well as the system Hamiltonian at point x, denoted by $H \in \mathcal{H}_2(\mathbb{C})$, in the Pauli basis, thereby omitting the position arguments for the sake of an easier notation:

$$\varrho = \varrho_0\sigma_0 + \vec{\varrho}\cdot\vec{\sigma}\,, \qquad H = h_0\sigma_0 + \frac{J_{ex}}{2}\vec{m}\cdot\vec{\sigma}\,. \tag{1.2.22}$$

Using the textbook definition of the spin operator,

$$\vec{S} := \frac{\hbar}{2}\vec{\sigma}\,, \tag{1.2.23}$$

one obtains from Eq. (4.2.12) that the mean value $\vec{s}$ of the spin of the electron in the state ϱ at position x is computed via

$$\vec{s}(x) := \mathrm{tr}\left(\varrho(x,x)\vec{S}\right) = \hbar\vec{\varrho}(x,x)\,. \tag{1.2.24}$$

We are now interested in the dynamics of $\vec{s}$ induced by the exchange coupling in the Hamiltonian (1.2.22). Starting from the von Neumann equation, using the invariance of the trace under cyclic permutation and, furthermore, applying the relations (4.2.3) it is straightforward to arrive at

$$\partial_t\vec{s}\Big|_{ex} = \frac{J_{ex}}{\hbar}\vec{m}\times\vec{s}\,. \tag{1.2.25}$$

Thus, a torque is exchanged if $\vec{s} = (s_1, s_2, s_3)$ has a component perpendicular to $\vec{m}$. Indeed, this is the case solely for spin-coherent states. The crucial point is that, by choosing $\vec{m}$ as the quantization axis, we made $\vec{m}$ the z-axis in spin space. Hence, in Eq. (1.2.22) one has $\vec{m} = (0,0,1)$. Therefore, no torque is exchanged if and only if s_1 and s_2 are zero, which is the case if the matrix (1.2.19) is diagonal, thus for incoherent states.

Let us elaborate briefly on the characteristic length scales of the spin-diffusion equation. The SDE is given by

$$\partial_t\vec{s} = D\Delta\vec{s} - \frac{\vec{s}}{\tau} - \frac{J_{ex}}{\hbar}\vec{s}\times\vec{m}\,. \tag{1.2.26}$$

[3]In case of a mixed state both eigenvalues can be non-zero.

This is a heuristic equation of motion for the spin accumulation generated at a non-magnetic / ferromagnetic interface, as introduced in section 1.1 and sketched in Fig. 1.1iii. Its first term on the right-hand-side describes the diffusion from the interface into the bulk. The second term accounts for spin-flip scattering and the third term models the torque due the magnetization of a ferromagnet. The steady-state solution of Eq. (1.2.26) satisfies

$$0 = D\Delta\vec{s} - \frac{\vec{s}}{\tau} - \frac{J_{ex}}{\hbar}\vec{s}\times\vec{m}\,. \tag{1.2.27}$$

Assuming that $\vec{m}$ does not depend on x, we introduce the parallel and the perpendicular component of the spin accumulation with respect to the constant magnetization direction $\vec{m}$:

$$s_{\parallel} := \vec{s}\cdot\vec{m}\,, \qquad \vec{s}_{\perp} := \vec{s} - (\vec{s}\cdot\vec{m})\vec{m}\,. \tag{1.2.28}$$

In the steady state, these quantities obey

$$0 = D\Delta s_{\parallel} - \frac{s_{\parallel}}{\tau}\,, \tag{1.2.29}$$

$$0 = D\Delta\vec{s}_{\perp} - \frac{\vec{s}_{\perp}}{\tau} - \frac{J_{ex}}{\hbar}\vec{s}_{\perp}\times\vec{m}\,. \tag{1.2.30}$$

From Eq (1.2.29) one obtains that the parallel component decays on the length scale

$$\lambda_{sdl} = \sqrt{\tau D}\,, \tag{1.2.31}$$

which is commonly referred to as the spin-diffusion length. Let $(\vec{e}_i)_{i=1}^3$ denote an orthonormal basis in $\mathbb{R}^3$ with $\vec{e}_3 = \vec{m}$. Hence, the perpendicular component $\vec{s}_{\perp}$ can be mapped into the complex plane,

$$\vec{s}_{\perp} = s_{\perp,1}\vec{e}_1 + s_{\perp,2}\vec{e}_2 \quad \mapsto \quad z = s_{\perp,1} + is_{\perp,2}\,. \tag{1.2.32}$$

For the vector product in Eq. (1.2.30) one obtains

$$\vec{s}_{\perp}\times\vec{m} = \begin{vmatrix} \vec{e}_1 & \vec{e}_2 & \vec{e}_3 \\ s_{\perp,1} & s_{\perp,2} & 0 \\ 0 & 0 & 1 \end{vmatrix} = s_{\perp,2}\vec{e}_1 - s_{\perp,1}\vec{e}_2 \quad \mapsto \quad -iz\,. \tag{1.2.33}$$

Therefore, Eq. (1.2.30) can be written as

$$0 = D\Delta z - \frac{z}{\tau} + i\frac{J_{ex}}{\hbar}z\,, \tag{1.2.34}$$

and one obtains the characteristic length scale of the decay of the perpendicular

component,

$$\lambda_\perp = \sqrt{\frac{\tau\hbar D}{|\hbar - iJ_{ex}\tau|}}\,. \tag{1.2.35}$$

According to [29], typical parameter values in cobalt are $\tau \approx 10^{-12}\,\mathrm{s}$, $D \approx 10^{-3}\,\mathrm{m^2 s^{-1}}$ and $J_{ex} \approx 0.2$ eV. Thus, the length scales evaluate to

$$\lambda_{sdl} \approx 100\,\mathrm{nm}\,, \qquad \lambda_\perp \approx 5\,\mathrm{nm}\,, \tag{1.2.36}$$

which shows that the action of the spin-transfer torque is concentrated in the vicinity of the non-magnetic / ferromagnetic interface.

References

[1] NF Mott. A discussion of the transition metals on the basis of quantum mechanics. *Proceedings of the Physical Society*, 47:571, 1935.

[2] I.A. Campbell, A. Fert, and A.R. Pomeroy. Evidence for two current conduction in iron. *Phil. Mag.*, 15:977, 1967.

[3] A. Fert and I.A. Campbell. Two-current conduction in nickel. *Phys. Rev. Lett.*, 21(16):1190, 1968.

[4] M.D. Stiles and J. Miltat. Spin-transfer torque and dynamics. *Top. Appl. Phys.*, 101:225, 2006.

[5] Chappert C., A. Fert, and F.N. Van Dau. The emergence of spin electronics in data storage. *Nature Mat.*, 6:813–823, 2007.

[6] I. Zutic, J. Fabian, and Das Sarma S. Spintronics: Fundamentals and applications. *Rev. Mod. Phys.*, 76(2):323–410, 2004.

[7] A. Fert and I.A. Campbell. Electrical resistivity of ferromagnetic nickel and iron based alloys. *J Phys. F: Metal Phys.*, 6(5):849, 1976.

[8] MN Baibich, JM Broto, A. Fert, F. Nguyen Van Dau, F. Petroff, P. Etienne, G. Creuzet, A. Friederich, and J. Chazelas. Giant magnetoresistance of(001) fe/(001) cr magnetic superlattices. *Physical Review Letters*, 61(21):2472–2475, 1988.

[9] G. Binasch, P. Grünberg, F. Saurenbach, W. Zinn, et al. Enhanced magnetoresistance in layered magnetic structures with antiferromagnetic interlayer exchange. *Physical Review B*, 39(7):4828–4830, 1989.

[10] A. Fert. Nobel lecture: Origin, development, and future of spintronics. *Rev. Mod. Phys.*, 80:1517, 2008.

[11] S.M. Thompson. The discovery, development and future of gmr: The nobel prize 2007. *J. Phys. D: Appl. Phys.*, 41:093001, 2007.

[12] J.A. Katine and E.E. Fullerton. Device implications of spin-transfer torques. *J. Magn. Magn. Mater.*, 320:1217–1226, 2008.

[13] G.A. Prinz. Magnetoelectronics applications. *J. Magn. Magn. Mater.*, 200:57–68, 1999.

[14] L. Berger. Emission of spin waves by a magnetic multilayer traversed by a current. *Phys. Rev. B*, 54(13):9353, 1996.

[15] J.C. Slonczewski. Current-driven excitation of magnetic multilayers. *J. Magn. Magn. Mater.*, 159:L1–L7, 1996.

[16] M.D. Stiles and A. Zangwill. Anatomy of spin-transfer torque. *Phys. Rev. B*, 66(014407), 2002.

[17] D.C. Ralph and M.D. Stiles. Spin transfer torques. *J. Magn. Magn. Mater.*, 320:1190–1216, 2008.

[18] J. Grollier, P. Boulenc, V. Cros, A. Hamzic, A. Vaurès, A. Fert, and G. Faini. Switching a spin valve back and forth by current-induced domain wall motion. *Appl. Phys. Lett.*, 83:509, 2003.

[19] J.A. Katine, F.J. Albert, R.A. Buhrman, E.B. Myers, and D.C. Ralph. Current-driven magnetization reversal and spin-wave excitations in co/cu/co pillars. *Phys. Rev. Lett.*, 84(14):3149, 2000.

[20] T. Kimura, Y. Otani, and P.M. Levy. Electrical control of the direction of spin accumulation. *Phys. Rev. Lett.*, 99(166601), 2007.

[21] S.I. Kiselev, J.C. Sankey, I.N. Krivorotov, N.C. Emley, R.J. Schoelkopf, R.A. Buhrman, and D.C. Ralph. Microwave oscillations of a nanomagnet driven by a spin-polarised current. *Nature*, 425(380), 2003.

[22] I.N. Krivorotov, N.C. Emley, J.C. Sankey, S.I. Kiselev, D.C. Ralph, and R.A. Buhrman. Time-domain measurements of nanomagnet dynamics driven by spin-transfer torques. *Science*, 307:228, 2005.

[23] M. Tsoi, A.G.M. Jansen, J. Bass, W.-C. Chiang, M. Seck, V. Tsoi, and P. Wyder. Excitation of a magnetic multilayer by an electric current. *Phys. Rev. Lett.*, 80(19):4281, 1998.

[24] F.J. Jedema, H.B. Heersche, A.T. Filip, J.J.A. Baselmans, and B.J. van Wees. Electrical detection of spin precession in a metallic mesoscopic spin valve. *Nature*, 416:713, 2002.

[25] G.S.D. Beach, M. Tsoi, and J.L. Erskine. Current-induced domain wall motion. *J. Magn. Magn. Mater.*, 320:1272–1281, 2008.

[26] E. Simanek. Spin accumulation and resistance due to a domain wall. *Phys. Rev. B*, 63:224412, 2001.

[27] G. Tatara, H. Kohno, and J. Shibata. Theory of domain wall dynamics under current. *J. Phys. Soc. Jap.*, 77(3):031003, 2008.

[28] A. Thiaville, Y. Nakatani, J. Miltat, and Y. Suzuki. Micromagnetic understanding of current-driven domain wall motion in patterned nanowires. *Eurphys. Lett.*, 69(6):990–996, 2005.

[29] S. Zhang, P.M. Levy, and A. Fert. Mechanisms of spin-polarized current-driven magnetization switching. *Phys. Rev. Lett.*, 88(23):236601–1, 2002.

[30] C.J. Garcia-Cervera and X.-P. Wang. Spin-polarized currents in ferromagnetic multilayers. *J. Comp. Phys.*, 224:699, 2007.

Chapter 2

Spin-Transfer Torques: self-consistent solution of the Spin-Diffusion Equation and the Landau-Lifshitz Equation

---

S. Possanner and *N. Ben Abdallah*,
published as a proceeding to the 2010 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD), 6-8 Sept. **2010**, p. 37

Abstract. We present a numerical scheme that allows for the self-consistent treatment of the Landau-Lifshitz equation and the spin-diffusion equation in one space dimension. The scheme is used to simulate magnetic precessions in ferromagnet / normal-metal multilayers that are traversed by strong currents and results are compared to recent experimental observations. The good qualitative and quantitative agreement shows that the diffusive exchange-torque model proposed by Zhang et al. [1] is a legitimate alternative to the ballistic interface-torque model commonly used to describe magnetization dynamics in spin valves.

2.1 Introduction

It has been shown that a current injected perpendicular to the planes of a ferromagnet / normal-metal multilayer can lead to dynamical magnetic states in thin soft-magnetic layers [2, 3]. Depending on the magnitude of the currrent, precession or complete reversal of the magnetization is observed. The underlying effect, often called spin-transfer-torque (STT), has been predicted theoretically by Berger [4] and Slonczewski [5] in 1996 and is the topic of ongoing research.

In a macroscopic continuum approach, the dynamics of the direction of magnetization in a ferromagnet (FM) of volume Ω, $\vec{m}(x,t):\ \Omega\times\mathbb{R}^+\to S^2$, is determined by the Landau-Lifshitz equation [6],

$$\partial_t\vec{m} = -\frac{\gamma\mu_0}{1+\alpha^2}\vec{m}\times\vec{H}_{eff} - \frac{\alpha\gamma\mu_0}{1+\alpha^2}\vec{m}\times(\vec{m}\times\vec{H}_{eff})\,, \tag{2.1.1}$$

where γ is the gyromagnetic ratio, μ_0 stands for the magnetic constant, α denotes the damping parameter and $\vec{H}_{eff}$ is the effective field which is proportional to the functional derivative of the free energy of the system with respect to $\vec{m}$. For the description of the STT in a thin magnetic layer of a FM/normal-metal (NM) multilayer structure (called spin-valve), usually one adds to (2.1.1) a torque-term of the form

$$\partial_t\vec{m}\Big|_{STT} = c_j\vec{m}\times(\vec{m}\times\vec{m}_F)\,, \tag{2.1.2}$$

where c_j is proportional to the current-density traversing the structure perpendicular to the layer planes and $\vec{m}_F\in S^2$ is the magnetization direction of a polarizing thick FM layer. Expression (2.1.2) is derived from ballistic transmission/reflection of spin-coherent Fermi-surface-states at a NM/FM interface [5, 7]. The model (2.1.2) is well established and has proven to be capable of explaining experimental observations, at least qualitatively [8].

An alternative/complementary approach towards STT has been presented by Zhang et. al [1], who assume that the exchange interaction between conduction electrons and core electrons is responsible for the observed magnetization dynamics. They state that one should treat self-consistently the evolution of the core magnetic moments, determined by (2.1.1), and the density of conduction-electron spins, $\vec{s}(x,t):\ \Omega\times\mathbb{R}^+\to\mathbb{R}^3$, determined, for example, by a diffusion equation,

$$\partial_t\vec{s} = D\Delta\vec{s} - \frac{\vec{s}}{\tau} - \frac{J_{ex}}{\hbar}\vec{s}\times\vec{m}\,, \tag{2.1.3}$$

where D is the diffusion constant, Δ denotes the Laplacian, τ stands for the spin-flip

scattering time and J_{ex} is the *sd* exchange constant. Naturally, due to the exchange coupling J_{ex}, the spin density $\vec{s}$ appears in the free energy of $\vec{m}$ and thus in the effective field, which, in the one-dimensional case, we assume to be

$$\vec{H}_{eff} = \vec{H}_{app} + \frac{2K_u}{\mu_0 M_s}(\vec{m}\cdot\vec{u})\vec{u} - M_s(\vec{m}\cdot\vec{w})\vec{w} + \\ + \frac{2A_{ex}}{\mu_0 M_s}\Delta\vec{m} + \frac{J_{ex}}{\mu_0 M_s}\vec{s}. \tag{2.1.4}$$

Here, $\vec{H}_{app}$ denotes the applied magnetic field, K_u the anisotropy constant, M_s the saturation magnetization, A_{ex} the exchange constant, $\vec{u}$ stands for the direction of the magnetic easy axis and $\vec{w}$ is the normal of the easy plane in thin FM layers (the $\vec{w}$-term is missing in thick layers).

Many experimental setups (e.g. in [3]) feature a 5-layer structure like N_{e1} / F_F / N / F / N_{e2}, where $\mathrm{N}_{e1(2)}$ denote NM-electrodes, F_F is a thick polarising FM layer, N is a NM spacer layer and F is the thin FM layer in which the interesting dynamics take place. To simulate spin-injection into such a system, one can impose a Dirichlet boundary condition $\vec{s}_{inj}$ on equation (2.1.3) at the $\mathrm{N}_{e1}/\mathrm{F}_F$ interface, regardless whether electrons flow from N_{e1} to N_{e2} or vice versa (it is the orientation of $\vec{s}_{inj}$ with respect to $\vec{m}_F$ that changes with the sign of the current). The vector $\vec{s}_{inj}$, by the theory of spin injection through an NM/FM junction [9], reads

$$\vec{s}_{inj} = -\vec{m}_F\beta\sqrt{\frac{\tau}{D}}\frac{j}{q}\,, \tag{2.1.5}$$

where β is a polarisation parameter which we choose to be 1 throughout this paper, j denotes the current density and q the proton charge. Thus, for $j > 0$ (particle density flow from N_{e1} to N_{e2}), one obtains a vector $\vec{s}_{inj}$ that is antiparallel to $\vec{m}_F$ and vice versa. At the other three interfaces of the five-layer structure one demands continuity of $\vec{s}$ and the spin current, $-D\partial_x\vec{s}$, and at infinity one imposes the homogenous Neumann condition $\partial_x\vec{s}(\infty) = 0$. Equation (2.1.3) with the boundary/interface conditions mentioned above has been studied analytically in the case of constant $\vec{m}$ [10, 11]. It is the purpose of this work to go beyond the case of $\vec{m}$=*const.* in the Zhang-Levy-Fert (ZLF) model [1] and solve the Landau-Lifshitz equation (2.1.1) and the spin-diffusion equation (2.1.3) self-consistently. The goal is to show that, with reasonable physical parameters, the diffusive ZLF-model is capable of reproducing experimental findings, thereby emphasizing its status as an alternative to the ballistic interface-torque picture, equation (2.1.2).

This paper is structured as follows: in section 2.2 we present a numerical scheme that solves self-consistently the system (2.1.1), (2.1.3) in multilayered structures in

one space dimension. In section 2.3 we apply this scheme to simulate magnetization dynamics in spin-valves that were investigated in [3]. Power spectra of the various oscillation regimes obtained (depending on applied current) are presented. Finally, in section 3.1.3 we discuss our findings in comparison with experiments and macrospin simulations of the Slonczewski model, equation (2.1.2).

2.2 Numerical scheme

Let $\vec{g} := \gamma\mu_0\vec{H}_{eff}/(1+\alpha^2)$, then the system (2.1.1), (2.1.3) can be written as

$$\partial_t \begin{pmatrix} \vec{m} \\ \vec{s} \end{pmatrix} = \begin{pmatrix} -\vec{m}\times\vec{g} - \alpha\vec{m}\times(\vec{m}\times\vec{g}) \\ D\Delta\vec{s} - \dfrac{\vec{s}}{\tau} - \dfrac{J_{ex}}{\hbar}\vec{s}\times\vec{m} \end{pmatrix} =: A(\vec{m},\vec{s})\,, \tag{2.2.1}$$

where we defined the non-linear operator A. This operator is split into four parts, $A = A_1 + A_2 + A_3 + A_4$, which read

$$A_1(\vec{m},\vec{s}) := \begin{pmatrix} 0 \\ D\Delta\vec{s} - \dfrac{\vec{s}}{\tau} - \dfrac{J_{ex}}{\hbar}\vec{s}\times\vec{m} \end{pmatrix}, \tag{2.2.2}$$

$$A_2(\vec{m},\vec{s}) := \begin{pmatrix} -\alpha\vec{m}\times[\vec{m}\times\vec{g}] \\ 0 \end{pmatrix}, \tag{2.2.3}$$

$$A_3(\vec{m},\vec{s}) := \begin{pmatrix} -\vec{m}\times\left[\vec{g} - \dfrac{2\gamma A_{ex}}{M_s(1+\alpha^2)}\Delta\vec{m}\right] \\ 0 \end{pmatrix}, \tag{2.2.4}$$

$$A_4(\vec{m},\vec{s}) := \begin{pmatrix} -\vec{m}\times\left[\dfrac{2\gamma A_{ex}}{M_s(1+\alpha^2)}\Delta\vec{m}\right] \\ 0 \end{pmatrix}. \tag{2.2.5}$$

In our finite-difference scheme, let $n\Delta t$ denote the discrete moments in time, Δt being the time step. Applying Strang operator splitting [12] to (2.2.2)-(2.2.5) we obtain the expression

$$(\vec{m}^{n+1},\vec{s}^{n+1}) = A_1^{\frac{1}{2}}A_2^{\frac{1}{2}}A_3^{\frac{1}{2}}A_4^{1}A_3^{\frac{1}{2}}A_2^{\frac{1}{2}}A_1^{\frac{1}{2}}(\vec{m}^{n},\vec{s}^{n})\,. \tag{2.2.6}$$

Here, $A_i^l(\vec{m}^n,\vec{s}^n)$ denotes an advancement in time of $(\vec{m}^n,\vec{s}^n)$ by $l\Delta t$ by solving a discretized form of the equation $\partial_t(\vec{m},\vec{s}) = A_i(\vec{m},\vec{s})$. The operator A_1 is treated with a Crank-Nicolson scheme that is second order in space. The appearing linear system is solved by Gauss-Seidel iteration and we use the fact that the solution to

the equation $\vec{x} + \vec{x} \times \vec{a} = \vec{y}$, where $\vec{x}, \vec{a}, \vec{y} \in \mathbb{R}^3$, is given by

$$\vec{x} = \frac{\vec{y} + (\vec{a} \cdot \vec{y})\vec{a} + \vec{a} \times \vec{y}}{1 + |a|^2} \,. \tag{2.2.7}$$

The damping operator A_2 is treated with a scheme developed in a work by E et al. [13]. For the operator A_3, denoting the expression in the square brackets in (2.2.4) by $f(\vec{m}, \vec{s})$, we implement a norm-conserving Crank-Nicolson scheme,

$$\frac{\vec{m}^{**} - \vec{m}^{*}}{l\Delta t} = -\frac{\vec{m}^{**} + \vec{m}^{*}}{2} \times f(\vec{m}^{*}, \vec{s}^{*}) \,, \tag{2.2.8}$$

which is linear in $\vec{m}^{**}$ and is solved by Gauss-Seidel iteration using (2.2.7). In case of the operator A_4 we treat the Laplace term semi-implicitly and linearize as follows,

$$\begin{aligned} \frac{\vec{m}_k^{**} - \vec{m}_k^{*}}{l\Delta t} &= -\frac{\vec{m}_k^{**} + \vec{m}_k^{*}}{2} \times \frac{\vec{m}_{k+1}^{**} - 2\vec{m}_k^{**} + \vec{m}_{k-1}^{**}}{\Delta x^2} \\ &= -\frac{\vec{m}_k^{**} + \vec{m}_k^{*}}{2} \times \frac{\vec{m}_{k+1}^{**} + 2\vec{m}_k^{*} + \vec{m}_{k-1}^{**}}{\Delta x^2} \,. \end{aligned} \tag{2.2.9}$$

Here, k stands for the spatial index and Δx denotes the grid spacing. The system (2.2.9) is solved by Gauss-Seidel iteration using (2.2.7).

2.3 Simulation of spin-valve under current

We simulate the spin transfer-torque in a four layer structure that reflects the experimental setup of [3]. The layer thicknesses are 40 nm FM / 10 nm NM / 3 nm FM / 3 μm NM. Current flows in the x-direction, where x=0 denotes the interface between an electrode and the 40 nm FM layer (F_F), and $\vec{s}(x{=}0, t) = \vec{s}_{inj}(t)$, i.e., the Dirichlet condition is allowed to change in time with the direction of $\vec{m}_F$, see eq. (2.1.5). In this setup, the dynamics inflicted on the thick layer magnetization have been found to be minor in the investigated parameter range. A time-independent vector $\vec{s}_{inj}$ does not account for a correct physical picture and changes the results drastically, causing the dynamics of $\vec{m}_F$ to dominate over those in the thin FM layer. The number of grid points in the respective layers is: 32 / 16 / 16 / 64. Further refinement did not cause a change in the results. The easy axis in the thick FM layer is assumed to be in the z-direction and the one in the thin FM layer is chosen to be tilted by 5° to that axis. In the FM layers, we use parameters associated to Cobalt [3, 8], $K_u = 2.41 \times 10^4\,\mathrm{Jm}^{-3}$, $M_s = 8 \times 10^5\,\mathrm{Am}^{-1}$, $A_{ex} = 2 \times 10^{-11}\,\mathrm{Jm}^{-1}$, $\tau = 10^{-12}\,\mathrm{s}$, $\alpha = 0.02$, $D = 10^{-3}\,\mathrm{m}^2\mathrm{s}^{-1}$ and $J_{ex} = 0.1\,\mathrm{eV}$ [1]. In the NM layers we use $D = 10^{-3}\,\mathrm{m}^2\mathrm{s}^{-1}$ and $\tau = 10^{-10}\,\mathrm{s}$. Homogenous Neumann conditions are applied on boundaries for the Landau-Lifshitz equation in each FM layer. Considering the

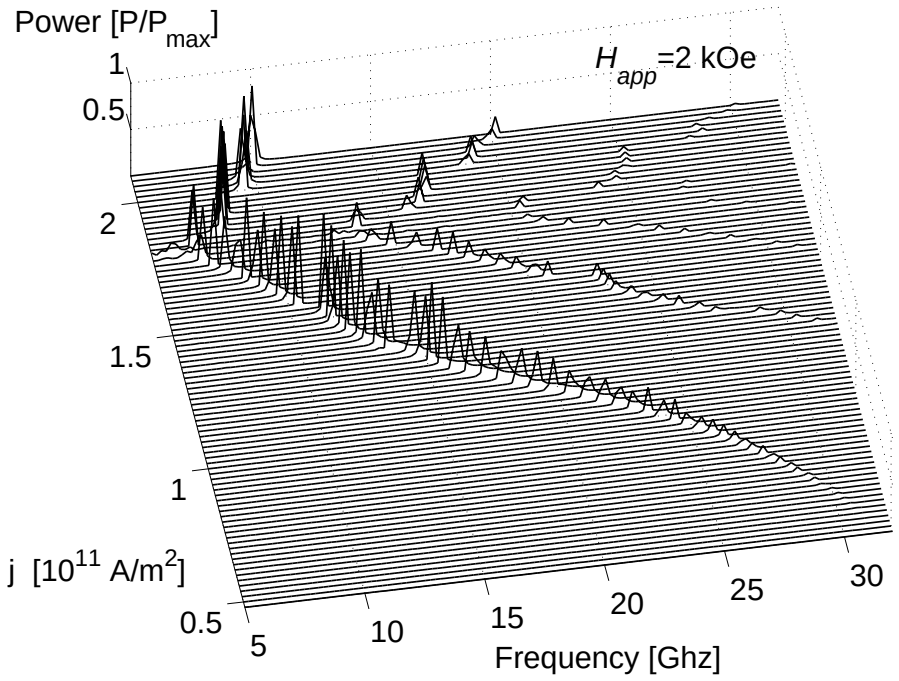


Figure 2.1: Simulated power spectra of magnetic oscillations in the investigated spin valve for different current densities j. P_{max} stands for the largest of all observed peaks.

spin-diffusion equation, the boundary/interface conditions discussed in section 2.1 are implemented.

In all simulations, a magnetic field $|\vec{H}_{app}| = 2\,\mathrm{kOe}$ was applied in the positive z-direction. Thus, the intitial state of the spin valve is that the magnetizations of both FM layers are almost parallel, $\vec{m}_F$ being perfectly aligned with $\vec{H}_{app}$ and $\vec{m}$ of the thin layer slightly tilted. We then inject a positive current, which, according to equation (2.1.5), leads to an $\vec{s}_{inj}$ that is antiparallel to $\vec{m}_F$. The spin density $\vec{s}$ will thus have a small perpendicular component to $\vec{m}$ in the thin FM layer, inflicting the so-called spin-transfer torque. At sufficiently large currents, one observes magnetic oscillations in different regimes (see below) in the thin FM layer. Note that if the easy axis of the thin FM layer would point in the z-direction too, no torque would occur in the zero-temperature ZLF-model studied here. Besides, the intra-layer exchange coupling in the thin FM layer is so strong that its magnetization distribution $\vec{m}(x,t)$ stays almost uniform for all t. In the thick FM layer, $\vec{m}_F(x,t)$ does not stay uniform but shows only very small deviations from $+\vec{z}$ (up to 3°) in the course of a simulation. In all simulations, a time step $\Delta t = 10^{-13}\,\mathrm{s}$ was used, smaller time steps did not lead to a change in the results.

In order to analyse our simulations, we assume that the deviation of the electrical resistance in a spin valve from its value in the parallel configuration is $\Delta R = \Delta R_{max}(1 - \hat{\vec{m}}_F \cdot \hat{\vec{m}})/2$ [8], where $\hat{\vec{m}}_F \cdot \hat{\vec{m}}$ denotes the dot product of the spatial mean

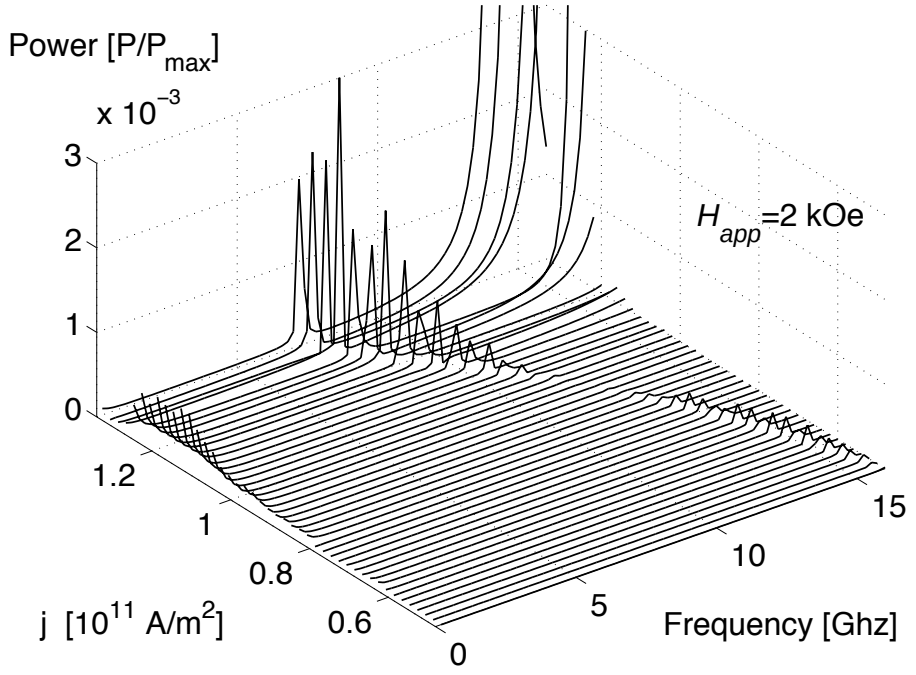
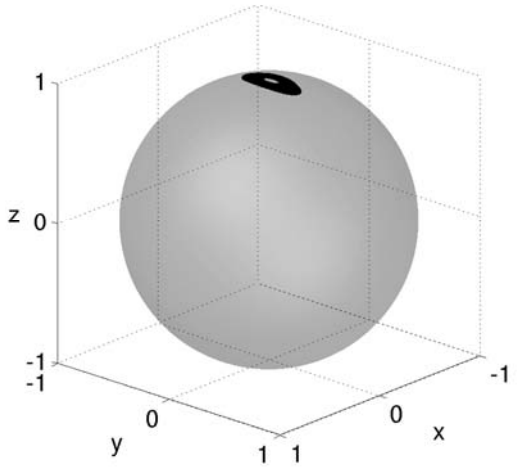


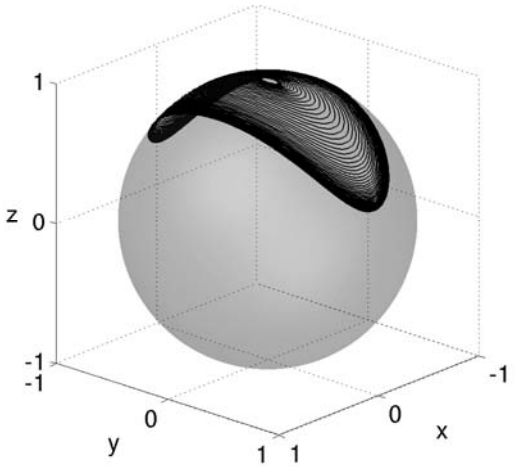
Figure 2.2: Simulated power spectra for current densities $j \in [0.48, 1.34] \times 10^{11}\,\mathrm{Am}^{-2}$ in the range of 0-16 Ghz.

values of the magnetizations in the two FM layers. Power spectra of the magnetic oscillations in the thin layer are obtained by performing a discrete Fourier transform of the simulated signals $\Delta R(t)$ (we use 5000 data points that are spaced by $\Delta t = 10^{-12}\,\mathrm{s}$) and by plotting the squared norm of the Fourier coefficients over the frequency.

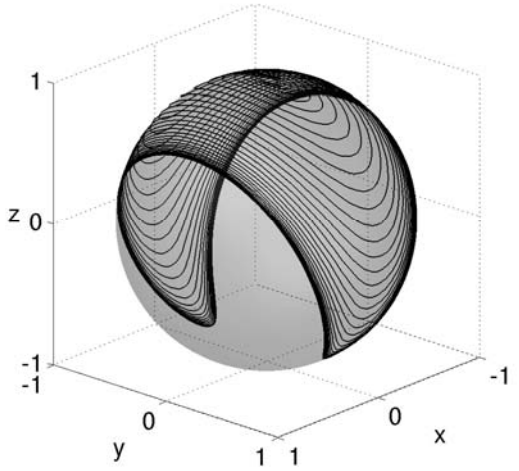
Figure 6.1 shows the power spectra for current densities in the interval $j \in [0.48, 2.10] \times 10^{11}\,\mathrm{Am}^{-2}$. The dominant peaks are the ones corresponding to $2f_0$, i.e., two times the actual oscillation frequency. They start to emerge at around $j = 0.6 \times 10^{11}\,\mathrm{Am}^{-2}$ at $f \approx 30\,\mathrm{Ghz}$ and shift down in frequency with increasing current. At currents larger than $j = 1.4 \times 10^{11}\,\mathrm{Am}^{-2}$, higher harmonics ($4f_0$, $6f_0$) emerge too. The power of the odd harmonics (f_0, $3f_0$, etc.) is three to five orders of magnitude smaller than the one of $2f_0$, as can be seen in Figure 6.2, where power spectra in the intervals $j \in [0.48, 1.34] \times 10^{11}\,\mathrm{Am}^{-2}$ and $f \in [0, 16]\,\mathrm{Ghz}$ are depicted. Oscillations start at a frequency $f_0 = 16\,\mathrm{Ghz}$ for $j = 0.48 \times 10^{11}\,\mathrm{Am}^{-2}$, then f_0 diminishes with increasing current before it emerges again at frequencies lower than 10 Ghz at about two orders of magnitude more powerful. During this transition, we observe a low-frequency background.



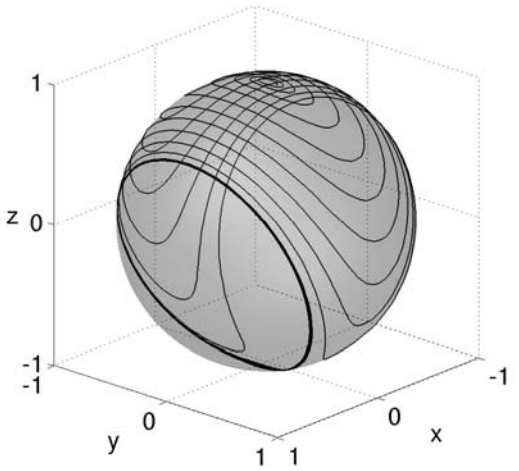
(i) $j = 0.5 \times 10^{11}\ \mathrm{Am}^{-2}$



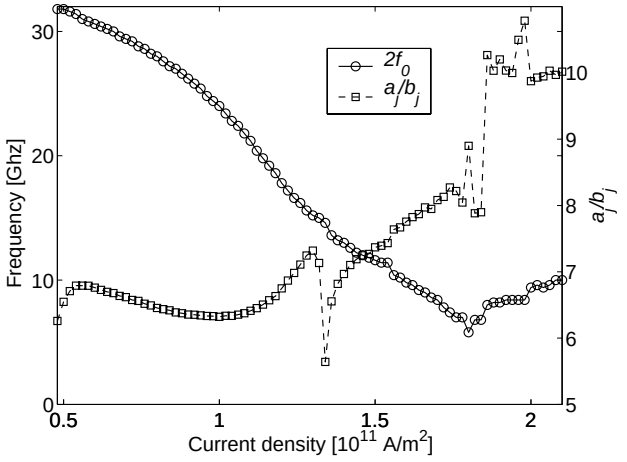
(ii) $j = 0.9 \times 10^{11}\ \mathrm{Am}^{-2}$



(iii) $j = 1.4 \times 10^{11}\ \mathrm{Am}^{-2}$



(iv) $j = 2.1 \times 10^{11}\ \mathrm{Am}^{-2}$



(v) Simulation results of two times the fundamental oscillation frequency f_0 and the ratio a_j/b_j as a function of current density j.

Figure 2.3: Obtained oscillation regimes of the mean value $\hat{\vec{m}}(t)$ in the thin FM layer: i) small angle precession, ii) large angle precession, iii) clamshell orbit, iv) out-of-plane precession.

We refer to [8] for a detailed interpretation of the obtained power spectra. The main reason for the occurrence of different harmonics is that $\hat{\vec{m}}$ moves along elliptical trajectories that are not symmetric with respect to the z-axis, as can be seen in Figure 6.3 where the different simulated oscillation regimes are depicted. We found a) small angle precessions, b) large angle precessions, c) clamshell orbits and d) out-of-plane precessions.

As derived by Zhang et al. [1], the effect of adding the $\vec{s}$-term to the effective field, equation (2.1.4), can be expressed in compact form by augmenting the Landau-Lifshitz equation by two terms,

$$\partial_t \vec{m}\Big|_{ZLF} = a_j \vec{m} \times (\vec{m} \times \vec{m}_F) + b_j \vec{m} \times \vec{m}_F \,. \tag{2.3.1}$$

This expression differs from the Slonczewski torque term (2.1.2) by the second term on the right-hand-side, the so-called effective field term. In our simulations, we determined the coefficients $a_j(x,t)$ and $b_j(x,t)$ in the thin FM layer, integrated them over the thin layer domain and then calculated the ratio of the time mean of these integrals. Results are depicted in Figure 6.4 for different current densities, along with the corresponding frequencies $2f_0$. The ratio a_j/b_j is found to be roughly 7 for small/large angle precessions and around 10 for out of plane precessions.

2.4 Discussion

As far as f_0 is concerned (Figure 6.2), our results show good qualitative and quantitative agreement with the experimental findings in [3]. However, the much stronger peaks of $2f_0$ (Figure 6.1) were not reported there, which is quite surprising. The different oscillation regimes (Figure 6.3) have been reported in macrospin simulations [6] with the Slonczewski spin-torque model (2.1.2). The similarity to the ZLF-model is understood by looking at the ratio a_j/b_j (Figure 6.4), which is between 6 and 10 in our simulations. An experimental value of 5.3 is reported by Zimmler et al. [14] for a 3 nm Cobalt layer.

2.5 Conclusion

We showed that the diffusive exchange-torque model (ZLF model) proposed by Zhang et al. [1] represents a legitimate alternative to the ballistic interface torque-model usually used to describe magnetization dynamics in spin valves. For the first time a numerical study of the coupled system Landau-Lifshitz equation / spin-diffusion equation has been carried out using reasonable physical parameters in

order to enable comparison with experimental data. The obtained simulation results encourage the study of the ZLF-model in greater detail, especially in 3D geometries.

Acknowledgment

The authors acknowledge support from the Marie Curie Early Stage Network DEASE: MEST-CT-2005-021122 funded by the European Union and from the project QUATRAIN funded by the french National agency for research (ANR 2007-2011).

References

[1] S. Zhang, P.M. Levy, and A. Fert. Mechanisms of spin-polarized current-driven magnetization switching. *Phys. Rev. Lett.*, 88(23):236601–1, 2002.

[2] M. Tsoi, A.G.M. Jansen, J. Bass, W.-C. Chiang, M. Seck, V. Tsoi, and P. Wyder. Excitation of a magnetic multilayer by an electric current. *Phys. Rev. Lett.*, 80(19):4281, 1998.

[3] S.I. Kiselev, J.C. Sankey, I.N. Krivorotov, N.C. Emley, R.J. Schoelkopf, R.A. Buhrman, and D.C. Ralph. Microwave oscillations of a nanomagnet driven by a spin-polarised current. *Nature*, 425(380), 2003.

[4] L. Berger. Emission of spin waves by a magnetic multilayer traversed by a current. *Phys. Rev. B*, 54(13):9353, 1996.

[5] J.C. Slonczewski. Current-driven excitation of magnetic multilayers. *J. Magn. Magn. Mater.*, 159:L1–L7, 1996.

[6] D.V. Berkov and J. Miltat. Spin-torque driven magnetization dynamics: Micromagnetic modeling. *J. Magn. Magn. Mater.*, 320:1238–1259, 2008.

[7] M.D. Stiles and A. Zangwill. Anatomy of spin-transfer torque. *Phys. Rev. B*, 66(014407), 2002.

[8] D.V. Berkov and N.L. Gorn. Magnetization precession due to a spin-polarized current in a thin nanoelement: Numerical simulation study. *Phys. Rev. B*, 72:094401, 2005.

[9] I. Zutic, J. Fabian, and Das Sarma S. Spintronics: Fundamentals and applications. *Rev. Mod. Phys.*, 76(2):323–410, 2004.

[10] A. Shpiro, P.M. Levy, and S. Zhang. Self-consistent treatment of nonequilibrium spin torques in magnetic multilayers. *Phys. Rev. B*, 67:104430, 2003.

[11] J. Guo, M.B.A. Jalil, and S.G. Tan. Efficient spin transfer torque in pseudo-spin-valve structure. *J. Appl. Phys.*, 103:07A718, 2008.

[12] W. Hundsdorfer and J.G. Verwer. *Numerical solution of time-dependent advection-diffusion-reaction equations*. Springer Verlag Berlin, 2003.

[13] W. E and Wang Xiao-Ping. Numerical methods for the landau-lifshitz equation. *SIAM J. Numer. Anal.*, 38(5):1647, 2000.

[14] M.A. Zimmler, Oezyilmaz B., W. Chen, and A.D. Kent. Current-induced effective magnetic fields in co/cu/co nanopillars. *Phys. Rev. B*, 70:184438, 2004.

Part II

Kinetic Modeling of Spin-Coherent Electron Transport

Chapter 3

Quantum Mechanics in Phase Space

Overview. It is the aim of this chapter to formulate the basic concepts of quantum mechanics in the phase space [1, 2]. This formulation has been particularly useful for studying the transition from the quantum- to the kinetic or mesoscopic level in transport theory [3, 4, 5]. The incorporation of spin-degrees of freedom is straightforward and will be of central interest in what follows [6, 7]. In a first step we shall introduce the Wigner-Weyl correspondence between operators acting in the system's state space and symbols in the phase space. The strong similarities between these symbols and observables in classical mechanics will be highlighted. Subsequently, the semi-classical scaling of quantum mechanical expressions is discussed in detail and we shall explicitly compute the first few terms of the Moyal product of operator symbols. The Moyal product will then be applied to perform the semi-classical limit in the von Neumann equation for a single particle with spin. In the particular case that the spin part of the system's Hamiltonian is small compared to its scalar part, the semi-classical limit results in the matrix-Vlasov equation. This equation determines the time evolution of a distribution matrix which plays the role of the distribution function in classical mechanics, with the difference that its matrix form allows for coherence of spin basis states in a particle's dynamics.

3.1 Weyl quantization and Wigner transform

The following is a collection of basic notations and properties needed to formulate quantum mechanics in the phase space. The state space of a single particle with spin is $(L^2(\mathbb{R}^d_x, \mathbb{C}))^n$ and it will be denoted simply by $(L^2)^n$. Here, d is the dimension of the position space and n stands for the dimension of the spin space, e.g. $n = 2$ for spin-1/2 particles and $n = 3$ for spin-1 particles. For $\phi, \psi \in (L^2)^n$, the definition of the scalar product

$$(\phi, \psi)_{(L^2)^n} := \sum_{i=1}^{n} \int \mathrm{d}x \, \phi_i(x) \overline{\psi}_i(x) \tag{3.1.1}$$

makes $(L^2)^n$ a Hilbert space with the norm

$$||\phi||_{(L^2)^n} := \sqrt{(\phi, \phi)_{(L^2)^n}} \, . \tag{3.1.2}$$

The Fourier transform of $\phi \in (L^2)^n$ and its inverse are defined component-wise,

$$\hat{\phi}_i(\xi) = \frac{1}{(2\pi)^{d/2}} \int \mathrm{d}x \, \phi_i(x) e^{-i\xi \cdot x} \, , \qquad \phi_i(x) = \frac{1}{(2\pi)^{d/2}} \int \mathrm{d}\xi \, \hat{\phi}_i(\xi) e^{ix \cdot \xi} \, , \tag{3.1.3}$$

where $\xi \in \mathbb{R}^d_\xi$ plays the role of the conjugate variable to the position x. We note the fundamental relations (in a weak sense)

$$\frac{1}{(2\pi)^d} \int \mathrm{d}\xi \, e^{\pm ix \cdot \xi} = \delta(x) \, , \qquad \frac{1}{(2\pi)^d} \int \mathrm{d}x \, e^{\pm ix \cdot \xi} = \delta(\xi) \, , \tag{3.1.4}$$

where $\delta(\cdot)$ stands for the Dirac delta distribution. The action of an operator A in $(L^2)^n$ can be defined by means of the integral kernel $\mathcal{A} : \mathbb{R}^{2d}_x \to \mathbb{C}^{n \times n}$, given component-wise by

$$[A\phi]_i(x) = \sum_{j=1}^{n} \int \mathrm{d}y \, \mathcal{A}_{ij}(x, y) \phi_j(y). \tag{3.1.5}$$

The corresponding adjoint operator $A^\dagger$ in $(L^2)^n$ is defined via the scalar product,

$$\begin{aligned} (A\phi, \psi)_{(L^2)^n} &= \sum_{i=1}^{n} \int \mathrm{d}x \, [A\phi]_i(x) \overline{\psi}_i(x) \\ &= \sum_{i,j=1}^{n} \int \mathrm{d}x \int \mathrm{d}y \, \mathcal{A}_{ij}(x, y) \phi_j(y) \overline{\psi}_i(x) \\ &:= (\phi, A^\dagger \psi)_{(L^2)^n} \, . \end{aligned} \tag{3.1.6}$$

We deduce from Eq. (3.1.6) that the integral kernel $\mathcal{A}^\dagger$ of the adjoint operator is defined by

$$\mathcal{A}^\dagger_{ij}(x,y) := \overline{\mathcal{A}}_{ji}(y,x)\,. \tag{3.1.7}$$

In the framework of pseudo-differential calculus [8], each operator kernel $\mathcal{A}(x,y)$ is assigned a symbol $a : \mathbb{R}^d_x \times \mathbb{R}^d_\xi \to \mathbb{C}^{n\times n}$ in the phase space via the mapping $\mathcal{W} : \mathcal{A}_{ij} \mapsto a_{ij}$, defined by

$$a_{ij}(x,\xi) = \mathcal{W}[\mathcal{A}_{ij}](x,\xi) := \int \mathrm{d}y\, \mathcal{A}_{ij}\left(x+\frac{y}{2}, x-\frac{y}{2}\right) e^{-iy\cdot\xi}. \tag{3.1.8}$$

The mapping (3.1.8) of operator kernels to phase space symbols is called the Wigner transform. It is easily verified that its inverse is given by

$$\mathcal{A}_{ij}(x,y) = \mathcal{W}^{-1}[a_{ij}](x,y) = \frac{1}{(2\pi)^d}\int \mathrm{d}\xi\, a_{ij}\left(\frac{x+y}{2},\xi\right) e^{i(x-y)\cdot\xi}. \tag{3.1.9}$$

Inserting the definition (3.1.7) into Eq. (3.1.8) results in

$$a^\dagger_{ij}(x,\xi) = \overline{a}_{ji}(x,\xi) \tag{3.1.10}$$

for the symbol $a^\dagger_{ij}$ corresponding to the adjoint operator $A^\dagger$. Thus, in the case that $A = A^\dagger$ is a self-adjoint operator, the matrix $a(x,\xi)$ defined component-wise in Eq. (3.1.8) is Hermitian at every point (x,ξ) in the phase space. Inserting the inverse of the Wigner transform (3.1.9) into Eq. (3.1.5) allows one to write the action of an operator A in $(L^2)^n$ by means of its phase space symbol a,

$$[A\phi]_i(x) = \frac{1}{(2\pi)^d}\sum_{j=1}^n \int \mathrm{d}y \int \mathrm{d}\xi\, a_{ij}\left(\frac{x+y}{2},\xi\right)\phi_j(y) e^{i(x-y)\cdot\xi}. \tag{3.1.11}$$

Equation (3.1.11) defines the mapping $\mathrm{Op}(a) = A$. This mapping is called the Weyl quantization rule; it provides a direct correspondence between functions in the phase space and operators in $(L^2)^n$.

Before applying the Wigner-Weyl calculus to quantum mechanics, we shall briefly discuss the basic physical units in which quantities will be displayed. The physical unit of a quantity q will be denoted by $[q]$. Let us start the discussion with a general wavefunction $\phi(x) \in (L^2)^n$. It is clear that the unit of its components is

$$[\phi_i] = \frac{1}{[x]^{1/2}}\,, \tag{3.1.12}$$

since $|\phi_i(x)|^2$ is interpreted as a probability distribution function in position space. A particle's momentum p is given by $p = \hbar\xi$, where $\hbar$ denotes the Planck constant

divided by 2π, since

$$[\hbar] = [xp]\,, \qquad\qquad [\xi] = \frac{1}{[x]}\,. \tag{3.1.13}$$

In order for the modulus square of the Fourier transform of $\phi_i(x)$ to be interpreted as a momentum distribution, one has to properly redefine the Fourier transform and its inverse (3.1.3), respectively,

$$\begin{aligned} \tilde{\phi}_i(p) &:= \frac{1}{(2\pi\hbar)^{d/2}} \int \mathrm{d}x\, \phi_i(x) e^{-\frac{i}{\hbar}x\cdot p}\,, \\ \phi_i(x) &:= \frac{1}{(2\pi\hbar)^{d/2}} \int \mathrm{d}p\, \tilde{\phi}_i(p) e^{\frac{i}{\hbar}x\cdot p}\,. \end{aligned} \tag{3.1.14}$$

Using Eq. (3.1.12) and Eq. (3.1.13) it is clear that the definition (3.1.14) results in

$$[\tilde{\phi}_i] = \frac{1}{[p]^{1/2}}\,, \tag{3.1.15}$$

which is the desired result. Now let $(\phi^k)_{k=1}^{\infty}$ stand for an orthonormal basis set in $(L^2)^n$, i.e.

$$||\phi^k||^2_{(L^2)^n} = 1\,, \qquad (\phi^k, \phi^l)_{(L^2)^n} = \delta_{kl}\,, \qquad \forall\, l, k \in \mathbb{N}\,, \tag{3.1.16}$$

where $\delta_{kl} = 1$ for $k = l$ and zero otherwise. The components $\mathcal{A}_{ij}$ of the intgral kernel of a Hermitian operator A in $(L^2)^n$ can be written as

$$\mathcal{A}_{ij}(x, y) = \sum_{k,l=1}^{\infty} c_A^{kl} \phi_i^k(x) \overline{\phi_j^l}(y)\,, \tag{3.1.17}$$

with the coefficients $c_A^{kl} = \overline{c}_A^{lk} \in \mathbb{C}$. From Eq. (3.1.12) it follows that

$$[\mathcal{A}_{ij}] = \frac{[c_A^{kl}]}{[x]}\,. \tag{3.1.18}$$

In particular, a single-particle quantum state (pure or mixed) is defined by its density operator $\rho : (L^2)^n \to (L^2)^n$. The corresponding operator kernel ϱ is called the density matrix; its components will be denoted by

$$\varrho_{ij}(x, y) = \sum_{k=1}^{\infty} \lambda_k \phi_i^k(x) \overline{\phi_j^k}(y)\,. \tag{3.1.19}$$

where $\lambda_k \in \mathbb{R}^+$ stands for the relative portion of the state $\phi^k \in (L^2)^n$ in the mixed

state (3.1.19) and one has $\sum_k \lambda_k = 1$. Moreover,

$$[\varrho_{ij}] = \frac{1}{[x]}\,. \tag{3.1.20}$$

The equivalents of (3.1.17) and (3.1.19) in the "physical" phase space $\mathbb{R}^d_x \times \mathbb{R}^d_p$ are defined via the Wigner transform (3.1.8) in the following manner [9]:

$$a_{ij}(x,p) := \mathcal{W}[\varrho_{ij}]\left(x, \frac{p}{\hbar}\right) = \int \mathrm{d}y\, \mathcal{A}_{ij}\left(x + \frac{y}{2}, x - \frac{y}{2}\right) e^{-\frac{i}{\hbar}y\cdot p}\,, \tag{3.1.21}$$

$$w_{ij}(x,p) := \frac{1}{(2\pi\hbar)^d}\mathcal{W}[\varrho_{ij}]\left(x, \frac{p}{\hbar}\right) = \frac{1}{(2\pi\hbar)^d}\int \mathrm{d}y\, \varrho_{ij}\left(x + \frac{y}{2}, x - \frac{y}{2}\right) e^{-\frac{i}{\hbar}y\cdot p}\,, \tag{3.1.22}$$

which yields

$$[a_{ij}] = [c_A^{kl}]\,, \qquad [w_{ij}] = \frac{1}{[xp]}\,. \tag{3.1.23}$$

The normalization factor $(2\pi\hbar)^{-d}$ is present only in the Wigner transform of the density matrix (3.1.22), but absent for all other operators. The reason for this will become transparent in the last paragraph of this section. The inverse of (3.1.21) and (3.1.22), respectively, is given by

$$\mathcal{A}_{ij}(x,y) = \frac{1}{(2\pi\hbar)^d}\int \mathrm{d}p\, a_{ij}\left(\frac{x+y}{2}, p\right) e^{\frac{i}{\hbar}(x-y)\cdot p}\,. \tag{3.1.24}$$

$$\varrho_{ij}(x,y) = \int \mathrm{d}p\, w_{ij}\left(\frac{x+y}{2}, p\right) e^{\frac{i}{\hbar}(x-y)\cdot p}\,. \tag{3.1.25}$$

The phase space symbols (3.1.22) are the components of the Wigner matrix $w = (w_{ij})_{i,j=1}^n$, which is a Hermitian matrix since ρ is self-adjoint. In view of Eq. (3.1.19) it will be useful to define the Wigner transform of two state vectors $\phi, \psi \in L^2$:

$$\mathcal{W}[\phi,\psi](x,p) := \frac{1}{(2\pi\hbar)^d}\int \mathrm{d}y\, \phi\left(x + \frac{y}{2}\right)\overline{\psi}\left(x - \frac{y}{2}\right) e^{-\frac{i}{\hbar}y\cdot p}\,. \tag{3.1.26}$$

From Eq. (3.1.19) we obtain

$$w_{ij} = \sum_{k=1}^{\infty} \lambda_k \mathcal{W}(\phi_i^k, \phi_j^k)\,. \tag{3.1.27}$$

Working in the Wigner representation (3.1.22) respectively (3.1.27) displays many analogies between quantum and classical systems. For instance, from Eq. (3.1.26)

one immediately obtains

$$\int \mathrm{d}p\, \mathcal{W}[\phi,\phi](x,p) = |\phi(x)|^2 \,. \tag{3.1.28}$$

Moreover, inserting the inverse Fourier transform (3.1.14) into Eq. (3.1.26) yields

$$\mathcal{W}[\phi,\psi] = \frac{1}{(2\pi\hbar)^{2d}} \int \mathrm{d}y \int \mathrm{d}q \int \mathrm{d}q'\, \tilde{\phi}(q)\overline{\tilde{\psi}}(q') e^{\frac{i}{\hbar}[(q+q')/2-p]\cdot y} e^{\frac{i}{\hbar}(q+q')\cdot x} \tag{3.1.29}$$

Performing now the variable transformation

$$\begin{aligned} q &= u + \frac{v}{2} & \qquad u &= \frac{q+q'}{2} \\ q' &= u - \frac{v}{2} & \qquad v &= q - q' \end{aligned} \tag{3.1.30}$$

and furthermore using the identity (3.1.4) results in

$$\mathcal{W}[\phi,\psi](x,p) := \frac{1}{(2\pi\hbar)^d} \int \mathrm{d}v\, \tilde{\phi}\left(p + \frac{v}{2}\right) \overline{\tilde{\psi}}\left(p - \frac{v}{2}\right) e^{\frac{i}{\hbar}v\cdot x} \,. \tag{3.1.31}$$

Thus, one obtains

$$\int \mathrm{d}x\, \mathcal{W}[\phi,\phi](x,p) = |\tilde{\phi}(p)|^2 \,, \tag{3.1.32}$$

which is again analogous to a classical phase space distribution function. It can be seen in Eq. (3.1.29) that the sign appearing in the exponential of the Fourier transform of operator kernels is reversed for the second argument; thus one defines

$$\begin{aligned} \tilde{\varrho}(p,q) &:= \frac{1}{(2\pi\hbar)^d} \int \mathrm{d}x \int \mathrm{d}y\, \rho(x,y) e^{-\frac{i}{\hbar}x\cdot p} e^{\frac{i}{\hbar}y\cdot q} \,, \\ \varrho(x,y) &:= \frac{1}{(2\pi\hbar)^d} \int \mathrm{d}p \int \mathrm{d}q\, \tilde{\rho}(p,q) e^{\frac{i}{\hbar}x\cdot p} e^{-\frac{i}{\hbar}y\cdot q} \,. \end{aligned} \tag{3.1.33}$$

In the density matrix approach with physical units, the expectation value $\langle A\rangle$ of an operator A is computed by

$$\langle A\rangle = \mathrm{tr}\,(\varrho A) = \sum_{ij} \int \mathrm{d}x \int \mathrm{d}z\, \varrho_{ij}(x,z) \mathcal{A}_{ji}(z,x). \tag{3.1.34}$$

Inserting here the inverse Wigner transforms (3.1.24) and (3.1.25) results in

$$\langle A\rangle = \frac{1}{(2\pi\hbar)^d} \sum_{ij} \int \mathrm{d}x\mathrm{d}z\mathrm{d}p\mathrm{d}q\, w_{ij}\left(\frac{x+z}{2}, p\right) e^{\frac{i}{\hbar}(x-z)\cdot p} a_{ji}\left(\frac{z+x}{2}, q\right) e^{\frac{i}{\hbar}(z-x)\cdot q} \tag{3.1.35}$$

Performing now the variable transformation

$$\begin{aligned} y &= \frac{x+z}{2} & x &= y + \frac{y'}{2} \\ y' &= x - z & z &= y - \frac{y'}{2} \end{aligned} \tag{3.1.36}$$

and furthermore using the identity (3.1.4) leads to

$$\langle A \rangle = \sum_{ij} \int \mathrm{d}x \int \mathrm{d}p \, w_{ij}(x,p) a_{ji}(x,p) \,, \tag{3.1.37}$$

which resembles the classical expectation value of the observable a in the phase space, given the distribution matrix w. Finally, using Eq. (3.1.24) one can restate the Weyl quantization (3.1.11) in physical units:

$$[A\phi]_i(x) = \frac{1}{(2\pi\hbar)^d} \sum_{j=1}^{n} \int \mathrm{d}y \int \mathrm{d}p \, a_{ij}\left(\frac{x+y}{2}, p\right) \phi_j(y) e^{\frac{i}{\hbar}(x-y)\cdot p}. \tag{3.1.38}$$

As a concluding remark we shall briefly elaborate on the normalization factor $(2\pi\hbar)^{-d}$ in Eq. (3.1.22), which is absent in Eq. (3.1.21) for operators other than the state operator. Its presence led to the relations (3.1.28) and (3.1.32); these relations together with Eq. (3.1.27) ensure the normalization of the state operator ρ:

$$\mathrm{tr}\,(\rho) = \sum_{i} \int \mathrm{d}x \int \mathrm{d}p \, w_{ii}(x,p) = 1 \,, \tag{3.1.39}$$

where $\mathrm{tr}\,(\cdot)$ denotes the operator trace and we used the orthonormality of the basis functions (3.1.16). The absence of $(2\pi\hbar)^{-d}$ in Eq. (3.1.21) is explained by the formulation of the Weyl quantization (3.1.38). The latter is obtained by inserting the inverse Wigner transform (3.1.24) into Eq. (3.1.5). Because of the presence of the normalization factor in the *inverse* Wigner transform one obtains the correspondence

$$p \mapsto -i\hbar\nabla_x \,, \tag{3.1.40}$$

which amounts to the usual definition of the momentum operator.

3.2 Semi-classical scaling

It is the aim of this section to make transparent the notion of the "semiclassical scaling" in the phase-space formulation of quantum mechanics. Let us thus reformulate the equations of the Wigner-Weyl calculus in a dimensionless form which

allows one to tune the role of quantum effects within a physical system by means of a small parameter ε. In general, these quantum effects are non-negligible if the wave-nature of quantum states is prominent, leading to coherence, diffraction or interference of states of the physical system. As in classical optics, these phenomena occur when wave functions are delocalized and when the size of obstacles is on the same scale as the wavelength attributed to state vectors. For instance, let ℓ_c stand for a characteristic length scale of the system, e.g. the scale on which a given potential varies considerably. Furthermore, let ξ_c denote a characteristic wave number of state vectors of the system such that the characteristic wave length is the *de Broglie* wave length $\lambda_{DB} = 2\pi/\xi_c$. The quantum and the classical regime can then be characterized as follows:

$$\text{quantum:} \quad \frac{\lambda_{DB}}{\ell_c} \approx 1\,, \qquad \text{classical:} \quad \frac{\lambda_{DB}}{\ell_c} \ll 1\,. \tag{3.2.1}$$

Associated to ξ_c is the characteristic momentum $p_c = \hbar\xi_c$; thus we introduce the semiclassical parameter ε as

$$\varepsilon = \frac{\hbar}{\ell_c p_c}\,, \tag{3.2.2}$$

which is also called the "scaled Planck constant". According to Eq. (3.2.1), the smaller ε becomes the more is the classical regime approached. Since $\ell_c p_c$ is a characteristic action of the system, one could also start from the characteristic energy E_c and time τ_c, respectively, and define $\varepsilon = \hbar/E_c\tau_c$. Indeed, for parabolic bands these two approaches are equivalent. Let us now introduce the dimensionless variables

$$x' = \frac{x}{\ell_c}\,, \qquad p' = \frac{p}{p_c}\,, \tag{3.2.3}$$

and let us furthermore rewrite the Wigner matrix (3.1.22) in these variables,

$$w_{ij}(x'\ell_c, p'p_c) = \left(\frac{\ell_c}{2\pi\hbar}\right)^d \int \mathrm{d}y'\, \varrho_{ij}\left[\left(x' + \frac{y'}{2}\right)\ell_c, \left(x' - \frac{y'}{2}\right)\ell_c\right] e^{-\frac{i\ell_c p_c}{\hbar} p'\cdot y'}\,. \tag{3.2.4}$$

Defining now the dimensionless quantities

$$\mathring{\varrho}_{ij}(x', y') := (\ell_c)^d \varrho_{ij}(x'\ell_c, y'\ell_c)\,, \qquad \mathring{w}_{ij}(x', y') := (\ell_c p_c)^d w_{ij}(x'\ell_c, p'p_c)\,, \tag{3.2.5}$$

and furthermore using the definition of the semiclassical parameter (3.2.2), from Eq. (3.2.4) one obtains the scaled version of the Wigner matrix (3.1.22) and its inverse:

$$\mathring{w}_{ij}(x',p') = \frac{1}{(2\pi\varepsilon)^d}\int \mathrm{d}y'\, \mathring{\varrho}_{ij}\left(x'+\frac{y'}{2}, x'-\frac{y'}{2}\right) e^{-\frac{i}{\varepsilon}p'\cdot y'}\,, \tag{3.2.6}$$

$$\mathring{\varrho}_{ij}(x',y') = \int \mathrm{d}p'\, \mathring{w}_{ij}\left(\frac{x'+y'}{2}, p'\right) e^{\frac{i}{\varepsilon}(x'-y')\cdot p'}\,. \tag{3.2.7}$$

For a general operator kernel $\mathcal{A}_{ij}$ with characteristic magnitude A_c (in appropriate physical units) one defines its dimensionless form and the corresponding Wigner representation as

$$\mathring{\mathcal{A}}_{ij}(x',y') := \frac{(\ell_c)^d}{A_c}\mathcal{A}_{ij}(x'\ell_c, y'\ell_c)\,, \qquad \mathring{a}_{ij}(x',p') := \frac{1}{A_c}a_{ij}(x'\ell_c, p'p_c)\,. \tag{3.2.8}$$

Then, inserting Eqs. (3.2.3) and (3.2.2) into Eqs. (3.1.21) and (3.1.24) results in

$$\mathring{a}_{ij}(x',p') = \int \mathrm{d}y'\, \mathring{\mathcal{A}}_{ij}\left(x'+\frac{y'}{2}, x'-\frac{y'}{2}\right) e^{-\frac{i}{\varepsilon}p'\cdot y'}\,, \tag{3.2.9}$$

$$\mathring{\mathcal{A}}_{ij}(x',y') = \frac{1}{(2\pi\varepsilon)^d}\int \mathrm{d}p'\, \mathring{a}_{ij}\left(\frac{x'+y'}{2}, p'\right) e^{\frac{i}{\varepsilon}(x'-y')\cdot p'}\,. \tag{3.2.10}$$

Finally, with the scaled version of the operator A in $(L^2)^n$ and the scaled version of the state vector $\phi \in (L^2)^n$, given by

$$\mathring{A} := \frac{A}{A_c}\,, \qquad \mathring{\phi}(x') := (\ell_c)^{d/2}\phi(x'\ell_c)\,, \qquad \tilde{\mathring{\phi}}(p') := (p_c)^{d/2}\tilde{\phi}(p'p_c)\,, \tag{3.2.11}$$

the Weyl quantization (3.1.38) transforms into

$$[\mathring{A}\mathring{\phi}]_i(x') = \frac{1}{(2\pi\varepsilon)^d}\sum_{j=1}^{n}\int \mathrm{d}y' \int \mathrm{d}p'\, \mathring{a}_{ij}\left(\frac{x'+y'}{2}, p'\right) \mathring{\phi}_j(y') e^{\frac{i}{\varepsilon}(x'-y')\cdot p'}. \tag{3.2.12}$$

It should be noted that from Eq. (3.2.11) one obtains that the Fourier transform and its inverse (3.1.14) can be written as

$$\begin{aligned} \tilde{\mathring{\phi}}_i(p') &= \frac{1}{(2\pi\varepsilon)^d}\int \mathrm{d}x'\, \mathring{\phi}_i(x') e^{-\frac{i}{\varepsilon}x'\cdot p'}\,, \\ \mathring{\phi}_i(x') &= \int \mathrm{d}p'\, \tilde{\mathring{\phi}}_i(p') e^{\frac{i}{\varepsilon}x'\cdot p'}\,. \end{aligned} \tag{3.2.13}$$

Putting the factor $1/(2\pi\varepsilon)^d$ in front of the Fourier transform, and not its inverse, has an important consequence: the Fourier transform of wavefunctions that are strongly varying in position space, i.e. $\mathring{\phi}(x') = \mathring{\phi}(x'/\varepsilon)$, is scale invariant, which means that

it does not depend on the magnitude of the parameter ε.

In summary, after easing the notation[1], we have obtained here the following scaled equations:

- Wigner transforms:

$$w_{ij}(x,\xi) = \frac{1}{(2\pi)^d}\int \mathrm{d}y\, \varrho_{ij}\left(x+\frac{\varepsilon y}{2}, x-\frac{\varepsilon y}{2}\right) e^{-iy\cdot\xi}\,, \tag{3.2.14a}$$

$$a_{ij}(x,\xi) = \varepsilon^d \int \mathrm{d}y\, \mathcal{A}_{ij}\left(x+\frac{\varepsilon y}{2}, x-\frac{\varepsilon y}{2}\right) e^{-iy\cdot\xi}\,. \tag{3.2.14b}$$

- Inverse Wigner transforms:

$$\varrho_{ij}(x,y) = \varepsilon^d \int \mathrm{d}\xi\, w_{ij}\left(\frac{x+y}{2}, \varepsilon\xi\right) e^{i(x-y)\cdot\xi}\,, \tag{3.2.15a}$$

$$\mathcal{A}_{ij}(x,y) = \frac{1}{(2\pi)^d}\int \mathrm{d}\xi\, a_{ij}\left(\frac{x+y}{2}, \varepsilon\xi\right) e^{i(x-y)\cdot\xi}\,. \tag{3.2.15b}$$

- Weyl quantizations:

$$[\rho\phi]_i(x) = \varepsilon^d \sum_{j=1}^{n}\int \mathrm{d}y \int \mathrm{d}\xi\, w_{ij}\left(\frac{x+y}{2}, \varepsilon\xi\right)\phi_j(y) e^{i(x-y)\cdot\xi}\,, \tag{3.2.16a}$$

$$[A\phi]_i(x) = \frac{1}{(2\pi)^d}\sum_{j=1}^{n}\int \mathrm{d}y \int \mathrm{d}\xi\, a_{ij}\left(\frac{x+y}{2}, \varepsilon\xi\right)\phi_j(y) e^{i(x-y)\cdot\xi}\,. \tag{3.2.16b}$$

3.3 The Moyal product

It is the aim of this section to derive a power series expression in the semiclassical parameter ε for the product of an arbitrary operator A in $(L^2)^n$ with the state operator ρ in $(L^2)^n$ in the Wigner picture. This semiclassical expansion is commonly referred to as the Moyal product [2]. It will serve as the basis for performing the semiclassical limit of the von Neumann equation, as demonstrated in the next section. At this point we assume that $a \in C^\infty(\mathbb{R}^d_x \times \mathbb{R}^d_\xi)$, where a denotes the symbol of A in the phase space, while for the Wigner matrix we assume $w \in H^0(\mathbb{R}^d_x \times \mathbb{R}^d_\xi)$. Let us investigate the action of the operator $C = A\rho$ on a state $\phi \in (L^2)^n$. For this we use Eq. (3.2.16) to obtain

$$\psi_j(y) = [\rho\phi]_j(y) = \varepsilon^d \sum_{k=1}^{n}\int \mathrm{d}z \int \mathrm{d}\eta\, w_{jk}\left(\frac{y+z}{2}, \varepsilon\eta\right)\phi_k(z) e^{i(y-z)\cdot\eta}\,, \tag{3.3.1}$$

[1] We omit the primes and the circles and moreover write greek letters for scaled momenta.

and furthermore to write

$$[A\psi]_i(x) = \frac{1}{(2\pi)^d}\sum_{j=1}^{n}\int \mathrm{d}y \int \mathrm{d}\xi\, a_{ij}\left(\frac{x+y}{2}, \varepsilon\xi\right)\psi_j(y)e^{i(x-y)\cdot\xi}\,. \tag{3.3.2}$$

Inserting Eq. (3.3.1) into Eq. (3.3.2) results in

$$[A\rho\phi]_i(x) = \left(\frac{\varepsilon}{2\pi}\right)^d \sum_{j,k=1}^{n}\int \mathrm{d}y\mathrm{d}z \int \mathrm{d}\eta\mathrm{d}\xi\, a_{ij}\left(\frac{x+y}{2}, \varepsilon\xi\right) w_{jk}\left(\frac{y+z}{2}, \varepsilon\eta\right)\phi_k(z) \times e^{i(y-z)\cdot\eta}e^{i(x-y)\cdot\xi}\,. \tag{3.3.3}$$

which yields the following components of the integral kernel $\mathcal{C}$ of the operator C:

$$\mathcal{C}_{ik}(x,z) = \left(\frac{\varepsilon}{2\pi}\right)^d \sum_{j=1}^{n}\int \mathrm{d}y \int \mathrm{d}\eta\mathrm{d}\xi\, a_{ij}\left(\frac{x+y}{2}, \varepsilon\xi\right) w_{jk}\left(\frac{y+z}{2}, \varepsilon\eta\right) \times e^{i(y-z)\cdot\eta}e^{i(x-y)\cdot\xi}\,. \tag{3.3.4}$$

The symbol c_{ij} in the phase space corresponding to (3.3.4) is computed via the transformation for the density matrix (3.2.14a) (this choice will become transparent in the next section):

$$\begin{aligned} c_{ik}(x,\theta) &= \frac{1}{(2\pi)^d}\int \mathrm{d}z\, \mathcal{C}_{ik}\left(x+\frac{\varepsilon z}{2}, x-\frac{\varepsilon z}{2}\right)e^{-iz\cdot\theta} \qquad (3.3.5)\\ &= \frac{\varepsilon^d}{(2\pi)^{2d}}\sum_{j=1}^{n}\int \mathrm{d}y\mathrm{d}z \int \mathrm{d}\eta\mathrm{d}\xi\, a_{ij}\left(\frac{x+y}{2}+\frac{\varepsilon z}{4}, \varepsilon\xi\right) w_{jk}\left(\frac{x+y}{2}-\frac{\varepsilon z}{4}, \varepsilon\eta\right) \\ &\qquad \times e^{-i(x-y-\frac{\varepsilon z}{2})\cdot\eta}e^{i(x-y+\frac{\varepsilon z}{2})\cdot\xi}e^{-iz\cdot\theta}\,. \end{aligned}$$

Let us now introduce the dummy variable v via

$$c_{ik}(x,\theta) = \frac{\varepsilon^d}{(2\pi)^{2d}}\sum_{j=1}^{n}\int \mathrm{d}v\mathrm{d}y\mathrm{d}z \int \mathrm{d}\eta\mathrm{d}\xi\, a_{ij}\left(\frac{v+y}{2}+\frac{\varepsilon z}{4}, \varepsilon\xi\right) w_{jk}\left(\frac{v+y}{2}-\frac{\varepsilon z}{4}, \varepsilon\eta\right) \times \delta(x-v)e^{-i(v-y-\frac{\varepsilon z}{2})\cdot\eta}e^{i(v-y+\frac{\varepsilon z}{2})\cdot\xi}e^{-iz\cdot\theta}\,. \tag{3.3.6}$$

and let us furthermore perform the variable transformations

$$\begin{aligned} a &= \frac{v+y}{2} & v &= a+\frac{b}{2}\\ b &= v-y & y &= a-\frac{b}{2}\,, \end{aligned} \tag{3.3.7}$$

which results in

$$c_{ik}(x,\theta) = \frac{\varepsilon^d}{(2\pi)^{2d}} \sum_{j=1}^{n} \int \mathrm{d}a\mathrm{d}b\mathrm{d}z \int \mathrm{d}\eta\mathrm{d}\xi\, a_{ij}\left(a + \frac{\varepsilon z}{4}, \varepsilon\xi\right) w_{jk}\left(a - \frac{\varepsilon z}{4}, \varepsilon\eta\right) \times \delta\left(x - a - \frac{b}{2}\right) e^{-i(b-\frac{\varepsilon z}{2})\cdot\eta} e^{i(b+\frac{\varepsilon z}{2})\cdot\xi} e^{-iz\cdot\theta}\,. \quad (3.3.8)$$

As a next step we write w_{jk} in its Fourier representation,

$$w_{jk}\left(a - \frac{\varepsilon z}{4}, \varepsilon\eta\right) = \frac{1}{(2\pi)^d} \int \mathrm{d}r \int \mathrm{d}\chi\, \tilde{w}_{jk}(\chi, r) e^{-i\left(a-\frac{\varepsilon z}{4}\right)\cdot\chi} e^{-i\varepsilon\eta\cdot r}\,, \quad (3.3.9)$$

where $\tilde{w}_{jk}$ denotes the Fourier transform of w_{jk}. Inserting Eq. (3.3.9) into Eq. (3.3.8) leads to

$$c_{ik}(x,\theta) = \frac{\varepsilon^d}{(2\pi)^{3d}} \sum_{j=1}^{n} \int \mathrm{d}a\mathrm{d}b\mathrm{d}r\mathrm{d}z \int \mathrm{d}\eta\mathrm{d}\xi\mathrm{d}\chi\, a_{ij}\left(a + \frac{\varepsilon z}{4}, \varepsilon\xi\right) \tilde{w}_{jk}(\chi, r) \times \delta\left(x - a - \frac{b}{2}\right) e^{-i\left(a-\frac{\varepsilon z}{4}\right)\cdot\chi} e^{-i(b-\frac{\varepsilon z}{2}+\varepsilon r)\cdot\eta} e^{i(b+\frac{\varepsilon z}{2})\cdot\xi} e^{-iz\cdot\theta}\,. \quad (3.3.10)$$

In what follows we shall use frequently the relations (3.1.4). Performing the integral with respect to η and subsequently with respect to b leads to the transformation $b \mapsto \varepsilon z/2 - \varepsilon r$, which results in

$$c_{ik}(x,\theta) = \frac{\varepsilon^d}{(2\pi)^{2d}} \sum_{j=1}^{n} \int \mathrm{d}a\mathrm{d}r\mathrm{d}z \int \mathrm{d}\xi\mathrm{d}\chi\, a_{ij}\left(a + \frac{\varepsilon z}{4}, \varepsilon\xi\right) \tilde{w}_{jk}(\chi, r) \times \delta\left(x - a - \frac{\varepsilon z}{4} + \frac{\varepsilon r}{2}\right) e^{-i\left(a-\frac{\varepsilon z}{4}\right)\cdot\chi} e^{i(\varepsilon z-\varepsilon r)\cdot\xi} e^{-iz\cdot\theta}\,. \quad (3.3.11)$$

From the integral with respect to a we obtain $a \mapsto x - \varepsilon z/4 + \varepsilon r/2$,

$$c_{ik}(x,\theta) = \frac{\varepsilon^d}{(2\pi)^{2d}} \sum_{j=1}^{n} \int \mathrm{d}r\mathrm{d}z \int \mathrm{d}\xi\mathrm{d}\chi\, a_{ij}\left(x + \frac{\varepsilon r}{2}, \varepsilon\xi\right) \tilde{w}_{jk}(\chi, r) \times e^{-i\left(x-\frac{\varepsilon z}{2}+\frac{\varepsilon r}{2}\right)\cdot\chi} e^{i(\varepsilon z-\varepsilon r)\cdot\xi} e^{-iz\cdot\theta}\,. \quad (3.3.12)$$

The integral with respect to z now yields a delta distribution,

$$c_{ik}(x,\theta) = \frac{\varepsilon^d}{(2\pi)^{d}} \sum_{j=1}^{n} \int \mathrm{d}r \int \mathrm{d}\xi\mathrm{d}\chi\, a_{ij}\left(x + \frac{\varepsilon r}{2}, \varepsilon\xi\right) \tilde{w}_{jk}(\chi, r) \times \delta\left(\frac{\varepsilon\chi}{2} + \varepsilon\xi - \theta\right) e^{-i\left(x+\frac{\varepsilon r}{2}\right)\cdot\chi} e^{-i\varepsilon r\cdot\xi}\,. \quad (3.3.13)$$

The variable transformation $\xi' = \varepsilon\xi$ and the subsequent integration with respect to ξ' results in the transformation $\varepsilon\xi \mapsto \theta - \varepsilon\chi/2$,

$$c_{ik}(x,\theta) = \frac{1}{(2\pi)^d} \sum_{j=1}^{n} \int \mathrm{d}r \int \mathrm{d}\chi \, a_{ij}\left(x + \frac{\varepsilon r}{2}, \theta - \frac{\varepsilon\chi}{2}\right) \tilde{w}_{jk}(\chi, r) e^{-ix\cdot\chi} e^{-i\theta\cdot r} \,. \tag{3.3.14}$$

Equation (3.3.14) is referred to as the Moyal product. From here on we shall use the following notation for the scaled Wigner transform (3.2.14a) and its inverse (3.2.15a):

$$\mathcal{W}_\varepsilon(\varrho) = w^{(\varepsilon)} \,, \qquad \mathcal{W}_\varepsilon^{-1}(w^{(\varepsilon)}) = \varrho \,. \tag{3.3.15}$$

Furthermore, the scaled version of the mapping $\mathrm{Op}(w) = \rho$ defined in Eq. (3.2.16a) (or in (3.1.11), respectively) and its inverse will be denoted by

$$\mathrm{Op}_\varepsilon(w^{(\varepsilon)}) = \rho \,, \qquad \mathrm{Op}_\varepsilon^{-1}(\rho) = w^{(\varepsilon)} \,. \tag{3.3.16}$$

Here, $w^{(\varepsilon)}$ denotes the scaled Wigner matrix with components $(w_{ij}^{(\varepsilon)})_{i,j=1}^n$. The superscript ε indicates that the semiclassical parameter is intended to go to zero. Hence, we introduce the following notation for the Moyal product (3.3.14):

$$\mathrm{Op}_\varepsilon^{-1}(A\rho) = a \#_\varepsilon w^{(\varepsilon)} \,. \tag{3.3.17}$$

Inserting into Eq. (3.3.14) the Taylor series expansion

$$a_{ij}\left(x + \frac{\varepsilon r}{2}, \theta - \frac{\varepsilon\chi}{2}\right) = a_{ij}(x,\theta) + \frac{\varepsilon}{2}\left[\nabla_x a_{ij}(x,\theta)\cdot r - \nabla_\theta a_{ij}(x,\theta)\cdot\chi\right] + \mathcal{O}(\varepsilon^2)$$

and subsequently using the Fourier representation (3.3.9) results in the semiclassical expansion of the Moyal product:

$$a \#_\varepsilon w^{(\varepsilon)} = a w^{(\varepsilon)} + i\frac{\varepsilon}{2}\{a, w^{(\varepsilon)}\}_{(x,\theta)} + \mathcal{O}(\varepsilon^2) \,, \tag{3.3.18}$$

where $\{\cdot,\cdot\}_{(x,\theta)}$ denotes the Poisson bracket in the phase space variables defined by

$$\{a, w^{(\varepsilon)}\}_{(x,\theta)} := \nabla_x a \cdot \nabla_\theta w^{(\varepsilon)} - \nabla_\theta a \cdot \nabla_x w^{(\varepsilon)} \,. \tag{3.3.19}$$

It can be readily verified that for reversed order of the operators A and ρ, one obtains

$$[\mathrm{Op}_\varepsilon^{-1}(\rho A)]_{ik} = \frac{1}{(2\pi)^d} \sum_{j=1}^{n} \int \mathrm{d}r \int \mathrm{d}\chi \, a_{ij}\left(x - \frac{\varepsilon r}{2}, \theta + \frac{\varepsilon\chi}{2}\right) \tilde{w}_{jk}(\chi, r) e^{-ix\cdot\chi} e^{-i\theta\cdot r} \,. \tag{3.3.20}$$

Therefore, the semiclassical expansion of the Moyal product (3.3.20) reads

$$w^{(\varepsilon)}\#_\varepsilon a = w^{(\varepsilon)}a + i\frac{\varepsilon}{2}\{w^{(\varepsilon)}, a\}_{(x,\theta)} + \mathcal{O}(\varepsilon^2)\,. \tag{3.3.21}$$

3.4 The matrix-Vlasov equation

The scaled version of the Wigner-Weyl calculus in quantum mechanics introduced in the course of the previous two sections allows for a transparent study of the transition from quantum dynamics to classical dynamics in the phase space. Here, in particular, we shall present a formal derivation of the matrix-Vlasov equation, which is an extension of the usual Vlasov equation to incorporate the spin-degree of freedom of particles on the mesoscopic level. This equation serves as a basis for the work presented in Chapter 4. Let $h : \mathbb{R}^d_x \times \mathbb{R}^d_\xi \to \mathbb{C}^{d\times d}$ stand for the phase space symbol of the Hamilton operator H in $(L^2)^n$ of a single particle with spin. On the quantum level and in the semiclassical scaling, the dynamics of the state operator ρ in $(L^2)^n$ describing this particle are given by by the scaled von Neumann equation,

$$\partial_t\rho + \frac{i}{\varepsilon}[H, \rho] = 0\,, \tag{3.4.1}$$

where $[H, \rho] = H\rho - \rho H$ denotes the commutator. By applying the operation $\mathrm{Op}^{-1}_\varepsilon(\cdot)$ defined in Eq. (3.3.16) we can map Eq. (3.4.1) into the phase space,

$$\partial_t w^{(\varepsilon)} + \frac{i}{\varepsilon}(h\#_\varepsilon w^{(\varepsilon)} - w^{(\varepsilon)}\#_\varepsilon h) = 0\,. \tag{3.4.2}$$

Using now the semiclassical expansion of the Moyal product, Eqs. (3.3.18) and (3.3.21), results in

$$\partial_t w^{(\varepsilon)} + \frac{i}{\varepsilon}[h, w^{(\varepsilon)}] - \frac{1}{2}\{h, w^{(\varepsilon)}\}_{(x,\xi)} + \frac{1}{2}\{w^{(\varepsilon)}, h\}_{(x,\xi)} + \mathcal{O}(\varepsilon) = 0\,. \tag{3.4.3}$$

For $\varepsilon \ll 1$ the commutator $[h, w^{(\varepsilon)}]$ is the dominant term. Hence in the limit $\varepsilon \to 0$ one obtains $[h, w^{(\varepsilon)}] = 0\ \forall t$ which leads to a system of n coupled scalar equations for the n possible spin states of the particle. A rigorous analysis of this case as well as of the following matrix case for spin-1/2 particles can be found in [6, 10]. The matrix form of $w^{(\varepsilon)}$ is retained in the limit $\varepsilon \to 0$ under the premiss that the symbol h scales as follows:

$$h(x, \xi) = h_c(x, \xi)\mathbb{1}_n + \varepsilon h_s(x, \xi) \qquad \forall (x, \xi) \in \mathbb{R}^d_x \times \mathbb{R}^d_\xi\,. \tag{3.4.4}$$

Here, $\mathbb{1}_n$ denotes the identity operator in $\mathbb{C}^{d\times d}$, $h_c(x,\xi) \in \mathbb{R}\,\forall(x,\xi)$ stands for the scalar part (or charge part) of the symbol h and $h_s := h - h_c\mathbb{1}_n$ denotes the spin part. Inserting Eq. (3.4.4) into Eq. (3.4.3) and taking the limit $\varepsilon \to 0$ results in

$$\partial_t w^{(0)} + i[h_s, w^{(0)}] - \{h_c, w^{(0)}\}_{(x,\xi)} = 0\,, \tag{3.4.5}$$

which is referred to as the matrix-Vlasov equation. The latter is a mesoscopic transport equation which allows for quantum superposition of different spin states at each point (x,ξ) in the phase space. This "spin-coherence" is reflected by non-vanishing off-diagonal elements in $w^{(0)}$. The notion of spin-coherence is described in detail for spin-1/2 electrons in the sections 4.2 and 4.3 of chapter 4.

References

[1] E. Wigner. On the quantum correction for thermodynamic equilibrium. *Physical Review*, 40(5):749, 1932.

[2] C. Zachos, D. Fairlie, and T. Curtright. *Quantum mechanics in phase space: an overview with selected papers*, volume 34. World Scientific Pub Co Inc, 2005.

[3] P.L. Lions and T. Paul. Sur les mesures de wigner. *Rev. Mat. Iberoamericana*, 9(3):553–618, 1993.

[4] PA Markowich, NJ Mauser, and F. Poupaud. A wigner-function approach to (semi) classical limits: Electrons in a periodic potential. *Journal of Mathematical Physics*, 35:1066, 1994.

[5] P. Bechouche, F. Poupaud, and J. Soler. Quantum transport and boltzmann operators. *Journal of statistical physics*, 122(3):417–436, 2006.

[6] P. Gérard, P.A. Markowich, N.J. Mauser, and F. Poupaud. Homogenization limits and wigner transforms. *Communications on Pure and Applied Mathematics*, 50(4):323–379, 1997.

[7] R. El Hajj. Diffusion models for spin transport derived from the spinor boltzmann equation. *Comm. Math. Sci. (to appear)*.

[8] L. Hörmander. The analysis of linear partial differential operators, vol. 3. pseudo-differential operators, volume 274 of grundlehren der mathematischen wissenschaften, 1985.

[9] L.E. Ballentine. *Quantum Mechanics - A Modern Development*. World Scientific Publishing, 1998.

[10] R. El Hajj. *Etude mathématique et numérique de modèles de transport: application à la spintronique*. PhD thesis, Institut de Mathématiques de Toulouse (IMT), Université Paul Sabatier, 2008.

Chapter 4

Diffusion Limit of a Generalized Matrix-Boltzmann Equation for Spin-Polarized Transport

S. Possanner and *C. Negulescu*,
published in *Kinetic and Related Models*, Vol. 4, p. 1159, **2011**

Abstract. The aim of the present paper is the mathematical study of a linear Boltzmann equation with different matrix collision operators, modelling the spin-polarized, semi-classical electron transport in non-homogeneous ferromagnetic structures. In the collision kernel, the scattering rate is generalized to a hermitian, positive-definite 2×2 matrix whose eigenvalues stand for the different scattering rates of, for example, spin-up and spin-down electrons in spintronic applications. We identify four possible structures of linear matrix collision operators that yield existence and uniqueness of a weak solution of the Boltzmann equation for a general Hamilton function. We are able to prove positive-(semi)definiteness of a solution for an operator that features an anti-symmetric structure of the gain respectively the loss term with respect to the occurring matrix products. Furthermore, in order to obtain matrix drift-diffusion equations, we perform the diffusion limit with one of the symmetric operators assuming parabolic spin bands with uniform band gap and in the case that the precession frequency of the spin distribution vector around the exchange field of the Hamiltonian scales with order ε^2. Numerical simulations of the here obtained macroscopic model were carried out in non-magnetic/ferromagnetic multilayer structures and for a magnetic Bloch domain wall. The results show that our model can be used to improve the understanding of spin-polarized transport in spintronics applications.

4.1 Introduction

The coupling between the spin- and the charge degree of freedom of an electron system is a growing research topic in physics and mathematics. What is simply called 'spintronics' (spin-electronics) has become a vast field with many promising applications and plenty of challenging problems to be solved. The field includes, for example, quantum computing [1, 2] (qubits), spin-dependent transport in ferromagnets [3] (giant magneto-resistance effect or GMR effect), semiconductor spintronics [4, 5] (spin field-effect transistors, magnetic resonant tunneling diodes) and spin-transfer torques in ferromagnets [6, 7, 8] (current-induced magnetic switching and domain wall motion). The benefits of spintronics lie in the fact that the magnetic state of a system can be changed by manipulating charges with electric fields, which can be handled rather easily and more precisely as compared to magnetic fields. In most of the spintronic applications, spin-polarized electron transport in solids plays a crucial role. By spin-polarized transport we mean that, in addition to the charge distribution of an electron system, it is necessary to keep track of its spin distribution to obtain a correct description. Magnetic impurities, a ferromagnetic environment, strong spin-orbit coupling [9, 10] or an applied magnetic field often require the spin-polarized treatment of transport in these systems.

The spin of an electron represents a two-state quantum system [11]. This means that once a direction is chosen in real space, the electron spin can be determined to be either parallel (spin-up) or anti-parallel (spin-down) to that direction. Prior to the measurement, a general spin state (or spin-coherent state) is a quantum superposition of the spin-up and spin-down basis states. The density matrix of a spin-coherent state is a hermitian 2×2 matrix, where the spin-coherence is represented by non-vanishing off-diagonal elements. From a mathematical point of view, spin systems resemble electron-hole systems in Graphene [12]. For such systems, there exist various matrix transport models on the microscopic (von Neumann and Wigner equation), the kinetic (Boltzmann equation [13, 14, 15, 16]) and the macroscopic level (drift-diffusion and fluiddynamic equations [17, 18, 16, 19], quantum drift-diffusion and quantum fluiddynamic equations [20, 21]). For the purpose of engineering spintronic devices, macroscopic models are very appealing. On the one hand, they enable efficient numerical simulations on the desired length scale ($10^1 - 10^3\,\mathrm{nm}$) and, on the other hand, they incorporate the scattering of electrons from phonons (non-zero temperature) and impurities (material imperfections). However, spin-coherent drift-diffusion models occurring in literature are still mostly heuristic. Recently, El Hajj and Ben Abdallah [22] introduced a spin-coherent collision operator in the linear

BGK approximation to obtain a matrix Boltzmann equation. They performed rigorous diffusion limits in various scalings to derive a number of matrix drift-diffusion models. The rigorous derivation of a spin-coherent collision operator from the microscopic scale is still an open problem.

In this work we address, at first, the kinetic level of spin-coherent transport. Our goal is to set up a matrix Boltzmann equation that incorporates spin-dependent scattering rates, or more precisely, that features a collision kernel with matrix-valued transition probabilities from momentum $\hbar k$ to $\hbar k'$. Such a kinetic equation can be viewed as a generalization of the model in [22] to spin-dependent mean-free paths. In [22], the scattering rates were scalar quantities (which yield one mean free path for both spin species) whereas in our case they are fully occupied hermitian, positive-definite 2×2 matrices. The eigenvalues of these matrices stand for the different scattering rates of spin-up and spin-down electrons (yielding two distinct respective mean free paths). The observation of spin-dependent electron resistances in ferromagnets [23, 24] was crucial for triggering the research on spintronics, therefore the generalization of scalar scattering rates to matrix-valued scattering rates is a logical step. The main problem, as compared to the scalar case, is to deal with the matrix products that will occur in the newly defined collision operator.

The effects of spin-dependent mean free paths on non-coherent spin-polarized transport (two-component models) have been studied comprehensively [25, 26, 27, 28, 29, 30]. However, to our knowledge, there exist no works on the consequences of spin-dependent scattering for spin-coherent electron systems. Our approach is to add, on the right-hand-side of the spin-coherent Vlasov equation, the four most simple types of linear matrix collision operators which preserve the hermiticity of the electron distribution matrix. We then apply the method of characteristics and a fixed point argument to check existence and uniqueness of a weak solution of the respective matrix Boltzmann equation. Additionally, using the maximum principle, we check the positive-(semi)definiteness of the solution. We identify one collision operator that satisfies the maximum principle. This operator features an anti-symmetric structure with respect to the matrix products in the gain and the loss term, respectively. The anti-symmetric collision operator is mass- but not spin-conserving. In the subsequent sections of the paper, we focus on a mass- and spin-conserving collision operator, which has a symmetric structure of the gain respectively the loss term. In contrast to the anti-symmetric operator, the symmetric collision operator describes only spin-conserving momentum scattering, no spin-flip processes. The spin-flip scattering is then described by a second collision operator. This strat-

egy permits to treat spin-conserving respectively spin-flip scattering on different timescales. However, the verification of the maximum principle for this two-operator approach remains an open problem.

In the second part of this paper we perform the diffusion limit in a scaled form of the matrix Boltzmann equation, using standard techniques [31] known from the scalar case. The necessary physical assumptions for this step are Boltzmann statistics and local thermal equilibrium (detailed balance) in each spin band. Additionally, we make the strong assumption of parabolic spin bands with uniform band gap, a model that is known in spintronic literature as the Stoner model [32, 33]. Relaxing this assumption should clearly be the topic of following works. However, even in the simple setting of the Stoner model, we obtain, in the macroscopic limit, a matrix drift-diffusion model that features a coupling between the charge- and the spin degree of freedom. The coupling we get is linear in the polarization p of the scattering rates ($0 \leq p < 1$).

The third part of this work contains some numerical studies of the derived spin-coherent drift-diffusion model in one-dimensional multilayer structures and for strongly varying magnetization on the scale of several nanometers (e.g. a magnetic domain wall). We use a standard Crank-Nicolson finite difference scheme to solve the four coupled drift-diffusion equations on a uniform grid. The results show that our model provides a new means for studying spin-polarized transport in arbitrary magnetic structures (e.g. non-collinear multilayers or strongly varying magnetization).

4.2 Some notations and Lemmas

As a help for the better understanding of this work, we start by introducing some relevant notations.

Definition 4.2.1 (**Pauli matrices**). *By $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ we denote the triple of the three Pauli matrices,*

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{4.2.1}$$

To be consistent, the 2×2 unit matrix is denoted by

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{4.2.2}$$

The Pauli matrices satisfy the following properties, for $k, l, m \in \{1, 2, 3\}$,

$$[\sigma_k, \sigma_l] = 2i \sum_m \varepsilon_{klm} \sigma_m \tag{4.2.3}$$

$$[\sigma_k, \sigma_l]_+ = 2\delta_{kl} \sigma_0 \tag{4.2.4}$$

$$\sigma_k \sigma_l = \delta_{kl} \sigma_0 + i \sum_m \varepsilon_{klm} \sigma_m. \tag{4.2.5}$$

Here, $[\sigma_k, \sigma_l] = \sigma_k \sigma_l - \sigma_l \sigma_k$ stands for the commutator of σ_k and σ_l, $[\sigma_k, \sigma_l]_+ = \sigma_k \sigma_l + \sigma_l \sigma_k$ denotes the anti-commutator, δ_{km} the Kronecker delta, ε_{klm} the Levi-Civita symbol, defined by

$$\varepsilon_{klm} = \begin{vmatrix} \delta_{k1} & \delta_{k2} & \delta_{k3} \\ \delta_{l1} & \delta_{l2} & \delta_{l3} \\ \delta_{m1} & \delta_{m2} & \delta_{m3} \end{vmatrix}. \tag{4.2.6}$$

From (4.2.3) one deduces, for $\vec{a}, \vec{b} \in \mathbb{R}^3$,

$$(\vec{a} \cdot \vec{\sigma})(\vec{b} \cdot \vec{\sigma}) = (\vec{a} \cdot \vec{b})\sigma_0 + i(\vec{a} \times \vec{b}) \cdot \vec{\sigma} \tag{4.2.7}$$

$$[\vec{a} \cdot \vec{\sigma}, \vec{b} \cdot \vec{\sigma}] = 2i(\vec{a} \times \vec{b}) \cdot \vec{\sigma}. \tag{4.2.8}$$

Definition 4.2.2 (**Matrix spaces**). *Let $\mathcal{H}_2(\mathbb{C})$ denote the vector space of hermitian 2×2 matrices. Associated with the Frobenius scalar product $\langle \cdot, \cdot \rangle_2$ and the corresponding norm $|| \cdot ||_2$,*

$$\langle A, B \rangle_2 := \operatorname{tr}(AB), \qquad ||A||_2 := \sqrt{\operatorname{tr}(A^2)} \qquad \forall A, B \in \mathcal{H}_2(\mathbb{C}), \tag{4.2.9}$$

the vector space $\mathcal{H}_2(\mathbb{C})$ becomes a Hilbert space, where $A = (a_{ij})$ and $B = (b_{ij})$ with $i, j \in \{1, 2\}$ and

$$\operatorname{tr}(AB) = \sum_{ij} a_{ij} b_{ji} \tag{4.2.10}$$

denotes the trace of AB. By $\mathcal{H}_2^{0,+}(\mathbb{C})$ we denote the subspace of $\mathcal{H}_2(\mathbb{C})$ containing positive semi-definite matrices and by $\mathcal{H}_2^{+}(\mathbb{C})$ we identify the subspace of positive definite matrices.

The space $\mathcal{H}_2(\mathbb{C})$ is spanned by the four matrices σ_0, σ_1, σ_2 and σ_3. In this basis, the coefficients of a matrix $A \in \mathcal{H}_2(\mathbb{C})$ are denoted by $a_0 \in \mathbb{R}$, $\vec{a} = (a_1, a_2, a_3) \in \mathbb{R}^3$. We have the following isomorphism between $\mathcal{H}_2(\mathbb{C})$ and $\mathbb{R}^4$,

$$A = a_0 \sigma_0 + \vec{a} \cdot \vec{\sigma} \ \in \mathcal{H}_2(\mathbb{C}) \quad \Longleftrightarrow \quad \tilde{A} = (a_0, \vec{a}) \ \in \mathbb{R}^4. \tag{4.2.11}$$

The coefficients of A in this Pauli basis are computed as follows,

$$a_0 = \frac{1}{2}\mathrm{tr}\,(A) \qquad ; \qquad \vec{a} = \frac{1}{2}\mathrm{tr}\,(\vec{\sigma}A)\,. \tag{4.2.12}$$

In the following, we call the space $\mathbb{R}^3$ of the coefficient vectors $\vec{a}$ of a matrix $A \in \mathcal{H}_2(\mathbb{C})$ in the Pauli basis the *spin space*. Only vectors in spin space are written with the overlined arrow symbol. The eigenvalues of $A \in \mathcal{H}_2(\mathbb{C})$ are given by $a_+ = a_0 + |\vec{a}|$ and $a_- = a_0 - |\vec{a}|$, respectively.

Lemma 4.2.3. *Let $A \in \mathcal{H}_2^+(\mathbb{C})$ with components $(a_0, \vec{a})$ in the Pauli basis, then*

$$A^{1/2} = \frac{1}{2}\left(\sqrt{a_0 + |\vec{a}|} + \sqrt{a_0 - |\vec{a}|}\right)\sigma_0 + \frac{1}{2}\left(\sqrt{a_0 + |\vec{a}|} - \sqrt{a_0 - |\vec{a}|}\right)\frac{\vec{a}}{|\vec{a}|}\cdot\vec{\sigma}. \tag{4.2.13}$$

Lemma 4.2.4. *Let $A, B \in \mathcal{H}_2(\mathbb{C})$. Then we have $AB + BA \in \mathcal{H}_2(\mathbb{C})$ and $ABA \in \mathcal{H}_2(\mathbb{C})$. Moreover, in the Pauli basis,*

$$\frac{1}{2}AB + \frac{1}{2}BA = (a_0 b_0 + \vec{a}\cdot\vec{b})\sigma_0 + (a_0\vec{b} + b_0\vec{a})\cdot\vec{\sigma}$$

and

$$\begin{aligned} A^{1/2}BA^{1/2} = {} & (a_0 b_0 + \vec{a}\cdot\vec{b})\sigma_0 + \\ & + \left[b_0\vec{a} + \left(a_0 - \sqrt{a_0^2 - |\vec{a}|^2}\right)\left(\vec{b}\cdot\frac{\vec{a}}{|\vec{a}|}\right)\frac{\vec{a}}{|\vec{a}|} + \vec{b}\sqrt{a_0^2 - |\vec{a}|^2}\right]\cdot\vec{\sigma}. \end{aligned}$$

Lemma 4.2.5. *For $A, B \in \mathcal{H}_2^{0,+}(\mathbb{C})$ we have $ABA \in \mathcal{H}_2^{0,+}(\mathbb{C})$. However, $AB + BA$ is not necessarily in $\mathcal{H}_2^{0,+}(\mathbb{C})$.*

Lemma 4.2.6 (**Trace properties**). *We have*

$$\mathrm{tr}\,(AB) \geq 0 \quad \forall A, B \in \mathcal{H}_2^{0,+}(\mathbb{C}) \tag{4.2.14}$$

$$\mathrm{tr}\left(A^2 + B^2\right) \geq 2\left|\mathrm{tr}\,(AB)\right| \quad \forall A, B \in \mathcal{H}_2(\mathbb{C}) \tag{4.2.15}$$

$$0 \leq \mathrm{tr}\left((AB)^2\right) \leq \mathrm{tr}\left(A^2B^2\right) \leq 2\mathrm{tr}\left(A^2\right)\mathrm{tr}\left(B^2\right) \quad \forall A \in \mathcal{H}_2(\mathbb{C}), B \in \mathcal{H}_2^{0,+}(\mathbb{C}). \tag{4.2.16}$$

Proof. Relation (4.2.15) follows from

$$\begin{aligned} 0 \leq \mathrm{tr}\left((A+B)^2\right) = \mathrm{tr}\left(A^2 + 2AB + B^2\right) \quad &\Longrightarrow \quad \mathrm{tr}\left(A^2 + B^2\right) \geq -2\mathrm{tr}\,(AB) \\ 0 \leq \mathrm{tr}\left((A-B)^2\right) = \mathrm{tr}\left(A^2 - 2AB + B^2\right) \quad &\Longrightarrow \quad \mathrm{tr}\left(A^2 + B^2\right) \geq 2\mathrm{tr}\,(AB)\,. \end{aligned}$$

To prove (4.2.16), let $A = (a_{ij})$ and $B = (b_{ij})$, then $\bar{a}_{ij}$ stands for the complex

conjugate of a_{ij}. Further, let θ be a unitary matrix, $\theta\bar{\theta}^t = \sigma_0$, such that $\bar{\theta}^t B\theta$ is diagonal. One has

$$\mathrm{tr}\left((AB)^2\right) = \mathrm{tr}\left((\theta\bar{\theta}^t A\theta\bar{\theta}^t B)^2\right) = \mathrm{tr}\left(\bar{\theta}^t A\theta\bar{\theta}^t B\theta\bar{\theta}^t A\theta\bar{\theta}^t B\theta\right) = \mathrm{tr}\left((A'B'_D)^2\right).$$

Here, $A' = \bar{\theta}^t A\theta = (a'_{ij})$ with $(a'_{ii}) \in \mathbb{R}$, and B'_D is a diagonal matrix with eigenvalues $(b'_{ii}) \in \mathbb{R}^+$. We obtain

$$\mathrm{tr}\left((AB)^2\right) = \mathrm{tr}\left((A'B'_D)^2\right) = \sum_{ik} a'_{ik}b'_{kk}a'_{ki}b'_{ii} = \sum_{ik} |a'_{ik}|^2 b'_{kk}b'_{ii} \geq 0, \tag{4.2.17}$$

which gives the first inequality in (4.2.16). The second inequality follows from

$$\mathrm{tr}\left((A'B'_D)^2\right) = \sum_{ik} |a'_{ik}|^2 b'_{kk}b'_{ii} \leq \sum_{ik} |a'_{ik}|^2 b'^2_{ii} = \mathrm{tr}\left(A'^2 B'^2_D\right) \tag{4.2.18}$$

The third inequality in (4.2.16) follows from the fact that A^2 and B^2 are positive hermitian matrices. □

4.3 Preliminaries

4.3.1 Spin-coherent semiclassical electrons

The state of a spin-coherent semi-classical electron system is characterized by the *distribution matrix* $F : \mathbb{R}^+ \times \mathbb{R}^d_x \times \mathbb{R}^d_k \to \mathcal{H}^{0,+}_2(\mathbb{C})$. Here, t denotes the time and x and k are the respective variables for position and momentum (more precisely, k stands for an electron's wave vector and $\hbar k$ for its crystal momentum). The phase space of the electron system is $\mathbb{R}^d_x \times \mathbb{R}^d_k = \mathbb{R}^{2d}$ where d is the number of space dimensions. In the Pauli basis the distribution matrix is written as $F = (\frac{1}{2}f_0, \vec{f})$. The coefficient $f_0(t, x, k) = \mathrm{tr}\,(F)$ is the scalar distribution function of the electrons ignoring the spin. The vector $\vec{f}(t, x, k) = \frac{1}{2}\mathrm{tr}\,(\vec{\sigma}F)$ represents the vector spin polarization of the electron system and $\hbar\vec{f}$ is the electron spin density at $(t, x, k) \in \mathbb{R}^+ \times \mathbb{R}^{2d}$. The two eigenvalues of F, denoted by $f_\pm = \frac{1}{2}f_0 \pm |\vec{f}|$, stand for the distribution functions of electrons with spin in the direction $+\vec{f}/|\vec{f}|$ and in the direction $-\vec{f}/|\vec{f}|$, respectively. These directions, determined by the distribution matrix itself, define the z-axis of a coordinate system in spin space in which F is diagonal. This coordinate frame defined by $\vec{f}/|\vec{f}|$ depends on $(t, x, k) \in \mathbb{R}^+ \times \mathbb{R}^{2d}$. Since it is our purpose to describe the spin-coherence of electrons with respect to a given field $\vec{\Omega}(t, x, k)$ in spin space, it is preferrable to work in a coordinate frame independent of $(t, x, k) \in \mathbb{R}^+ \times \mathbb{R}^{2d}$ and to keep track of the direction $\vec{f}/|\vec{f}|$ therein.

The energy density of the system in the state F is computed from the *Hamilton matrix* $H : \mathbb{R}^+ \times \mathbb{R}^{2d} \to \mathcal{H}_2(\mathbb{C})$. We write this matrix as $H = H_b + H_{so}$, where H_b is called the *band matrix* and H_{so} denotes the *spin-orbit matrix*, given by

$$H_b(t,x,k) = h_0(t,x,k)\sigma_0 + \lambda(t,x,k)\vec{\Omega}(t,x)\cdot\vec{\sigma} \tag{4.3.1}$$

$$H_{so}(t,x,k) = \vec{h}_{so}(t,x,k)\cdot\vec{\sigma}. \tag{4.3.2}$$

Here, $h_0 : \mathbb{R}^+ \times \mathbb{R}^{2d} \to \mathbb{R}$, $\lambda : \mathbb{R}^+ \times \mathbb{R}^{2d} \to \mathbb{R}$, $\vec{\Omega} : \mathbb{R}^+ \times \mathbb{R}^{2d} \to S^2$ (the unit sphere in $\mathbb{R}^3$) and $\vec{h}_{so} : \mathbb{R}^+ \times \mathbb{R}^{2d} \to \mathbb{R}^3$. The eigenvalues of H_b read $h_{b,\uparrow} = h_0 + |\lambda|$ and $h_{b,\downarrow} = h_0 - |\lambda|$. For fixed t and x, they represent the two different transport bands eligible for spin-coherent electrons. We refer to the band $h_{b,\uparrow}$ as the *up-band* and to $h_{b,\downarrow}$ as the *down-band*, respectively. The band gap is given by $2|\lambda|$. The unit vector $\vec{\Omega}$ shall play the role of the local direction of magnetization in a ferromagnet, therefore it depends on t and x but not on the momentum k of the electrons. The distribution functions $f_\uparrow$ and $f_\downarrow$ in the up- and in the down-band are given by the orthogonal projection $\Pi_{\uparrow/\downarrow} : \mathcal{H}_2(\mathbb{C}) \to \mathbb{R}$ of F on the eigenspace associated to the respective eigenvalue of H_b,

$$f_\uparrow = \Pi_\uparrow(F) = \left\langle \frac{1}{2}(\sigma_0 + \vec{\Omega}\cdot\vec{\sigma}), F \right\rangle_2 = \frac{1}{2}f_0 + \vec{\Omega}\cdot\vec{f} \tag{4.3.3}$$

$$f_\downarrow = \Pi_\downarrow(F) = \left\langle \frac{1}{2}(\sigma_0 - \vec{\Omega}\cdot\vec{\sigma}), F \right\rangle_2 = \frac{1}{2}f_0 - \vec{\Omega}\cdot\vec{f}. \tag{4.3.4}$$

A possible absence of spin-coherence is reflected by

$$\pm\vec{f} \parallel \vec{\Omega} \quad \Longleftrightarrow \quad [F, H_b] = 0. \tag{4.3.5}$$

In this case, the equations (4.3.3) and (4.3.4) yield $f_\uparrow = \frac{1}{2}f_0 \pm |\vec{f}|$ and $f_\downarrow = \frac{1}{2}f_0 \mp |\vec{f}|$ where $|\vec{f}|$ is the usual scalar spin polarization in two-component models.

The matrix H_{so} stands for contributions to the electron energy arising from the spin-orbit coupling that electrons experience when moving through a crystal lattice. The field $\vec{h}_{so}$ may, for example, contain terms like the Elliott-Yafet, the D'yakonov-Perel' or the Rashba spin-orbit couplings [4]. The separation of H into H_b and H_{so} was made in order to single out the field $\vec{\Omega}$, which plays a central role in the theory developed in this work. The total Hamilton matrix reads

$$H = H_b + H_{so} = h_0\sigma_0 + \vec{h}\cdot\vec{\sigma}, \tag{4.3.6}$$

where $\vec{h} = (\lambda\vec{\Omega} + \vec{h}_{so})$ is called the *pseudo-exchange field*.

Given a distribution matrix F, the energy $E : \mathbb{R}^+ \times \mathbb{R}^{2d} \to \mathbb{R}$ of the system at

(t, x, k) is obtained by

$$E(t,x,k) = \langle H, F \rangle_2 = h_0(t,x,k) f_0(t,x,k) + 2\vec{h}(t,x,k) \cdot \vec{f}(t,x,k). \tag{4.3.7}$$

The ballistic dynamics of the spin-coherent semi-classical electron system is described by the matrix Vlasov equation [34, **?**],

$$\begin{cases} \partial_t F + \dfrac{1}{\hbar}\left(\nabla_k h_0 \cdot \nabla_x F - \nabla_x h_0 \cdot \nabla_k F\right) + \dfrac{i}{\hbar}[F, \vec{h} \cdot \vec{\sigma}] = 0 \\ F(t=0,x,k) = F_{in}(x,k) \\ F_{in}(x,k) \in \mathcal{H}_2^{0,+}(\mathbb{C}) \qquad \forall (x,k). \end{cases} \tag{4.3.8}$$

The commutator $[F, \vec{h} \cdot \vec{\sigma}]$ describes the precession of the spin polarization $\vec{f}$ of the electron system around the pseudo-exchange field $\vec{h}$. Equation (4.3.8) is obtained by passing to the limit $\varepsilon \to 0$ (where ε stands for the scaled Planck constant $\hbar$) in the Schrödinger equation with the Hamiltonian (4.3.6), for the case where the modulus $|\vec{h}|$ of the pseudo-exchange field scales with order ε.[1] Thus, equation (4.3.8) is merely the correct semi-classical equation for electrons in a weak exchange field $\vec{h}$. This is the case, for example, if the band gap λ is small compared to the Fermi energy of the electron system. To take into account scattering processes (non-ballistic transport), we shall add collision terms at the right-hand-side of the matrix Vlasov equation (4.3.8).

4.3.2 Spin-coherent collision operators

For a spin-polarized electron system, there are two kinds of possible collision processes, namely the spin-conserving and the spin-flip collisions. Spin-conserving collisions drive the velocity distribution of the electrons towards thermal equilibrium, i.e., the Fermi-Dirac or the Maxwellian distribution, depending on the used statistics. On the other hand, spin-flip collisions will relax the electron spin density towards its thermal equilibrium field, which is, at each point $x \in \mathbb{R}^d_x$, parallel to the local pseudo-exchange field $\vec{h}$ defined in (4.3.6). In ferromagnets, the two processes happen at very different timescales [17], spin-conserving collisions occurring at a much higher frequency than spin-flip collisions.

In the present paper, we shall write seperate collision operators for each of the

[1] The fact that $\hbar$ still appears in (4.3.8) is due to rescaling to physical variables after the limiting procedure.

two processes. Spin-flip processes will be modeled by a relaxation term [22],

$$\frac{1}{\tau_{sf}} Q_{sf}(F) := \frac{1}{\tau_{sf}} \left(\frac{1}{2} \mathrm{tr}\,(F)\, \sigma_0 - F \right), \tag{4.3.9}$$

where $\tau_{sf} \in \mathbb{R}^+$ is the average time between two subsequent spin-flip collisions. The focus of this work is on spin-conserving momentum scattering of electrons with (magnetic) impurities. The essential point is that impurity potentials may look different for spin-up and spin-down electrons and that, additionally, the latter feature different density of states in spin-polarized materials [35]. Both effects lead to spin-dependent momentum scattering rates. Our goal is to construct a collision operator that, on the one hand, describes the impurity scattering of spin-coherent electrons and, on the other hand, takes into account spin-dependent collision rates. Moreover, the new *spin-coherent collision operator* must satisfy the necessary mathematical properties to yield a well-defined theory. We state four requirements on such an operator:

1. incorporate spin-dependent scattering rates,

2. yield a two-component Boltzmann model if

$$[F(t,x,k), H_b(t,x,k)] = 0 \qquad \forall (t,x,k) \in \mathbb{R}^+ \times \mathbb{R}^{2d},$$

3. be a map with range in $\mathcal{H}_2(\mathbb{C})$,

4. conserve the positive-(semi)definiteness of F,

$$F(t=0,x,k) \in \mathcal{H}_2^{0,+}(\mathbb{C}) \quad \Longrightarrow \quad F(t,x,k) \in \mathcal{H}_2^{0,+}(\mathbb{C}) \quad \forall (t,x,k) \in \mathbb{R}^+ \times \mathbb{R}^{2d}.$$

Let $S : \mathbb{R}^+ \times \mathbb{R}^{3d} \to \mathcal{H}_2^+(\mathbb{C})$ denote the *scattering matrix*, in the Pauli basis written as

$$S(t,x,k,k') = s_0(t,x,k,k')\sigma_0 + \vec{s}(t,x,k,k') \cdot \vec{\sigma}. \tag{4.3.10}$$

This matrix shall describe the rate at which spin-coherent electrons scatter from k to k' due to collisions with (magnetic) impurities. The eigenvalues of S, denoted by $s_\uparrow$ and $s_\downarrow$, shall stand for the respective (scaled) scattering rates of electrons in the up-band and in the down-band. An electron distribution F consists only of non-coherent spin-up and spin-down states if it commutes with the band matrix H_b, c.f. (4.3.5). In this case, the eigenvalues of F should scatter at the rates $s_\uparrow$ and $s_\downarrow$,

respectively. Therefore, it is necessary that

$$[S(t,x,k,k'), H_b(t,x,k)] = 0 \qquad \forall (t,x,k,k') \in \mathbb{R}^+ \times \mathbb{R}^{3d}. \tag{4.3.11}$$

We construct a linear spin-coherent collision operator $Q_{ij}(F)$ as a sum of a gain term $Q_i^+(F)$ and a loss term $Q_j^-(F)$,

$$\frac{1}{\tau_c} Q_{ij}(F) := \frac{1}{\tau_c} Q_i^+(F) - \frac{1}{\tau_c} Q_j^-(F) \qquad i,j \in \{1,2\}, \tag{4.3.12}$$

where we defined $\tau_c \in \mathbb{R}^+$, the time between two subsequent spin-conserving collision processes. The two basic possible structures of the gain and loss term, respectively, read

$$Q_1^+(F) := \int_{\mathbb{R}^d_{k'}} \left(\frac{1}{2} S'F' + \frac{1}{2} F'S' \right) \mathrm{d}k' \tag{4.3.13}$$

$$Q_2^+(F) := \int_{\mathbb{R}^d_{k'}} S'^{1/2} F' S'^{1/2} \mathrm{d}k', \tag{4.3.14}$$

$$Q_1^-(F) := \frac{1}{2} \Lambda F + \frac{1}{2} F \Lambda \tag{4.3.15}$$

$$Q_2^-(F) := \int_{\mathbb{R}^d_{k'}} S^{1/2} F S^{1/2} \mathrm{d}k', \tag{4.3.16}$$

where we denoted $F' = F(t,x,k')$, $S' = S(t,x,k',k)$, and $\Lambda = \int S \mathrm{d}k'$. The gain and loss terms are chosen such that they conserve the hermiticity of the electron distribution function F. We stress that, according to (5.2.3), there are four possible spin-coherent collision operators that fall into two categories, namely the symmetric operators Q_{11} respectively Q_{22}, and the anti-symmetric operators Q_{12} respectively Q_{21}. At first the different structures of gain and loss term in the anti-symmetric operators may seem counter-intuitive or unphysical. However, in the theory of open quantum systems [9], the well-known Lindblad equation features a product structure similar to the operator Q_{21}. It is easily seen that the symmetric operators Q_{11} and Q_{22} are mass- and spin-conserving,

$$\int_{\mathbb{R}^d_k} Q_{11}(F) \mathrm{d}k = \int_{\mathbb{R}^d_k} Q_{22}(F) \mathrm{d}k = 0, \tag{4.3.17}$$

whereas the anti-symmetric operators Q_{21} and Q_{12} are just mass-conserving,

$$\int_{\mathbb{R}^d_k} Q_{12}(F)\mathrm{d}k = -\int_{\mathbb{R}^d_k} Q_{21}(F)\mathrm{d}k \neq 0 \tag{4.3.18}$$

$$\int_{\mathbb{R}^d_k} \mathrm{tr}\left(Q_{12}(F)\right)\mathrm{d}k = \int_{\mathbb{R}^d_k} \mathrm{tr}\left(Q_{21}(F)\right)\mathrm{d}k = 0. \tag{4.3.19}$$

Therefore, the anti-symmetric operators contribute to the spin-flip scattering on the time scale τ_c, which contradicts our assumption that spin flip processes should happen on the timescale τ_{sf}.

4.4 The model

Adding the collision operators (4.3.9) and (5.2.3) on the right-hand side of (4.3.8), one obtains the following generalized matrix Boltzmann equation,

$$\begin{cases} \partial_t F + \dfrac{1}{\hbar}\left(\nabla_k h_0 \cdot \nabla_x F - \nabla_x h_0 \cdot \nabla_k F\right) + \dfrac{i}{\hbar}[F, \vec{h}\cdot\vec{\sigma}] = \dfrac{1}{\tau_c}Q_{ij}(F) + \dfrac{1}{\tau_{sf}}Q_{sf}(F) \\ F(t=0,x,k) = F_{in}(x,k) \\ F_{in}(x,k) \in \mathcal{H}_2^{0,+}(\mathbb{C}) \qquad \forall (x,k). \end{cases} \tag{4.4.1}$$

Here, h_0 and $\vec{h}$ are the components of the Hamilton matrix (4.3.6). Equation (4.4.1) has been investigated by el Hajj [22] for the case that S, occurring in $Q_{ij}(F)$ and defined in (4.3.10), is a scalar, $S = s_0\sigma_0$. The case $S \in \mathcal{H}_2^+(\mathbb{C})$ is a new problem which is the topic of this work.

4.4.1 Further assumptions

Let us summarize in this section the physical hypothesis we need for the further development. Generalizations to these assumptions shall be treated in forthcoming works.

Assumption 4.4.1 (**Shifted parabolic bands**). *We treat the case of two parabolic transport bands. Moreover, the band gap between the two spin bands does not depend on the momentum k. Thus, in equation (4.3.1) for H_b, we assume*

$$\begin{cases} h_0 = \dfrac{\hbar^2|k|^2}{2m} + V(t,x) \\ \lambda = \lambda(t,x), \end{cases} \tag{4.4.2}$$

where m is the effective mass of the electrons, $\lambda : \mathbb{R}^+ \times \mathbb{R}^d_x \to \mathbb{R}$ is the k-independent band gap and $V : \mathbb{R}^+ \times \mathbb{R}^d_x \to \mathbb{R}$ is an external potential energy.

Assumption 4.4.1 is known in the physics literature as the *Stoner model* [36]. It is a crude simplification of the problem, since transport bands in ferromagnets often do not have parabolic shape. However, the Stoner model is still a basic tool used to understand electron properties in ferromagnets [32, 33]. In this paper, we investigate the effects of spin-dependent scattering in the framework of the Stoner model and leave the case of more complicated bandstructures open for future work.

Assumption 4.4.2 (**Boltzmann statistics**). *In thermal equilibrium, the distribution matrix has the form $F_{eq} = c \exp(-\beta H_{th})$ where $c \in \mathbb{R}^+$ is a normalization constant, $\beta = 1/k_B T$ is the inverse of the thermal energy and H_{th} denotes the Hamilton matrix (4.3.6) of the system without externally applied electric or magnetic fields.*

From the assumptions 4.4.1 and 4.4.2 we deduce

$$F_{eq}(x,k) = N(x)\mathcal{M}_\beta(k) \qquad \text{with} \qquad [N(x), H_{th}(x,k)] = 0 \quad \forall (x,k) \in \mathbb{R}^{2d}, \tag{4.4.3}$$

where $N \in \mathcal{H}_2(\mathbb{C})$ is a hermitian matrix and $\mathcal{M}_\beta$ stands for the scalar Maxwellian at thermal energy β^{-1},

$$\mathcal{M}_\beta(k) = \left(\frac{\beta\hbar^2}{2\pi m}\right)^{d/2} \exp\left(-\frac{\beta\hbar^2|k|^2}{2m}\right). \tag{4.4.4}$$

Assumption 4.4.3 (**Detailed balance**). *Let $\sigma(A)$ denote the ordered spectrum of $A \in \mathcal{H}_2(\mathbb{C})$. We assume local thermal equilibrium in each band,*

$$\sigma(SF_{eq}) = \sigma(S'F'_{eq}). \tag{4.4.5}$$

This assumption implies that spin-conserving momentum scattering occurs at a much faster timescale than spin-flip scattering, so that equilibrium is established in each band separately before the whole system reaches equilibrium.

Under the assumption 4.4.3, from (4.4.3) one obtains

$$s_{\uparrow(\downarrow)}\mathcal{M}_\beta = s'_{\uparrow(\downarrow)}\mathcal{M}'_\beta \quad \Longrightarrow \quad \frac{s_\uparrow}{s_\downarrow} = \frac{s'_\uparrow}{s'_\downarrow} \qquad \forall k, k'. \tag{4.4.6}$$

In (4.4.6) we see that the ratio of the scattering rates for spin-up and spin-down electrons must not depend on k. We deduce $s_\uparrow = C(t,x)s_\downarrow$ where $C \in \mathbb{R}^+$. Therefore,

the scattering matrix S can be written as

$$S(t,x,k,k') = \alpha(t,x,k,k')P(t,x) \tag{4.4.7}$$

$$P(t,x) = \sigma_0 + p(t,x)\vec{\Omega}(t,x)\cdot\vec{\sigma}, \tag{4.4.8}$$

where $\alpha \in \mathbb{R}^+$ denotes the scattering rate from k to k' at (t,x) and $P : \mathbb{R}^+ \times \mathbb{R}^d_x \to \mathcal{H}_2(\mathbb{C})$ is called the *polarization matrix*. Note that, because of (4.3.11), the direction of P in spin space has to be $\vec{\Omega}$, the direction of the local magnetization. Moreover, the parameter p, which satisfies $0 \le p(t,x) < 1$, represents the spin-polarization of the scattering rates, whose ratio $C(t,x)$ is now given by

$$s_\uparrow = \frac{1+|p(t,x)|}{1-|p(t,x)|}s_\downarrow. \tag{4.4.9}$$

Further, by inserting the eigenvalues of S written in (4.4.7) into (4.4.6), one can define the function ϕ which is symmetric in k and k' as

$$\phi(t,x,k,k') = \frac{\alpha(t,x,k,k')}{\mathcal{M}_\beta(k')} = \frac{\alpha(t,x,k',k)}{\mathcal{M}_\beta(k)} = \phi(t,x,k',k). \tag{4.4.10}$$

4.4.2 Symmetric collision operator for the Stoner model

In the present paper, we shall perform the diffusion limit with the symmetric operator Q_{22} and denote it simply by $Q(F)$,

$$Q(F) := Q_{22}(F). \tag{4.4.11}$$

From (4.4.7), (4.4.8) and (4.4.10), one deduces the collision operator (4.4.11) in the Stoner model,

$$Q(F) = P^{1/2}K(F)P^{1/2}, \tag{4.4.12}$$

where

$$K(F)(k) := K^+(F)(k) - K^-(F)(k) = \mathcal{M}_\beta \int_{\mathbb{R}^d_{k'}} \phi F' \mathrm{d}k' - F\lambda. \tag{4.4.13}$$

Here, $\lambda = \int_{\mathbb{R}^d_{k'}} \phi \mathcal{M}'_\beta \mathrm{d}k'$ denotes the collision frequency.

4.4.3 Scaled model

The next step is to scale (4.4.1) in a way suitable for performing the diffusion limit. The main assumption is that the time scales τ_c and τ_{sf} are very different,

$$\tau_c \ll \tau_{sf} \qquad \Longrightarrow \qquad \varepsilon := \sqrt{\frac{\tau_c}{\tau_{sf}}} \ll 1. \tag{4.4.14}$$

Here ε is a small parameter intended to go to zero. Let $\bar{v} = (\beta m)^{-1/2}$ denote the thermal velocity of the electrons. The length scale $\bar{l}$ we choose is the geometric average of the two occurring mean free paths $l_c = \tau_c \bar{v}$ and $l_{sf} = \tau_{sf}\bar{v}$, respectively,

$$\bar{l} = \sqrt{l_c l_{sf}}. \tag{4.4.15}$$

The characteristic time, momentum and energy scales are chosen as

$$\bar{t} = \tau_{sf} \qquad\qquad \bar{k} = \frac{m\bar{v}}{\hbar} \qquad\qquad \bar{E} = \beta^{-1}. \tag{4.4.16}$$

Applying the scaling (4.4.15)-(4.4.16) to (4.4.1) with the collision operator $Q_{ij}(F) = Q_{22}(F) = Q(F)$, $Q(F)$ given by (4.4.12), then multiplying by τ_{sf} and subsequently inserting (4.4.14) leads to the diffusion-scaled matrix Boltzmann equation (now in dimensionless variables t, x and k). The scaling of the pseudo-exchange field $\vec{h} = (\lambda\vec{\Omega} + \vec{h}_{so})$ is crucial for performing the diffusion limit. In this work, we assume the weak coupling $\lambda/\bar{E} = \mathcal{O}(\varepsilon^2)$ and $|\vec{h}_{so}|/\bar{E} = \mathcal{O}(\varepsilon^2)$. Thus, under the hypothesis of the Stoner model (4.4.2), the scaled Hamilton matrix (4.3.6) reads

$$\begin{cases} H^\varepsilon(t,x,k) = \left(\dfrac{|k|^2}{2} + \hat{V}(t,x)\right)\sigma_0 + \varepsilon^2 \hat{\vec{h}}(t,x,k)\cdot\vec{\sigma} \\ \hat{\vec{h}}(t,x,k) := \vec{\Omega}(t,x) + \hat{\vec{h}}_{so}(t,x,k), \end{cases} \tag{4.4.17}$$

where we defined $\hat{V} := V/\bar{E}$ and $\hat{\vec{h}}_{so} = \vec{h}_{so}/\bar{E}$. Using the scaled Hamilton matrix (4.4.17) leads to the following scaled version of (4.4.1),

$$\begin{cases} \partial_t F^\varepsilon + \dfrac{1}{\varepsilon}\mathcal{T}(F^\varepsilon) + i[F^\varepsilon, \hat{\vec{h}}\cdot\vec{\sigma}] = \dfrac{1}{\varepsilon^2}Q(F^\varepsilon) + Q_{sf}(F) \\ F(t=0,x,k) = F_{in}(x,k) \\ F_{in}(x,k) \in \mathcal{H}_2^{0,+}(\mathbb{C}) \qquad \forall (x,k), \end{cases} \tag{4.4.18}$$

where the transport operator $\mathcal{T}(F^\varepsilon)$ is defined by

$$\mathcal{T}(F^\varepsilon) = k\cdot\nabla_x F^\varepsilon - \nabla_x \hat{V}\cdot\nabla_k F^\varepsilon \tag{4.4.19}$$

and the scaled collision operator $Q(F^\varepsilon)$ is given by (4.4.12) and (4.4.13) where $\mathcal{M}_\beta$ is to be replaced by the scaled Maxwellian,

$$\mathcal{M}(k) = \left(\frac{1}{2\pi}\right)^{d/2} \exp\left(-\frac{|k|^2}{2}\right). \tag{4.4.20}$$

The aim of the present work is to go to the limit $\varepsilon \to 0$ in equation (4.4.18) in order to obtain a macroscopic model, more suitable for numerical simulations. Therefore, we shall make a Hilbert ansatz $F^\varepsilon = F^0 + \varepsilon F^1 + \varepsilon^2 F^2 + \dots$ of the solution and sort the appearing terms in powers of ε. The obtained equations read

$$Q(F^0) = 0 \tag{4.4.21}$$

$$Q(F^1) = \mathcal{T}(F^0) \tag{4.4.22}$$

$$Q(F^2) = \partial_t F^0 + \mathcal{T}(F^1) + i[F^0, \hat{\vec{h}} \cdot \vec{\sigma}] - Q_{sf}(F^0). \tag{4.4.23}$$

Let us summarize now the main steps of this work.

4.4.4 Contents of the paper

Section 4.5 deals with the analysis of the generalized matrix Boltzmann equation (4.4.1). We prove existence and uniqueness of a weak solution. Moreover, it is shown that the solution F satisfies the maximum principle when the anti-symmetric collision operator $Q_{21}(F)$ is used in the Boltzmann equation. Section 4.6 contains the analysis of the collision operator $Q(F)$ appearing in (4.4.18) in a proper mathematical framework. In section 4.7 we present the main theorem of this work, namely the diffusion limit $\varepsilon \to 0$ in (4.4.18). Section 6.5 contains some numerical results of the macroscopic matrix drift-diffusion equations obtained in the diffusion limit. Some implications for physical applications such as spin-transfer torque devices or domain wall dynamics in ferromagnets are discussed briefly.

4.5 Existence, positive-definiteness and uniqueness of a weak solution

Let us first start by studying the matrix Boltzmann equation (4.4.1), in particular proving the existence, positive-(semi)definiteness and uniqueness of a weak solution. Without loss of generality the constants $\hbar$ and τ_{sf} are set to one. Let us introduce the following Hilbert space:

Definition 4.5.1 (**Hilbert space**). *By $\mathbb{L}^2_{\mathcal{M}}$ we denote the following space,*

$$\mathbb{L}^2_{\mathcal{M}} := \left\{ F : \mathbb{R}^{2d} \to \mathcal{H}_2(\mathbb{C}) \,\middle|\, \int_{\mathbb{R}^d_x} \int_{\mathbb{R}^d_k} ||F||_2^2 \mathcal{M}^{-1} \mathrm{d}k \mathrm{d}x < \infty \right\} \tag{4.5.1}$$

associated with the scalar product and the corresponding norm

$$(F,G)_{\mathbb{L}^2_{\mathcal{M}}} := \int_{\mathbb{R}^d_x} \int_{\mathbb{R}^d_k} \langle F, G \rangle_2 \, \mathcal{M}^{-1} \mathrm{d}k \mathrm{d}x \qquad ||F||_{\mathbb{L}^2_{\mathcal{M}}} = \sqrt{(F,F)_{\mathbb{L}^2_{\mathcal{M}}}}, \tag{4.5.2}$$

where $\mathcal{M}$ stands for the scaled Maxwellian (4.4.20).

Assumption 4.5.2. *Let $h_0 \in L^\infty(0,T; W^{2,\infty}_{loc}(\mathbb{R}^{2d}))$, $\vec{h} \in (L^\infty_{loc}([0,T) \times \mathbb{R}^{2d}))^3$ and let us define the transport operator $\mathcal{T}_{h_0} : D(\mathcal{T}_{h_0}) \to \mathbb{L}^2_{\mathcal{M}}$ by*

$$\mathcal{T}_{h_0}(F) := \nabla_k h_0 \cdot \nabla_x F - \nabla_x h_0 \cdot \nabla_k F \tag{4.5.3}$$

and where the definition domain and norm are given by

$$D(\mathcal{T}_{h_0}) := \{ F \in \mathbb{L}^2_{\mathcal{M}} \,|\, \mathcal{T}_{h_0}(F) \in \mathbb{L}^2_{\mathcal{M}} \} \tag{4.5.4}$$

$$||F||^2_{D(\mathcal{T})} := ||F||^2_{\mathbb{L}^2_{\mathcal{M}}} + ||\mathcal{T}_{h_0}(F)||^2_{\mathbb{L}^2_{\mathcal{M}}}. \tag{4.5.5}$$

Assumption 4.5.3. *The scattering matrix S defined in (4.3.10) is chosen in such a way that (5.2.3) is a well-defined linear operator, $Q_{ij} : \mathbb{L}^2_{\mathcal{M}} \to \mathbb{L}^2_{\mathcal{M}}$, satisfying*

$$\exists\, c > 0 \, s.t. \qquad ||Q_{ij}(F)||^2_{\mathbb{L}^2_{\mathcal{M}}} \leq c \, ||F||^2_{\mathbb{L}^2_{\mathcal{M}}} \qquad \forall F \in \mathbb{L}^2_{\mathcal{M}}. \tag{4.5.6}$$

An example is given in section 4.6.

Definition 4.5.4 (**Weak solution**). *Let $\hbar = \tau_{sf} = 1$ and $F_{in} \in \mathbb{L}^2_{\mathcal{M}}$. For a fixed time $T > 0$, a function $F \in L^2(0,T; \mathbb{L}^2_{\mathcal{M}})$ is called a weak solution of (4.4.1) if it satisfies*

$$\begin{aligned} -\int_0^T \int \mathrm{d}x \int \mathrm{d}k \, \langle \partial_t \Psi, F \rangle_2 \, \mathrm{d}t - \int_0^T \int \mathrm{d}x \int \mathrm{d}k \, \langle \mathcal{T}_{h_0}(\Psi), F \rangle_2 \, \mathrm{d}t + \\ + i \int_0^T \int \mathrm{d}x \int \mathrm{d}k \, \left\langle \Psi, [F, \vec{h} \cdot \vec{\sigma}] \right\rangle_2 \mathrm{d}t = \int_0^T \int \mathrm{d}x \int \mathrm{d}k \, \langle \Psi, Q_{ij}(F) \rangle_2 \, \mathrm{d}t + \\ + \int_0^T \int \mathrm{d}x \int \mathrm{d}k \, \langle \Psi, Q_{sf}(F) \rangle_2 \, \mathrm{d}t + \int \mathrm{d}x \int \mathrm{d}k \, \langle \Psi, F_{in} \rangle_2 \end{aligned} \tag{4.5.7}$$

for all test functions $\Psi \in C^1_c([0,T) \times \mathbb{R}^{2d}, \mathcal{H}_2(\mathbb{C}))$.

Proposition 4.5.5 (**Existence/Uniqueness**). *Let $T > 0$ be fixed. Under the*

assumptions 4.5.2, 4.5.3 and with $F_{in} \in \mathbb{L}^2_{\mathcal{M}}$, the matrix Boltzmann equation (4.4.1) admits a unique weak solution $F \in L^\infty(0,T;\mathbb{L}^2_{\mathcal{M}})$.

Proof. Let us define the fixed point map

$$\mathcal{F} : L^2(0,T;\mathbb{L}^2_{\mathcal{M}}) \to L^2(0,T;\mathbb{L}^2_{\mathcal{M}}) \qquad ; \qquad F^{old} \mapsto F^{new}, \tag{4.5.8}$$

where F^{new} is a solution of

$$\begin{cases} \partial_t F^{new} + \mathcal{T}_{h_0}(F^{new}) + i[F^{new}, \vec{h}\cdot\vec{\sigma}] - \\ \qquad\qquad - Q_{sf}(F^{new}) + Q_j^-(F^{new}) = Q_i^+(F^{old}) \\ F^{new}(t=0,x,k) = F_{in}(x,k). \end{cases} \tag{4.5.9}$$

The first step is to show that $\mathcal{F}$ is well-defined. For this take $F^{old} \in L^2(0,T;\mathbb{L}^2_{\mathcal{M}})$ and denote by $G_i := Q_i^+(F^{old}) \in L^2(0,T;\mathbb{L}^2_{\mathcal{M}})$. Let us use the decomposition of each matrix $F \in \mathcal{H}_2(\mathbb{C})$ in the Pauli basis $\{\sigma_0,\sigma_1,\sigma_2,\sigma_3\}$, which means $F = \frac{1}{2}f_0 + \vec{f}\cdot\vec{\sigma}$ and where we denote the coefficients by $\tilde{F} = (\frac{1}{2}f_0, \vec{f})$. Using this decomposition, system (4.5.9) now writes

$$\begin{cases} \partial_t \tilde{F} + \mathcal{T}_{h_0}(\tilde{F}) + (\mathbf{A} + \mathbf{D}_j)\tilde{F} = \tilde{G}_i \\ \tilde{F}(t=0,x,k) = \tilde{F}_{in}(x,k). \end{cases} \tag{4.5.10}$$

Here, $\mathbf{A} \in M_4(\mathbb{R})$ is the 4×4 matrix representation of the operator $i[F,\vec{h}\cdot\vec{\sigma}] - Q_{sf}(F)$,

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & -2h_3 & 2h_2 \\ 0 & 2h_3 & -1 & -2h_1 \\ 0 & -2h_2 & 2h_1 & -1 \end{pmatrix},$$

and $\mathbf{D}_j \in M_4(\mathbb{R})$ is the matrix corresponding to the loss term Q_j^-, c.f. (4.3.15) and (4.3.16), with $\Lambda \in \mathcal{H}_2(\mathbb{C}) \Leftrightarrow \tilde{\Lambda} = (\lambda_0, \vec{\lambda})$,

$$\mathbf{D}_1 = \begin{pmatrix} \lambda_0 & \lambda_1 & \lambda_2 & \lambda_3 \\ \lambda_1 & \lambda_0 & 0 & 0 \\ \lambda_2 & 0 & \lambda_0 & 0 \\ \lambda_3 & 0 & 0 & \lambda_0 \end{pmatrix},$$

$$\mathbf{D}_2 = \begin{pmatrix} \lambda_0 & \lambda_1 & \lambda_2 & \lambda_3 \\ \lambda_1 & r + \frac{\lambda_1^2}{|\vec{\lambda}|^2}(\lambda_0 - r) & \frac{\lambda_1\lambda_2}{|\vec{\lambda}|^2}(\lambda_0 - r) & \frac{\lambda_1\lambda_3}{|\vec{\lambda}|^2}(\lambda_0 - r) \\ \lambda_2 & \frac{\lambda_2\lambda_1}{|\vec{\lambda}|^2}(\lambda_0 - r) & r + \frac{\lambda_2^2}{|\vec{\lambda}|^2}(\lambda_0 - r) & \frac{\lambda_2\lambda_3}{|\vec{\lambda}|^2}(\lambda_0 - r) \\ \lambda_3 & \frac{\lambda_3\lambda_1}{|\vec{\lambda}|^2}(\lambda_0 - r) & \frac{\lambda_3\lambda_2}{|\vec{\lambda}|^2}(\lambda_0 - r) & r + \frac{\lambda_3^2}{|\vec{\lambda}|^2}(\lambda_0 - r) \end{pmatrix},$$

where $r = (\lambda_0^2 - |\vec{\lambda}|^2)^{1/2}$. The matrix $\mathbf{D}_2$ is obtained straightforwardly from Lemmas 4.2.4 and 4.2.3. Now let $\mathcal{Z}_{t,x,k}(s) = (\mathcal{X}(s), \mathcal{K}(s))$ denote the characteristics in the phase space of (4.5.10) starting at the point $(x,k) \in \mathbb{R}^{2d}$ at time t. Its components satisfy the following system of equations,

$$\begin{cases} \dfrac{\partial \mathcal{X}(s)}{\partial s} = \nabla_k h_0(s, \mathcal{X}(s), \mathcal{K}(s)) \\ \dfrac{\partial \mathcal{K}(s)}{\partial s} = -\nabla_x h_0(s, \mathcal{X}(s), \mathcal{K}(s)) \\ \mathcal{X}(t) = x \qquad \mathcal{K}(t) = k. \end{cases} \tag{4.5.11}$$

Defining now, for each fixed $(x_0, k_0) \in \mathbb{R}^{2d}$ the function

$$g(t, x_0, k_0) := \tilde{F}(t, \mathcal{Z}_{0,x_0,k_0}(t))\,, \quad \forall t \in [0,T]\,,$$

one gets the system

$$\begin{cases} \frac{d}{dt} g(t, x_0, k_0) = \\ \qquad\qquad = -\left(\mathbf{A} + \mathbf{D}_j\right)(t, \mathcal{Z}_{0,x_0,k_0}(t))\, g(t, x_0, k_0) + \tilde{G}_i(t, \mathcal{Z}_{0,x_0,k_0}(t)) \\ g(0, x_0, k_0) = \tilde{F}_{in}(x_0, k_0)\,. \end{cases} \tag{4.5.12}$$

Denoting the "evolution matrix" by $R(\cdot; s) : \mathbb{R}^+ \to M_4(\mathbb{C})$, which represents for each $0 \le s \le T$ the unique solution of the following homogeneous system

$$\begin{cases} \frac{d}{dt} R(t; s) = -\left(\mathbf{A} + \mathbf{D}_j\right)(t, \mathcal{Z}_{0,x_0,k_0}(t))\, R(t; s)\,, \quad \forall t \in [s, T] \\ R(s; s) = Id\,, \end{cases}$$

and which satisfies for $0 \le s \le t \le T$

$$||R(t; s)|| \le \exp\left\{ \int_s^t ||(\mathbf{A} + \mathbf{D}_j)(\tau, \mathcal{Z}_{0,x_0,k_0}(\tau)||\, d\tau \right\} \le C\,, \tag{4.5.13}$$

where $C > 0$ is a constant independent on s, t, x_0, k_0 and $||\cdot||$ is an operator norm in $\mathcal{L}(\mathbb{C}^4)$, the solution of (4.5.12) can be written as

$$g(t, x_0, k_0) = R(t; 0)\, \tilde{F}_{in}(x_0, k_0) + \textstyle\int_0^t R(t; s)\, \tilde{G}_i(s, \mathcal{Z}_{0,x_0,k_0}(s))\, ds\,. \tag{4.5.14}$$

Remarking now that $\tilde{F}(t,x,k) = g(t, \mathcal{Z}_{t,x,k}(0))$, one has the Duhamel formula

$$\tilde{F}(t,x,k) = R(t;0)\,\tilde{F}_{in}(\mathcal{Z}_{t,x,k}(0)) + \\ + \int_0^t R(t;s)\,\tilde{G}_i(s, \mathcal{Z}_{t,x,k}(s))\,ds\,, \quad \forall (t,x,k) \in [0,T]\times\mathbb{R}^{2d}\,, \quad (4.5.15)$$

which is a solution of (4.5.10) respectively (4.5.9) and therefore of (4.4.1). The goal is now to prove that the fixed point map $\mathcal{F}$ is a contraction, admitting thus a unique fixed point $F \in L^2([0,T],\mathbb{L}^2_{\mathcal{M}})$. From equation (4.5.15) we know that a solution $F \in L^\infty(0,T,\mathbb{L}^2_{\mathcal{M}})$ satisfies the following estimate,

$$||F^{new}(t,\cdot,\cdot)||_{\mathbb{L}^2_{\mathcal{M}}} \leq C\,||F_{in}||_{\mathbb{L}^2_{\mathcal{M}}} + \\ + C\int_0^t \left|\left|Q_i^+(F^{old})(\tau,\cdot,\cdot)\right|\right|_{\mathbb{L}^2_{\mathcal{M}}} \mathrm{d}\tau \quad \forall t\in[0,T]. \quad (4.5.16)$$

Squaring gives the following estimate in the L^∞-norm,

$$||F^{new}||^2_{L^\infty(0,T;\mathbb{L}^2_{\mathcal{M}})} \leq 2C\,||F_{in}||^2_{\mathbb{L}^2_{\mathcal{M}}} + 2TC||Q_i^+(F^{old})||^2_{L^2(0,T;\mathbb{L}^2_{\mathcal{M}})},$$

yielding $F^{new} \in L^2(0,T,\mathbb{L}^2_{\mathcal{M}})$. To prove that $\mathcal{F}$ is a contraction we introduce the following norm in $L^2(0,T,\mathbb{L}^2_{\mathcal{M}})$,

$$||G||^2_\delta := \int_0^T e^{-\delta t}\,||G(t,\cdot,\cdot)||^2_{\mathbb{L}^2_{\mathcal{M}}}\,\mathrm{d}t \qquad\qquad \forall G \in L^2(0,T,\mathbb{L}^2_{\mathcal{M}}),$$

where the parameter $\delta > 0$ shall be specified later. We estimate

$$\begin{aligned}
||\mathcal{F}(F_1^{old}) - \mathcal{F}(F_2^{old})||^2_\delta &= ||F_1^{new} - F_2^{new}||^2_\delta \\
&= \int_0^T e^{-\delta t}\,||F_1^{new}(t) - F_2^{new}(t)||^2_{\mathbb{L}^2_{\mathcal{M}}}\,\mathrm{d}t \\
&\leq 2C\int_0^T e^{-\delta t}\int_0^t \left|\left|Q_i^+(F_1^{old})(s) - Q_i^+(F_2^{old})(s)\right|\right|^2_{\mathbb{L}^2_{\mathcal{M}}}\,\mathrm{d}s\mathrm{d}t \\
&= 2C\int_0^T\int_s^T e^{-\delta t}\left|\left|Q_i^+(F_1^{old})(s) - Q_i^+(F_2^{old})(s)\right|\right|^2_{\mathbb{L}^2_{\mathcal{M}}}\,\mathrm{d}t\mathrm{d}s \\
&\leq 2C\int_0^T \left|\left|F_1^{old}(s) - F_2^{old}(s)\right|\right|^2_{\mathbb{L}^2_{\mathcal{M}}} \frac{e^{-\delta s} - e^{-\delta T}}{\delta}\mathrm{d}s \\
&\leq \frac{2C}{\delta}||F_1^{old} - F_2^{old}||^2_\delta,
\end{aligned}$$

yielding that the parameter δ can be chosen in such a manner that $\mathcal{F}$ is a contraction. Therefore $\mathcal{F}$ admits a unique fixed point in $L^2([0,T],\mathbb{L}^2_{\mathcal{M}})$, solution of the matrix Boltzmann equation (4.4.1). □

Proposition 4.5.6 (Positive-(semi)definiteness). *Let $T > 0$ be fixed. Further let $\phi \in L^\infty([0,T]\times\mathbb{R}^{2d})$ and assume 4.5.2 and 4.5.3. The matrix Boltzmann equation (4.4.1) with the collision operator $Q_{21}(F)$ conserves the positive-(semi)definiteness of a weak solution $F \in L^\infty(0,T;\mathbb{L}^2_{\mathcal{M}})$,*

$$F_{in}(x,k) \in \mathcal{H}_2^{0,+}(\mathbb{C})\forall(x,k)\in\mathbb{R}^{2d} \Longrightarrow F(t,x,k)\in\mathcal{H}_2^{0,+}(\mathbb{C})\forall(t,x,k)\in[0,T]\times\mathbb{R}^{2d}.$$

Proof. To prove this Proposition we shall show that for $F_{in}\in\mathcal{H}_2^{0,+}(\mathbb{C})$, the smallest eigenvalue of F, denoted $f_- = \frac{1}{2}f_0 - |\vec{f}|$, satisfies $f_-(t,x,k)\geq 0\,\forall\,(t,x,k)\in\mathbb{R}^+\times\mathbb{R}^{2d}$. For this, let us find the equation of motion satisfied by f_-. Starting from

$$\partial_t F + \mathcal{T}_{h_0}(F) + i[F,\vec{h}\cdot\vec{\sigma}] - Q_{sf}(F) + Q_j^-(F) = G_i \tag{4.5.17}$$

where we defined $G_i(t,x,k) := Q_i^+(F^{old})(t,x,k)$, $G_i(t,x,k)\in\mathcal{H}_2(\mathbb{C})\,\forall(t,x,k)\in\mathbb{R}^+\times\mathbb{R}^{2d}$, and taking the trace, we get

$$\partial_t f_0 + \mathcal{T}_{h_0}(f_0) + \mathrm{tr}\left(Q_j^-(F)\right) = \mathrm{tr}\left(G_i\right). \tag{4.5.18}$$

For this we recall:

$$F = \frac{1}{2}f_0\sigma_0 + \vec{f}\cdot\vec{\sigma} \qquad ; \qquad f_0 = \mathrm{tr}\,(F) \tag{4.5.19}$$

$$[F,\vec{h}\cdot\vec{\sigma}] = [\vec{f}\cdot\vec{\sigma},\vec{h}\cdot\vec{\sigma}] = 2i(\vec{f}\times\vec{h})\cdot\vec{\sigma}. \tag{4.5.20}$$

Multiplying now equation (4.5.17) with $\frac{1}{2}\vec{\sigma}$, taking the trace and further taking the scalar product with $\vec{f}/|\vec{f}|$ permits to get an equation for $|\vec{f}|$,

$$\partial_t|\vec{f}| + \mathcal{T}_{h_0}(|\vec{f}|) + |\vec{f}| + \frac{1}{2}\mathrm{tr}\left(\vec{\sigma}Q_j^-(F)\right)\cdot\frac{\vec{f}}{|\vec{f}|} = \frac{1}{2}\mathrm{tr}\left(\vec{\sigma}G_i\right)\cdot\frac{\vec{f}}{|\vec{f}|} \tag{4.5.21}$$

Subtracting (4.5.21) from (4.5.18) multiplied by $\frac{1}{2}$ gives, for $f_- = \frac{1}{2}f_0 - |\vec{f}|$,

$$\partial_t f_- + \mathcal{T}_{h_0}(f_-) - |\vec{f}| + \Pi_-(Q_j^-(F)) = \Pi_-(G_i) \tag{4.5.22}$$

where the operator $\Pi_-:\mathcal{H}_2(\mathbb{C})\to\mathbb{R}$ is the projection on the smallest eigenvalue f_- of $F\in\mathcal{H}_2(\mathbb{C})$. We shall now apply the maximum principle to (4.5.22). For this, we rewrite the equation in the form

$$\begin{cases} \partial_t f_- + \mathcal{T}_{h_0}(f_-) + \omega_{ij}f_- = \gamma_{ij} \\ f_-(t=0,x,k) = f_{-,in}(x,k), \end{cases} \tag{4.5.23}$$

where the coefficients $\omega_{ij} = \omega_{ij}(t,x,k) \in \mathbb{R}$ and $\gamma_{ij} = \gamma_{ij}(t,x,k) \in \mathbb{R}\,\forall(t,x,k) \in \mathbb{R}^+ \times \mathbb{R}^{2d}$ are computed in Appendix 4.A for $i,j \in \{1,2\}$. Now let $\mathcal{Z}_{t,x,k}(s) = (\mathcal{X}(s),\mathcal{K}(s))$ denote the characteristics in the phase space of (4.5.23) defined by (4.5.11). Using the Duhamel formula, we get the following identity for f_-

$$f_-(t,x,k) = \exp\left(-\int_0^t \omega_{ij}(s,\mathcal{Z}_{t,x,k}(s))\mathrm{d}s\right) f_{-,in}(\mathcal{Z}_{t,x,k}(0)) + \\ + \int_0^t \exp\left(-\int_\tau^t \omega_{ij}(s,\mathcal{Z}_{t,x,k}(s))\mathrm{d}s\right)\gamma_{ij}(\tau,\mathcal{Z}_{t,x,k}(\tau))\mathrm{d}\tau. \tag{4.5.24}$$

We see that $f_-(t,x,k) \geq 0 \forall(t,x,k) \in \mathbb{R}\times\mathbb{R}^{2d}$ is satisfied if $\gamma_{ij}(t,x,k) \geq 0 \forall(t,x,k) \in \mathbb{R}\times\mathbb{R}^{2d}$. The coefficients γ_{ij} for the respective choice of the gain and the loss term are written in (4.A.9)-(4.A.12). Lemma 4.2.5 yields $F^{old} \in \mathcal{H}_2^{0,+}(\mathbb{C}) \Rightarrow G_2 = Q_2^+(F^{old}) \in \mathcal{H}_2^{0,+}(\mathbb{C})$ and thus $\Pi_-(G_2) \geq 0$. However, $G_1 = Q_1^+(F^{old})$ is not necessarily a positive-(semi)definite matrix if $F^{old} \in \mathcal{H}_2^{0,+}(\mathbb{C})$. Additionally, we observe that one needs an estimate for the term $\Pi_-(G_2)$ in (4.A.12) in order to check the maximum principle for Q_{22}. This shall be done in a forthcoming work. In conclusion, we have

$$F^{old}(t,x,k) \in \mathcal{H}_2^{0,+}(\mathbb{C}) \quad \Longrightarrow \quad \gamma_{21}(t,x,k) \geq 0 \qquad \forall(t,x,k) \in \mathbb{R}\times\mathbb{R}^{2d}, \tag{4.5.25}$$

yielding that the collision operator Q_{21} guarantees that a solution F of (4.4.1) satisfies the maximum principle. □

4.6 Properties of the collision operator Q(F)

In this section we analyze the spin-coherent collision operator $Q(F)$ occurring in the scaled Boltzmann equation (4.4.18). By $Im\,Q$ and $Ker\,Q$ we denote the image and the kernel of Q. The variables t and x shall be considered as parameters in this section and will often be omitted.

Definition 4.6.1. *By $\mathcal{L}^2_{\mathcal{M}}$ we denote the following Hilbert space,*

$$\mathcal{L}^2_{\mathcal{M}} := \left\{ F : \mathbb{R}^d_k \to \mathcal{H}_2(\mathbb{C}) \,\middle|\, \int_{\mathbb{R}^d_k} ||F||_2^2 \mathcal{M}^{-1}\mathrm{d}k < \infty \right\}, \tag{4.6.1}$$

equipped with the scalar product and the corresponding norm

$$(F,G)_{\mathcal{L}^2_{\mathcal{M}}} := \int_{\mathbb{R}^d_k} \langle F,G\rangle_2 \,\mathcal{M}^{-1}\mathrm{d}k \qquad ||F||_{\mathcal{L}^2_{\mathcal{M}}} = \sqrt{(F,F)_{\mathcal{L}^2_{\mathcal{M}}}}, \tag{4.6.2}$$

where $\mathcal{M}$ stands for the scaled Maxwellian (4.4.20).

Assumption 4.6.2. *The polarization matrix $P : \mathbb{R}^+ \times \mathbb{R}^d_x \to \mathcal{H}_2^+(\mathbb{C})$, written in the Pauli basis as $P(t,x) = \sigma_0 + p(t,x)\vec{\Omega}(t,x)\cdot\vec{\sigma}$, is a hermitian, positive-definite matrix with $|\vec{\Omega}| = 1$. Its eigenvalues are $p_\uparrow = 1+|p|$ and $p_\downarrow = 1-|p|$ where $0 \leq |p| < 1$. The scattering rate $\phi \in L^\infty(\mathbb{R}^{2d})$ is symmetric in k and k' and is bounded from above and below,*

$$0 < \phi_1 \leq \phi(k,k') \leq \phi_2 < \infty \qquad \forall k, k'. \tag{4.6.3}$$

Proposition 4.6.3 (**Spin-coherent collision operator**)**.** *Under assumption 4.6.2, the spin-coherent collision operator $Q : \mathcal{L}^2_\mathcal{M} \to \mathcal{L}^2_\mathcal{M}$ written in (4.4.12), with the scaled Maxwellian (4.4.20), satisfies the following properties:*

i) $Q : \mathcal{L}^2_\mathcal{M} \to \mathcal{L}^2_\mathcal{M}$ is a linear, self-adjoint, continuous and non-positive operator.

ii) Conservation of mass and spin:

$$\int_{\mathbb{R}^d_k} Q(F)\mathrm{d}k = 0 \qquad \forall F \in \mathcal{L}^2_\mathcal{M}. \tag{4.6.4}$$

iii) The kernel of Q has the form

$$Ker\, Q = \{F \in \mathcal{L}^2_\mathcal{M} \,|\, \exists\, N \in \mathcal{H}_2(\mathbb{C})\ s.t.\ F = N\mathcal{M}\} \tag{4.6.5}$$

and we have

$$(Ker\, Q)^\perp = \left\{ F \in \mathcal{L}^2_\mathcal{M} \,\middle|\, \int_{\mathbb{R}^d_k} F\mathrm{d}k = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \right\}. \tag{4.6.6}$$

iv) Let $\mathcal{P} : \mathcal{L}^2_\mathcal{M} \to Ker\, Q$ be the orthogonal projection operator on $Ker\, Q$. Then we have the coercitivity relation

$$\exists\, d > 0\ s.t. \qquad -(Q(F), F)_{\mathcal{L}^2_\mathcal{M}} \geq d\, ||F - \mathcal{P}(F)||_{\mathcal{L}^2_\mathcal{M}} \qquad \forall F \in \mathcal{L}^2_\mathcal{M}. \tag{4.6.7}$$

v) The image of Q is closed and we have $Im\, Q = (Ker\, Q)^\perp$. Further, the equation $Q(F) = G$ has a solution in $\mathcal{L}^2_\mathcal{M}$ if and only if $G \in Im\, Q$. The solution is moreover unique in $(Ker\, Q)^\perp$.

Proof of Proposition 4.6.3. First we show that $Q : \mathcal{L}^2_\mathcal{M} \to \mathcal{L}^2_\mathcal{M}$ is a well defined operator, i.e. $F \in \mathcal{L}^2_\mathcal{M} \Rightarrow Q(F) \in \mathcal{L}^2_\mathcal{M}$. For this, we shall show

$$\exists\, c > 0\ s.t. \qquad ||Q(F)||^2_{\mathcal{L}^2_\mathcal{M}} \leq c\, ||F||^2_{\mathcal{L}^2_\mathcal{M}} \qquad \forall F \in \mathcal{L}^2_\mathcal{M}. \tag{4.6.8}$$

Using Lemma 4.2.6 and $\mathrm{tr}\,(P)^2 < 1$ one obtains

$$||Q(F)||^2_{\mathcal{L}^2_{\mathcal{M}}} = \int_{\mathbb{R}^d_k} \mathrm{tr}\left(Q^2(F)\right)\mathcal{M}^{-1}\mathrm{d}k \leq \int_{\mathbb{R}^d_k} \mathrm{tr}\left(P^2K^2\right)\mathcal{M}^{-1}\mathrm{d}k \tag{4.6.9}$$
$$\leq 2\int_{\mathbb{R}^d_k} \mathrm{tr}\,(P)^2\,\mathrm{tr}\,(K)^2\,\mathcal{M}^{-1}\mathrm{d}k \leq 2\int_{\mathbb{R}^d_k} \mathrm{tr}\,(K)^2\,\mathcal{M}^{-1}\mathrm{d}k.$$

By using $\mathrm{tr}\,(K^2) = \mathrm{tr}\,((K^+ - K^-)^2) \leq 2\mathrm{tr}\,((K^+)^2 + (K^-)^2)$, which follows from (4.2.15), in (4.6.9), then applying the Cauchy-Schwartz inequality and Assumption 4.6.2, one obtains

$$\begin{aligned}
||Q(F)||^2_{\mathcal{L}^2_{\mathcal{M}}} &\leq 4\int_{\mathbb{R}^d_k} \mathcal{M}\mathrm{tr}\left(\left(\int_{\mathbb{R}^d_{k'}} \phi F'\mathrm{d}k'\right)^2\right)\mathrm{d}k + 4\int_{\mathbb{R}^d_k} \lambda^2\mathrm{tr}\left(F^2\right)\mathcal{M}^{-1}\mathrm{d}k \\
&\leq 4\int_{\mathbb{R}^d_k} \mathcal{M}\mathrm{tr}\left(\left(\int_{\mathbb{R}^d_{k'}} \phi\frac{F'}{\sqrt{\mathcal{M}'}}\sqrt{\mathcal{M}'}\mathrm{d}k'\right)^2\right)\mathrm{d}k + 4\phi_2^2\int_{\mathbb{R}^d_k} \mathrm{tr}\left(F^2\right)\mathcal{M}^{-1}\mathrm{d}k \\
&\leq 4\int_{\mathbb{R}^d_k} \mathcal{M}\mathrm{tr}\left(\int_{\mathbb{R}^d_{k'}} \phi^2\frac{F'^2}{\mathcal{M}'}\mathrm{d}k'\int_{\mathbb{R}^d_{k'}} \mathcal{M}'\mathrm{d}k'\right)\mathrm{d}k + 4\phi_2^2\int_{\mathbb{R}^d_k} \mathrm{tr}\left(F^2\right)\mathcal{M}^{-1}\mathrm{d}k \\
&\leq 4\phi_2^2\int_{\mathbb{R}^d_{k'}} \mathrm{tr}\left(F'^2\right)\mathcal{M}'^{-1}\mathrm{d}k' + 4\phi_2^2\int_{\mathbb{R}^d_k} \mathrm{tr}\left(F^2\right)\mathcal{M}^{-1}\mathrm{d}k \\
&= 8\phi_2^2\,||F||^2_{\mathcal{L}^2_{\mathcal{M}}},
\end{aligned}$$

which proves (4.6.8). Let us continue with the proof of Proposition 4.6.3.

i) The linearity of $Q : \mathcal{L}^2_{\mathcal{M}} \to \mathcal{L}^2_{\mathcal{M}}$ is obvious. Since we already proved the inequality (4.6.8) we know that Q is a bounded operator and therefore continuous. The self-adjointness follows from

$$\begin{aligned}
(Q(F),G)_{\mathcal{L}^2_{\mathcal{M}}} &= \mathrm{tr}\int_{\mathbb{R}^d_k} Q(F)\frac{G}{\mathcal{M}}\mathrm{d}k = \mathrm{tr}\int_{\mathbb{R}^d_k}\int_{\mathbb{R}^d_{k'}} \phi P^{1/2}(\mathcal{M}F' - \mathcal{M}'F)P^{1/2}\frac{G}{\mathcal{M}}\mathrm{d}k'\mathrm{d}k \\
&= -\frac{1}{2}\mathrm{tr}\int_{\mathbb{R}^d_k}\int_{\mathbb{R}^d_{k'}} \phi\mathcal{M}\mathcal{M}'P^{1/2}\left(\frac{F'}{\mathcal{M}'} - \frac{F}{\mathcal{M}}\right)P^{1/2}\left(\frac{G'}{\mathcal{M}'} - \frac{G}{\mathcal{M}}\right)\mathrm{d}k'\mathrm{d}k.
\end{aligned}$$

Lemma 4.2.6 gives the non-positivity of $Q : \mathcal{L}^2_{\mathcal{M}} \to \mathcal{L}^2_{\mathcal{M}}$ by regarding

$$(Q(F),F)_{\mathcal{L}^2_{\mathcal{M}}} = -\frac{1}{2}\int_{\mathbb{R}^d_k}\int_{\mathbb{R}^d_{k'}} \phi\mathcal{M}\mathcal{M}'\mathrm{tr}\left(\left(P^{1/2}\left(\frac{F'}{\mathcal{M}'} - \frac{F}{\mathcal{M}}\right)\right)^2\right)\mathrm{d}k'\mathrm{d}k \leq 0. \tag{4.6.10}$$

ii) is trivial by integrating (4.4.13) over k.

iii) Assume that F lies in the Kernel of Q, $F \in Ker\, Q \Rightarrow Q(F) = 0$
$\Rightarrow (Q(F), F)_{\mathcal{L}^2_{\mathcal{M}}} = 0$. One obtains

$$(Q(F), F)_{\mathcal{L}^2_{\mathcal{M}}} = -\frac{1}{2}\int_{\mathbb{R}^d_k}\int_{\mathbb{R}^d_{k'}} \phi \mathcal{M}\mathcal{M}' \mathrm{tr}\left(\left(P^{1/2}\left(\frac{F'}{\mathcal{M}'} - \frac{F}{\mathcal{M}}\right)\right)^2\right) \mathrm{d}k'\mathrm{d}k = 0$$
$$\forall F \in Ker\, Q.$$

From equation (4.2.17) we deduce that, for any $F \in Ker\, Q$, the above expression is zero if and only if $F'/\mathcal{M}' = F/\mathcal{M} = 0\, \forall k, k'$, because P is strictly positive definite. This condition is fulfilled if and only if $F = N\mathcal{M}$ with $N \in \mathcal{H}_2(\mathbb{C})$ independent of k and k'. Conversely, it is verified easily from (4.4.13) that $F = N\mathcal{M}$ implies $Q(F) = 0 \Rightarrow F \in Ker\, Q$.

By definition $(Ker\, Q)^\perp = \left\{ F \in \mathcal{L}^2_{\mathcal{M}} \,\middle|\, (F, G)_{\mathcal{L}^2_{\mathcal{M}}} = 0 \ \forall G \in Ker\, Q \right\}$. This leads to

$$(F, N\mathcal{M})_{\mathcal{L}^2_{\mathcal{M}}} = \mathrm{tr}\left(N\int_{\mathbb{R}^d_k} F \mathrm{d}k\right) = 0 \ \forall N \in \mathcal{H}_2(\mathbb{C})$$
$$\Longrightarrow \quad \int_{\mathbb{R}^d_k} F \mathrm{d}k = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \ \forall F \in (Ker\, Q)^\perp.$$

iv) Since $Q : \mathcal{L}^2_{\mathcal{M}} \to \mathcal{L}^2_{\mathcal{M}}$ is linear and continuous, $Ker\, Q \subset \mathcal{L}^2_{\mathcal{M}}$ is closed. It follows the existence of an orthogonal projection $\mathcal{P} : \mathcal{L}^2_{\mathcal{M}} \to Ker\, Q$ and further $\mathcal{L}^2_{\mathcal{M}} = Ker\, Q \oplus^\perp (Ker\, Q)^\perp$ where $\oplus^\perp$ denotes the direct orthogonal sum with respect to $(\cdot,\cdot)_{\mathcal{L}^2_{\mathcal{M}}}$. We want to show that

$$\exists\, d > 0\ s.t. \qquad -(Q(F), F)_{\mathcal{L}^2_{\mathcal{M}}} \geq d\, ||F - \mathcal{P}(F)||_{\mathcal{L}^2_{\mathcal{M}}} \qquad \forall F \in \mathcal{L}^2_{\mathcal{M}}. \tag{4.6.11}$$

The case $F \in Ker\, Q$ is trivial because then $Q(F) = 0$ and $\mathcal{P}(F) = F$. Let $F \in (Ker\, Q)^\perp$. Then we have to show that $-(Q(F), F)_{\mathcal{L}^2_{\mathcal{M}}} \geq c\, ||F||_{\mathcal{L}^2_{\mathcal{M}}}$. Equation (4.6.10) yields

$$-(Q(F), F)_{\mathcal{L}^2_{\mathcal{M}}} = \frac{1}{2}\int_{\mathbb{R}^d_k}\int_{\mathbb{R}^d_{k'}} \phi \mathcal{M}\mathcal{M}' \mathrm{tr}\left(\left(P^{1/2}\left(\frac{F'}{\mathcal{M}'} - \frac{F}{\mathcal{M}}\right)\right)^2\right) \mathrm{d}k'\mathrm{d}k. \tag{4.6.12}$$

From equation (4.2.17) and the fact that the smallest eigenvalue of P reads $p_\downarrow =$

$1-|p|$ (Assumption 4.6.2), we deduce

$$\phi \mathrm{tr}\left(\left(P^{1/2}\left(\frac{F'}{\mathcal{M}'}-\frac{F}{\mathcal{M}}\right)\right)^2\right) \geq \phi_1\left(1-|p|\right)\mathrm{tr}\left(\left(\frac{F'}{\mathcal{M}'}-\frac{F}{\mathcal{M}}\right)^2\right).$$

Inserting this relation into (4.6.12) yields

$$\begin{aligned}
-\left(Q(F),F\right)_{\mathcal{L}^2_{\mathcal{M}}} &\geq \frac{1}{2}\phi_1\left(1-|p|\right)\mathrm{tr}\int_{\mathbb{R}^d_k}\int_{\mathbb{R}^d_{k'}}\mathcal{M}\mathcal{M}'\left(\frac{F'}{\mathcal{M}'}-\frac{F}{\mathcal{M}}\right)^2 \mathrm{d}k'\mathrm{d}k = \\
&= \frac{1}{2}\phi_1\left(1-|p|\right)\mathrm{tr}\int_{\mathbb{R}^d_k}\int_{\mathbb{R}^d_{k'}}\mathcal{M}\mathcal{M}'\left(\frac{F'^2}{\mathcal{M}'^2}-\frac{F'F}{\mathcal{M}'\mathcal{M}}-\frac{FF'}{\mathcal{M}\mathcal{M}'}+\frac{F^2}{\mathcal{M}^2}\right)\mathrm{d}k'\mathrm{d}k = \\
&= \phi_1\left(1-|p|\right)\mathrm{tr}\int_{\mathbb{R}^d_k}\frac{F^2}{\mathcal{M}}\mathrm{d}k,
\end{aligned}$$

where the last line follows from (4.6.6).

v) First we show that $Im\,Q$ is closed. Let $G_n := Q(F_n)$ be a sequence with $F_n \in \mathcal{L}^2_{\mathcal{M}}$ and $G_n \in Im\,Q$ so that $G_n \to G \in \mathcal{L}^2_{\mathcal{M}}$ as $n \to \infty$. One has to show that $G \in Im\,Q$, i.e. there exists $F \in \mathcal{L}^2_{\mathcal{M}}$ s.t. $Q(F) = G$. To any sequence $F_n \in \mathcal{L}^2_{\mathcal{M}}$ one can construct $H_n := F_n - \mathcal{P}F_n$ with $Q(H_n) = Q(F_n) = G_n$. One has $H_n \in (Ker\,Q)^\perp$ and therefore the coercitivity relation (4.6.7) yields

$$-\left(Q(H_n-H_m),H_n-H_m\right)_{\mathcal{L}^2_{\mathcal{M}}} \geq c\,||H_n-H_m||^2_{\mathcal{L}^2_{\mathcal{M}}}.$$

We also have

$$||Q(H_n)-Q(H_m)||_{\mathcal{L}^2_{\mathcal{M}}}\,||H_n-H_m||_{\mathcal{L}^2_{\mathcal{M}}} \geq -\left(Q(H_n-H_m),H_n-H_m\right)_{\mathcal{L}^2_{\mathcal{M}}}$$

and therefore

$$\frac{1}{c}\,||G_n-G_m||_{\mathcal{L}^2_{\mathcal{M}}} \geq ||H_n-H_m||_{\mathcal{L}^2_{\mathcal{M}}}.$$

Since $\{G_n\}_{n\in\mathbb{N}}$ is convergent (and thus a Cauchy sequence in $\mathcal{L}^2_{\mathcal{M}}$) one obtains that $\{H_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathcal{L}^2_{\mathcal{M}}$. But $\mathcal{L}^2_{\mathcal{M}}$ is complete and therefore $H_n \to H \in \mathcal{L}^2_{\mathcal{M}}$ as $n \to \infty$. Because Q is continuous, $Q(H_n) \to Q(H)$ as $n \to \infty$ and we assumed that $Q(H_n) = G_n \to G \in \mathcal{L}^2_{\mathcal{M}}$. One obtains $Q(H) = G$ with $H \in \mathcal{L}^2_{\mathcal{M}}$.

We now prove that $Q(F) = G$ has a unique solution $F \in (Ker\,Q)^\perp$. It is clear that for $G \in Im\,Q$ there exists a solution $F \in \mathcal{L}^2_{\mathcal{M}}$. Let $F \in \mathcal{L}^2_{\mathcal{M}}$ be such a solution, then $(F - \mathcal{P}F) \in (Ker\,Q)^\perp$ is also a solution. Suppose that there are two solutions

$F_a, F_b \in (Ker\, Q)^\perp$, $F_a \neq F_b$ such that $Q(F_a) = Q(F_b) = G$. One obtains

$$Q(F_a - F_b) = 0 \quad \Rightarrow \quad F_a - F_b \in Ker\, Q \cap (Ker\, Q)^\perp = \{0\} \quad \Rightarrow \quad F_a = F_b.$$

□

4.7 Diffusion limit

In this section we shall finally investigate the diffusion limit $\varepsilon \to 0$ of the matrix Boltzmann equation (4.4.18) in order to obtain a macroscopic model, which is used in the next section for some numerical experiments.

Theorem 4.7.1 (Diffusion limit). *Let $\mathcal{T}_d : \mathcal{H}_2(\mathbb{C}) \to (\mathcal{H}_2(\mathbb{C}))^d$ denote the following transport operator:*

$$\mathcal{T}_d(N) := (-\nabla_x \hat{V}(x) - \nabla_x) N(x).$$

Under the assumption 4.6.2, in the limit $\varepsilon \to 0$, the solution F^ε of the matrix Boltzmann equation (4.4.18) converges weakly to $F^0 = N(t,x)\mathcal{M}(k)$ where $N(t,x) \in \mathcal{H}_2^{0,+}(\mathbb{C})\, \forall (t,x) \in \mathbb{R}^+ \times \mathbb{R}_x^d$ and $\mathcal{M}$ stands for the scalar Maxwellian (4.4.20). Moreover, N satisfies the following drift-diffusion equation,

$$\partial_t N + \nabla_x \cdot \mathcal{J} + i \left[N, (\vec{\Omega} + \hat{\vec{g}}_{so}) \cdot \vec{\sigma} \right] - \frac{1}{2} \mathrm{tr}\,(N)\, \sigma_0 + N = 0. \tag{4.7.1}$$

The current density $\mathcal{J} \in (\mathcal{H}_2(\mathbb{C}))^d$ reads

$$\mathcal{J} = \mathbb{D} A(N), \tag{4.7.2}$$

where the diffusion matrix $\mathbb{D} \in \mathbb{R}^{d \times d}$ is given by

$$\mathbb{D} = \int_{\mathbb{R}_k^d} k \otimes \theta(k) \mathrm{d}k, \tag{4.7.3}$$

with $\theta(k) \in \mathbb{R}^d\, \forall k \in \mathbb{R}_k^d$ being the unique solution of

$$\int_{\mathbb{R}_{k'}^d} \phi(\mathcal{M}\theta' - \mathcal{M}'\theta) \mathrm{d}k' = -k\mathcal{M} \tag{4.7.4}$$

that satisfies $\int_{\mathbb{R}_k^d} \theta \mathrm{d}k = 0$. The matrix $A(N)(t,x) \in (\mathcal{H}_2(\mathbb{C}))^d\, \forall (t,x) \in \mathbb{R}^+ \times \mathbb{R}_x^d$ is given by

$$A(N) = P^{-1/2} \mathcal{T}_d(N) P^{-1/2}. \tag{4.7.5}$$

The term $\hat{\vec{g}}_{so}$ in the commutator of (4.7.1) stems from the spin-orbit contribution to the pseudo-exchange field,

$$\hat{\vec{g}}_{so}(t,x) = \int_{\mathbb{R}^d_k} \mathcal{M}(k)\hat{\vec{h}}_{so}(t,x,k)\mathrm{d}k. \tag{4.7.6}$$

Remark 4.7.2. *By expressing $\mathcal{J}$ given in (4.7.2) in the Pauli basis, $\mathcal{J} = \left(\frac{1}{2}j_0, \vec{j}\right)$ with $j_0 \in \mathbb{R}^d_x$ and $\vec{j} \in \mathbb{R}^d_x \times \mathbb{R}^3$, one obtains a coupling between $n_0 \in \mathbb{R}^+$ and $\vec{n} \in \mathbb{R}^3$, the charge- and the spin degree of freedom, respectively. Writing $N = \left(\frac{1}{2}n_0, \vec{n}\right)$ and for simplicity assuming that $\hat{\vec{g}}_{so} = 0$, the charge and spin currents read*

$$j_0 = \frac{\mathbb{D}}{1-|p|^2}\left[\mathcal{T}_d(n_0) - 2p\mathcal{T}_d(\vec{n})\cdot\vec{\Omega}\right] \tag{4.7.7}$$

$$\vec{j} = \frac{\mathbb{D}}{1-|p|^2}\left[\sqrt{1-|p|^2}\mathcal{T}_d(\vec{n}) + (1-\sqrt{1-|p|^2})(\mathcal{T}_d(\vec{n})\cdot\vec{\Omega})\vec{\Omega} - \frac{p}{2}\mathcal{T}_d(n_0)\vec{\Omega}\right]. \tag{4.7.8}$$

The respective parallel and transverse components of the spin-current with respect to the local magnetization $\vec{\Omega}$ read

$$\vec{j}\cdot\vec{\Omega} = \frac{\mathbb{D}}{1-|p|^2}\left[\mathcal{T}_d(\vec{n})\cdot\vec{\Omega} - \frac{p}{2}\mathcal{T}_d(n_0)\right] \tag{4.7.9}$$

$$\vec{j} - (\vec{j}\cdot\vec{\Omega})\vec{\Omega} = \frac{\mathbb{D}\sqrt{1-|p|^2}}{1-|p|^2}\left[\mathcal{T}_d(\vec{n}) - (\mathcal{T}_d(\vec{n})\cdot\vec{\Omega})\vec{\Omega}\right]. \tag{4.7.10}$$

4.7.1 Formal approach

Here, we present a formal proof of Theorem 4.7.1. We consider the equations obtained from the Hilbert ansatz for F^ε written in (4.4.21)-(4.4.23). Proposition 4.6.3 yields

$$Q(F^0) = 0 \quad \Longrightarrow \quad F^0(t,x,k) = N(t,x)\mathcal{M}(k) \qquad N : \mathbb{R}^+ \times \mathbb{R}^d_x \to \mathcal{H}_2(\mathbb{C}). \tag{4.7.11}$$

Moreover,

$$\int_{\mathbb{R}^d_k} \mathcal{T}(F^0)\mathrm{d}k = \nabla_x N \cdot \int_{\mathbb{R}^d_k} k\mathcal{M}\mathrm{d}k - N\nabla_x V \cdot \int_{\mathbb{R}^d_k} \nabla_k \mathcal{M}\mathrm{d}k = 0. \tag{4.7.12}$$

Therefore $\mathcal{T}(F^0) \in (Ker\, Q)^\perp = Im\, Q$ which gives the existence and uniqueness of $F^1 \in (Ker\, Q)^\perp$ s.t. $Q(F^1) = \mathcal{T}(F^0)$. From (4.4.22) one obtains

$$Q(F^1) = \mathcal{T}(F^0) = \nabla_x N \cdot k\mathcal{M} - N\nabla_x V \cdot \nabla_k \mathcal{M} = -\mathcal{T}_d(N)\cdot k\mathcal{M}. \tag{4.7.13}$$

To solve (4.7.13) we make the Ansatz $F^1 = A(t,x)\cdot\theta(k)$ with $A(t,x)\in(\mathcal{H}_2(\mathbb{C}))^d$ and $\theta(k)\in\mathbb{R}^d$ for $(t,x,k)\in\mathbb{R}^+\times\mathbb{R}^{2d}$. We obtain

$$Q(A\cdot\theta) = P^{1/2}AP^{1/2}\cdot\int_{\mathbb{R}^d_{k'}}\phi(\mathcal{M}\theta' - \mathcal{M}'\theta)\mathrm{d}k'. \tag{4.7.14}$$

From Proposition 4.6.3 we get that

$$\int_{\mathbb{R}^d_{k'}}\phi(\mathcal{M}\theta' - \mathcal{M}'\theta)\mathrm{d}k' = -k\mathcal{M} \tag{4.7.15}$$

has a unique solution $\theta\in(Ker\,Q)^\perp$ that satisfies $\int_{\mathbb{R}^d_k}\theta(k)\mathrm{d}k = 0$. It follows that $F^1 = A\cdot\theta$ is the unique solution of $Q(F^1) = -\mathcal{T}_d(N)\cdot k\mathcal{M}$ if the matrix A satisfies

$$P^{1/2}AP^{1/2} = \mathcal{T}_d(N). \tag{4.7.16}$$

Integration of equation (4.4.23) with respect to k now yields

$$\partial_t N + \int_{\mathbb{R}^d_k}\mathcal{T}(A\cdot\theta)\mathrm{d}k + i[N,(\vec{\Omega}+\hat{\vec{g}}_{so})\cdot\vec{\sigma}] - \frac{1}{2}\mathrm{tr}\,(N)\,\sigma_0 + N = 0, \tag{4.7.17}$$

where

$$\hat{\vec{g}}_{so}(t,x) = \int_{\mathbb{R}^d_k}\mathcal{M}(k)\hat{\vec{h}}_{so}(t,x,k)\mathrm{d}k. \tag{4.7.18}$$

Now we shall define, for $k\in\mathbb{R}^d$ and $\theta\in\mathbb{R}^d$, the tensor product $k\otimes\theta\in\mathbb{R}^{d\times d}$ as

$$(k\otimes\theta)_{ij} := k_i\theta_j \qquad ; \qquad i,j\in\{1,\dots,d\}. \tag{4.7.19}$$

Moreover, for $A\in(\mathcal{H}_2(\mathbb{C}))^d$, we define $A := (A_1,\dots,A_d)$ where the components $A_i\in\mathcal{H}_2(\mathbb{C})$ for $i\in\{1,\dots,d\}$. Then we shall use the following notation for the gradient with respect to $u\in\mathbb{R}^d$ of $A\in(\mathcal{H}_2(\mathbb{C}))^d$, $\nabla_u : (\mathcal{H}_2(\mathbb{C}))^d\to(\mathcal{H}_2(\mathbb{C}))^{d\times d}$,

$$(\nabla_u A)_{ij} := \partial_{u_j}A_i \qquad ; \qquad i,j\in\{1,\dots,d\}. \tag{4.7.20}$$

Finally, for $b\in\mathbb{R}^{d\times d}$ and $C\in(\mathcal{H}_2(\mathbb{C}))^{d\times d}$, $b:C\in\mathcal{H}_2(\mathbb{C})$ denotes the Frobenius product

$$b:C = \sum_{ij}b_{ij}C_{ji} \qquad ; \qquad i,j\in\{1,\dots,d\}. \tag{4.7.21}$$

With the preceeding definitions, the integral appearing in equation (4.7.17) can now be written as

$$\int_{\mathbb{R}^d_k} \mathcal{T}(A\cdot\theta)\mathrm{d}k = \int_{\mathbb{R}^d_k} [(k\otimes\theta) : \nabla_x A - A\cdot\nabla_k\theta\nabla_x\hat{V}]\mathrm{d}k = \\ = \nabla_x \cdot \left[\left(\int_{\mathbb{R}^d_k} k\otimes\theta\mathrm{d}k\right) A\right], \tag{4.7.22}$$

where $\nabla_k\theta \in \mathbb{R}^{d\times d}$ is defined by (4.7.20). Equation (4.7.22) leads to the definition of the diffusion matrix $\mathbb{D} \in \mathbb{R}^{d\times d}$ and the current density $\mathcal{J} \in (\mathcal{H}_2(\mathbb{C}))^d$, respectively,

$$\mathbb{D} := \int_{\mathbb{R}^d_k} k\otimes\theta\mathrm{d}k \tag{4.7.23}$$

$$\mathcal{J} := \mathbb{D}A, \tag{4.7.24}$$

where we mean

$$\mathcal{J}_i = (\mathbb{D}A)_i = \sum_j \mathbb{D}_{ij}A_j \qquad ; \qquad i,j \in \{1,\ldots,d\} \tag{4.7.25}$$

for the components $\mathcal{J}_i \in \mathcal{H}_2(\mathbb{C})$ of $\mathcal{J}$.

4.8 Numerical results

In this section we present some numerical solutions of the spin-coherent drift-diffusion equations (4.7.1)-(4.7.6). We will consider the one-dimensional case, $d = 1$, for different multilayer structures. The multilayers consist of alternating non-magnetic (N) and ferromagnetic (F) layers, respectively. An N-layer is characterized by $\vec{\Omega} = 0$ (no magnetization) in its domain, thus having no spin polarization of scattering rates, leading to $P = \sigma_0$ in (4.7.5). By contrast, the F-layers feature non-vanishing magnetization, $\vec{\Omega} \neq 0$, and non-vanishing spin polarization of scattering rates, $0 < p < 1$ in (4.4.8). In order to focus on the effects of $p \neq 0$ in ferromangets, we do not take into account spin-orbit couplings and assume $\hat{\vec{g}}_{so} = 0$. Moreover, to solve for θ in (4.7.4), we assume that $\phi = 1/\tau_c = const.$ is the same in every layer. This leads to $\theta = \tau_c k\mathcal{M}$ and from (4.7.3), after rescaling, one obtains the diffusion coefficient

$$\mathbb{D} = \frac{\tau_c k_B T}{m}, \tag{4.8.1}$$

where T is the temperature of the electron system and k_B stands for the Boltzmann constant. Therefore, the rescaled, one-dimensional version of (4.7.1)-(4.7.6) we consider now reads

$$\begin{cases} \partial_t N + \mathbb{D}\partial_x A(N) + i\frac{\gamma}{\hbar}\left[N, \vec{\Omega}\cdot\vec{\sigma}\right] - \frac{1}{\tau_{sf}}\left(\frac{1}{2}\mathrm{tr}\,(N)\,\sigma_0 - N\right) = 0 \\ A(N) = P^{-1/2}\left(-\frac{\partial_x V}{k_B T}N - \partial_x N\right)P^{-1/2}. \end{cases} \tag{4.8.2}$$

In the following let us denote $\eta = \sqrt{1-|p|^2}$. Applying Remark 4.7.2 to (4.8.2), the system of equations to be solved becomes

$$\begin{cases} \partial_t n_0 + \partial_x j_0 = 0 \\ \partial_t \vec{n} + \partial_x \vec{j} - \frac{2\gamma}{\hbar}\vec{n}\times\vec{\Omega} + \frac{1}{\tau_{sf}}\vec{n} = 0 \\ j_0 = \frac{\mathbb{D}}{\eta^2}\left[-\frac{\partial_x V}{k_B T}n_0 - \partial_x n_0 - 2p\left(-\frac{\partial_x V}{k_B T}\vec{n} - \partial_x\vec{n}\right)\cdot\vec{\Omega}\right] \\ \vec{j} = \frac{\mathbb{D}}{\eta^2}\Bigg\{\eta\left(-\frac{\partial_x V}{k_B T}\vec{n} - \partial_x\vec{n}\right) + (1-\eta)\left[\left(-\frac{\partial_x V}{k_B T}\vec{n} - \partial_x\vec{n}\right)\cdot\vec{\Omega}\right]\vec{\Omega} - \\ \qquad -\frac{p}{2}\left(-\frac{\partial_x V}{k_B T}n_0 - \partial_x n_0\right)\vec{\Omega}\Bigg\}, \end{cases} \tag{4.8.3}$$

where n_0 is the electron charge density and $\vec{n}$ is the non-equilibrium spin density. Initial and boundary conditions are specified for each of the investigated problems separately in the respective subsections. At interfaces between domains with different sets of parameters we require continuity of the densities n_0 respectively $\vec{n}$ and of the currents j_0 respectively $\vec{j}$. We use a standard Crank-Nicolson finite difference scheme to solve the system (4.8.3) in a three-layer and in a five-layer structure. In addition, charge and spin transport through a magnetic domain wall, which is essentially a rapid change of the direction of magnetization $\vec{\Omega}$ over some nanometers, is investigated. For all simulations, in the equations (4.8.3), we set $\mathbb{D} = 10^{-3}\,\mathrm{m^2 s^{-1}}$, $k_B T = 0.025\,\mathrm{eV}$ and $\tau_{sf} = 10^{-12}\,\mathrm{s}$. Moreover, we set $2\gamma\tau_{sf}/\hbar = 4.0$ for the simulations of the five-layer structure, c.f. section 4.8.2, and $2\gamma\tau_{sf}/\hbar = 20.0$ for the domain wall simulations, c.f. section 4.8.3. These parameter values are in the range of the parameters cited in [17]. The injected charge density is always $n_0 = 1.0$.

4.8.1 Three-layer system: $N/F/N$

The first system we investigate is composed of three layers, each of which has a thickness of 400 nm, thus the total thickness is $L = 1200\,\mathrm{nm}$. The structure is non-

metal/ferromagnet/non-metal where the interfaces are located at $x_1 = 400\,\text{nm}$ and $x_2 = 800\,\text{nm}$. In the three different domains, we choose the following parameters:

$$x \in (0, x_1]: \quad \vec{\Omega}(x) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad p(x) = 0, \tag{4.8.4}$$

$$x \in (x_1, x_2]: \quad \vec{\Omega}(x) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad p(x) = const., \tag{4.8.5}$$

$$x \in (x_2, L]: \quad \vec{\Omega}(x) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad p(x) = 0. \tag{4.8.6}$$

We apply a constant electric field that is given by $-\Delta V/L$. The system (4.8.3) is subjected to the following initial and boundary conditions:

$$\begin{cases} n_0(t=0,x) = 1.0 \quad \forall x \in (0, L] \\ \vec{n}(t=0,x) = 0 \quad \forall x \in (0, L] \\ n_0(t, x=0) = 1.0 \quad \forall t \\ n_0(t, x=L) = 1.0 \quad \forall t \\ \partial_x \vec{n}(t,x)\Big|_{x=0} = 0 \quad \forall t \\ \partial_x \vec{n}(t,x)\Big|_{x=L} = 0 \quad \forall t. \end{cases} \tag{4.8.7}$$

The system (4.8.3), (4.8.7) is solved on an equally spaced grid with 50 points in each layer. The time step was set to $0.005\,\tau_{sf}$. Grid spacing and time step were chosen such that further refinement did not change the results. We conducted two series of simulations. The results, representing steady state solutions obtained after running the system (4.8.3) for long enough time, are depicted in Figures 4.1 and 4.2, respectively. In the first series, we set $\Delta V = -1.0\,\text{V}$ and vary the parameter p of the scattering polarization in the ferromagnet. The discontinuity of p (and $\vec{\Omega}$, respectively) at the interfaces between the magnetic/non-magnetic layers acts as a source of non-equilibrium spin-polarization in z-direction as soon as a voltage is applied. This so-called 'spin injection' is a well known property of magnetic/non-magnetic multilayers [26]. In our new spin-coherent model, it arises from the particular form of the charge current density j_0 in (4.8.3). Consider the N/F interface at $x = x_1$ in Figure 4.1. Under an applied bias $\partial_x V < 0$, when there is no initial spin polarization of the current ($\vec{n} = 0$) and $n_0(x) = const.\,\forall\, x$, the 'inflow' $j_0^-(x_1)$ into the interface from $x < x_1$ is smaller than the 'outflow' $j_0^+(x_1)$ into $x > x_1$ because $\eta^2 = 1 - p^2 < 1$

when $p > 0$. In steady state we have $j_0^-(x_1) = j_0^+(x_1)$, thus the discontinuity of η has been compensated by a decrease of the charge density n_0 in the ferromagnet, which, by the equation for the spin current $\vec{j}$ in (4.8.3), leads to a non-equilibrium spin polarization $\vec{n}$ at $x = x_1$. This is a rather simplified yet intuitive explanation of what happens at the interfaces, however, the actual coupling between n_0 and $\vec{n}$ is more complicated and is only obtained from the self-consistent solution of (4.8.3). The broadening of the peaks of $\vec{n} = (0, 0, n_3)$ at the interfaces is due to spin diffusion, that is, the created spin density decreases exponentially away from the interfaces on the scale of the spin diffusion length, which is of the order of 100 nm. The peaks at the interfaces are asymmetric because the electric field drives the electrons carrying the non-equilibrium spin polarization from left to right. Curves with higher peaks of n_3 correspond to larger values of p and steeper slopes of n_0 at $x = x_1$. Note that in the domain $x > x_2$, one finds a significant reduction of the charge density n_0 for large scattering polarizations, $0.6 < p < 1$. In the second series of simulations,

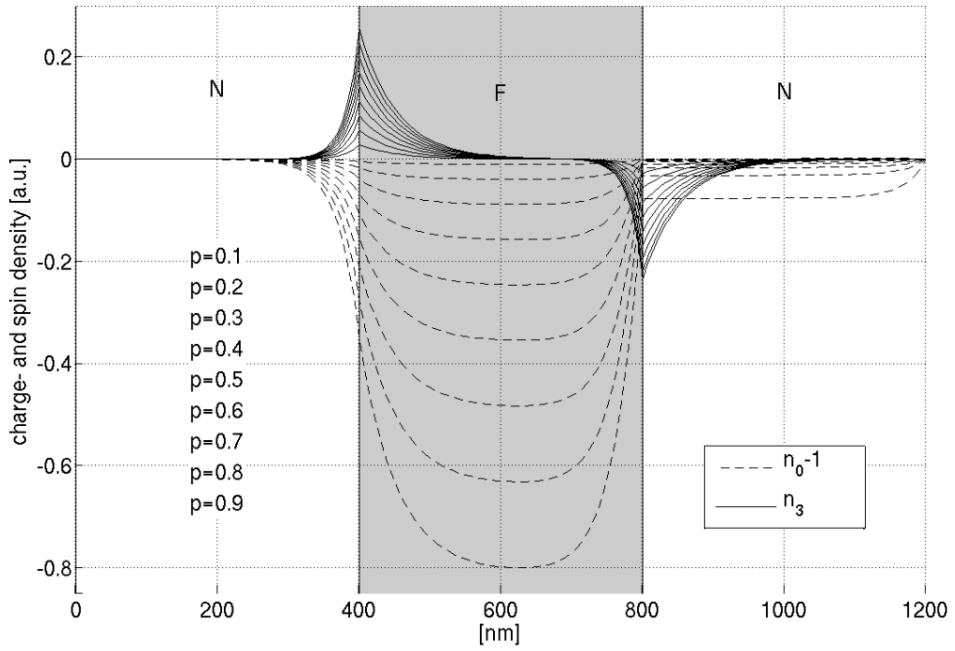


Figure 4.1: Simulated charge density $n_0 - 1$ and non-equilibrium spin density $\vec{n} = (0, 0, n_3)$ in a three-layer structure (non-magnet/ferromagnet/non-magnet) with applied voltage $\Delta V = -1.0\,\mathrm{V}$ for different values of the scattering polarization parameter p in the ferromagnet. The magnetization in the F-layer is in the z-direction, $\vec{\Omega} = (0, 0, 1)$.

depicted in Figure 4.2, we set $p = 0.33$ in the ferromagnetic domain and vary the applied voltage ΔV. Again, curves with higher peaks of n_3 correspond to larger values of ΔV and steeper slopes of n_0 at $x = x_1$. With increasing applied bias, the peaks become more and more asymmetric and the decay length (spin-diffusion length) of n_3 becomes larger.

4.8.2 Five-layer system: $N/F_1/N/F_2/N$

The second system under consideration is a five-layer system with total thickness $L = 900\,\text{nm}$. Sandwiched between two non-magnetic contact layers (200 nm each), we put two magnetic layers with a thickness of 200 nm and different directions of magnetization $\vec{\Omega}_1$ and $\vec{\Omega}_2$, respectively, which are separated by a thin non-magnetic spacer layer of 50 nm. We have interfaces at $x_1 = 200\,\text{nm}$, $x_2 = 400\,\text{nm}$, $x_3 = 450\,\text{nm}$ and $x_4 = 650\,\text{nm}$. The five domains have the following properties:

$$x \in (0, x_1]: \quad \vec{\Omega}(x) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad p(x) = 0, \tag{4.8.8}$$

$$x \in (x_1, x_2]: \quad \vec{\Omega}_1(x) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad p_1(x) = const., \tag{4.8.9}$$

$$x \in (x_2, x_3]: \quad \vec{\Omega}(x) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad p(x) = 0, \tag{4.8.10}$$

$$x \in (x_3, x_4]: \quad \vec{\Omega}_2(x) = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad p_2(x) = const., \tag{4.8.11}$$

$$x \in (x_4, L]: \quad \vec{\Omega}(x) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad p(x) = 0. \tag{4.8.12}$$

The system (4.8.3), (4.8.7) is solved on a grid with 30 points in the 200 nm layers and 10 points in the 50 nm spacer layer. The time step was set to $0.01\,\tau_{sf}$. Grid spacing and time step were chosen such that further refinement did not change the results. The applied voltage is $\Delta V = -1.0\,\text{V}$ and the scattering polarization in both magnetic layers is $p = p_1 = p_2 = 0.33$. The steady state of this system is depicted in Figure 4.3. As in the three-layer case, c.f. section 4.8.1, the interfaces x_1-x_4 act as sources of non-equilibrium spin polarization. In the 5-layer setup, the F_1-layer leads to a spin polarization in z-direction whereas the F_2-layer causes a spin polarization in y-direction. Moreover, the non-magnetic spacer layer is thin enough such that a non-vanishing component of $\vec{n}$ perpendicular to $\vec{\Omega}_2$ arrives at the interface $x_3 = 450\,\text{nm}$. The perpendicular component then rotates around $\vec{\Omega}_2$ and decays on a length scale that is determined by the strength of the exchange coupling $2\gamma\tau_{sf}/\hbar$ which was set to 4.0 in this simulation.

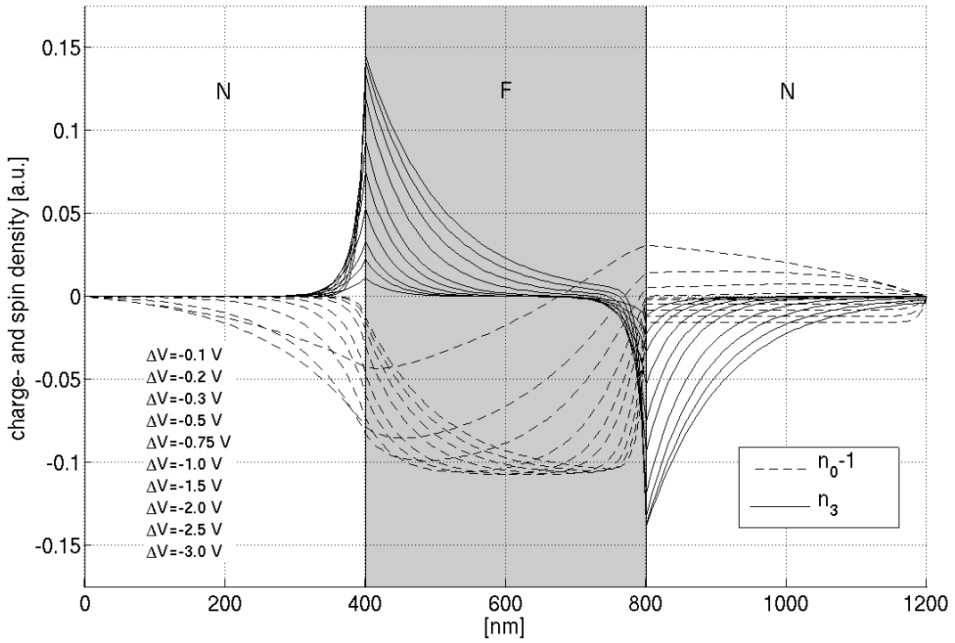


Figure 4.2: Simulated charge density $n_0 - 1$ and non-equilibrium spin density $\vec{n} = (0, 0, n_3)$ in a three-layer structure (non-magnet/ferromagnet/non-magnet) for different applied voltages ΔV. The scattering polarization in the ferromagnet is $p = 0.33$, the magnetization in the F-layer points in the z-direction, $\vec{\Omega} = (0, 0, 1)$.

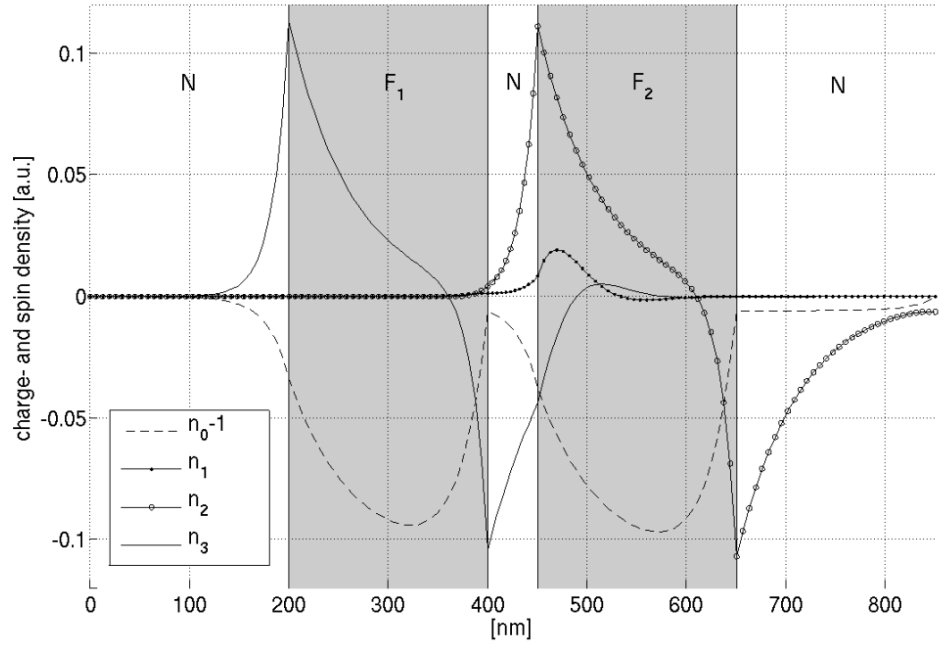


Figure 4.3: Simulated charge density $n_0 - 1$ and non-equilibrium spin density $\vec{n} = (n_1, n_2, n_3)$ in a five-layer structure (non-magnet/ferromagnet/non-magnet/ferromagnet/non-magnet) with applied voltage $\Delta V = -1.0\,\mathrm{V}$. The scattering polarization in the two ferromagnetic layers is $p = 0.33$. In the F_1-layer the magnetization points in the z-direction, $\vec{\Omega}_1 = (0, 0, 1)$, whereas in the F_2-layer the magnetization points in the y-direction, $\vec{\Omega}_2 = (0, 1, 0)$.

A perpendicular component of the spin density with respect to the local magnetization can lead to magnetization dynamics if the current density j_0 is large enough [37, 38, 39]. This so-called 'spin-transfer torque' is the subject of ongoing research in the field of microelectronics [6, 7]. The spin-coherent drift-diffusion model developed in this work, when coupled to an equation of motion for $\vec{\Omega}$ (e.g. the Landau-Lifshitz equation [40]), is a possible approach towards a better understanding of these processes. As is demonstrated in Figure 4.3, our model has the advantage that it accounts for creation of non-equilibrium spin polarization at *each* interface in a multilayer structure.

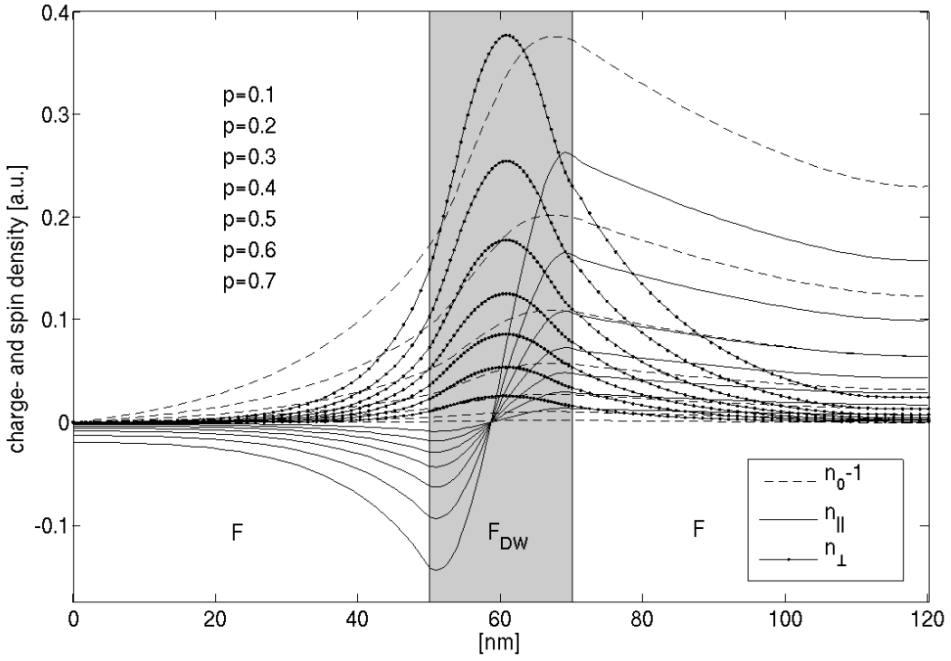


Figure 4.4: Simulated charge density $n_0 - 1$ and the parallel and perpendicular components, $n_{\parallel}$ and $n_{\perp}$, respectively, of the non-equilibrium spin density $\vec{n}$ with respect to the local magnetization $\vec{\Omega}(x)$ for different values of the scattering polarization parameter p. The respective magnetizations in the two F-domains are constant and anti-parallel and a domain wall was realized in the F_{DW}-domain, where $\vec{\Omega}(x)$ is given by (4.8.14). The applied voltage is $\Delta V = -0.2\,\mathrm{V}$.

4.8.3 Magnetic domain wall

The third system we consider is a ferromagnet with a thickness $L = 120\,\mathrm{nm}$. At $x = 0$, the magnetization $\vec{\Omega}$ points in the $+z$-direction whereas at $x = L$ it points in the $-z$-direction. In the center of the ferromagnet, between $x_1 = 50\,\mathrm{nm}$ and $x_2 = 70\,\mathrm{nm}$, we place a small region F_{DW} in which the magnetization $\vec{\Omega}$ rotates from $+z$ to $-z$ without acquiring an $x-$component, $\vec{\Omega}_1(x) = 0\,\forall\, x \in (x_1, x_2]$. The region

F_{DW} thus models a magnetic Bloch (domain) wall [8]. More precisely, we have:

$$x \in (0, x_1]: \quad \vec{\Omega}(x) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad p(x) = const., \tag{4.8.13}$$

$$x \in (x_1, x_2]: \quad \vec{\Omega}(x) = \begin{pmatrix} 0 \\ \sin\left[\frac{\pi(x-x_1)}{x_2-x_1}\right] \\ \cos\left[\frac{\pi(x-x_1)}{x_2-x_1}\right] \end{pmatrix}, \quad p(x) = const., \tag{4.8.14}$$

$$x \in (x_2, L]: \quad \vec{\Omega}(x) = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}, \quad p(x) = const. \tag{4.8.15}$$

The scattering polarization is assumed to be the same in all domains. In this section, we solve the system (4.8.3), (4.8.7) but associated with von Neumann conditions for the charge density n_0 at $x = L$,

$$\partial_x n_0(t, x)\Big|_{x=L} = 0 \quad \forall t. \tag{4.8.16}$$

The von Neumann condition was chosen so that the electron charge density can evolve freely at the right boundary of the domain. The grid has 40 points in each layer and the time step was set to $\tau_{sf} \cdot 10^{-5}$. Grid spacing and time step were chosen such that further refinement did not change the results. We conducted two series of simulations, the results being depicted in Figures 4.4 and 4.5, respectively. We plot the respective parallel and perpendicular components of $\vec{n}$ with respect to the local magnetization $\vec{\Omega}$, given by $n_\parallel = \vec{n} \cdot \vec{\Omega}$ and $n_\perp = |\vec{n} - (\vec{n} \cdot \vec{\Omega})\vec{\Omega}|$. Figure 4.4 displays the case where $\Delta V = -0.2\,\mathrm{V}$ and the scattering polarization p is modulated between 0.1 and 0.7. For increasing p, one observes increasing $n_\perp$ and stronger variations of $n_\parallel$ in the F_{DW} domain. Moreover, large p leads to a significant difference between the charge densities on the two sides of the domain wall.

In a second series, we set $p = 0.2$ and modulate the applied voltage ΔV between $-0.1\,\mathrm{V}$ and $-1.0\,\mathrm{V}$. The obtained results are displayed in Figure 4.5. Similar to the previous case, larger applied voltages lead to larger values of $n_\perp$ and stronger variations of $n_\parallel$ in the F_{DW} domain. However, at $\Delta V \approx -0.7\,\mathrm{V}$, we observe a saturation of the maximum value of the perpendicular component $n_\perp$, hence it stops increasing when the applied bias is increased further. In contrast, $n_\parallel$ and the offset of n_0 between the left and the right side of the domain wall still increase, but at a lower rate. This can be seen from the two curves for $\Delta V = -0.75\,\mathrm{V}$ and $\Delta V = -1.0\,\mathrm{V}$. To this point, we have not yet extracted the explanation for this behavior from (4.8.3).

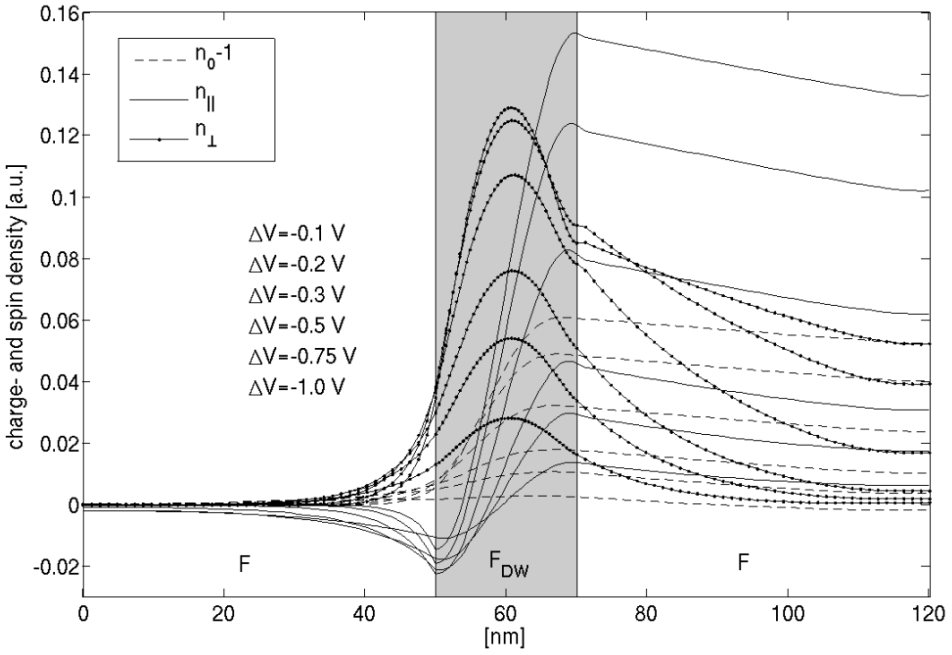


Figure 4.5: Simulated charge density $n_0 - 1$ and the parallel and perpendicular components, $n_{\|}$ and $n_{\perp}$, respectively, of the non-equilibrium spin density $\vec{n}$ with respect to the local magnetization $\vec{\Omega}(x)$ for different applied voltages ΔV. The respective magnetizations in the two F-domains are constant and anti-parallel and a domain wall was realized in the F_{DW}-domain, where $\vec{\Omega}(x)$ is given by (4.8.14). The scattering polarization is $p = 0.2$.

4.9 Conclusions

In the present work, the authors introduced four spin-coherent collision operators that yield a mathematically well-posed matrix Boltzmann equation, describing the spin-coherent electron transport in ferromagnetic structures and which incorporates spin-dependent scattering rates. Existence and uniqueness of a weak solution to this equation were shown in Proposition 4.5.5. Moreover, the maximum principle was verified for the anti-symmetric collision operator Q_{21} in Proposition 4.5.6. Assuming parabolic spin bands with momentum-independent band gap (Stoner model), further assuming Boltzmann statistics and applying the condition of detailed balance, the symmetric collision operator Q_{22} was investigated from a rigorous mathematical point of view (Proposition 4.6.3). We then performed the diffusion limit in the scaled matrix Boltzmann equation, the small parameter ε^2 being the ratio of the respective time scales of spin conserving and spin altering collision processes. The obtained spin-coherent drift-diffusion equation (Theorem 4.7.1) contains a coupling between the charge- and the spin degree of freedom of the electron system that is

linear in the polarization p of the scattering rates. The new macroscopic model was applied to simulate spin-polarized transport in three different one-dimensional structures, namely a three- and a five-layer magnetic/non-magnetic multilayer and a magnetic domain wall. The simulations show that our model can improve the understanding of spin-polarized electron transport, which is important in spintronic research fields such as spin-transfer torque devices and current-induced domain wall motion.

4.A Computation of the coefficients ω_{ij} and γ_{ij}

Here, we shall compute the coefficients ω_{ij} respectively γ_{ij}, $i,j \in \{1,2\}$, appearing in equation (4.5.23) for the eigenvalue f_- of the distribution matrix F. Starting from (4.5.22) and looking at (4.5.18) and (4.5.21) we deduce

$$\Pi_-(Q_j^-(F)) = \frac{1}{2}\mathrm{tr}\left(Q_j^-(F)\right) - \frac{1}{2}\mathrm{tr}\left(\vec{\sigma}Q_j^-(F)\right)\cdot\frac{\vec{f}}{|\vec{f}|}. \tag{4.A.1}$$

The loss terms $Q_j^-(F)$ are defined in (4.3.15)-(4.3.16). We recall that $\Lambda = \int S\mathrm{d}k'$, $\Lambda = \lambda_0\sigma_0 + \vec{\lambda}\cdot\vec{\sigma}$ and $F = \frac{1}{2}f_0\sigma_0 + \vec{f}\cdot\vec{\sigma}$ in the Pauli basis. From Lemma 4.2.4 we deduce

$$\frac{1}{2}\mathrm{tr}\left(Q_1^-(F)\right) = \frac{1}{2}\mathrm{tr}\left(Q_2^-(F)\right) = \frac{1}{2}\lambda_0 f_0 + \vec{\lambda}\cdot\vec{f}. \tag{4.A.2}$$

Moreover, we have

$$\frac{1}{2}\mathrm{tr}\left(\vec{\sigma}Q_1^-(F)\right)\cdot\frac{\vec{f}}{|\vec{f}|} = \lambda_0|\vec{f}| + \frac{1}{2}f_0\vec{\lambda}\cdot\frac{\vec{f}}{|\vec{f}|} \tag{4.A.3}$$

$$\begin{aligned}\frac{1}{2}\mathrm{tr}\left(\vec{\sigma}Q_2^-(F)\right)\cdot\frac{\vec{f}}{|\vec{f}|} &= \frac{1}{2}f_0\vec{\lambda}\cdot\frac{\vec{f}}{|\vec{f}|} + |\vec{f}|\int_{\mathbb{R}^d_{k'}}\sqrt{s_0^2-|\vec{s}|^2}\mathrm{d}k' + \\ &\quad + |\vec{f}|\int_{\mathbb{R}^d_{k'}}\left(s_0-\sqrt{s_0^2-|\vec{s}|^2}\right)\left(\frac{\vec{f}}{|\vec{f}|}\cdot\frac{\vec{s}}{|\vec{s}|}\right)^2\mathrm{d}k'.\end{aligned} \tag{4.A.4}$$

Let us introduce the angle η between $\vec{\lambda}$ and $\vec{f}$ via

$$\cos(\eta) = \frac{\vec{\lambda}}{|\vec{\lambda}|}\cdot\frac{\vec{f}}{|\vec{f}|} = \frac{\vec{s}}{|\vec{s}|}\cdot\frac{\vec{f}}{|\vec{f}|}, \tag{4.A.5}$$

where the second equality is a consequence of (4.3.11), stating that the direction $\vec{s}/|\vec{s}|$ of S must not depend on k'. Then, inserting (4.A.2) and (4.A.3) respectively

(4.A.4) into (4.A.1), a straightforward calculation yields

$$\Pi_-(Q_1^-(F)) = f_-\left(\lambda_0 - |\vec{\lambda}|\cos(\eta)\right) \tag{4.A.6}$$

$$\Pi_-(Q_2^-(F)) = f_-\left(\lambda_0 - |\vec{\lambda}|\cos(\eta)\right) + |\vec{f}|\sin^2(\eta)\int_{\mathbb{R}^d_{k'}}\left(s_0 - \sqrt{s_0^2 - |\vec{s}|^2}\right)\mathrm{d}k' \tag{4.A.7}$$

Inserting (4.A.6) respectively (4.A.7) into (4.5.22) we obtain the coefficients ω_{ij} and γ_{ij}, $i,j \in \{1,2\}$, defined in (4.5.23),

$$\omega_{11} = \omega_{12} = \omega_{21} = \omega_{22} = \frac{1}{\tau_c}\left(\lambda_0 - |\vec{\lambda}|\cos(\eta)\right) \tag{4.A.8}$$

$$\gamma_{11} = \frac{1}{\tau_c}\Pi_-(G_1) + \frac{1}{\tau_{sf}}|\vec{f}| \tag{4.A.9}$$

$$\gamma_{21} = \frac{1}{\tau_c}\Pi_-(G_2) + \frac{1}{\tau_{sf}}|\vec{f}| \tag{4.A.10}$$

$$\gamma_{12} = \frac{1}{\tau_c}\Pi_-(G_1) + |\vec{f}|\left(\frac{1}{\tau_{sf}} - \frac{1}{\tau_c}\sin^2(\eta)\int_{\mathbb{R}^d_{k'}}\left(s_0 - \sqrt{s_0^2 - |\vec{s}|^2}\right)\mathrm{d}k'\right) \tag{4.A.11}$$

$$\gamma_{22} = \frac{1}{\tau_c}\Pi_-(G_2) + |\vec{f}|\left(\frac{1}{\tau_{sf}} - \frac{1}{\tau_c}\sin^2(\eta)\int_{\mathbb{R}^d_{k'}}\left(s_0 - \sqrt{s_0^2 - |\vec{s}|^2}\right)\mathrm{d}k'\right). \tag{4.A.12}$$

Acknowledgments

The authors would like to send a special and warm thank to their PhD advisor, Naoufel Ben Abdallah, from whom they learned a lot and who was also at the origin of the present work. This work has been supported by the Austrian Science Fund, Vienna, under the contract number P21326-N16 and by the ANR QUATRAIN ("Quantum transport in nanoscale structures").

References

[1] D. Loss and D.P. DiVincenzo. Quantum computation with quantum dots. *Phys. Rev. A*, 57:120–126, 1998.

[2] M. Wenin and W. P otz. Optimal control of a single qubit by direct inversion. *Phys. Rev. A*, 74:022319, 2006.

[3] A. Fert. Nobel lecture: Origin, development, and future of spintronics. *Rev. Mod. Phys.*, 80:1517, 2008.

[4] I. Zutic, J. Fabian, and Das Sarma S. Spintronics: Fundamentals and applications. *Rev. Mod. Phys.*, 76(2):323–410, 2004.

[5] C. Ertler, A. Matos-Abiague, M. Gmitra, M. Turek, and J. Fabian. Perspectives in spintronics: magnetic resonant tunneling, spin-orbit coupling, and gamnas. *J. Phys. Conf. Ser.*, 129:012021, 2008.

[6] M.D. Stiles and J. Miltat. Spin-transfer torque and dynamics. *Top. Appl. Phys.*, 101:225, 2006.

[7] J.A. Katine and E.E. Fullerton. Device implications of spin-transfer torques. *J. Magn. Magn. Mater.*, 320:1217–1226, 2008.

[8] G.S.D. Beach, M. Tsoi, and J.L. Erskine. Current-induced domain wall motion. *J. Magn. Magn. Mater.*, 320:1272–1281, 2008.

[9] M.I. Dyakonov, editor. *Spin Physics in Semiconductors*. Springer-Verlag Berlin Heidelberg, 2008.

[10] H.A. Engel, E.I. Rashba, and B.I. Halperin. *Theory of Spin Hall Effects in Semiconductors*. Handbook of Magnetism and Advanced Magnetic Materials. John Wiley & Sons, Ltd, 2007.

[11] L.E. Ballentine. *Quantum Mechanics - A Modern Development*. World Scientific Publishing, 1998.

[12] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, M. I. Katsnelson, I. V. Grigorieva, S. V. Dubonos, and A. A. Firsov. Two-dimensional gas of massless dirac fermions in graphene. *Nature*, 438:197–200, 2005.

[13] E. Simanek. Spin accumulation and resistance due to a domain wall. *Phys. Rev. B*, 63:224412, 2001.

[14] J. Zhang, P.M. Levy, S. Zhang, and V. Antropov. Identification of transverse spin currents in noncollinear magnetic structures. *Phys. Rev. Lett.*, 93(256602), 2004.

[15] F. Piéchon and A. Thiaville. Spin transfer torque in continuous textures: Semiclassical boltzmann approach. *Phys. Rev. B*, 75:174414, 2007.

[16] L. Barletti and G. Frosali. Diffusive limit of the two-band k.p model for semiconductors. *J. Stat. Phys.*, 139(2):280–306, 2010.

[17] S. Zhang, P.M. Levy, and A. Fert. Mechanisms of spin-polarized current-driven magnetization switching. *Phys. Rev. Lett.*, 88(23):236601–1, 2002.

[18] S. Saikin. A drift-diffusion model for spin-polarized transport in a two-dimensional non-degenerate electron gas controlled by spin?orbit interaction. *J. Phys.: Condens. Matter*, 16:5071–5081, 2010.

[19] L. Barletti and F. Méhats. Quantum electronic transport in graphene: A kinetic and fluid-dynamic approach. *Mathematical Methods in Applied Sciences*, 34(7):807–818, 2010.

[20] O. Morandi and F. Sch urrer. Wigner model for quantum transport in graphene. *J. Phys. A (accepted)*, 2010.

[21] L. Barletti and F. Méhats. Quantum drift-diffusion modeling of spin transport in nanostructures. *J. Math. Phys.*, 51:053304, 2010.

[22] R. El Hajj. Diffusion models for spin transport derived from the spinor boltzmann equation. *Comm. Math. Sci. (to appear).*

[23] I.A. Campbell, A. Fert, and A.R. Pomeroy. Evidence for two current conduction in iron. *Phil. Mag.*, 15:977, 1967.

[24] A. Fert and I.A. Campbell. Electrical resistivity of ferromagnetic nickel and iron based alloys. *J Phys. F: Metal Phys.*, 6(5):849, 1976.

[25] M. Johnson and R.H. Silsbee. Interfacial charge-spin coupling: Injection and detection of spin magnetization in metals. *Phys. Rev. Lett.*, 55(17):1790, 1985.

[26] P.C. van Son, H. van Kempen, and P. Wyder. Boundary resistance of the ferromagnetic-nonferromagnetic metal interface. *Phys. Rev. Lett.*, 58(21):2271, 1987.

[27] R.Q. Hood and R.M. Falicov. Boltzmann-equation approach to the negative magnetoresistance of ferromagnetic-normal-metal multilayers. *Phys. Rev. B*, 46(13):8287, 2007.

[28] T. Valet and A. Fert. Theory of the perpendicular magnetoresistance in magnetic multilayers. *Phys. Rev. B*, 48(10):7099, 1993.

[29] J. Xiao, A. Zangwill, and M.D. Stiles. A numerical method to solve the boltzmann equation for a spin valve. *Eur. Phys. J. B*, 59:415–427, 2007.

[30] L. Villegas-Lelovsky. Hydrodynamic model for spin-polarized electron transport in semiconductors. *J. Appl. Phys.*, 101:053707, 2007.

[31] F. Poupaud. Diffusion approximation of the linear semiconductor equation: analysis of boundary layers. *Asympt. Anal.*, 4:293–317, 1991.

[32] O. Gunnarsson. Band model for magnetism of transition metals in the spindensity-functional formalism. *J. Phys. F: Metal Phys.*, 6(4):587, 1976.

[33] W. Nolting and A. Ramakanth. *Quantum Theory of Magnetism.* Springer-Verlag Berlin Heidelberg, 2009.

[34] R. El Hajj. *Etude mathématique et numérique de modèles de transport: application à la spintronique.* PhD thesis, Institut de Mathématiques de Toulouse (IMT), Université Paul Sabatier, 2008.

[35] C. Vouille, A. Barthélémy, F. Elokan Mpondo, A. Fert, P. A. Schroeder, S. Y. Hsu, A. Reilly, and R. Loloee. Microscopic mechanisms of giant magnetoresistance. *Phys. Rev. B*, 60(9):6710, 1999.

[36] D.C. Ralph and M.D. Stiles. Spin transfer torques. *J. Magn. Magn. Mater.*, 320:1190–1216, 2008.

[37] J.C. Slonczewski. Current-driven excitation of magnetic multilayers. *J. Magn. Magn. Mater.*, 159:L1–L7, 1996.

[38] L. Berger. Emission of spin waves by a magnetic multilayer traversed by a current. *Phys. Rev. B*, 54(13):9353, 1996.

[39] M. Tsoi, A.G.M. Jansen, J. Bass, W.-C. Chiang, M. Seck, V. Tsoi, and P. Wyder. Excitation of a magnetic multilayer by an electric current. *Phys. Rev. Lett.*, 80(19):4281, 1998.

[40] D.V. Berkov and J. Miltat. Spin-torque driven magnetization dynamics: Micromagnetic modeling. *J. Magn. Magn. Mater.*, 320:1238–1259, 2008.

Part III

Dissipation and Decoherence in Open Quantum Systems

Chapter 5

Non-Markovian Quantum Dynamics from Environmental Relaxation

S. Possanner and *B. A. Stickler*,
published in *Physical Review A*, Vol. 85, Nr. 6, p. 062115, **2012**

Abstract. We consider the dynamics of composite quantum systems in the particular case that the state operator relaxes towards the Born approximation. For this we augment the von Neumann equation by a relaxation operator imposing a finite relaxation time τ_r. Under the premise that the relaxation is the dominant process we obtain a hierarchy of non-Markovian master equations. The latter arises from an expansion of the total state operator in powers of the relaxation time τ_r. In the Born-Markov limit $\tau_r \to 0$ the Lindblad master equation is recovered. Higher order contributions enable a systematic treatment of correlations and non-Markovian dynamics in a recursive manner.

5.1 Introduction

The notion of quantum dissipation and decoherence arising from system-environment coupling is becoming increasingly important in many branches of physics such as quantum computation [1], quantum optics [2], or semiconductor spintronics [3]. The progress in atomic and molecular interferometry made over the last decade [4, 5, 6] enables the testing of these important concepts of the theory of open quantum systems [7, 8, 9]. The latter is the most prominent tool for tackling such fundamental problems as the collapse of the wave function during measurement [10, 11] or the transition between the micro- and the macroscopic world in general [12].

The peculiar nature of quantum states (coherent, delocalized, correlated, entangled) makes the treatment of non-equilibrium processes considerably more complicated than in the classical case. The usual approach is to start from a closed quantum system consisting of interacting degrees of freedom A and B. The state operator ρ of the composite system AB undergoes unitary (Hamiltonian) time evolution,

$$\partial_t \rho = -i[H, \rho] \,, \tag{5.1.1}$$

where H denotes the system's Hamiltonian, the square brackets $[\cdot, \cdot]$ stand for the commutator and we set the reduced Planck constant $\hbar = 1$. In the composite state space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$, the most general form of the Hamiltonian H reads

$$H = H_A \otimes \mathbb{1}_B + \mathbb{1}_A \otimes H_B + H_I \,, \tag{5.1.2}$$

where the operator subscript $A(B)$ indicates an operator acting in $\mathcal{H}_A$ ($\mathcal{H}_B$), $\mathbb{1}_A$ ($\mathbb{1}_B$) denotes the identity and the operator H_I accounts for the interactions between A and B. Taking the partial trace, $\mathrm{tr}_B(\cdot)$, over the subsystem B in the von Neumann equation (5.1.1) yields the exact equation of motion for the "relevant" degrees of freedom A, i.e.

$$\partial_t \rho_A = -i[H_A, \rho_A] - \mathrm{tr}_B \left(i[H_I, \rho] \right) \,, \tag{5.1.3}$$

where we introduced the reduced state operator ρ_A via

$$\rho_A := \mathrm{tr}_B(\rho) \,. \tag{5.1.4}$$

In general, the reduced equation of motion (5.1.3) is an integro-differential equation, featuring memory effects in B that cause the second term on the right-hand-side to be non-local in time. It describes the subsystem A as an open quantum system that exchanges energy with the environment B. In the special case of Markovian time

evolution, memory effects become negligible and equation (5.1.3) takes on the form

$$\partial_t \rho_A = \mathcal{L}\rho_A \,. \tag{5.1.5}$$

Here, the operator $\mathcal{L}$ is the infinitesimal generator of a dynamical semigroup [13, 14, 15]. In its most general form $\mathcal{L}$ is given by [16]

$$\begin{aligned} \mathcal{L}\rho_A &= -i[H_A^{eff}, \rho_A] \\ &+ \sum_{k=0}^{K-1} \Gamma_k \left(L_k^\dagger \rho_A L_k - \frac{1}{2} L_k^\dagger L_k \rho_A - \frac{1}{2} \rho_A L_k^\dagger L_k \right) . \end{aligned} \tag{5.1.6}$$

where H_A^{eff} is an effective Hamiltonian, $\Gamma_k \geq 0$ are transition rates (channels) and L_k is an operator basis in the K-dimensional space [1] of hermitian operators in $\mathcal{H}_A$. Equation (5.1.5) is commonly referred to as a master equation of Lindblad form, or Lindblad master equation. The second term on the right hand side of the generator (5.1.6) may account for quantum decoherence as well as dissipation in A due to interactions with its environment B. Master equations of the Lindblad form (5.1.5) are frequently encountered in various fields of quantum physics, in particular in the context of quantum Brownian motion or quantum optics [18, 19, 20, 21, 22, 23, 24, 25, 26].

The Lindblad master equation is obtained by performing Markovian approximations to the exact dynamics (5.1.3). This usually means that a typical parameter α of the composite system such as the correlation time, mass ratio or timescale ratio, tends towards zero or infinity [7, 8, 9]. The dynamics (5.1.5) are, therefore, only exact in the respective limiting case, which might not necessarily be a good approximation of the physical system considered. It is, thus, desirable to study the corrections to the Markovian case (5.1.5) which arise when the limiting parameter mentioned above is small, but not zero (or large, but still finite). One expects to obtain non-Markovian corrections which account for correlations between system A and environment B. The enhanced model will be more difficult to treat, but it should still be much less involved than a full treatment of the composite system AB.

Over the last decade, considerable effort has been put into the derivation of non-Markovian corrections to the Lindblad master equation (5.1.5) [27, 28, 29, 30, 31, 32, 33, 34]. Two well-established approaches proved to be particularly fruitful, i.e. the projection operator technique and the time-convolutionless projection operator method. The projection operator technique results in the Nakajima-Zwanzig equation [35, 36, 37] which is an exact equation for open quantum systems and its

[1] To the knowledge of the authors, rigorous proofs for the existence of (5.1.6) in the case $K = \infty$ and for unbounded Hamiltonians H are still lacking.

solution is comparably difficult to the solution of Eq. (5.1.1). A series expansion of the Nakajima-Zwanzig integral kernel yields non-Markovian evolution equations which are non-local in time [7]. This drawback is remedied by elimination of the non-locality in time with the help of a back propagator as developed by Shibata et al. [38, 39]. The resulting equation is referred to as the time-convolutionless master equation and it provides the means for the derivation of time-local, non-Markovian contributions to Eq. (5.1.5) in ascending orders of the coupling strength between degrees of freedom A and B.

It might be interesting to note that the projection techniques described above have been motivated by a technical point of view. The aim is to eliminate from the von Neumann equation the irrelevant degrees of freedom B without employing any further assumptions on the dynamics of the physical system. Subsystem B is usually described by an arbitrary reference state χ_B, which is why a physical interpretation of the results obtained appears to be difficult. Nevertheless, these methods are exact, but difficult to treat in the general case.

In this work we present an alternative approach towards non-Markovian contributions to Eq. (5.1.5). This approach is based on a particular physical picture and is closely related to the diffusion limit of the linear Boltzmann equation in classical kinetic theory [40, 41, 42]. In our approach, the non-Markovianity arises from the relaxation of parts of the environment towards an equilibrium state χ_B on a finite timescale τ_r. By explicitly accounting for this relaxation process by means of a relaxation operator Q in Eq. (5.1.1), we use a Hilbert expansion technique to derive a hierarchy of master equations for subsystem A. In the limit $\tau_r \to 0$, we retrieve the Lindblad master equation (5.1.5). It has to be emphasized that by introducing the operator Q we depart from the exact description of the system's dynamics. However, this approach as well as the resulting equations of motion follow a clear physical picture and, therefore, allow for an easy interpretation.

The paper is organized as follows. In section 6.2 we specify the physical picture of our approach. Moreover, we introduce the relaxation operator Q and a scaled version of the resulting equation of motion for the state operator ρ of the composite system AB. In section 5.3 we employ a Hilbert expansion of ρ and derive a hierarchy of master equations for the reduced state operator ρ_A. Section 5.4 contains a discussion of the results obtained. The paper is summarized in section 5.5 and a short outlook for possible future work is presented. A mathematical analysis of the relaxation operator Q as well as the proof of existence and uniqueness of solutions of the equation of motion for ρ can be found in the appendices 5.A and 5.B, respectively. In App. C we explicitly compute the second order contribution to the hierarchy of master equations obtained.

5.2 Physical model and scaling

A common approximation to the state operator ρ of a composite system AB in which subsystem A obeys Markovian dynamics is

$$\rho(t) = \rho_A(t) \otimes \chi_B \,, \tag{5.2.1}$$

where $\rho_A(t)$ is the solution of Eq. (5.1.5) and χ_B is some reference state in the environment B. This approximation is known as the Born approximation. It clearly depends on the physical system whether or not the state (5.2.1) represents a good approximation to the exact solution of Eq. (5.1.1). For systems where this is not the case, it might be desirable to have corrections to the Born approximation that can be expanded in orders of a typical parameter α which is zero in the Markovian limit. In order to achieve this, let us regard the Born approximation as a sort of equilibrium state of the composite system AB and let τ_r denote the corresponding relaxation time. We shall explicitly account for the relaxation of ρ towards the Born approximation by rewriting the equation of motion (5.1.1) as

$$\partial_t \rho = -i[H, \rho] + \frac{1}{\tau_r} Q(\rho) \,. \tag{5.2.2}$$

Here we introduced the relaxation operator Q as

$$Q(\rho) := \mathrm{tr}_B \left(\rho\right) \otimes \chi_B - \rho \,, \tag{5.2.3}$$

where $\mathrm{tr}_B \left(\chi_B\right) = 1$ and we remark that

$$\mathrm{tr}_B \left(Q(\rho)\right) = 0 \quad \forall \, \rho \,. \tag{5.2.4}$$

In what follows the limit $\tau_r \to 0$ in Eq. (5.2.2) will be denoted as the Born-Markov limit. Hence, taking in Eq. (5.2.2) the partial traces over degrees of freedom A and B, respectively, yields

$$\partial_t \rho_A = -i[H_A, \rho_A] - \mathrm{tr}_B \left(i[H_I, \rho]\right) \,, \tag{5.2.5}$$

$$\partial_t \rho_B = -i[H_B, \rho_B] - \mathrm{tr}_A \left(i[H_I, \rho]\right) + \frac{\chi_B - \rho_B}{\tau_r} \,, \tag{5.2.6}$$

where $\mathrm{tr}_A \left(\rho\right) = \rho_B$ is the reduced state operator of the environment B. Although Eqs. (5.1.3) and (5.2.5) might seem to be identical on a first glance, the total state operator ρ will be different in these two equations, because of the introduction of the relaxation operator Q in Eq. (5.2.2). It depends on the particular situation whether

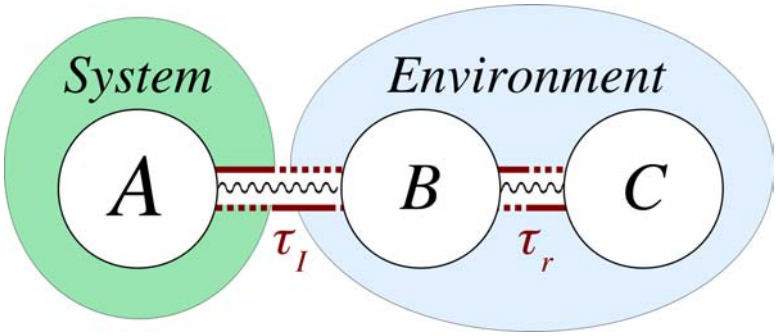


Figure 5.1: (Color online) Schematic representation of an observed "System" A that interacts with an "Environment" that consists of two parts B and C, respectively. Subsystem B interacts with A on a time scale τ_I. Moreover, on a timescale τ_r, B may exchange energy (information) with subsystem C, which is assumed to be completely isolated from the observed system A.

the Hamiltonian H_B contains interactions within the environment B or whether these interactions have been absorbed into the relaxation term Eq. (5.2.6).

We point out that Eq. (5.2.2) does not conserve the total energy of the composite system AB, which is, consequently, not a closed quantum system. Eq. (5.2.2) rather resembles a configuration in which subsystem B is coupled to a third subsystem C, which can be regarded as isolated from A. This situation is sketched in Figure 5.1. Hence, the environment of A is a composite system BC. In this case Eq. (5.2.2) results from tracing out the degrees of freedom C from the total equation of motion for the composite system ABC.

The remaining effect of subsystem C is that it relaxes the state operator ρ_B to a particular equilibrium state χ_B on a timescale τ_r. In the case that C is a reservoir, i.e. features an infinite number of degrees of freedom, χ_B could be the minimizer of a certain entropy functional in B. For instance, system A could contain the conduction band electrons in a semiconductor, whereas system B describes the lattice phonons coupled to an external heat bath C. On the other hand, one could imagine that a probe C prepares the state χ_B with a mean frequency $1/\tau_r$. Such a scenario could be realized by two interacting spins, where one of the two spins is constantly monitored and prepared to be in state χ_B. Another possible scenario could be a composite quantum system, where subsystem A interacts solely with a part of the total environment due to short range iteractions.

If the state χ_B is a pure state the corresponding state of the composite system AB must be uncorrelated [43], i.e. of the form (5.2.1). In writing Eq. (5.2.2), we presuppose that even for a mixed state χ_B, the coupling of C to B leads to decorrelations in AB. Thus, correlations between A and B due to the interaction H_I are gradually destroyed on a timescale τ_r by the coupling of B to C.

The aim of the following sections is to find approximate solutions to Eq. (5.2.2) in cases where the time scale τ_r is small compared to all other relevant timescales of the system. For technical reasons which will become clear in the next section, let us introduce the mean-field operator H_A^{mf} acting in $\mathcal{H}_A$,

$$H_A^{mf} := \mathrm{tr}_B \left(H_I \chi_B \right) . \tag{5.2.7}$$

We define furthermore,

$$\begin{aligned} \widetilde{H}_I &:= H_I - H_A^{mf} \otimes \mathbb{1}_B \,, \\ \widetilde{H}_A &:= H_A + H_A^{mf} \,, \\ \widetilde{H}_{AB} &:= \widetilde{H}_A \otimes \mathbb{1}_B + \mathbb{1}_A \otimes H_B \,, \end{aligned} \tag{5.2.8}$$

and rewrite Eq. (5.2.2) with the help of the definitions (5.2.8):

$$\partial_t \rho = -i[\widetilde{H}_{AB}, \rho] - i[\widetilde{H}_I, \rho] + \frac{1}{\tau_r} Q(\rho) \,. \tag{5.2.9}$$

As a next step we present a scaled version of Eq. (5.2.9) which is appropriately suited for the Born-Markov limit. For this suppose one can define a timescale τ_{AB} induced by $\widetilde{H}_{AB}$ as well as a timescale τ_I induced by $\widetilde{H}_I$. The former timescale is a characteristic for the evolution of the isolated, mean-field-corrected subsystems A and B, respectively, whereas the latter is a characteristic for the mean-field-corrected interaction between A and B. The introduction of a typical parameter $\alpha \ll 1$ via

$$\frac{\tau_I}{\tau_{AB}} = \mathcal{O}(\alpha) \quad , \quad \frac{\tau_r}{\tau_{AB}} = \mathcal{O}(\alpha^2) \,, \tag{5.2.10}$$

and of the timescale τ_{AB} to describe the dynamics,

$$t' = \frac{t}{\tau_{AB}} \,, \tag{5.2.11}$$

yields

$$\partial_{t'} \rho = -i[\widetilde{H}_{AB}, \rho] - \frac{i}{\alpha}[\widetilde{H}_I, \rho] + \frac{1}{\alpha^2} Q(\rho) \,. \tag{5.2.12}$$

Equation (5.2.12) corresponds to the equation of motion (5.2.9) for the composite system AB in the Born-Markov scaling. We remark that since $\alpha \ll 1$, Eq. (5.2.12) implies strong interactions between system A and environment B while the relaxation towards the Born approximation is the dominant process.

5.3 Derivation of master equations

5.3.1 Hilbert expansion of the state operator

It is the aim of this section to search for an approximate solution $\rho^{(\alpha)}(t')$ of Eq. (5.2.12) with initial condition $\rho^{(\alpha)}(0) = \rho_i^{(\alpha)}$. For small values of α this approximate solution is supposed to be close to the exact solution ρ. In what follows we write t instead of t' for the scaled time (5.2.11). Thus, we consider the following initial value problem,

$$\begin{aligned} &\partial_t \rho^{(\alpha)} = -i[\widetilde{H}_{AB}, \rho^{(\alpha)}] - \frac{i}{\alpha}[\widetilde{H}_I, \rho^{(\alpha)}] + \frac{1}{\alpha^2} Q(\rho^{(\alpha)}) \,, \\ &\rho^{(\alpha)}(0) = \rho_i^{(\alpha)} \,. \end{aligned} \tag{5.3.1}$$

The first question of interest is whether or not the initial value problem (5.3.1) has a unique solution. The proof of existence and uniqueness of a solution $\rho^{(\alpha)}(t)$ on a finite time interval $[0, T]$ to the initial value problem (5.3.1) is given in App. 5.B.

Let us proceed with the approximate solution of Eq. (5.3.1). We shall employ a series expansion of the solution in powers of α, thus assuming $\rho^{(\alpha)}$ to be analytic in α within a certain radius around $\alpha = 0$. By inserting the Hilbert expansion into Eq. (5.3.1),

$$\rho^{(\alpha)} = \sum_{n=0}^{\infty} \alpha^n \rho_n \,, \tag{5.3.2}$$

subsequently multiplying by α^2 and sorting the terms in orders of α, one obtains the following system of equations

$$Q(\rho_0) = 0 \,, \tag{5.3.3a}$$

$$Q(\rho_1) = i[\widetilde{H}_I, \rho_0] \,, \tag{5.3.3b}$$

$$Q(\rho_2) = \partial_t \rho_0 + i[\widetilde{H}_{AB}, \rho_0] + i[\widetilde{H}_I, \rho_1] \,, \tag{5.3.3c}$$

$$Q(\rho_3) = \partial_t \rho_1 + i[\widetilde{H}_{AB}, \rho_1] + i[\widetilde{H}_I, \rho_2] \,, \tag{5.3.3d}$$

$$Q(\rho_n) = \partial_t \rho_{n-2} + i[\widetilde{H}_{AB}, \rho_{n-2}] + i[\widetilde{H}_I, \rho_{n-1}] \,, \tag{5.3.3e}$$

$$\text{for } n \geq 4 \,.$$

We remark that even though Eqs. (5.3.3c) and (5.3.3d) are of the general form (5.3.3e), they have been written explicitly for the purpose of a better understanding of the concepts elaborated in this section.

Regarding Eqs. (5.3.3), the question immediately arises whether or not the system is well-posed, i.e. whether or not the right-hand-sides of Eqs. (5.3.3) lie in

the image of the operator Q, such that a solution $\rho^{(\alpha)}$ of the form (5.3.2) can be obtained, at least in principle. It is therefore necessary to investigate the operator Q, defined in Eq. (5.2.3), in more detail. We note in passing that Q is very similar to one of the projection operators used in the projection operator techniques mentioned in the introduction [35, 36, 38, 39]. However, strictly speaking it is not a projection operator since $Q^2 = -Q$.

For the subsequent analysis, let us introduce the following notations:

- $\mathscr{H}$: space of hermitian operators in $\mathcal{H}$.
- $\mathscr{H}_{A,B}$: space of hermitian operators in $\mathcal{H}_{A,B}$.

Moreover, let $D(Q) \subset \mathscr{H}$ stand for the domain of Q, thus the operator Q is a mapping

$$Q : D(Q) \to \mathscr{H} . \tag{5.3.4}$$

We assume that $D(Q)$ is a linear space (a detailed analysis of the operator Q can be found in App. 5.A). Here, we briefly repeat the main results of App. 5.A needed in what follows:

(i) Let $Ker\, Q$ denote the kernel of Q. One has

$$D(Q) = Ker\, Q \oplus (Ker\, Q)^{\perp} , \tag{5.3.5}$$

where $(Ker\, Q)^{\perp}$ denotes the space orthogonal to the kernel of Q. Hence any $X \in D(Q)$ can be decomposed into

$$X = X^{Ker} + X^{\perp} , \tag{5.3.6}$$

where $X^{Ker} \in Ker\, Q$ and $X^{\perp} \in (Ker\, Q)^{\perp}$.

(ii) For $X^{Ker} \in Ker\, Q$ one has

$$X^{Ker} = X_A \otimes \chi_B , \quad X_A \in \mathscr{H}_A. \tag{5.3.7}$$

(iii) For $X^{\perp} \in (Ker\, Q)^{\perp}$ one has

$$\mathrm{tr}_B \left(X^{\perp} \right) = 0 . \tag{5.3.8}$$

(iv) Let $Im\, Q$ denote the image of Q. One has

$$Im\, Q = (Ker\, Q)^{\perp} . \tag{5.3.9}$$

(v) The equation $Q(X) = Y$ is well-posed (and thus has a solution) if $Y \in (Ker\, Q)^{\perp}$. Moreover, it has a unique solution in $(Ker\, Q)^{\perp}$ denoted $X^{\perp}$. It follows immediately from Eq. (5.3.8) that this solution is given by

$$X^{\perp} = -Y\,. \tag{5.3.10}$$

We begin now with the investigation of well-posedness of Eqs. (5.3.3). We use Eqs. (5.3.6) and (5.3.7) to decompose each term ρ_n of the Hilbert expansion,

$$\rho_n = \rho_A^{(n)} \otimes \chi_B + \rho_n^{\perp}\,. \tag{5.3.11}$$

Moreover, we note the important property

$$\mathrm{tr}_B\left(i[\widetilde{H}_I, \rho_A^{(n)} \otimes \chi_B]\right) = 0 \qquad \forall\, \rho_A^{(n)} \in \mathscr{H}_A\,, \tag{5.3.12}$$

which is a consequence of the introduction of the mean-field operator, c.f. Eq. (5.2.8). Let us take the trace over the degrees of freedom B in Eqs. (5.3.3) and let us, furthermore, use the property (5.3.12) to obtain

$$0 = \mathrm{tr}_B\left(i[\widetilde{H}_I, \rho_0^{\perp}]\right)\,, \tag{5.3.13a}$$

$$0 = \partial_t \rho_A^{(0)} + i[\widetilde{H}_A, \rho_A^{(0)}] + \mathrm{tr}_B\left(i[\widetilde{H}_I, \rho_1^{\perp}]\right)\,, \tag{5.3.13b}$$

$$0 = \partial_t \rho_A^{(1)} + i[\widetilde{H}_A, \rho_A^{(1)}] + \mathrm{tr}_B\left(i[\widetilde{H}_I, \rho_2^{\perp}]\right)\,, \tag{5.3.13c}$$

$$0 = \partial_t \rho_A^{(n-2)} + i[\widetilde{H}_A, \rho_A^{(n-2)}] + \mathrm{tr}_B\left(i[\widetilde{H}_I, \rho_{n-1}^{\perp}]\right)\,, \quad \text{for } n \geq 4\,, \tag{5.3.13d}$$

where we omitted the result $0 = 0$ obtained from Eq. (5.3.3a). From property (v) of the relaxation operator Q it is clear that the system (5.3.3) is well-posed if and only if Eqs. (5.3.13) are fulfilled. In what follows we shall present an inductive proof that this can be indeed achieved. Furthermore, we shall prove that a system consisting of the first $N \in \mathbb{N}$ equations (5.3.13) is closed and that its solution can be computed recursively from Eqs. (5.3.3).

We know a priori that Eq. (5.3.3a) is well-posed and that its solution is obtained as

$$\rho_0 = \rho_A^{(0)} \otimes \chi_B\,, \qquad \rho_0^{\perp} = 0\,. \tag{5.3.14}$$

Assuming Eqs. (5.3.3b) to (5.3.3e) are also well-posed, we can employ Eq. (5.3.10)

to determine their unique solutions $\rho_n^\perp \in (Ker\, Q)^\perp$ as

$$\rho_1^\perp = -i[\widetilde{H}_I, \rho_0]\,, \tag{5.3.15a}$$

$$\rho_2^\perp = -\partial_t\rho_0 - i[\widetilde{H}_{AB}, \rho_0] - i[\widetilde{H}_I, \rho_1]\,, \tag{5.3.15b}$$

$$\rho_3^\perp = -\partial_t\rho_1 - i[\widetilde{H}_{AB}, \rho_1] - i[\widetilde{H}_I, \rho_2]\,, \tag{5.3.15c}$$

$$\rho_n^\perp = -\partial_t\rho_{n-2} - i[\widetilde{H}_{AB}, \rho_{n-2}] - i[\widetilde{H}_I, \rho_{n-1}]\,, \tag{5.3.15d}$$

$$\text{for } n \geq 4\,.$$

From Eq. (5.3.14) we deduce that Eq. (5.3.13a) is fulfilled trivially and, thus, Eq. (5.3.3b) is well-posed. This enables us to insert the result (5.3.15a) into Eq. (5.3.13b) to obtain

$$\partial_t\rho_A^{(0)} = -i[\widetilde{H}_A, \rho_A^{(0)}] - \mathrm{tr}_B\left([\widetilde{H}_I, [\widetilde{H}_I, \rho_A^{(0)} \otimes \chi_B]]\right)\,. \tag{5.3.16}$$

Equation (5.3.16) is a master equation of Lindblad form, as will be elaborated later in more detail in subsection 5.3.2. The second term on the right-hand-side of Eq. (5.3.16) is the dissipative part; thus, let us define the "dissipator" $\mathcal{D} : D(\mathcal{D}) \subset \mathscr{H}_A \to \mathscr{H}_A$,

$$\mathcal{D}(X_A) := -\mathrm{tr}_B\left([\widetilde{H}_I, [\widetilde{H}_I, X_A \otimes \chi_B]]\right)\,. \tag{5.3.17}$$

Since $X_A \otimes \chi_B = X^{Ker} \in Ker\, Q$, the operator $\mathcal{D}$ can also be viewed as a mapping from $Ker\, Q$ to $\mathscr{H}_A$. Using the short notation (5.3.17), equation (5.3.16) reads

$$\partial_t\rho_A^{(0)} = -i[\widetilde{H}_A, \rho_A^{(0)}] + \mathcal{D}(\rho_A^{(0)})\,. \tag{5.3.18}$$

For now we suppose the Lindblad master equation (5.3.18) to have a unique solution. This assumption is sufficient for completing the inductive proof of well-posedness of Eqs. (5.3.3), as will become transparent in the remainder of this subsection.

A first consequence of well-posedness of the Lindblad equation (5.3.18) is that Eq. (5.3.3c) is also well-posed and, thus, that its unique solution $\rho_2^\perp \in (Ker\, Q)^\perp$ given in Eq. (5.3.15b) is valid. Inserting this into Eq. (5.3.13c) results in

$$\begin{aligned}\partial_t\rho_A^{(1)} = &- i[\widetilde{H}_A, \rho_A^{(1)}] \\ &+ \mathrm{tr}_B\left(i[\widetilde{H}_I, \partial_t\rho_0 + i[\widetilde{H}_{AB}, \rho_0] + i[\widetilde{H}_I, \rho_1]]\right)\,.\end{aligned} \tag{5.3.19}$$

This equation can be simplified by use of Eq. (5.3.14), the property (5.3.12), and the result (5.3.15a) which yields

$$\partial_t\rho_A^{(1)} = -i[\widetilde{H}_A, \rho_A^{(1)}] + \mathcal{D}(\rho_A^{(1)}) + \mathcal{S}_1\,, \tag{5.3.20}$$

with

$$\begin{aligned}\mathcal{S}_1 = &- \mathrm{tr}_B \left(i[\widetilde{H}_I, \rho_A^{(0)} \otimes [H_B, \chi_B] \right) \\ &+ \mathrm{tr}_B \left(i[\widetilde{H}_I, [\widetilde{H}_I, [\widetilde{H}_I, \rho_A^{(0)} \otimes \chi_B]]] \right) .\end{aligned} \tag{5.3.21}$$

We remark that the first two terms on the right-hand-side of Eq. (5.3.20) form exactly the Lindblad generator from equation (5.3.18). The additional term $\mathcal{S}_1$ does not depend on $\rho_A^{(1)}$ and, thus, can be viewed as a well-defined, local source term. Hence, Eq. (5.3.20) has a unique solution. We deduce that Eq. (5.3.3d) is well-posed and its unique solution $\rho_3^{\perp} \in (Ker\, Q)^{\perp}$ given in Eq. (5.3.15c) is valid. One can already see the evolving pattern that will result in the well-posedness of the entire system (5.3.3). In order to complete the proof we shall proceed by induction. Therefore, suppose that Eqs. (5.3.3e) are well-posed up to order $n-1$. The solution to the $(n-1)$-th order equation is then written as

$$\rho_{n-1} = \rho_A^{(n-1)} \otimes \chi_B + \rho_{n-1}^{\perp} . \tag{5.3.22}$$

Due to Eq. (5.3.15d), $\rho_{n-1}^{\perp}$ is given by

$$\rho_{n-1}^{\perp} = -\partial_t \rho_{n-3} - i[\widetilde{H}_{AB}, \rho_{n-3}] - i[\widetilde{H}_I, \rho_{n-2}] . \tag{5.3.23}$$

The aim is now to specify under which condition the n-th order Eq. (5.3.3e) is also well-posed. From Eq. (5.3.13d) one deduces that this condition reads

$$\partial_t \rho_A^{(n-2)} = -i[\widetilde{H}_A, \rho_A^{(n-2)}] - \mathrm{tr}_B \left(i[\widetilde{H}_I, \rho_{n-1}^{\perp}] \right) . \tag{5.3.24}$$

Inserting Eq. (5.3.23) into Eq. (5.3.24) yields

$$\begin{aligned}\partial_t \rho_A^{(n-2)} = &- i[\widetilde{H}_A, \rho_A^{(n-2)}] + \mathrm{tr}_B \left(i[\widetilde{H}_I, \partial_t \rho_{n-3}] \right) \\ &- \mathrm{tr}_B \left([\widetilde{H}_I, [\widetilde{H}_{AB}, \rho_{n-3}]] \right) \\ &- \mathrm{tr}_B \left([\widetilde{H}_I, [\widetilde{H}_I, \rho_{n-2}]] \right) .\end{aligned} \tag{5.3.25}$$

Again we employ the decompositions

$$\rho_{n-2} = \rho_A^{(n-2)} \otimes \chi_B + \rho_{n-2}^{\perp} , \tag{5.3.26}$$

$$\rho_{n-3} = \rho_A^{(n-3)} \otimes \chi_B + \rho_{n-3}^{\perp} , \tag{5.3.27}$$

and profit from the fact that $\rho_{n-2}^{\perp} \in (Ker\, Q)^{\perp}$ is uniquely defined by Eq. (5.3.10),

with the result

$$\rho_{n-2}^{\perp} = -\partial_t \rho_{n-4} - i[\widetilde{H}_{AB}, \rho_{n-4}] - i[\widetilde{H}_I, \rho_{n-3}]\,. \tag{5.3.28}$$

We note that we were able to obtain Eqs. (5.3.23), (5.3.26) and (5.3.27) because we supposed Eq. (5.3.3e) to be well-posed up to order $n-1$. Moreover,

$$\partial_t \rho_{n-3} = \partial_t \rho_A^{(n-3)} \otimes \chi_B + \partial_t \rho_{n-3}^{\perp}\,, \tag{5.3.29}$$

and thus property (5.3.12) yields

$$\mathrm{tr}_B\left(i[\widetilde{H}_I, \partial_t \rho_{n-3}]\right) = \mathrm{tr}_B\left(i[\widetilde{H}_I, \partial_t \rho_{n-3}^{\perp}]\right)\,. \tag{5.3.30}$$

The decompositions (5.3.26) and (5.3.27) are applied to Eq. (5.3.25) and one obtains, also using Eq. (5.3.30),

$$\begin{aligned}
\partial_t \rho_A^{(n-2)} =& - i[\widetilde{H}_A, \rho_A^{(n-2)}] + \mathrm{tr}_B\left(i[\widetilde{H}_I, \partial_t \rho_{n-3}^{\perp}]\right) \\
& - \mathrm{tr}_B\left([\widetilde{H}_I, [\widetilde{H}_{AB}, \rho_A^{(n-3)} \otimes \chi_B + \rho_{n-3}^{\perp}]]\right) \\
& - \mathrm{tr}_B\left([\widetilde{H}_I, [\widetilde{H}_I, \rho_{n-2}^{\perp}]]\right) \\
& - \mathrm{tr}_B\left([\widetilde{H}_I, [\widetilde{H}_I, \rho_A^{(n-2)} \otimes \chi_B]]\right)\,.
\end{aligned} \tag{5.3.31}$$

In the last term on the right-hand-side one can introduce the definition (5.3.17) of the dissipator $\mathcal{D}$ in order to obtain, finally

$$\partial_t \rho_A^{(n-2)} = -i[\widetilde{H}_A, \rho_A^{(n-2)}] + \mathcal{D}(\rho_A^{(n-2)}) + \mathcal{S}_{n-2}\,, \tag{5.3.32}$$

where $\mathcal{S}_{n-2}$, $n \geq 4$, is given by

$$\begin{aligned}
\mathcal{S}_{n-2} =& \mathrm{tr}_B\left(i[\widetilde{H}_I, \partial_t \rho_{n-3}^{\perp}]\right) \\
& - \mathrm{tr}_B\left([\widetilde{H}_I, \rho_A^{(n-3)} \otimes [H_B, \chi_B]]\right) \\
& - \mathrm{tr}_B\left([\widetilde{H}_I, [\widetilde{H}_{AB}, \rho_{n-3}^{\perp}]]\right) \\
& + \mathrm{tr}_B\left([\widetilde{H}_I, [\widetilde{H}_I, \partial_t \rho_{n-4}]]\right) \\
& + \mathrm{tr}_B\left(i[\widetilde{H}_I, [\widetilde{H}_I, [\widetilde{H}_{AB}, \rho_{n-4}]]]\right) \\
& + \mathrm{tr}_B\left(i[\widetilde{H}_I, [\widetilde{H}_I, [\widetilde{H}_I, \rho_{n-3}]]]\right)\,.
\end{aligned} \tag{5.3.33}$$

Here $\rho_{n-2}^{\perp}$ was expressed with the help of Eq. (5.3.28). Equation (5.3.32) is called the $(n-2)$-th order master equation for $n \geq 4$. It arises solely from the requirement

that Eq. (5.3.3e) of order n is well-posed. Besides the Lindblad generator, which has already been found in the zeroth and first order master equations (5.3.18) and (5.3.20), respectively, Eq. (5.3.32) comprises the additional source term $\mathcal{S}_{n-2}$. This term depends solely on operators ρ_k obtained from Eqs. (5.3.3e) of order $k < n-2$. Therefore, under the premise that the Lindblad master equation (5.3.18) yields sufficiently well-behaved solutions, Eqs. (5.3.32) are solvable up to arbitrary order n, which proves the well-posedness of Eqs. (5.3.3). It follows, moreover, from the particular form of the source term $\mathcal{S}_{n-2}$ that the first $n-2$ equations (5.3.32) form a closed system of equations, in which solutions can be computed recursively. For $n = 4$, the source term (5.3.33) is evaluated in App. 5.C.

5.3.2 Lindblad master equation

We shall briefly elaborate on the Lindblad master equation (5.3.18). This equation will also be called zeroth order master equation. Recalling that $\widetilde{H}_I = H_I - H_A^{mf} \otimes \mathbb{1}_B$, a straightforward calculation results in the following form of the dissipator (5.3.17),

$$\begin{aligned} &\mathcal{D}(\rho_A^{(0)}) = \mathrm{tr}_B \left(2H_I\rho_0 H_I - H_I^2\rho_0 - \rho_0 H_I^2\right) \\ &-2H_A^{mf}\rho_A^{(0)}H_A^{mf} + \left(H_A^{mf}\right)^2\rho_A^{(0)} + \rho_A^{(0)}\left(H_A^{mf}\right)^2 . \end{aligned} \tag{5.3.34}$$

We note that the interaction Hamiltonian H_I can be written in the form [7]

$$H_I = \sum_i A_i \otimes B_i \,, \tag{5.3.35}$$

where $A_i \in \mathscr{H}_A$ and $B_i \in \mathscr{H}_B$. Therefore, the mean-field operator reads

$$H_A^{mf} = \sum_i A_i \mathrm{tr}_B \left(B_i \chi_B\right) \,. \tag{5.3.36}$$

By inserting relations (5.3.35) and (5.3.36) into Eq. (5.3.34), one obtains

$$\mathcal{D}(\rho_A^{(0)}) = \sum_{ij} \Gamma_{ij} \left(2A_j\rho_A^{(0)}A_i - A_iA_j\rho_A^{(0)} - \rho_A^{(0)}A_iA_j\right) \,, \tag{5.3.37}$$

where the coefficients Γ_{ij} are defined as

$$\Gamma_{ij} = \langle B_iB_j\rangle_{\chi_B} - \langle B_i\rangle_{\chi_B}\langle B_j\rangle_{\chi_B} \,. \tag{5.3.38}$$

Here we made use of the standard definition of correlation functions

$$\langle B_iB_j\rangle_{\chi_B} = \mathrm{tr}_B \left(B_iB_j\chi_B\right) \,, \tag{5.3.39}$$

and expectation values,

$$\langle B_i \rangle_{\chi_B} = \mathrm{tr}_B \left(B_i \chi_B \right) . \tag{5.3.40}$$

In fact, the coefficients (5.3.38) stand for the covariance of B_i and B_j in the state χ_B. In summary, the zeroth order master equation for the reduced system A can be written as

$$\begin{aligned} \partial_t \rho_A^{(0)} = & - i \left[H_A, \rho_A^{(0)} \right] - i \sum_i \langle B_i \rangle_{\chi_B} \left[A_i, \rho_A^{(0)} \right] \\ & + \sum_{ij} \Gamma_{ij} \left(2 A_j \rho_A^{(0)} A_i - A_i A_j \rho_A^{(0)} - \rho_A^{(0)} A_i A_j \right) . \end{aligned} \tag{5.3.41}$$

This equation can be transformed into the Lindblad form (5.1.5) by expanding the operators A_i, A_j in an appropriate basis $L_k \in \mathscr{H}_A$. The second term on the right-hand-side of Eq. (5.3.41) represents an energy shift induced by the mean-field approximation of the interaction between system A and environment B. We emphasize that this energy shift has to occur in the zeroth order equation of the reduced system, because otherwise the Hilbert expansion (5.3.2) would result in an ill-posed equation (5.3.3b). For the same reason the coefficients (5.3.38) stand for the covariance of B_i and B_j, rather than their correlation.

5.4 Discussion

Let us briefly summarize what has been accomplished so far. The goal of the present work was to find approximate solutions $\rho^{(\alpha)}$ to Eq. (5.3.1) in the case that the parameter α is small but not zero. For this, we invoked a Hilbert expansion of the form

$$\rho^{(\alpha)} = \sum_{n=0}^{\infty} \alpha^n \left(\rho_A^{(n)} \otimes \chi_B + \rho_n^{\perp} \right) , \tag{5.4.1}$$

where $\rho_A^{(n)} \otimes \chi_B \in Ker\, Q$ and $\rho_n^{\perp} \in (Ker\, Q)^{\perp}$. We note that in this representation, the reduced state operator reads

$$\mathrm{tr}_B \left(\rho^{(\alpha)} \right) := \rho_A^{(\alpha)} = \sum_{n=0}^{\infty} \alpha^n \rho_A^{(n)} . \tag{5.4.2}$$

After inserting the ansatz (5.4.1) into Eq. (5.3.1) we required equality of the left- and the right-hand-side of the equation in each power α^n. The further requirement of well-posedness of the resulting equations (5.3.3) gave rise to the following hierarchy

of master equations for the $\rho_A^{(n)}$:

$$\partial_t \rho_A^{(0)} = \mathcal{L}\rho_A^{(0)}\,, \tag{5.4.3a}$$
$$\partial_t \rho_A^{(1)} = \mathcal{L}\rho_A^{(1)} + \mathcal{S}_1\,, \tag{5.4.3b}$$
$$\partial_t \rho_A^{(n-2)} = \mathcal{L}\rho_A^{(n-2)} + \mathcal{S}_{n-2} \quad \text{for } n \geq 4\,. \tag{5.4.3c}$$

Here, $\mathcal{L}$ is the generator of a dynamical semigroup of the Lindblad form (5.1.6), specified in Eq. (5.3.41). The source terms $\mathcal{S}_1$ to $\mathcal{S}_{n-2}$ are given by Eqs. (5.3.21) and (5.3.33), respectively. Under the initial conditions

$$\rho_A^{(0)}(0) = \rho_{A,i}^{(0)} \quad \in \mathscr{H}_A\,, \tag{5.4.4a}$$
$$\rho_A^{(1)}(0) = \rho_{A,i}^{(1)} \quad \in \mathscr{H}_A\,, \tag{5.4.4b}$$
$$\rho_A^{(n-2)}(0) = \rho_{A,i}^{(n-2)} \quad \in \mathscr{H}_A \quad \text{for } n \geq 4\,, \tag{5.4.4c}$$

the formal solution of Eqs. (5.4.3) can be obtained via Duhamel's formula

$$\rho_A^{(0)}(t) = e^{\mathcal{L}t}\rho_{A,i}^{(0)}\,, \tag{5.4.5a}$$
$$\rho_A^{(1)}(t) = e^{\mathcal{L}t}\rho_{A,i}^{(1)} + \int_0^t ds\, e^{\mathcal{L}(t-s)}\mathcal{S}_1(s)\,, \tag{5.4.5b}$$
$$\rho_A^{(n-2)}(t) = e^{\mathcal{L}t}\rho_{A,i}^{(n-2)} + \int_0^t ds\, e^{\mathcal{L}(t-s)}\mathcal{S}_{n-2}(s) \quad \text{for } n \geq 4\,, \tag{5.4.5c}$$

where $e^{\mathcal{L}t}$ with $t \geq 0$ denotes the dynamical semigroup generated by $\mathcal{L}$. Under the assumption that the power series (5.4.2) converges for all $t \geq 0$ (which is reasonable for small α), one can perform the sum in the results (5.4.5) in order to obtain

$$\rho_A^{(\alpha)}(t) = e^{\mathcal{L}t}\rho_{A,i}^{(\alpha)} + \int_0^t ds\, e^{\mathcal{L}(t-s)}\mathcal{S}^{(\alpha)}(s)\,. \tag{5.4.6}$$

Here we defined the initial values $\rho_{A,i}^{(\alpha)}$ and the operator $\mathcal{S}^{(\alpha)}$ as

$$\rho_{A,i}^{(\alpha)} := \sum_{n=0}^{\infty} \alpha^n \rho_{A,i}^{(n)}\,, \tag{5.4.7}$$
$$\mathcal{S}^{(\alpha)} := \sum_{n=1}^{\infty} \alpha^n \mathcal{S}_n\,, \tag{5.4.8}$$

where we made use of $\mathcal{S}_0 = 0$. The integral on the right-hand-side of Eq. (5.4.6) makes the non-Markovianity of the time evolution of the reduced state (5.4.2) trans-

parent, since $\mathcal{S}^{(\alpha)}$ depends on $\rho_A^{(\alpha)}$ in a rather complicated way. In the Born-Markov limit $\alpha \to 0$, the term $\mathcal{S}^{(\alpha)}$ vanishes and one recovers the Markovian dynamics for the reduced system induced by the Lindblad generator $\mathcal{L}$.

In writing the formal solution (5.4.6), we note that the total state operator $\rho^{(\alpha)}$ of the composite system has been determined entirely. This follows from the fact that the terms $\rho_n^{\perp}$ in Eq. (5.4.1) are uniquely defined by Eqs. (5.3.15). Therefore, if the power series (5.4.1) is convergent for all $t \geq 0$ (which is reasonable for small α), it represents the unique solution of the initial value problem (5.3.1). Let us focus briefly on the "orthogonal" terms $\rho_n^{\perp}$. With the definition

$$\rho^{\perp} := \sum_{n=1}^{\infty} \alpha^n \rho_n^{\perp} \,, \tag{5.4.9}$$

where we used that $\rho_0^{\perp} = 0$, we point out that the contribution (5.4.9) to the solution (5.4.1) is traceless,

$$\mathrm{tr}\left(\rho^{\perp}\right) = 0 \,. \tag{5.4.10}$$

Therefore, it solely describes correlations between system A and environment B. Moreover, it is obvious that the power series (5.4.9) vanishes in the Born-Markov limit $\alpha \to 0$, thereby confirming the absence of correlations in the Markovian time evolution.

At first glance it might seem that nothing has been gained because the evaluation of Eqs. (5.4.2) and (5.4.9) requires the calculation of an infinite number of terms. However, provided that these power series converge, their benefits can be found in the fact that one can successively approach the exact solution $\rho^{(\alpha)}$ of Eq. (5.3.1) until a desired accuracy has been reached. For instance, truncating the series (5.4.1) after two terms appears to be a valid approximation in the case that $\alpha \ll 1$. In general, once the Lindblad master equation (5.3.41) has been solved for $\rho_A^{(0)}$, the source terms $\mathcal{S}_n$ and thus the higher order corrections $\rho_A^{(n)}$ and $\rho_n^{\perp}$ can be computed recursively in order to achieve the desired accuracy. In this way, correlations between system A and environment B can be incorporated rather easily in the reduced dynamics for A.

5.5 Conclusion

In this work we employed a Hilbert expansion in order to obtain approximate solutions of a von Neumann equation which was augmented by a relaxation operator Q. This operator relaxes the state operator of a composite quantum system AB towards the Born approximation on a timescale τ_r. This approach resulted in the

hierarchy of master equations (5.4.3) for the reduced state operator ρ_A. In zeroth order, which accounts for the exact dynamics in the Born-Markov limit $\tau_r \to 0$, a master equation of Lindblad form was recovered. The transition rates derived are exactly the covariance functions between different components of the environmental part of the interaction Hamiltonian H_I in the reference state χ_B. Moreover, the discussed approach allows to systematically incorporate correlations and non-Markovian effects in the reduced system dynamics. These effects can be calculated recursively from the solution to the Lindblad master equation, as was achieved in Eqs. (5.4.5). Such an approach might be advantageous in physical systems for which the Born approximation is nearly justified or for which a full treatment on the basis of projection techniques is far to complex.

We point out that the non-Markovian quantum dynamics derived here follow from a transparent physical picture, namely the relaxation of the total state operator towards the Born approximation. This appears to be a reasonable scenario, for instance, if the environment contains degrees of freedom that are completely isolated from the observed system, c.f. Fig. 5.1. Nevertheless, the results obtained are merely valid under three assumptions:

- the coupling of different degrees of freedom (B and C) of the environment diminish correlations between observed system and environment,
- the Born relaxation is by far the fastest process in the composite system,
- strong system-environment interaction (singular coupling scaling).

The derived model could be enhanced rather easily by replacing the relaxation operator Q, introduced in Eq. (5.2.3), by a more sophisticated dissipative term for subsystem B; for instance by a Lindblad dissipator. However, this would result in minor changes only, because the hierarchy (5.3.13) is a general result which does not depend on the particular form of the relaxation operator Q. This hierarchy is a mere consequence of the complete isolation of the observed system A from the environmental reservoir (or probe) C, which manifests itself in the relation $\mathrm{tr}_B\left(Q(\rho)\right) = 0$.

Note added in proof. We are thankful to A. Arnold for pointing out that the proof in Appendix 5.B of existence and uniqueness of a solution to the initial value problem (5.3.1) is achieved by Theorem 1.1 of Chapter 3 in [44].

5.A Analysis of the operator Q

Defining for $X, Y \in \mathscr{H}$ the scalar product

$$(X, Y) := \mathrm{tr}\left(\chi^{-1/2} X \chi^{-1/2} Y\right) , \tag{5.A.1}$$

where $\chi \in \mathscr{H}$ satisfies

$$\chi = \chi_A \otimes \chi_B \,, \quad \mathrm{tr}\,(\chi) = 1 \,, \quad (\chi v, v)_{\mathcal{H}} > 0 \quad \forall v \in \mathcal{H} \setminus \{0\} \,, \tag{5.A.2}$$

the space $\mathscr{H}$ becomes a Hilbert space [2] with the associated norm

$$||X|| = \sqrt{(X, X)} \quad \forall X \in \mathscr{H} \,. \tag{5.A.3}$$

In (5.A.2), $(\cdot, \cdot)_{\mathcal{H}}$ denotes the scalar product in $\mathcal{H}$. Note that

$$\begin{aligned} (X, Y) &= (Y, X) \,, \\ (X, X) &\geq 0 \,, \end{aligned} \tag{5.A.4}$$

$\forall X, Y \in \mathscr{H}$. The positive-definiteness of (5.A.1) follows from

$$\begin{aligned} (X, X) &= \mathrm{tr}\left(\chi^{-1/4} X \chi^{-1/4} \chi^{-1/4} X \chi^{-1/4}\right) \\ &= \mathrm{tr}\left(Y^2\right) \geq 0 \,, \end{aligned} \tag{5.A.5}$$

where $Y = \chi^{-1/4} X \chi^{-1/4} \in \mathscr{H}$. Moreover, (5.A.1) is linear in both arguments. We shall further define a scalar product in $\mathscr{H}_A$,

$$(X_A, Y_A)_A := \mathrm{tr}_A \left(\chi_A^{-1/2} X_A \chi_A^{-1/2} Y_A\right) \,, \tag{5.A.6}$$

and denote the corresponding norm by

$$||X_A||_A = \sqrt{(X_A, X_A)_A} \,, \tag{5.A.7}$$

$\forall X_A \in \mathscr{H}_A$ [3]. For the operator Q defined in Eq. (5.2.3),

$$Q(X) = \mathrm{tr}_B\,(X) \otimes \chi_B - X \,, \quad X \in D(Q) \subset \mathscr{H} \,, \tag{5.A.8}$$

one has the following properties:

(i) Q is linear, bounded, self-adjoint and non-positive.

(ii) For $X \in Ker\, Q$ (the kernel of Q) we have

$$X = X_A \otimes \chi_B \,, \quad X_A \in \mathscr{H}_A. \tag{5.A.9}$$

[2] In what follows we denote by $\mathscr{H}$ the space of Hermitian operators in $\mathcal{H}$ for which $||X|| < \infty$.
[3] We denote by $\mathscr{H}_A$ the space of Hermitian operators in $\mathcal{H}_A$ for which $||X||_A < \infty$.

Moreover, we have

$$(Ker\, Q)^{\perp} = \{X \in D(Q) \,|\, \mathrm{tr}_B\,(X) = 0\}\ . \tag{5.A.10}$$

(iii) Let $\mathcal{P} : \mathscr{H} \to Ker\, Q$ denote the orthogonal projection operator onto $Ker\, Q$. Then there exists $d > 0$ such that

$$-(Q(X), X) \geq d||X - \mathcal{P}(X)||^2 \tag{5.A.11}$$

$\forall X \in D(Q)$.

(iv) The image of Q, denoted by $Im\, Q$, is closed and we have

$$Im\, Q = (Ker\, Q)^{\perp}\,. \tag{5.A.12}$$

Further, the equation $Q(X) = Y$ has a solution in $D(Q)$ if and only if $Y \in Im\, Q$. The solution is moreover unique in $(Ker\, Q)^{\perp}$.

In the following we denote,

$$X_A := \mathrm{tr}_B\,(X)\ , \tag{5.A.13}$$

and

$$Q_A(X) = \mathrm{tr}_B\,(Q(X)) = 0\,, \quad \forall X \in D(Q)\,. \tag{5.A.14}$$

Moreover, in what follows we shall frequently make use of the identity

$$(X_A \otimes \chi_B, Y) = (X_A, Y_A)_A\,. \tag{5.A.15}$$

We shall now prove the above statements.

(i) The linearity of Q is obvious. Let us show that Q is a well-defined, bounded operator. For this one needs a constant $c > 0$ such that

$$||Q(X)||^2 \leq c||X||^2\,, \quad \forall X \in D(Q)\,. \tag{5.A.16}$$

Using the identities (5.A.14) and (5.A.15) one obtains

$$\begin{aligned} ||Q(X)||^2 &= (Q(X), Q(X)) \\ &= (X_A \otimes \chi_B, Q(X)) - (X, Q(X)) \\ &= (X_A, Q_A(X))_A - (X, Q(X)) \\ &= -(X, X_A \otimes \chi_B) + (X, X) \\ &= -||X_A||_A^2 + ||X||^2 \\ &\leq ||X||^2\,, \end{aligned} \tag{5.A.17}$$

which proves (5.A.16). Therefore Q is bounded (and thus also continous). The self-adjointness follows from

$$\begin{aligned}(Q(X),Y) &= (X_A\otimes\chi_B,Y)-(X,Y)\\ &= (X_A,Y_A)_A-(X,Y)\\ &= (Y_A,X_A)_A-(Y,X)\\ &= (Q(Y),X)\\ &= (X,Q(Y))\,.\end{aligned}\tag{5.A.18}$$

In order to prove the non-positivity of Q we estimate the term $||X_A||_A^2$ using Eq. (5.A.15) together with the Cauchy-Schwarz inequality,

$$\begin{aligned}||X_A||_A^2 &= (X,X_A\otimes\chi_B)\\ &\leq ||X||\,||X_A\otimes\chi||\\ &= ||X||\,||X_A||_A\\ &\leq ||X||^2\,.\end{aligned}\tag{5.A.19}$$

It follows the non-positivity of Q,

$$(Q(X),X) = ||X_A||_A^2-||X||^2\leq ||X||^2-||X||^2=0\,.\tag{5.A.20}$$

We remark that Eq. (5.A.19) implies

$$X\in\mathscr{H}\Rightarrow \mathrm{tr}_B\,(X)\in\mathscr{H}_A\,.\tag{5.A.21}$$

(ii) For $X\in Ker\,Q$ we have $(Q(X),X)=0$ which can be written as

$$\begin{aligned}0 &= ||X_A||_A^2-||X||^2\\ &= (X,X_A\otimes\chi_B)-(X,X)\,.\end{aligned}\tag{5.A.22}$$

The solutions of Eq. (5.A.22) are given by $X=X_A\otimes\chi_B$, $X_A\in\mathscr{H}_A$ arbitrary. Conversely,

$$Q(X_A\otimes\chi_B)=X_A\otimes\chi_B-X_A\otimes\chi_B=0\,.\tag{5.A.23}$$

Moreover, for $Y\in(Ker\,Q)^\perp$, we have $(X,Y)=0$ for $X\in Ker\,Q$ and thus

$$(X_A\otimes\chi_B,Y)=(X_A,Y_A)=0\quad\forall\,X_A\in\mathscr{H}_A\,.\tag{5.A.24}$$

Since Eq. (5.A.24) must hold for arbitrary $X_A \in \mathscr{H}_A$ we conclude

$$Y_A = \mathrm{tr}_B\,(Y) = 0 \quad \forall\, Y \in (Ker\, Q)^{\perp}\,. \tag{5.A.25}$$

(iii) We now prove the coercitivity relation (5.A.11). For $X \in Ker\, Q$ this relation is fulfilled trivially because $Q(X) = 0$ and $\mathcal{P}(X) = X$. Now suppose $X \in (Ker\, Q)^{\perp}$. Then, according to Eq. (5.A.25),

$$-(Q(X), X) = -||X_A||_A^2 + ||X||^2 = ||X||^2\,, \tag{5.A.26}$$

which completes the coercitivity proof.

(iv) First we show that the image of Q is closed. Let X_n be a sequence in $D(Q)$ and let J_n be a sequence in $Im\, Q$ such that $Q(X_n) = J_n$. Moreover, let $J_n \to J$ as $n \to \infty$. We have to prove that $J \in Im\, Q$, i.e. that there exists $X \in D(Q)$ such that $Q(X) = J$. To any sequence $X_n \in D(Q)$ one can construct a corresponding sequence $Y_n \in (Ker\, Q)^{\perp}$ by setting $Y_n = X_n - \mathcal{P}(X_n)$ and one has $Q(X_n) = Q(Y_n) = J_n$. The coercitivity relation then yields

$$-(Q(Y_n - Y_m), Y_n - Y_m) \geq d||Y_n - Y_m||^2 \qquad \forall n, m \in \mathbb{N}\,. \tag{5.A.27}$$

In addition, from the Cauchy-Schwarz inequality we have

$$||Q(Y_n) - Q(Y_m)||\,||Y_n - Y_m|| \geq -(Q(Y_n - Y_m), Y_n - Y_m)\,, \tag{5.A.28}$$

and consequently

$$\frac{1}{d}||Q(Y_n) - Q(Y_m)|| = \frac{1}{d}||J_n - J_m|| \geq ||Y_n - Y_m||\,. \tag{5.A.29}$$

Since J_n is a Cauchy sequence in $Im\, Q$ we obtain that Y_n is a Cauchy sequence in $D(Q)$. By assumption $D(Q)$ is complete and therefore $Y_n \to Y \in D(Q)$. We already proved that Q is continuous, i.e. $Q(Y_n) \to Q(Y)$. One obtains $Q(Y) = J$ with $Y \in D(Q)$. Thus, the image of Q is closed and we have $Im\, Q = (Ker\, Q)^{\perp}$.

We finally prove that the equation $Q(X) = Y$ has a unique solution $X \in (Ker\, Q)^{\perp}$. It is obvious that there exists a solution if $Y \in Im\, Q$. Let X be such a solution, then $X - \mathcal{P}(X) \in (Ker\, Q)^{\perp}$ is also a solution. Assume that there are two solutions $X_1, X_2 \in (Ker\, Q)^{\perp}$ such that $Q(X_1) = Q(X_2) = Y$. Then

$$Q(X_1) - Q(X_2) = Q(X_1 - X_2) = 0\,. \tag{5.A.30}$$

It follows that $X_1 - X_2 \in Ker\, Q \cap (Ker\, Q)^{\perp} = \{0\}$ and therefore $X_1 = X_2$.

5.B Existence and uniqueness

In this section we demonstrate the existence and uniqueness of a solution to the initial value problem (5.3.1) on the basis of a fixed point argument. Let $0 < T < \infty$ and $\alpha = 1$ for convenience (the proof holds for $\alpha > 0$). The following proposition holds:

- Let Q denote the operator defined in Eq. (5.2.3). Furthermore, let $H \in \mathscr{H}$ such that the Liouville operator $L(\cdot) = -i[H, \cdot]$ in $\mathscr{H}$ is the infinitesimal generator of a bounded one-parameter semigroup e^{Lt} in $\mathscr{H}$, i.e. one has $C > 0$ such that
$$||e^{Lt}\rho||^2_{\mathscr{H}} \leq C||\rho||^2_{\mathscr{H}} \quad \forall \rho \in \mathscr{H}\,, t \in [0,T]\,, \tag{5.B.1}$$
where $||\cdot||_{\mathscr{H}}$ denotes the norm (5.A.3). Then, the intial value problem
$$\begin{cases} \partial_t \rho - L(\rho) = Q(\rho) \\ \rho(0) = \rho_i \quad \in \mathscr{H} \end{cases} \tag{5.B.2}$$
admits a unique solution $\rho \in L^2([0,T], \mathscr{H})$.

In order to prove this claim we denote the norm and the corresponding scalar product in $L^2([0,T], \mathscr{H})$ by
$$||\rho||_{L^2} = \int_0^T dt\, ||\rho(t)||_{\mathscr{H}}\,, \tag{5.B.3}$$
and
$$(\rho, \sigma)_{L^2} = \int_0^T dt\, (\rho(t), \sigma(t))_{\mathscr{H}}\,, \tag{5.B.4}$$
respectively. Let us now define the following fixed point map,
$$\mathcal{F}:\, L^2([0,T], \mathscr{H}) \to L^2([0,T], \mathscr{H})\,, \quad \sigma \mapsto \rho\,, \tag{5.B.5}$$
where ρ is the solution of
$$\begin{cases} \partial_t \rho - L(\rho) + \rho = \mathrm{tr}_B\,(\sigma) \otimes \chi_B \\ \rho(0) = \rho_i \quad \in \mathscr{H}\,. \end{cases} \tag{5.B.6}$$
In the following, we denote $S_\sigma(t) = \mathrm{tr}_B\,(\sigma(t)) \otimes \chi_B$. The formal solution of (5.B.6) reads
$$\mathcal{F}(\sigma)(t) = \rho(t) = e^{\mathcal{T}t}\rho(0) + \int_0^t ds\, e^{\mathcal{T}(t-s)} S_\sigma(s)\,, \tag{5.B.7}$$
where $\mathcal{T} := L - \mathbb{1}$. The proof is performed in two steps. First we demonstrate that the mapping (5.B.5)-(5.B.7) is well-defined and we prove then that the mapping is a

contraction, possessing thus a unique fixed point, solution of equation (5.B.2). From Eq. (5.B.1) it follows that

$$||e^{\mathcal{T}t}\rho||^2_{\mathscr{H}} = e^{-2t}||e^{Lt}\rho||^2_{\mathscr{H}} \leq C||\rho||^2_{\mathscr{H}} \quad \forall t \in [0,T]\,. \tag{5.B.8}$$

Therefore, Eq. (5.B.7) yields

$$\begin{aligned} ||\rho(t)||_{\mathscr{H}} &\leq \sqrt{C}||\rho(0)||_{\mathscr{H}} + \sqrt{C}\int_0^t ds\,||S_\sigma(s)||_{\mathscr{H}}\,, \\ &\leq \sqrt{C}||\rho(0)||_{\mathscr{H}} + \sqrt{C}\int_0^T ds\,||S_\sigma(s)||_{\mathscr{H}}\,, \\ &\leq \sqrt{C}||\rho(0)||_{\mathscr{H}} + \sqrt{C}\int_0^T ds\,||\sigma(s)||_{\mathscr{H}}\,, \end{aligned} \tag{5.B.9}$$

where the last inequality follows from Eq. (5.A.19),

$$||S_\sigma(s)||_{\mathscr{H}} = ||\sigma_A(s)||_A \leq ||\sigma(s)||_{\mathscr{H}}\,. \tag{5.B.10}$$

Integration of (5.B.9) over t results in

$$||\rho||_{L^2} \leq T\sqrt{C}||\rho(0)||_{\mathscr{H}} + T\sqrt{C}||\sigma||_{L^2}\,, \tag{5.B.11}$$

which proves that the fixed point map (5.B.5) is well-defined.

In order to show that (5.B.7) is a contraction we introduce the following norm in $L^2([0,T],\mathscr{H})$,

$$||\rho||^2_\delta := \int_0^T dt\, e^{-\delta t}||\rho(t)||^2_{\mathscr{H}}\,. \tag{5.B.12}$$

For $\mathcal{F}$ to be contractive it is required that

$$||\mathcal{F}(\sigma_1) - \mathcal{F}(\sigma_2)||^2_\delta \leq k||\sigma_1 - \sigma_2||^2_\delta\,, \quad k<1\,. \tag{5.B.13}$$

One obtains

$$\begin{aligned} &||\mathcal{F}(\sigma_1) - \mathcal{F}(\sigma_2)||^2_\delta = \\ &= \int_0^T dt\, e^{-\delta t}\int_0^t ds\,||e^{\mathcal{T}(t-s)}[S_{\sigma_1}(s) - S_{\sigma_2}(s)]||^2_{\mathscr{H}} \\ &\leq C\int_0^T dt\, e^{-\delta t}\int_0^t ds\,||S_{\sigma_1}(s) - S_{\sigma_2}(s)||^2_{\mathscr{H}} \\ &= C\int_0^T ds\int_s^T dt\, e^{-\delta t}||S_{\sigma_1}(s) - S_{\sigma_2}(s)||^2_{\mathscr{H}} \\ &\leq \frac{C}{\delta}\int_0^T ds\, e^{-\delta s}||S_{\sigma_1}(s) - S_{\sigma_2}(s)||^2_{\mathscr{H}}\,. \end{aligned} \tag{5.B.14}$$

Due to the inequality (5.A.19) one has

$$\begin{aligned} ||S_{\sigma_1}(s) - S_{\sigma_2}(s)||^2_{\mathscr{H}} &= ||\mathrm{tr}_B\left(\sigma_1(s) - \sigma_2(s)\right) \otimes \chi_B||^2_{\mathscr{H}} \\ &= ||\mathrm{tr}_B\left(\sigma_1(s) - \sigma_2(s)\right)||^2_A \\ &\leq ||\sigma_1(s) - \sigma_2(s)||^2_{\mathscr{H}}\,, \end{aligned} \tag{5.B.15}$$

and this results in

$$||\mathcal{F}(\sigma_1) - \mathcal{F}(\sigma_2)||^2_\delta \leq \frac{C}{\delta}||\sigma_1 - \sigma_2||^2_\delta\,. \tag{5.B.16}$$

Here, δ can be chosen in such a way that relation (5.B.13) is fulfilled. Thus $\mathcal{F}$ is a contraction, its unique fixed point is the solution of Eq. (5.B.2).

5.C Second-order contribution

We compute the source term (5.3.33) for $n = 4$. For this, we need the solutions of Eqs. (5.3.3a) and (5.3.3b), namely ρ_0 and $\rho_1^\perp$, given by (5.3.14) and (5.3.15a), respectively. It follows that

$$\begin{aligned} \partial_t\rho_0 &= \partial_t\rho_A^{(0)} \otimes \chi_B \\ &= -i[\widetilde{H}_A, \rho_A^{(0)}] \otimes \chi_B + \mathcal{D}(\rho_A^{(0)}) \otimes \chi_B\,, \end{aligned} \tag{5.C.1}$$

and

$$\begin{aligned} \partial_t\rho_1^\perp &= -i[\widetilde{H}_I, \partial_t\rho_0] \\ &= -[\widetilde{H}_I, [\widetilde{H}_A, \rho_A^{(0)}] \otimes \chi_B] - i[\widetilde{H}_I, \mathcal{D}(\rho_A^{(0)}) \otimes \chi_B]\,. \end{aligned} \tag{5.C.2}$$

We now compute all the terms appearing on the right-hand side of Eq. (5.3.33). Whenever possible, we make use of definition (5.3.17) of the operator $\mathcal{D}$ in order to simplify the notation.

1. With the help of Eq. (5.C.2) the first term on the right-hand-side of Eq. (5.3.33)

$$\begin{aligned} \mathrm{tr}_B\left(i[\widetilde{H}_I, \partial_t\rho_1^\perp]\right) &= -\mathrm{tr}_B\left([\widetilde{H}_I, [\widetilde{H}_I, i[\widetilde{H}_A, \rho_A^{(0)}] \otimes \chi_B]]\right) \\ &\quad + \mathrm{tr}_B\left([\widetilde{H}_I, [\widetilde{H}_I, \mathcal{D}(\rho_A^{(0)}) \otimes \chi_B]]\right) \\ &= \mathcal{D}(i[\widetilde{H}_A, \rho_A^{(0)}]) - \mathcal{D}^2(\rho_A^{(0)})\,, \end{aligned} \tag{5.C.3}$$

where $\mathcal{D}^2(\cdot) = \mathcal{D}(\mathcal{D}(\cdot))$.

2. The second term results in

$$-\operatorname{tr}_B\left([\widetilde{H}_I,[\widetilde{H}_{AB},\rho_1^{\perp}]]\right)=\operatorname{tr}_B\left(i[\widetilde{H}_I,[\widetilde{H}_{AB},[\widetilde{H}_I,\rho_0]]]\right). \tag{5.C.4}$$

3. In the third term we set $\rho_A^{(n-3)}\to\rho_A^{(1)}$.

4. Equation (5.C.1) helps in evaluating term four to

$$\begin{aligned}\operatorname{tr}_B\left([\widetilde{H}_I,[\widetilde{H}_I,\partial_t\rho_0]]\right)&=\\&=-\operatorname{tr}_B\left([\widetilde{H}_I,[\widetilde{H}_I,i[\widetilde{H}_A,\rho_A^{(0)}]\otimes\chi_B]]\right)\\&\quad+\operatorname{tr}_B\left([\widetilde{H}_I,[\widetilde{H}_I,\mathcal{D}(\rho_A^{(0)})\otimes\chi_B]]\right)\\&=\mathcal{D}(i[\widetilde{H}_A,\rho_A^{(0)}])-\mathcal{D}^2(\rho_A^{(0)})\,.\end{aligned} \tag{5.C.5}$$

5. We get for term five

$$\begin{aligned}\operatorname{tr}_B\left(i[\widetilde{H}_I,[\widetilde{H}_I,[\widetilde{H}_{AB},\rho_0]]]\right)&=\\&=\operatorname{tr}_B\left([\widetilde{H}_I,[\widetilde{H}_I,i[\widetilde{H}_A,\rho_A^{(0)}]\otimes\chi_B]]\right)\\&\quad+\operatorname{tr}_B\left(i[\widetilde{H}_I,[\widetilde{H}_I,\rho_A^{(0)}\otimes[H_B,\chi_B]]]\right)\\&=-\mathcal{D}(i[\widetilde{H}_A,\rho_A^{(0)}])\\&\quad+\operatorname{tr}_B\left(i[\widetilde{H}_I,[\widetilde{H}_I,\rho_A^{(0)}\otimes[H_B,\chi_B]]]\right).\end{aligned} \tag{5.C.6}$$

6. Finally, term six becomes

$$\begin{aligned}\operatorname{tr}_B\left(i[\widetilde{H}_I,[\widetilde{H}_I,[\widetilde{H}_I,\rho_1]]]\right)&=\\&=\operatorname{tr}_B\left([\widetilde{H}_I,[\widetilde{H}_I,[\widetilde{H}_I,[\widetilde{H}_I,\rho_A^{(0)}\otimes\chi_B]]]]\right)\\&\quad+\operatorname{tr}_B\left(i[\widetilde{H}_I,[\widetilde{H}_I,[\widetilde{H}_I,\rho_A^{(1)}\otimes\chi_B]]]\right).\end{aligned} \tag{5.C.7}$$

Adding the results of Eqs. (5.C.3)-(5.C.7) gives the desired result for $\mathcal{S}_2$:

$$\begin{aligned}\mathcal{S}_2 = &-2\mathcal{D}^2(\rho_A^{(0)}) + \mathcal{D}(i[\widetilde{H}_A, \rho_A^{(0)}]) \\ &+ \mathrm{tr}_B\left(i[\widetilde{H}_I, [\widetilde{H}_{AB}, [\widetilde{H}_I, \rho_A^{(0)} \otimes \chi_B]]]\right) \\ &- \mathrm{tr}_B\left([\widetilde{H}_I, \rho_A^{(1)} \otimes [H_B, \chi_B]\right) \\ &+ \mathrm{tr}_B\left(i[\widetilde{H}_I, [\widetilde{H}_I, \rho_A^{(0)} \otimes [H_B, \chi_B]]]\right) \\ &+ \mathrm{tr}_B\left([\widetilde{H}_I, [\widetilde{H}_I, [\widetilde{H}_I, [\widetilde{H}_I, \rho_A^{(0)} \otimes \chi_B]]]]\right) \\ &+ \mathrm{tr}_B\left(i[\widetilde{H}_I, [\widetilde{H}_I, [\widetilde{H}_I, \rho_A^{(1)} \otimes \chi_B]]]\right) .\end{aligned} \tag{5.C.8}$$

5.4 Acknowledgments

The authors are very grateful to E. Schachinger for carefully reading the manuscript. We also thank W. Pötz and C. Negulescu for precious input and fruitful discussions. The first author acknowledges the support from the Austrian Science Fund, Vienna, under the contract number P21326-N16; the second author was supported by the Austrian Science Fund (FWF): P221290-N16.

References

[1] Michael Nielsen and Isaac Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, 2000.

[2] D.F. Walls and G.J. Milburn. *Quantum optics*. Springer verlag, 2008.

[3] D.D. Awschalom, D. Loss, and N. Samarth. *Semiconductor Spintronics and Quantum Computation*. Series on Nanoscience and Technology. Springer-Verlag Berlin, 2002.

[4] K. Hornberger, S. Uttenthaler, B. Brezger, L. Hackermüller, M. Arndt, and A. Zeilinger. Collisional decoherence observed in matter wave interferometry. *Physical review letters*, 90(16):160401, 2003.

[5] L. Hackermüller, K. Hornberger, B. Brezger, A. Zeilinger, and M. Arndt. Decoherence of matter waves by thermal emission of radiation. *Nature*, 427(6976):711–714, 2004.

[6] A.D. Cronin, J. Schmiedmayer, and D.E. Pritchard. Optics and interferometry with atoms and molecules. *Reviews of Modern Physics*, 81(3):1051, 2009.

[7] H.P. Breuer and F. Petruccione. *The theory of open quantum systems*, volume 28. Oxford University Press Oxford, 2002.

[8] U. Weiss. *Quantum dissipative systems*, volume 13. World Scientific Pub Co Inc, 2008.

[9] A. Rivas and S. Huelga. *Open Quantum Systems: An Introduction*. Springer Verlag, 2011.

[10] M. Arndt, K. Hornberger, and A. Zeilinger. Probing the limits of the quantum world. *Physics world*, 18(2005):35–40, 2005.

[11] R. Alicki. Search for a border between classical and quantum worlds. *Physical Review A*, 65(3):034104, 2002.

[12] E. Joos and H.D. Zeh. The emergence of classical properties through interaction with the environment. *Zeitschrift für Physik B Condensed Matter*, 59(2):223–243, 1985.

[13] A. Kossakowski. On quantum statistical mechanics of non-hamiltonian systems. *Reports on Mathematical Physics*, 3(4):247–274, 1972.

[14] E.B. Davies. Markovian master equations. *Communications in mathematical Physics*, 39(2):91–110, 1974.

[15] V. Gorini, A. Frigerio, M. Verri, A. Kossakowski, and ECG Sudarshan. Properties of quantum markovian master equations. *Reports on Mathematical Physics*, 13(2):149–173, 1978.

[16] G. Lindblad. On the generators of quantum dynamical semigroups. *Communications in Mathematical Physics*, 48(2):119–130, 1976.

[17] To the knowledge of the authors, rigorous proofs for the existence of (6) in the case $K = \infty$ and for unbounded Hamiltonians H are still lacking.

[18] L. Diosi. Quantum master equation of a particle in a gas environment. *EPL (Europhysics Letters)*, 30:63, 1995.

[19] C.W. Gardiner and P. Zoller. Quantum kinetic theory: A quantum kinetic master equation for condensation of a weakly interacting bose gas without a trapping potential. *Physical Review A*, 55(4):2902, 1997.

[20] PJ Dodd and JJ Halliwell. Decoherence and records for the case of a scattering environment. *Physical Review D*, 67(10):105018, 2003.

[21] B. Vacchini. Completely positive quantum dissipation. *Physical Review Letters*, 84(7):1374–1377, 2000.

[22] B. Vacchini. Translation-covariant markovian master equation for a test particle in a quantum fluid. *Journal of Mathematical Physics*, 42:4291, 2001.

[23] K. Hornberger. Master equation for a quantum particle in a gas. *Physical review letters*, 97(6):60601, 2006.

[24] K. Hornberger. Monitoring approach to open quantum dynamics using scattering theory. *EPL (Europhysics Letters)*, 77:50007, 2007.

[25] B. Vacchini and K. Hornberger. Quantum linear boltzmann equation. *Physics Reports*, 478(4):71–120, 2009.

[26] M. Knap, E. Arrigoni, W. von der Linden, and J.H. Cole. Emission characteristics of laser-driven dissipative coupled-cavity systems. *Physical Review A*, 83(2):023821, 2011.

[27] A. Royer. Reduced dynamics with initial correlations, and time-dependent environment and hamiltonians. *Physical review letters*, 77(16):3272–3275, 1996.

[28] S.M. Barnett and S. Stenholm. Hazards of reservoir memory. *Physical Review A*, 64(3):33808, 2001.

[29] A.A. Budini. Stochastic representation of a class of non-markovian completely positive evolutions. *Physical Review A*, 69(4):042107, 2004.

[30] Y.J. Yan and R.X. Xu. Quantum mechanics of dissipative systems. *Annu. Rev. Phys. Chem.*, 56:187–219, 2005.

[31] H.P. Breuer, J. Gemmer, and M. Michel. Non-markovian quantum dynamics: Correlated projection superoperators and hilbert space averaging. *Physical Review E*, 73(1):016139, 2006.

[32] H.P. Breuer. Non-markovian generalization of the lindblad theory of open quantum systems. *Physical Review A*, 75(2):022103, 2007.

[33] K.L. Liu and H.S. Goan. Non-markovian entanglement dynamics of quantum continuous variable systems in thermal environments. *Physical Review A*, 76(2):022312, 2007.

[34] D. Chruscinski and A. Kossakowski. Non-markovian quantum dynamics: Local versus nonlocal. *Physical review letters*, 104(7):70406, 2010.

[35] S. Nakajima. On quantum theory of transport phenomena—steady diffusion—. *Progress of Theoretical Physics*, 20:948–959, 1958.

[36] R. Zwanzig. Ensemble method in the theory of irreversibility. *The Journal of Chemical Physics*, 33:1338, 1960.

[37] I. Prigogine and R. Landshoff. Non-equilibrium statistical mechanics. *Physics Today*, 16:76, 1963.

[38] F. Shibata, Y. Takahashi, and N. Hashitsume. A generalized stochastic liouville equation. non-markovian versus memoryless master equations. *Journal of Statistical Physics*, 17(4):171–187, 1977.

[39] F. Shibata and T. Arimitsu. Expansion formulas in nonequilibrium statistical mechanics. *J. Phys. Soc. Jap*, 49:891–897, 1980.

[40] C. Cercignani. *The Boltzmann equation and its applications*, volume 67. Springer, 1988.

[41] N.B. Abdallah and P. Degond. On a hierarchy of macroscopic models for semiconductors. *Journal of Mathematical Physics*, 37(7):3306–3333, 1996.

[42] S. Possanner and C. Negulescu. Diffusion limit of a generalized matrix boltzmann equation for spin-polarized transport. *Kinetic and Related Models*, 4(4):1159–1191, December 2011.

[43] L.E. Ballentine. *Quantum Mechanics - A Modern Development*. World Scientific Publishing, 1998.

[44] A. Pazy. *Semigroups of linear operators and applications to partial differential equations*, volume 44. Springer-Verlag, 1983.

[45] In what follows we denote by $\mathscr{H}$ the space of Hermitian operators in $\mathcal{H}$ for which $||X|| < \infty$.

[46] We denote by $\mathscr{H}_A$ the space of Hermitian operators in $\mathcal{H}_A$ for which $||X||_A < \infty$.

Part IV

Numerical Study of a Quantum-Diffusive Spin Model

Chapter 6

Numerical Study of a Quantum-Diffusive Spin Model for Two-Dimensional Electron Gases

---

S. Possanner, *L. Barletti*, *F. Méhats* and *C. Negulescu*,
draft version, **2012**

Abstract. We investigate the time evolution of spin densities in a two-dimensional electron gas subjected to Rashba spin-orbit coupling on the basis of the quantum drift-diffusive model derived in Ref. [1]. This model assumes the electrons to be in a quantum equilibrium state in the form of a Maxwellian operator. The resulting quantum drift-diffusion equations for spin-up and spin-down densities are coupled in a non-local manner via two spin chemical potentials (Lagrange multipliers) and via off-diagonal elements of the equilibrium spin density and spin current matrices, respectively. We present two space-time discretizations of the model which comprise also the Poisson equation in order to account for electron-electron interactions. In a first step pure time discretization is applied in order to prove the well-posedness of the two schemes, both of which are based on a functional formalism to treat the non-local relations between spin densities. We then use the fully space-time discrete schemes to simulate the time evolution of a Rashba electron gas in a typical transistor geometry. Finite difference approximations are first order in time and second order in space. The discrete functionals introduced are minimized with the help of a conjugate gradient-based algorithm, where the Newton method is applied in order to find the respective line minima.

6.1 Introduction

The purpose of this paper is the numerical study of the quantum diffusive model for a spin-orbit system introduced in Ref. [1]. The derivation of this model is based on the quantum maximum entropy principle [3, 4] applied to a spin-orbit Hamiltonian of Rashba type [2]:

$$H = \begin{pmatrix} -\frac{\hbar^2}{2}\Delta + V & \alpha\hbar(\partial_x - i\partial_y) \\ -\alpha\hbar(\partial_x + i\partial_y) & -\frac{\hbar^2}{2}\Delta + V \end{pmatrix}. \tag{6.1.1}$$

Here, (x, y) are the spatial coordinates of the 2-dimensional region where the electrons are assumed to be confined, α is the Rashba constant and V is a potential term which may consist of an "external" part (representing e.g. a gate or an applied potential) and a self-consistent part, accounting for Coulomb interactions in the mean-field approximation.

The Rashba effect [2, 10] is a spin-orbit interaction undergone by electrons that are confined in an asymmetric 2-dimensional well (here, perpendicular to the z direction). The spin vector has a precession around a direction in the plane (x, y), perpendicular to the electron momentum $p = (p_x, p_y)$, the precession speed being $\alpha|p|$. Since it does not involve built-in magnetic fields, and hence may be implemented by means of standard silicon technologies, the Rashba effect is expected to be a fundamental ingredient for the realization of the so-called S-FET (Spin Field Effect Transistor) [10], a "spintronic" device in which the information is carried by the electron spin rather than by the electronic current (as in the usual electronic devices). The use of the electron spin, as an additional degree of freedom, may lead to electronic devices of higher speed and lower power consumption. The purpose of this work shall be to contribute to the understanding of how the Rashba effect can be employed in order to control the spin transport in these devices.

In standard electronic device modeling, fluid models and quantum fluid models are expected to be a very useful tool for designers. As compared to kinetic models, they are much more treatable and flexible from the numerical point of view. It is therefore desirable to extend the fluid description to the spinorial case. The existing drift-diffusive models for spin systems can be classified into two categories: the two-component drift-diffusion models and the spin-polarized or matrix based models. Both models have been used in practice, however their mathematical derivation is still at the very beginning (see Refs. [8, 9] for the classical approach and Ref. [1] for the quantum and semi-classical cases) .

In Ref. [1], in particular, a two-component quantum diffusive model for a 2-dimensional population of electrons with Rashba interaction is derived. Let us

summarize briefly the derivation of this model, which will be considered from the numerical point of view in the present work.

The starting point is the von Neumann equation (i.e. the Schrödinger equation for mixed states) for the Hamiltonian (6.1.1), endowed with a collisional term of BGK type

$$i\hbar\partial_t \varrho(t) = [H, \varrho(t)] + \frac{i}{\tau}\left(\varrho_{eq} - \varrho(t)\right),$$

where $\varrho(t) = (\varrho_{ij}(t))$ is the 2×2 density operator, representing the time-dependent mixed state of the system, and τ is the relaxation time. According to the theory developed in Refs. [3, 4] , the local equilibrium state ϱ_{eq} is chosen as the maximizer of a free energy-like functional, subject to the constraint of sharing with $\varrho(t)$ the local moments we are interested in, here the spin-up and spin-down (with respect to the z direction) electron densities n_1, n_2 (or, equivalently, the total electronic density $n_1 + n_2$ and the polarization $n_1 - n_2$). Then, the maximizer, which has the form of a Maxwellian operator, contains as many Lagrange multipliers (chemical potentials) as the chosen moments. These multipliers furnish the degrees of freedom necessary to satisfy the constraint equations. In our case, therefore, the local equilibrium state contains two chemical potential, A_1 and A_2, which depend on n_1 and n_2 through the constraint equations. The rigorous proof of relizability of the quantum Maxwellian associated to a given density and current has been obtained in Refs. [5, 6] for a scalar (i.e. non spinorial) Hamiltonian. By assuming $\tau \ll 1$ and applying the Chapman-Enskog method, the von Neumann equation leads in the limit to the "quantum drift-diffusive" system (6.2.1) for the unknown densities n_1 and n_2. Apart from the chemical potentials A_1 and A_2, which depend on n_1 and n_2 through the constraint, the system also contains some extra moments, namely the off-diagonal density n_{21} and currents J^x_{21}, J^y_{21}, which are computed via the equilibrium state and which depend on n_1 and n_2 as well. Note that, with respect to the original Hamiltonian (6.1.1), we shall work with a scaled version (see the Hamiltonian (6.2.4), which contains also the chemical potential) in which ε is the scaled Planck constant and α is rescaled as $\varepsilon\alpha$. This is, therefore, a semiclassical scaling with the additional assumption of small Rashba constant. Of course, the parameter ε is unimportant as long as we are not interested in the semiclassical behavior but becomes relevant when we look for some semiclassical approximation for small ε.

In summary, the diffusive equations (6.2.1), coupled to Eqs. (6.2.3)–(6.2.7) which represent the equilibrium state and the constraints, and associated with the Poisson equation (6.2.2) for the self-consistent potential, constitute the quantum diffusive model we are going to analyze numerically in this work. Needless to say, the model (6.2.1)–(6.2.7) is rather implicit and involved, and requires a very careful numerical treatment. The aim of the present paper is thus to present two discrete versions of

(6.2.1)–(6.2.7) suitable for time-resolved simulation of the spin populations n_1 and n_2 in a spatially confined, two-dimensional electron gas. In both schemes the finite-difference approximations of the occurring derivatives are first order in time and second order in space. At the core of the numerical study of the present model is the minimization of a functional that either maps from $\mathbb{R}^{3P}$ to $\mathbb{R}$ (in the first scheme) or from $\mathbb{R}^{2P}$ to $\mathbb{R}$ (in the second scheme), where P is the number of points on the space grid. We present an algorithm that uses a combination of the conjugate gradient method and the Newton method in order to find the minimum of the respective functional at each time step. The developed numerical schemes are used to compute the equilibrium spin densities in a common transistor geometry which features a spin-dependent potential barrier.

The paper is organized as follows. In Section 6.2, the continuous model is introduced and is endowed with suitable initial and boundary conditions. In Sec. 6.3 we perform two different time discretizations of the continuous model and prove the well-posedness of each of the two schemes. Then, in Sec. 6.4 two fully discrete schemes (i.e. both in time and space) are introduced and analyzed as well. Finally, Sec. 6.5 is devoted to numerical experiments.

6.2 The quantum spin drift-diffusion model

Let us start with the formulation of the quantum diffusive model introduced in section 6.1. The model describes the evolution of the spin-up and the spin-down densities n_1 and n_2, respectively, of a two-dimensional electron gas by means of the following quantum drift-diffusion equations:

$$\begin{aligned}
&\partial_t n_1 + \nabla\cdot(n_1\nabla(A_1 - V_s)) + \\
&\qquad + \alpha(A_1 - A_2)\,\mathcal{R}\mathrm{e}(\mathcal{D}n_{21}) - 2\alpha\,\mathcal{R}\mathrm{e}(n_{21}\mathcal{D}(A_2 - V_s)) - \\
&\qquad - \frac{2\alpha}{\varepsilon}(A_1 - A_2)\,\mathcal{I}\mathrm{m}(J^x_{21} - iJ^y_{21}) = 0\,, \\
&\partial_t n_2 + \nabla\cdot(n_2\nabla(A_2 - V_s)) + \\
&\qquad + \alpha(A_1 - A_2)\,\mathcal{R}\mathrm{e}(\mathcal{D}n_{21}) + 2\alpha\,\mathcal{R}\mathrm{e}(n_{21}\mathcal{D}(A_1 - V_s)) + \\
&\qquad + \frac{2\alpha}{\varepsilon}(A_1 - A_2)\,\mathcal{I}\mathrm{m}(J^x_{21} - iJ^y_{21}) = 0\,.
\end{aligned} \tag{6.2.1}$$

Here, $\nabla = (\partial_x, \partial_y)$, $\mathcal{D} = \partial_x - i\partial_y$, A_1 and A_2 denote the two chemical potentials (Lagrange multipliers), V_s stands for the self-consistent potential arising from the electron-electron interaction and n_{21}, J^x_{21} and J^y_{21} are off-diagonal elements of the spin-density matrix and the spin-current matrix written in (6.2.6) and (6.2.7), respectively. The parameter $\alpha > 0$ denotes the scaled Rashba constant and $\varepsilon > 0$

stands for the scaled Planck constant (for details regarding the scaling we refer to [1]). We use the following notations for the unknowns in (6.2.1),

$$N = \begin{pmatrix} n_1 & \overline{n}_{21} \\ n_{21} & n_2 \end{pmatrix}, \quad J_{21} = \begin{pmatrix} J_{21}^x \\ J_{21}^y \end{pmatrix}, \quad A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}, \quad V_s \,.$$

The self-consistent potential V_s is determined by the Poisson equation,

$$-\gamma^2 \Delta V_s = n_1 + n_2 \,, \tag{6.2.2}$$

where $\gamma > 0$ is proportional to the occurring Debye length. The system (6.2.1)-(6.2.2) is closed through the fact that the electrons are assumed to be in a quantum local equilibrium state at all times. This constraint allows one to relate the chemical potential A to the spin densities n_1 and n_2 as well as to the spin-mixing quantities n_{21} and J_{21}, respectively. More precisely, if $H(A)$ denotes the system Hamiltonian, the equilibrium state operator is given by

$$\varrho_{eq} = \exp(-H(A)), \tag{6.2.3}$$

where $\exp(\cdot)$ here denotes the operator exponential. In the present case, the Hamiltonian is given by

$$H(A) : D(H) \subset (L^2(\Omega))^2 \to (L^2(\Omega))^2 \,, \quad D(H) \subset (H^2(\Omega))^2 \,,$$

$$H(A) = \begin{pmatrix} -\frac{\varepsilon^2}{2}\Delta + V_{ext,1} + A_1 & \varepsilon^2\alpha(\partial_x - i\partial_y) \\ -\varepsilon^2\alpha(\partial_x + i\partial_y) & -\frac{\varepsilon^2}{2}\Delta + V_{ext,2} + A_2 \end{pmatrix}, \tag{6.2.4}$$

where $\Omega \subset \mathbb{R}^2$ denotes the bounded domain where the electrons are assumed to be confined. Moreover, we introduced two external, time-independent potentials $V_{ext,1}(x)$ and $V_{ext,2}(x)$ for the spin-up and the spin-down electrons, respectively. Assuming that $H(A)$ has a purely eigenvalue spectrum, the eigenvalues and the eigenvectors of $H(A)$, denoted by $\lambda_l(A)$ and $\psi_l(A) = (\psi_l^1(A), \psi_l^2(A))$, respectively, and solutions of

$$H(A)\psi_l(A) = \lambda_l(A)\psi_l(A) \,, \tag{6.2.5}$$

link the chemical potentials to the spin-density matrix N and to the spin-current matrix J,

$$N = \sum_l e^{-\lambda_l} \begin{pmatrix} |\psi_l^1|^2 & \psi_l^1 \overline{\psi_l^2} \\ \psi_l^2 \overline{\psi_l^1} & |\psi_l^2|^2 \end{pmatrix} = \begin{pmatrix} n_1 & \overline{n}_{21} \\ n_{21} & n_2 \end{pmatrix}, \tag{6.2.6}$$

$$\begin{aligned} J &= -\frac{i\varepsilon}{2} \sum_l e^{-\lambda_l} \begin{pmatrix} \overline{\psi_l^1}\nabla\psi_l^1 - \psi_l^1\nabla\overline{\psi_l^1} & \overline{\psi_l^2}\nabla\psi_l^1 - \psi_l^1\nabla\overline{\psi_l^2} \\ \overline{\psi_l^1}\nabla\psi_l^2 - \psi_l^2\nabla\overline{\psi_l^1} & \overline{\psi_l^2}\nabla\psi_l^2 - \psi_l^2\nabla\overline{\psi_l^2} \end{pmatrix} \\ &= \begin{pmatrix} J_1 & \overline{J}_{21} \\ J_{21} & J_2 \end{pmatrix}. \end{aligned} \tag{6.2.7}$$

The formulas (6.2.6) and (6.2.7) are the standard textbook expressions for the spin-density and the spin-current, respectively, corresponding to the density operator (6.2.3). The system (6.2.1)-(6.2.2) is now closed through the non-local relations $N(A)$ and $J(A)$, given by Eqs. (6.2.4)-(6.2.7). It is not yet clear whether these relations are invertible, i.e. if it is possible to compute $A(n_1, n_2)$. Hence, the equations (6.2.1) can be viewed as evolution equations for the chemical potentials A_1 and A_2 rather than for the spin densities n_1 and n_2. Indeed, the two time-discretizations of the system (6.2.1)-(6.2.7), which will be developed in section 6.3, represent these two possible viewpoints regarding the evolution equations (6.2.1).

Let $\Omega \subset \mathbb{R}^2$ be a bounded regular domain with boundary $\partial\Omega$. All the unknowns N, J_{21}, A and V_s of the system (6.2.1)-(6.2.7) are defined on the domain $[0,T] \times \Omega$. The index of the eigenvalues and the eigenvectors of $H(A)$ is $l \in \mathbb{N}$. We shall impose Dirichlet boundary conditions for the eigenvectors ψ_l,

$$\psi_l(x) = 0 \quad \text{for} \quad x \in \partial\Omega\,,$$

hence the current across the domain boundary $\partial\Omega$ is zero. As we will briefly show at the end of this section, the Hamiltonian (6.2.4) is not hermitian in $(L^2(\Omega))^2$ when imposing Neumann conditions on the wavefunctions $\psi \in (H^2(\Omega))^2$. The study of this problem as well as the implementation of transparent boundary conditions will be the topic of a forthcoming work. The self-consistent potential V_s is supplemented with Dirichlet conditions too,

$$V_s(x) = 0 \quad \text{for} \quad x \in \partial\Omega\,.$$

The chemical potentials A_1 and A_2 are allowed to move freely at the boundary,

therefore we take Neumann conditions,

$$\nabla(A_1(x) - V_s(x)) \cdot \nu(x) = 0 \quad \text{for} \quad x \in \partial\Omega\,,$$
$$\nabla(A_2(x) - V_s(x)) \cdot \nu(x) = 0 \quad \text{for} \quad x \in \partial\Omega\,.$$

Here, $\nu(x)$ denotes the outward normal to the boundary $\partial\Omega$ at x. Considering the initial conditions, one has two choices depending on the point of view of the evolution equations (6.2.1). Since we do not know whether or not (6.2.6) is invertible, the safe approach is to provide initial data for the chemical potentials. However, from the viewpoint of device modeling, it is more appealing to start from initial spin densities. We shall take the latter approach and assume that $n_1(0,x)$ and $n_2(0,x)$ are smooth and bounded.

In summary, we have the following quantum spin-drift-diffusion model,

$$\partial_t n_1 + \nabla \cdot (n_1 \nabla(A_1 - V_s)) + \alpha(A_1 - A_2)\,\mathcal{R}\mathrm{e}(\mathcal{D} n_{21}) \tag{6.2.8}$$
$$- 2\alpha\,\mathcal{R}\mathrm{e}(n_{21}\mathcal{D}(A_2 - V_s)) - \frac{2\alpha}{\varepsilon}(A_1 - A_2)\,\mathcal{I}\mathrm{m}(J_{21}^x - iJ_{21}^y) = 0\,,$$

$$\partial_t n_2 + \nabla \cdot (n_2 \nabla(A_2 - V_s)) + \alpha(A_1 - A_2)\,\mathcal{R}\mathrm{e}(\mathcal{D} n_{21}) \tag{6.2.9}$$
$$+ 2\alpha\,\mathcal{R}\mathrm{e}(n_{21}\mathcal{D}(A_1 - V_s)) + \frac{2\alpha}{\varepsilon}(A_1 - A_2)\,\mathcal{I}\mathrm{m}(J_{21}^x - iJ_{21}^y) = 0\,,$$

$$-\gamma^2 \Delta V_s = n_1 + n_2\,, \tag{6.2.10}$$

$$H(A)\psi_l(A) = \lambda_l(A)\psi_l(A)\,, \tag{6.2.11}$$

$$N = \sum_l e^{-\lambda_l(A)} \begin{pmatrix} |\psi_l^1(A)|^2 & \psi_l^1(A)\,\overline{\psi_l^2}(A) \\ \psi_l^2(A)\,\overline{\psi_l^1}(A) & |\psi_l^2(A)|^2 \end{pmatrix}, \tag{6.2.12}$$

$$J_{21} = -\frac{i\varepsilon}{2} \sum_l e^{-\lambda_l(A)} \left(\overline{\psi_l^1}(A) \nabla \psi_l^2(A) - \psi_l^2(A) \nabla \overline{\psi_l^1}(A) \right), \tag{6.2.13}$$

where the Hamiltonian $H(A)$ is given by (6.2.4), and supplemented with the follow-

ing initial- and boundary conditions,

$$
\begin{aligned}
&n_1(t=0,x) = n_1^0(x)\,, \quad n_2(t=0,x) = n_2^0(x) \quad \text{for} \quad x \in \Omega\,,\\
&V_s(x) = 0 \quad \text{for} \quad x \in \partial\Omega\,,\\
&\psi_l(x) = 0 \quad \text{for} \quad x \in \partial\Omega\,,\\
&\nabla(A_1(x) - V_s(x)) \cdot \nu(x) = 0 \quad \text{for} \quad x \in \partial\Omega\,,\\
&\nabla(A_2(x) - V_s(x)) \cdot \nu(x) = 0 \quad \text{for} \quad x \in \partial\Omega\,.
\end{aligned}
\tag{6.2.14}
$$

6.2.1 Hermiticity of the Hamiltonian

We briefly show that the Hamiltonian (6.2.4) is not hermitian in $(L^2(\Omega))^2$ when imposing Neumann boundary conditions on the wave functions $\psi \in (H^2(\Omega))^2$. Let us start with

$$
(H(A)\psi, \chi)_{(L^2(\Omega))^2} = \int_\Omega (\overline{\chi^1}, \overline{\chi^2}) \begin{pmatrix} -\frac{\varepsilon^2}{2}\,\Delta\psi^1 + (V_{ext,1} + A_1)\psi^1 + \varepsilon^2\alpha\mathcal{D}\psi^2 \\ -\frac{\varepsilon^2}{2}\,\Delta\psi^2 + (V_{ext,2} + A_2)\psi^2 - \varepsilon^2\alpha\mathcal{D}\psi^1 \end{pmatrix} dx\,.
$$

where $(\cdot,\cdot)_{(L^2(\Omega))^2}$ denotes the scalar product in $(L^2(\Omega))^2$. Specifically, let us look at the Rashba coupling terms,

$$
\begin{aligned}
\int_\Omega (\overline{\chi^1}\mathcal{D}\psi^2 - \overline{\chi^2}\mathcal{D}\psi^1)\,dx = &- \int_\Omega (\psi^2\mathcal{D}\overline{\chi^1} - \psi^1\mathcal{D}\overline{\chi^2})\,dx + \\
&+ \int_{\partial\Omega} \overline{\chi^1}\psi^2(1,-i)\cdot\nu(x)\,d\sigma - \int_{\partial\Omega} \overline{\chi^2}\psi^1(1,-i)\cdot\nu(x)\,d\sigma
\end{aligned}
\tag{6.2.15}
$$

Here, the boundary terms do not vanish when imposing Neumann conditions. However, if we considered the problem in the whole space $\Omega = \mathbb{R}^2$, the boundary terms would vanish and the Hamiltonian would be hermitian. Considering the problem in the whole $\mathbb{R}^2$ means, from the numerical point of view, imposing transparent boundary conditions for ψ. This would be the topic of a future paper.

6.3 Semi-discretization in time

In this section we take a first step towards a full space-time discretization of the system (6.2.8)-(6.2.14) by discretizing the time domain. The purpose of the semi-discretization is two-fold. Firstly, since the space discretization of the present two-dimensional spin model is quite involved, the functional formalism which will be applied in this work becomes much more transparent in the semi-discrete case than in the fully discrete case. Secondly, in contrast to the continuous case given by Eqs.

(6.2.8)-(6.2.14), existence and uniqueness of solutions of the semi-discrete system can be proven. Two different semi-discretizations will be presented. The first one was studied in [7] for a scalar quantum diffusive model (without the Rashba spin-orbit coupling). We shall use some of the techniques elaborated in [7] and apply them to the present spin model. The second semi-discrete scheme is an explicit one which relies heavily on the ability to invert the relation (6.2.12). Its benefits lie in the fact that, when passing to the full discretization, its treatment is far less involved as compared to the first scheme.

We remark here that the following identities applied in Eqs. (6.2.8)-(6.2.9) will be helpful in the subsequent analysis,

$$\begin{aligned}(A_1 - A_2)\mathcal{D}n_{21}^k - 2n_{21}^k\mathcal{D}(A_2) &= \mathcal{D}(n_{21}^k(A_1 - A_2)) - n_{21}^k\mathcal{D}(A_1 + A_2)\,,\\ (A_1 - A_2)\mathcal{D}n_{21}^k + 2n_{21}^k\mathcal{D}(A_1) &= \mathcal{D}(n_{21}^k(A_1 - A_2)) + n_{21}^k\mathcal{D}(A_1 + A_2)\,.\end{aligned} \tag{6.3.1}$$

6.3.1 A first semi-discrete system

Suppose $T > 0$ and let $t \in [0, T]$ with the discretization

$$t_k = k\Delta t\,, \quad k \in \{0, 1, \ldots, K\}\,, \quad \Delta t := \frac{T}{K}\,.$$

Then, inspired by [7], we choose the following time-discretization of the continuous problem (6.2.8)-(6.2.13),

$$\begin{aligned}\frac{n_1(A^{k+1}) - n_1^k}{\Delta t} &+ \nabla\cdot(n_1^k\nabla(A_1^{k+1} - V_s^{k+1})) + \alpha\,\mathcal{R}\mathrm{e}[\mathcal{D}(n_{21}^k(A_1^{k+1} - A_2^{k+1}))]\\ &- \alpha\,\mathcal{R}\mathrm{e}[n_{21}^k\mathcal{D}(A_1^{k+1} + A_2^{k+1} - 2V_s^{k+1})]\\ &- \frac{2\alpha}{\varepsilon}(A_1^{k+1} - A_2^{k+1})\,\mathcal{I}\mathrm{m}(J_x^{21,k} - iJ_y^{21,k}) = 0\,,\end{aligned} \tag{6.3.2}$$

$$\begin{aligned}\frac{n_2(A^{k+1}) - n_2^k}{\Delta t} &+ \nabla\cdot(n_2^k\nabla(A_2^{k+1} - V_s^{k+1})) + \alpha\,\mathcal{R}\mathrm{e}[\mathcal{D}(n_{21}^k(A_1^{k+1} - A_2^{k+1}))]\\ &+ \alpha\,\mathcal{R}\mathrm{e}[n_{21}^k\mathcal{D}(A_1^{k+1} + A_2^{k+1} - 2V_s^{k+1})]\\ &+ \frac{2\alpha}{\varepsilon}(A_1^{k+1} - A_2^{k+1})\,\mathcal{I}\mathrm{m}(J_x^{21,k} - iJ_y^{21,k}) = 0\,,\end{aligned} \tag{6.3.3}$$

$$-\gamma^2\Delta V_s^{k+1} = n_1(A^{k+1}) + n_2(A^{k+1})\,, \tag{6.3.4}$$

$$H(A^{k+1})\psi_l^{k+1} = \lambda_l^{k+1}\psi_l^{k+1}\,, \tag{6.3.5}$$

$$n_1(A^{k+1}) = \sum_l e^{-\lambda_l^{k+1}} |\psi_l^{1,k+1}|^2 \,, \quad n_2(A^{k+1}) = \sum_l e^{-\lambda_l^{k+1}} |\psi_l^{2,k+1}|^2 \,. \tag{6.3.6}$$

In this scheme one searches for the unknowns (A^{k+1}, V_s^{k+1}), given (N^k, J_{21}^k). The main difficulty concerning the solution of this system are the non-local relations (6.3.5)-(6.3.6). We shall thus construct a mapping $(A, V_s) \in (H^1(\Omega, \mathbb{R}))^3 \mapsto \mathcal{F}(A, V_s) \in \mathbb{R}$ whose unique minimum (A^{k+1}, V_s^{k+1}) is the solution of Eqs. (6.3.2)-(6.3.6). Once A^{k+1} and the eigenvalues λ_l^{k+1} respectively eigenvectors ψ_l^{k+1} are known, Eqs. (6.2.12)-(6.2.13) can be used to compute (N^{k+1}, J_{21}^{k+1}) and the process can be repeated. Let us thus introduce the two functionals

$$\mathcal{G} : (L^2(\Omega, \mathbb{R}))^2 \to \mathbb{R} \,, \quad \mathcal{F} : (H^1(\Omega, \mathbb{R}))^3 \to \mathbb{R} \,,$$

defined by

$$\mathcal{G}(A) := \sum_l e^{-\lambda_l(A)} \,, \qquad A \in (L^2(\Omega, \mathbb{R}))^2 \,, \tag{6.3.7}$$

where $\lambda_l(A)$ are the eigenvalues of the Hamiltonian (6.2.4), and

$$\mathcal{F}(A, V_s) = \mathcal{G}(A) + \mathcal{F}_1(A, V_s) + \mathcal{F}_2(A, V_s) + \mathcal{F}_3(A, V_s) + \mathcal{F}_4(A) \,, \tag{6.3.8}$$

where

$$\mathcal{F}_1(A, V_s) := \frac{\Delta t}{2} \int_\Omega n_1^k |\nabla(A_1 - V_s)|^2 \, dx + \frac{\Delta t}{2} \int_\Omega n_2^k |\nabla(A_2 - V_s)|^2 \, dx \,, \tag{6.3.9}$$

$$\mathcal{F}_2(A, V_s) := \frac{\gamma^2}{2} \int_\Omega |\nabla V_s|^2 \, dx + (n_1^k, A_1 - V_s) + (n_2^k, A_2 - V_s) \,, \tag{6.3.10}$$

$$\mathcal{F}_3(A, V_s) := \alpha \Delta t \, \mathcal{R}\mathrm{e} \left\{ \int_\Omega n_{21}^k (A_1 - A_2) \mathcal{D}(A_1 + A_2 - 2V_s) \, dx \right\} \,, \tag{6.3.11}$$

$$\mathcal{F}_4(A) := \frac{\alpha \Delta t}{\varepsilon} \, \mathcal{I}\mathrm{m} \left\{ \int_\Omega (A_1 - A_2)^2 (J_x^{21,k} - i J_y^{21,k}) \, dx \right\} \,. \tag{6.3.12}$$

The first and second Gateaux derivative of the functionals (6.3.7)-(6.3.12) can be found in appendix 6.B and 6.C, respectively. One can immediately see that a solution (A^{k+1}, V_s^{k+1}) of the semi-discrete system (6.3.2)-(6.3.6) satisfies

$$\mathrm{d}\mathcal{F}(A^{k+1}, V_s^{k+1})(\delta A, \delta V_s) = 0 \qquad \forall \, (\delta A, \delta V_s) \in (H^1(\Omega, \mathbb{R}))^3 \,,$$

and inversely. Thus, it remains to show that $\mathcal{F}$ has a unique extremum (minimum). This can be achieved in two steps and is detailed in appendix 6.C. First we show that $\mathcal{F}$ is strictly convex. Then it is sufficient to show that $\mathcal{F}$ is coercive to obtain the existence and uniqueness of the extremum (A^{k+1}, V_s^{k+1}), solution of the system (6.3.2)-(6.3.6).

6.3.2 A second semi-discrete system

We suggest here an alternative way to discretize in time the quantum drift-diffusion model (6.2.8)-(6.2.14). It is based on the point of view that one advances the spin densities in time, rather than the chemical potentials. We shall implement an explicit forward Euler scheme:

$$\begin{aligned}\frac{n_1^{k+1}-n_1^k}{\Delta t} &+ \nabla\cdot(n_1^k\nabla(A_1^k-V_s^k)) + \alpha\,\mathcal{R}\mathrm{e}\{\mathcal{D}(n_{21}^k(A_1^k-A_2^k))\} \\ &- \alpha\,\mathcal{R}\mathrm{e}(n_{21}^k\mathcal{D}(A_1^k+A_2^k-2V_s^k)) - \frac{2\alpha}{\varepsilon}(A_1^k-A_2^k)\,\mathcal{I}\mathrm{m}(J_{21}^{x,k}-iJ_{21}^{y,k}) \\ &= 0\,,\end{aligned} \tag{6.3.13}$$

$$\begin{aligned}\frac{n_2^{k+1}-n_2^k}{\Delta t} &+ \nabla\cdot(n_2^k\nabla(A_2^k-V_s^k)) + \alpha\,\mathcal{R}\mathrm{e}\{\mathcal{D}[n_{21}^k(A_1^k-A_2^k)]\} \\ &+ \alpha\,\mathcal{R}\mathrm{e}[n_{21}^k\mathcal{D}(A_1^k+A_2^k-2V_s^k)] + \frac{2\alpha}{\varepsilon}(A_1^k-A_2^k)\,\mathcal{I}\mathrm{m}(J_{21}^{x,k}-iJ_{21}^{y,k}) \\ &= 0\,,\end{aligned} \tag{6.3.14}$$

$$-\gamma^2\Delta V_s^k = n_1^k + n_2^k\,, \tag{6.3.15}$$

$$H(A^k)\psi_l^k = \lambda_l^k\psi_l^k\,, \tag{6.3.16}$$

$$N = \sum_l e^{-\lambda_l^k}\begin{pmatrix} |\psi_l^{1,k}|^2 & \psi_l^{1,k}\,\overline{\psi_l^{2,k}} \\ \psi_l^{2,k}\,\overline{\psi_l^{1,k}} & |\psi_l^{2,k}|^2 \end{pmatrix}, \tag{6.3.17}$$

$$J_{21}^k = -\frac{i\varepsilon}{2}\sum_l e^{-\lambda_l^k}\left(\overline{\psi_l^{1,k}}\nabla\psi_l^{2,k} - \psi_l^{2,k}\nabla\overline{\psi_l^{1,k}}\right). \tag{6.3.18}$$

In this case, given the spin-densities (n_1^k, n_2^k), one first uses the Poisson equation (6.3.15) to get V_s^k, then inverts the formulas (6.3.16)-(6.3.17) in order to get the chemical potentials (A_1^k, A_2^k). Finally one advances in time, using the drift-diffusion equations (6.3.13)-(6.3.14) in order to get the new spin densities (n_1^{k+1}, n_2^{k+1}) and one repeats the steps. The inversion of the non-local relation (6.3.16)-(6.3.17) can be achieved by minimizing the functional $\mathcal{G}_n : (L^2(\Omega,\mathbb{R}))^2 \to \mathbb{R}$, defined by

$$\mathcal{G}_n(A) := \mathcal{G}(A) + \int_\Omega n_1^k A_1\,dx + \int_\Omega n_2^k A_2\,dx \tag{6.3.19}$$

Indeed, the first derivative of this functional reads

$$\begin{aligned} \mathrm{d}\mathcal{G}_n(A)(\delta A) = &- \sum_l e^{-\lambda_l(A)} \int_\Omega \left(|\psi_l^1(A)|^2 \delta A_1 + |\psi_l^2(A)|^2 \delta A_2 \right) dx \\ &+ \int_\Omega n_1^k \delta A_1 \, dx + \int_\Omega n_2^k \delta A_2 \, dx \,. \end{aligned} \tag{6.3.20}$$

As shown in appendix 6.B, the functional $\mathcal{G}_n$ is strictly convex and coercive.

Remark 6.3.1. *The two semi-discrete systems presented in this section conserve the total mass* $(n_1 + n_2)$ *because of the particular choice of Dirichlet boundary conditions for the eigenvectors* ψ_l *of the Hamiltonian* (6.2.4)*. This can be obtained by integrating the sum of the semi-discrete drift-diffusion equations for* n_1 *and* n_2*, Eqs.* (6.3.2)-(6.3.3) *or* (6.3.13)-(6.3.14)*, respectively, over the domain* Ω*. The remaining boundary term is of the form*

$$\int_{\partial\Omega} n_{21}(A_1 - A_2)(1, -i) \cdot \nu(x) \, d\sigma \,,$$

which does not vanish for Neumann boundary conditions. This is in accordance with the result obtained in subsection 6.2.1, where we showed that Neumann conditions for ψ_l *lead to a non-hermitian Hamiltonian* (6.2.4) *in* $(L^2(\Omega))^2$*.*

6.4 Fully discrete system

This section is devoted to the full discretization of the continuous spin QDD model (6.2.8)-(6.2.14). The time discretization was done in the previous section, now we focus on the space discretization. Let $x \in \Omega = [0,1] \times [0,1]$ with the discretization

$$\begin{gathered} x_{ij} = \left((j-1)\Delta x \,, (i-1)\Delta y \right), \quad j \in \{1, 2, \ldots, M\}, \quad i \in \{1, 2, \ldots, N\}, \\ \Delta x := \frac{1}{M-1}, \qquad \Delta y := \frac{1}{N-1}. \end{gathered}$$

For functions $f(x)$ on Ω we write $f(x_{ij}) = f_{ij}$. A function $f(x)$ that is subjected to homogenous Dirichlet boundary conditions on $\partial\Omega$ satisfies

$$f_{1j} = f_{Nj} = 0 \quad \forall\, j \in \{1, 2, \ldots, M\}, \quad f_{i1} = f_{iM} = 0 \quad \forall\, i \in \{1, 2, \ldots, N\}.$$

We introduce the following index transformation,

$$(i, j) \mapsto p \qquad \forall \quad i \in \{2, \ldots, N-1\}, \; j \in \{2, \ldots, M-1\},$$

defined by

$$p = (N-2)(j-2) + i - 1\,, \quad p = 1, \ldots, P\,, \qquad P := (N-2)(M-2)\,.$$

For discrete functions $(f_{ij})_{i,j=2}^{N-1,M-1}$ in Ω the following vector notation will be implemented:

$$\hat{f} := (f_p)_{p=1}^{P} \in \mathbb{K}^P\,, \tag{6.4.1}$$

where $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$. The corresponding euclidean scalar product is denoted by

$$(\hat{f}, \hat{g})_P = \Delta x \Delta y \sum_p f_p \overline{g}_p = \Delta x \Delta y \sum_{i=2}^{N-1} \sum_{j=2}^{M-1} f_{ij} \overline{g}_{ij}\,.$$

6.4.1 A first fully discrete system (scheme 1)

The discretization matrices used in the following are defined in Appendix 6.D. In view of the boundary conditions (6.2.14), we choose the following space discretization of the semi-discrete system (6.3.2)-(6.3.6),

$$\begin{aligned}
&\frac{\hat{n}_1(\hat{A}_1^{k+1}, \hat{A}_2^{k+1}) - \hat{n}_1^k}{\Delta t} - \frac{2\alpha}{\varepsilon}(\hat{A}_1^{k+1} - \hat{A}_2^{k+1}) \circ \mathcal{I}\mathrm{m}(\hat{J}_x^{21,k} - i\hat{J}_y^{21,k}) \qquad (6.4.2)\\
&- \frac{1}{2}(D_x^+)^T[\hat{n}_1^k \circ D_x^+(\hat{A}_1^{k+1} - \hat{V}_s^{k+1})] - \frac{1}{2}(D_x^-)^T[\hat{n}_1^k \circ D_x^-(\hat{A}_1^{k+1} - \hat{V}_s^{k+1})]\\
&- \frac{1}{2}(D_y^+)^T[\hat{n}_1^k \circ D_y^+(\hat{A}_1^{k+1} - \hat{V}_s^{k+1})] - \frac{1}{2}(D_y^-)^T[\hat{n}_1^k \circ D_y^-(\hat{A}_1^{k+1} - \hat{V}_s^{k+1})]\\
&- \alpha\,\mathcal{R}\mathrm{e}\left\{\widetilde{D}_x^T[\hat{n}_{21}^k \circ (\hat{A}_1^{k+1} - \hat{A}_2^{k+1})]\right\} + \alpha\,\mathcal{R}\mathrm{e}\left\{i\widetilde{D}_y^T[\hat{n}_{21}^k \circ (\hat{A}_1^{k+1} - \hat{A}_2^{k+1})]\right\}\\
&- \alpha\,\mathcal{R}\mathrm{e}\left\{\hat{n}_{21}^k \circ [\widetilde{D}_x(\hat{A}_1^{k+1} + \hat{A}_2^{k+1} - 2\hat{V}_s^{k+1})]\right\}\\
&+ \alpha\,\mathcal{R}\mathrm{e}\left\{i\hat{n}_{21}^k \circ [\widetilde{D}_y(\hat{A}_1^{k+1} + \hat{A}_2^{k+1} - 2\hat{V}_s^{k+1})]\right\} = 0\,,
\end{aligned}$$

$$\frac{\hat{n}_2(\hat{A}_1^{k+1},\hat{A}_2^{k+1})-\hat{n}_2^k}{\Delta t}+\frac{2\alpha}{\varepsilon}(\hat{A}_1^{k+1}-\hat{A}_2^{k+1})\circ\mathcal{I}\mathrm{m}(\hat{J}_x^{21,k}-i\hat{J}_y^{21,k}) \tag{6.4.3}$$
$$-\frac{1}{2}(D_x^+)^T[\hat{n}_2^k\circ D_x^+(\hat{A}_2^{k+1}-\hat{V}_s^{k+1})]-\frac{1}{2}(D_x^-)^T[\hat{n}_2^k\circ D_x^-(\hat{A}_2^{k+1}-\hat{V}_s^{k+1})]$$
$$-\frac{1}{2}(D_y^+)^T[\hat{n}_2^k\circ D_y^+(\hat{A}_2^{k+1}-\hat{V}_s^{k+1})]-\frac{1}{2}(D_y^-)^T[\hat{n}_2^k\circ D_y^-(\hat{A}_2^{k+1}-\hat{V}_s^{k+1})]$$
$$-\alpha\,\mathcal{R}\mathrm{e}\left\{\widetilde{D}_x^T[\hat{n}_{21}^k\circ(\hat{A}_1^{k+1}-\hat{A}_2^{k+1})]\right\}+\alpha\,\mathcal{R}\mathrm{e}\left\{i\widetilde{D}_y^T[\hat{n}_{21}^k\circ(\hat{A}_1^{k+1}-\hat{A}_2^{k+1})]\right\}$$
$$+\alpha\,\mathcal{R}\mathrm{e}\left\{\hat{n}_{21}^k\circ[\widetilde{D}_x(\hat{A}_1^{k+1}+\hat{A}_2^{k+1}-2\hat{V}_s^{k+1})]\right\}$$
$$-\alpha\,\mathcal{R}\mathrm{e}\left\{i\hat{n}_{21}^k\circ[\widetilde{D}_y(\hat{A}_1^{k+1}+\hat{A}_2^{k+1}-2\hat{V}_s^{k+1})]\right\}=0\,,$$

$$-\gamma^2\Delta_{dir}\hat{V}_s^{k+1}=\hat{n}_1(\hat{A}_1^{k+1},\hat{A}_2^{k+1})+\hat{n}_2(\hat{A}_1^{k+1},\hat{A}_2^{k+1})\,, \tag{6.4.4}$$

$$H(\hat{A}_1^{k+1},\hat{A}_2^{k+1})\begin{pmatrix}\hat{\psi}_l^{1,k+1}\\ \hat{\psi}_l^{2,k+1}\end{pmatrix}=\lambda_l^{k+1}\begin{pmatrix}\hat{\psi}_l^{1,k+1}\\ \hat{\psi}_l^{2,k+1}\end{pmatrix}, \tag{6.4.5}$$

$$\hat{n}_1(\hat{A}_1^{k+1},\hat{A}_2^{k+1})=\sum_l e^{-\lambda_l^{k+1}}\hat{\psi}_l^{1,k+1}\circ\overline{\hat{\psi}_l^{1,k+1}}\,, \tag{6.4.6}$$
$$\hat{n}_2(\hat{A}_1^{k+1},\hat{A}_2^{k+1})=\sum_l e^{-\lambda_l^{k+1}}\hat{\psi}_l^{2,k+1}\circ\overline{\hat{\psi}_l^{2,k+1}}\,, \tag{6.4.7}$$

$$\hat{n}_{21}^k=\sum_l e^{-\lambda_l^k}\hat{\psi}_l^{2,k}\circ\overline{\hat{\psi}_l^{1,k}}\,, \tag{6.4.8}$$

$$\hat{J}_{21}^{x,k}=-\frac{i\varepsilon}{2}\sum_l e^{-\lambda_l^k}\left[D_x(\hat{\psi}_l^{2,k})\circ\overline{\hat{\psi}_l^{1,k}}-\hat{\psi}_l^{2,k}\circ D_x(\overline{\hat{\psi}_l^{1,k}})\right], \tag{6.4.9}$$
$$\hat{J}_{21}^{y,k}=-\frac{i\varepsilon}{2}\sum_l e^{-\lambda_l^k}\left[D_y(\hat{\psi}_l^{2,k})\circ\overline{\hat{\psi}_l^{1,k}}-\hat{\psi}_l^{2,k}\circ D_y(\overline{\hat{\psi}_l^{1,k}})\right]. \tag{6.4.10}$$

Here, the operator $\circ$ symbolizes the component by component multiplication of two vectors and the Hamiltonian $H(\hat{A}^{k+1})$ is given by

$$H(\hat{A}_1^{k+1}, \hat{A}_2^{k+1}) = \\ = \begin{pmatrix} -\frac{\varepsilon^2}{2}\Delta_{dir} + \mathrm{dg}(\hat{V}_{ext,1} + \hat{A}_1^{k+1}) & \varepsilon^2\alpha(D_x - iD_y) \\ -\varepsilon^2\alpha(D_x + iD_y) & -\frac{\varepsilon^2}{2}\Delta_{dir} + \mathrm{dg}(\hat{V}_{ext,2} + \hat{A}_2^{k+1}) \end{pmatrix},$$

where $\mathrm{dg}(\hat{f})$ stands for a diagonal $P \times P$ matrix where the diagonal elements are the components f_p of $\hat{f}$. The scheme (6.4.2)-(6.4.10) is consistent with the continuous model (6.2.8)-(6.2.14). It is of first order in time and of second order in space. Due to its rather implicit nature, the scheme (6.4.2)-(6.4.10) is not subjected to any stability condition. The solution $(\hat{A}_1^{k+1}, \hat{A}_2^{k+1}, \hat{V}_s^{k+1})$ of the system (6.4.2)-(6.4.10) is the minimizer of the following discrete functional $\widehat{\mathcal{F}}(\hat{A}_1, \hat{A}_2, \hat{V}_s) : \mathbb{R}^{3P} \to \mathbb{R}$,

$$\begin{aligned} \widehat{\mathcal{F}}(\hat{A}_1, \hat{A}_2, \hat{V}_s) : = \widehat{\mathcal{G}}(\hat{A}_1, \hat{A}_2) + \widehat{\mathcal{F}}_1(\hat{A}_1, \hat{A}_2, \hat{V}_s) + \\ + \widehat{\mathcal{F}}_2(\hat{A}_1, \hat{A}_2, \hat{V}_s) + \widehat{\mathcal{F}}_3(\hat{A}_1, \hat{A}_2, \hat{V}_s) + \widehat{\mathcal{F}}_4(\hat{A}_1, \hat{A}_2)\,, \end{aligned} \tag{6.4.11}$$

where

$$\widehat{\mathcal{G}}(\hat{A}_1, \hat{A}_2) := \sum_{l=1}^{2P} e^{-\lambda_l(\hat{A}_1, \hat{A}_2)}\,, \tag{6.4.12}$$

$$\begin{aligned} \widehat{\mathcal{F}}_1(\hat{A}_1, \hat{A}_2, \hat{V}_s) := \frac{\Delta t}{4} \Big[& (\hat{n}_1^k \circ D_x^+(\hat{A}_1 - \hat{V}_s), D_x^+(\hat{A}_1 - \hat{V}_s))_P \\ & + (\hat{n}_1^k \circ D_x^-(\hat{A}_1 - \hat{V}_s), D_x^-(\hat{A}_1 - \hat{V}_s))_P + (\hat{n}_1^k \circ D_y^+(\hat{A}_1 - \hat{V}_s), D_y^+(\hat{A}_1 - \hat{V}_s))_P \\ & + (\hat{n}_1^k \circ D_y^-(\hat{A}_1 - \hat{V}_s), D_y^-(\hat{A}_1 - \hat{V}_s))_P + (\hat{n}_2^k \circ D_x^+(\hat{A}_2 - \hat{V}_s), D_x^+(\hat{A}_2 - \hat{V}_s))_P \\ & + (\hat{n}_2^k \circ D_x^-(\hat{A}_2 - \hat{V}_s), D_x^-(\hat{A}_2 - \hat{V}_s))_P + (\hat{n}_2^k \circ D_y^+(\hat{A}_2 - \hat{V}_s), D_y^+(\hat{A}_2 - \hat{V}_s))_P \\ & + (\hat{n}_2^k \circ D_y^-(\hat{A}_2 - \hat{V}_s), D_y^-(\hat{A}_2 - \hat{V}_s))_P \Big]\,, \end{aligned} \tag{6.4.13}$$

$$\begin{aligned} \widehat{\mathcal{F}}_2(\hat{A}_1, \hat{A}_2, \hat{V}_s) : = & (\hat{n}_1^k, \hat{A}_1 - \hat{V}_s)_P + (\hat{n}_2^k, \hat{A}_2 - \hat{V}_s)_P \\ & + \frac{\gamma^2}{2} \left[(D_x^b \hat{V}_s, D_x^b \hat{V}_s)_P + (D_y^b \hat{V}_s, D_y^b \hat{V}_s)_P \right] \\ & + \frac{\Delta y}{\Delta x} \sum_{i=1}^{N} V_{s,iM}^2 + \frac{\Delta x}{\Delta y} \sum_{j=1}^{M} V_{s,Nj}^2 \end{aligned} \tag{6.4.14}$$

$$\begin{aligned}\widehat{\mathcal{F}}_3(\hat{A}_1,\hat{A}_2,\hat{V}_s) := \alpha\Delta t\,\mathcal{R}\mathrm{e}\Big[&\Big(\hat{n}_{21}^k\circ(\hat{A}_1-\hat{A}_2),\widetilde{D}_x(\hat{A}_1+\hat{A}_2-2\hat{V}_s)\Big)_P\\ &-i\Big(\hat{n}_{21}^k\circ(\hat{A}_1-\hat{A}_2),\widetilde{D}_y(\hat{A}_1+\hat{A}_2-2\hat{V}_s)\Big)_P\Big]\end{aligned} \tag{6.4.15}$$

$$\widehat{\mathcal{F}}_4(\hat{A}_1,\hat{A}_2) := \frac{\alpha\Delta t}{\varepsilon}\,\mathcal{I}\mathrm{m}\left[\Big((\hat{A}_1-\hat{A}_2)\circ(\hat{A}_1-\hat{A}_2),\hat{J}_x^{21,k}-i\hat{J}_y^{21,k}\Big)_P\right]. \tag{6.4.16}$$

Using the relation

$$\begin{aligned}-(\hat{V}_s,\Delta_{dir}\hat{V}_s)_P = &(D_x^b\hat{V}_s,D_x^b\hat{V}_s)_P+(D_y^b\hat{V}_s,D_y^b\hat{V}_s)_P\\ &+\frac{\Delta y}{\Delta x}\sum_{i=1}^{N}V_{s,iM}^2+\frac{\Delta x}{\Delta y}\sum_{j=1}^{M}V_{s,Nj}^2\,,\end{aligned} \tag{6.4.17}$$

it can be readily verified that a solution $(\hat{A}_1^{k+1},\hat{A}_2^{k+1},\hat{V}_s^{k+1})$ of (6.4.2)-(6.4.10) satisfies

$$\mathrm{d}\widehat{\mathcal{F}}(\hat{A}_1^{k+1},\hat{A}_2^{k+1},\hat{V}_s^{k+1})(\delta\hat{A},\delta\hat{V}_s)=0\qquad\forall\,(\delta\hat{A}_1,\delta\hat{A}_2,\delta\hat{V}_s)\in\mathbb{R}^{3P}\,.$$

6.4.2 A second fully discrete system (scheme 2)

We chose the following space discretization of the forward Euler scheme (6.3.13)-(6.3.18):

$$\begin{aligned}&\frac{\hat{n}_1^{k+1}-\hat{n}_1^k}{\Delta t}-\frac{2\alpha}{\varepsilon}(\hat{A}_1^k-\hat{A}_2^k)\circ\mathcal{I}\mathrm{m}(\hat{J}_x^{21,k}-i\hat{J}_y^{21,k})\\ &-\frac{1}{2}(D_x^+)^T[\hat{n}_1^k\circ D_x^+(\hat{A}_1^k-\hat{V}_s^k)]-\frac{1}{2}(D_x^-)^T[\hat{n}_1^k\circ D_x^-(\hat{A}_1^k-\hat{V}_s^k)]\\ &-\frac{1}{2}(D_y^+)^T[\hat{n}_1^k\circ D_y^+(\hat{A}_1^k-\hat{V}_s^k)]-\frac{1}{2}(D_y^-)^T[\hat{n}_1^k\circ D_y^-(\hat{A}_1^k-\hat{V}_s^k)]\\ &-\alpha\,\mathcal{R}\mathrm{e}\left\{\widetilde{D}_x^T[\hat{n}_{21}^k\circ(\hat{A}_1^k-\hat{A}_2^k)]\right\}+\alpha\,\mathcal{R}\mathrm{e}\left\{i\widetilde{D}_y^T[\hat{n}_{21}^k\circ(\hat{A}_1^k-\hat{A}_2^k)]\right\}\\ &-\alpha\,\mathcal{R}\mathrm{e}\left\{\hat{n}_{21}^k\circ[\widetilde{D}_x(\hat{A}_1^k+\hat{A}_2^k-2\hat{V}_s^k)]\right\}+\alpha\,\mathcal{R}\mathrm{e}\left\{i\hat{n}_{21}^k\circ[\widetilde{D}_y(\hat{A}_1^k+\hat{A}_2^k-2\hat{V}_s^k)]\right\}\\ &=0\,,\end{aligned} \tag{6.4.18}$$

$$
\begin{aligned}
&\frac{\hat{n}_2^{k+1}-\hat{n}_2^k}{\Delta t}+\frac{2\alpha}{\varepsilon}(\hat{A}_1^k-\hat{A}_2^k)\circ\mathcal{I}\mathrm{m}(\hat{J}_x^{21,k}-i\hat{J}_y^{21,k})\\
&-\frac{1}{2}(D_x^+)^T[\hat{n}_2^k\circ D_x^+(\hat{A}_2^k-\hat{V}_s^k)]-\frac{1}{2}(D_x^-)^T[\hat{n}_2^k\circ D_x^-(\hat{A}_2^k-\hat{V}_s^k)]\\
&-\frac{1}{2}(D_y^+)^T[\hat{n}_2^k\circ D_y^+(\hat{A}_2^k-\hat{V}_s^k)]-\frac{1}{2}(D_y^-)^T[\hat{n}_2^k\circ D_y^-(\hat{A}_2^k-\hat{V}_s^k)]\\
&-\alpha\,\mathcal{R}\mathrm{e}\left\{\widetilde{D}_x^T[\hat{n}_{21}^k\circ(\hat{A}_1^k-\hat{A}_2^k)]\right\}+\alpha\,\mathcal{R}\mathrm{e}\left\{i\widetilde{D}_y^T[\hat{n}_{21}^k\circ(\hat{A}_1^k-\hat{A}_2^k)]\right\}\\
&+\alpha\,\mathcal{R}\mathrm{e}\left\{\hat{n}_{21}^k\circ[\widetilde{D}_x(\hat{A}_1^k+\hat{A}_2^k-2\hat{V}_s^k)]\right\}-\alpha\,\mathcal{R}\mathrm{e}\left\{i\hat{n}_{21}^k\circ[\widetilde{D}_y(\hat{A}_1^k+\hat{A}_2^k-2\hat{V}_s^k)]\right\}\\
&=0\,,
\end{aligned}
\tag{6.4.19}
$$

$$
-\gamma^2\Delta_{dir}\hat{V}_s^k=\hat{n}_1^k+\hat{n}_2^k\,,
\tag{6.4.20}
$$

$$
H(\hat{A}_1^k,\hat{A}_2^k)\begin{pmatrix}\hat{\psi}_l^{1,k}\\ \hat{\psi}_l^{2,k}\end{pmatrix}=\lambda_l^k\begin{pmatrix}\hat{\psi}_l^{1,k}\\ \hat{\psi}_l^{2,k}\end{pmatrix}\,,
\tag{6.4.21}
$$

$$
\hat{n}_1^k=\sum_l e^{-\lambda_l^k}\hat{\psi}_l^{1,k}\circ\overline{\hat{\psi}_l^{1,k}}\,,\qquad \hat{n}_2^k=\sum_l e^{-\lambda_l^k}\hat{\psi}_l^{2,k}\circ\overline{\hat{\psi}_l^{2,k}}\,,
\tag{6.4.22}
$$

$$
\hat{n}_{21}^k=\sum_l e^{-\lambda_l^k}\hat{\psi}_l^{2,k}\circ\overline{\hat{\psi}_l^{1,k}}\,,
\tag{6.4.23}
$$

$$
\hat{J}_{21}^{x,k}=-\frac{i\varepsilon}{2}\sum_l e^{-\lambda_l^k}\left[D_x(\hat{\psi}_l^{2,k})\circ\overline{\hat{\psi}_l^{1,k}}-\hat{\psi}_l^{2,k}\circ D_x(\overline{\hat{\psi}_l^{1,k}})\right]\,,
\tag{6.4.24}
$$

$$
\hat{J}_{21}^{y,k}=-\frac{i\varepsilon}{2}\sum_l e^{-\lambda_l^k}\left[D_y(\hat{\psi}_l^{2,k})\circ\overline{\hat{\psi}_l^{1,k}}-\hat{\psi}_l^{2,k}\circ D_y(\overline{\hat{\psi}_l^{1,k}})\right]\,.
\tag{6.4.25}
$$

Here, the Hamiltonian H is the same discrete Hamiltonian as in the first fully discrete system. Clearly, the scheme (6.4.18)-(6.4.25) is consistent with the continuous model (6.2.8)-(6.2.14). It is of first order in time and of second order in space. A drawback of the explicit nature of the forward Euler scheme (6.3.13)-(6.3.18) is that its full discretization is not unconditionally stable, as compared to the implicit scheme

presented in the previous subsection. Rather, the space-time grid must be chosen in such a way that a CFL condition is fulfilled.

The solution of this scheme requires the inversion of the non-local relation (6.4.21)-(6.4.22) at each time step. For this let us define the discrete version $\widehat{\mathcal{G}}_n : \mathbb{R}^{2P} \to \mathbb{R}$ of (6.3.19),

$$\widehat{\mathcal{G}}_n(\hat{A}_1, \hat{A}_2) := \widehat{\mathcal{G}}(\hat{A}_1, \hat{A}_2) + (\hat{n}_1^k, \hat{A}_1)_P + (\hat{n}_2^k, \hat{A}_2)_P \,. \tag{6.4.26}$$

Indeed, the first derivative of this functional is given by

$$\begin{aligned} \mathrm{d}\widehat{\mathcal{G}}_n(\hat{A}_1, \hat{A}_2)(\delta\hat{A}_1, \delta\hat{A}_2) = &\left(-\sum_l e^{-\lambda_l^k} \hat{\psi}_l^{1,k} \circ \overline{\hat{\psi}_l^{1,k}} + \hat{n}_1^k, \delta\hat{A}_1 \right)_P \\ &+ \left(-\sum_l e^{-\lambda_l^k} \hat{\psi}_l^{2,k} \circ \overline{\hat{\psi}_l^{2,k}} + \hat{n}_2^k, \delta\hat{A}_2 \right)_P . \end{aligned} \tag{6.4.27}$$

It should be noted that numerical tests conducted in Sec. 6.5 proved that the scheme presented in this subsection was better suited for a numerical solution of the spin QDD model than a Lax-Friedrichs scheme. This was attributed to the sparse space discretization of the domain Ω, which led to a considerable loss of mass in the Lax-Friedrichs scheme. A more refined grid would thus make a Lax-Friedrichs discretization of Eqs. (6.3.13)-(6.3.18) an interesting alternative to Eqs. (6.4.18)-(6.4.25); however this was beyond the capacity of the accessible computational resources.

6.4.3 Initialization of scheme 1

As was briefly mentioned in Sec. 6.2, a natural way to initialize the system (6.4.2)-(6.4.10) would be to start from given initial chemical potentials $\hat{A}_1^0$ and $\hat{A}_2^0$, compute the corresponding spin- and current densities and subsequently begin the iteration. However, from an experimental point of view it is more appealing to start from the initial spin densities $\hat{n}_1^0$ and $\hat{n}_2^0$. The problem in the latter approach is the lack of information about the initial spin-mixing quantities $\hat{n}_{21}^0$, $\hat{J}_{21}^{x,0}$ and $\hat{J}_{21}^{y,0}$, which are not directly related to the spin densities. At $t = t^0$ it is thus necessary to do a half step of scheme 2, which means to minimize the functional (6.4.26) in order to obtain the chemical potentials corresponding to the initial spin densities $\hat{n}_1^0$ and $\hat{n}_2^0$. One can then proceed according to scheme 1.

6.5 Numerical results

This section deals with the numerical study of the two fully discrete schemes which were introduced in the previous section. The developed algorithms were implemented in the FORTRAN 90 language. Eigenvalue problems were solved using the routine 'zheev.f90' from the LAPACK library. The solution of scheme 1, equations (6.4.2)-(6.4.10), was achieved by minimizing the discrete functional (6.4.11) at each time step t_k, $k > 0$. At t_0 the system was initialized as detailed in Subsec. 6.4.3. Each minimization problem was solved by a conjugate gradient method in the parameter space $\mathbb{R}^{3P}$ (or $\mathbb{R}^{2P}$ for scheme 2, respectively). We denote vectors in the parameter space by capital letters $X \in \mathbb{R}^{3P}$, $X = (\hat{A}_1, \hat{A}_1, \hat{V}_s)$, and by ∇_X we denote the gradient in the parameter space. In what follows the dot '$\cdot$' stands for the usual euclidean scalar product in $\mathbb{R}^{3P}$. In order to find the line minimum of $\nabla_X \widehat{\mathcal{F}} \cdot Y_n$, where Y_n denotes the search direction ($|Y_n| = 1$) during the n-th step of the conjugate gradient scheme, a Newton method was employed. The derivative of $\nabla_X \widehat{\mathcal{F}} \cdot Y_n$ in the direction Y_n was computed numerically with a forward discretization and the small step size $\varepsilon_{NT} = 10^{-3}$,

$$(\nabla_X \widehat{\mathcal{F}}(X) \cdot Y_n)' \approx \frac{\nabla_X \widehat{\mathcal{F}}(X + \varepsilon_{NT} Y_n) \cdot Y_n - \nabla_X \widehat{\mathcal{F}}(X) \cdot Y_n}{\varepsilon_{NT}} .$$

The same method was applied to the functional $\widehat{\mathcal{G}}_n$ in scheme 2. The Newton method was considered converged when $|\nabla_X \widehat{\mathcal{F}}(X) \cdot Y_n| < 10^{-10}$. We established two convergence criteria for the conjugate gradient method. On the one hand, we demanded that the total mass was conserved up to a factor 10^{-4}. On the other hand, using the notations $X = (x_i)_{i=1}^{3P}$ and $\nabla_X \widehat{\mathcal{F}} = (\partial_{x_i} \widehat{\mathcal{F}})_{i=1}^{3P}$, we demanded that

$$\max_i |\partial_{x_i} \widehat{\mathcal{F}}| < 10^{-3} .$$

Again, the same criteria were applied for the functional $\widehat{\mathcal{G}}_n$ in scheme 2. The time evolution was assumed to be converged if $|n_{1(2)}^{k+1} - n_{1(2)}^{k}|/\Delta t$ was less than 10^{-1} at each grid point.

Our aim is to test the developed numerical schemes in a typical transistor geometry, depicted in Fig. 6.1. We expect to obtain equilibrium charge- and spin-distributions for such a device. The source electrode of the transistor is located in the upper left corner of the domain, being held at a fixed potential-value $V_{ext,S} = 0$. The drain electrode is opposite to the source in the upper right corner with a fixed potential-value $V_{ext,D} = -2.0$. The gate electrode, held at the fixed potential $V_{ext,G} = -3.0$, is centered at $x = 0.5$ at the upper boundary of the domain. The tran-

sistor environment described above manifests itself in the two spin-up respectively spin-down external potentials $V_{ext,1}$ and $V_{ext,2}$. Prior to starting our simulations, these potentials were computed from the Laplace equation using a Gauss-Seidel scheme, where the fixed values $V_{ext,S}$, $V_{ext,D}$, $V_{ext,G}$ entered as Dirichlet boundary conditions (Neumann conditions were used at non-electrode portions of the domain boundary). On top of that we added a potential barrier of height 2.0 and thickness 0.1, which is centered at $x = 0.5$, and which exists only for spin-up electrons (index '1'). The barrier was thus added to $V_{ext,1}$ only. The potentials $V_{ext,1} + V_s$ and $V_{ext,2} + V_s$ at the starting time t_0 are depicted in Fig. 6.1.

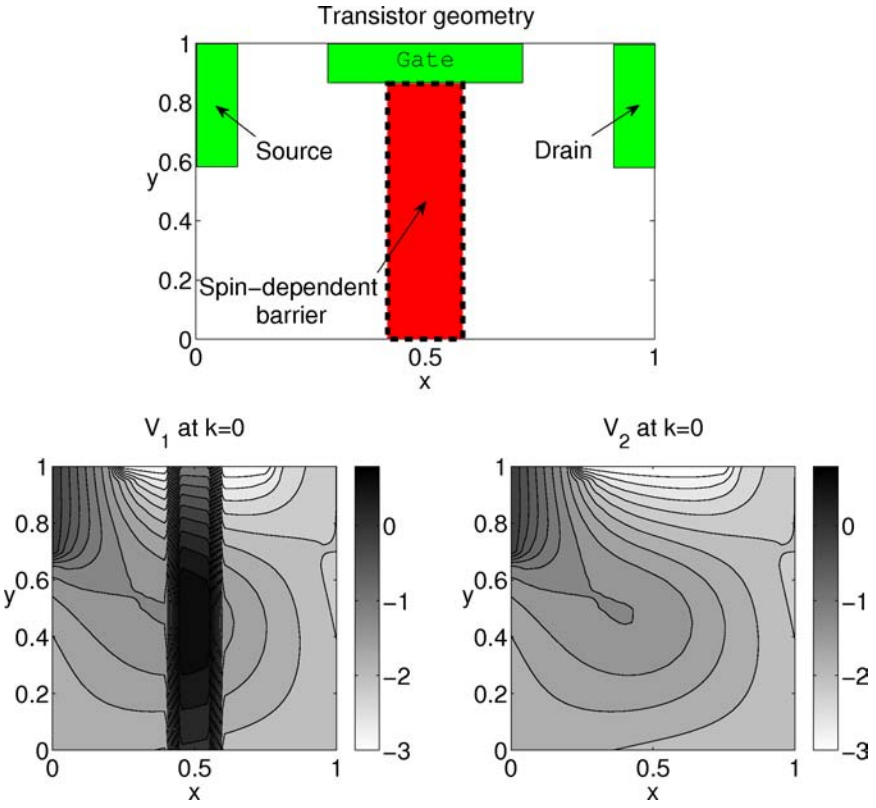


Figure 6.1: Schematic of the transistor geometry used in the simulations and the initial potentials $V_1 = V_{ext,1} + V_s$ and $V_2 = V_{ext,2} + V_s$ in that geometry at $t = t^0$.

Once the starting potentials $V_{ext,1}$ and $V_{ext,2}$ have been determined, we are interested in the evolution of given initial spin distributions n_1^0 and n_2^0 in the prescribed transistor environment. For the initial spin densities we choose two Gaussians centered at $(x, y) = (0.5, 0.5)$,

$$\begin{aligned} n_1^0(x,y) &= \frac{1}{0.12\pi}(1.0 + pol)\exp\left(-\frac{(x-0.5)^2}{0.06} - \frac{(y-0.5)^2}{0.06}\right), \\ n_2^0(x,y) &= \frac{1}{0.12\pi}(1.0 - pol)\exp\left(-\frac{(x-0.5)^2}{0.06} - \frac{(y-0.5)^2}{0.06}\right). \end{aligned} \tag{6.5.1}$$

Here, *pol* denotes the parameter of the initial spin polarization which was set to $pol = 0.5$. The initial data for n_1 and n_2 were discretized according to the conventions at the beginning of section 6.4. The initial total mass of the system was 1.0. The parameters of the space-discretization (for scheme 1 and for scheme 2) were chosen as

$$N = 21\,, \quad M = 21\,, \quad \Delta x = 0.05\,, \quad \Delta y = 0.05\,.$$

Employing the initial conditions (6.5.1), the numerical solution of scheme 1 and scheme 2, respectively, was carried out for values

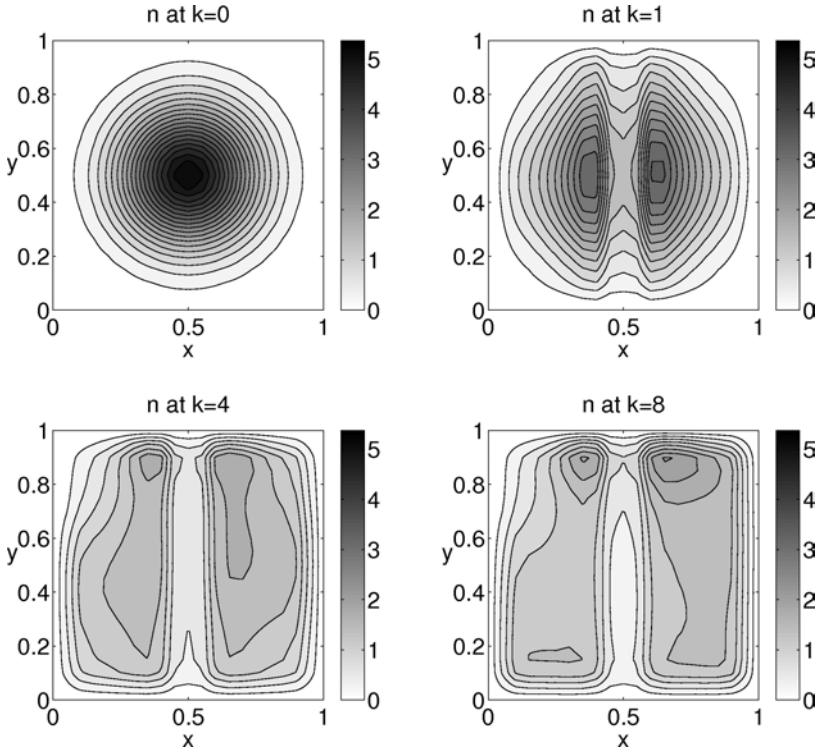
$$\alpha = 0.1\,, \qquad \varepsilon = 0.1\,, \tag{6.5.2}$$

of the scaled Rashba constant α and the semiclassical parameter ε, respectively. The respective time steps were

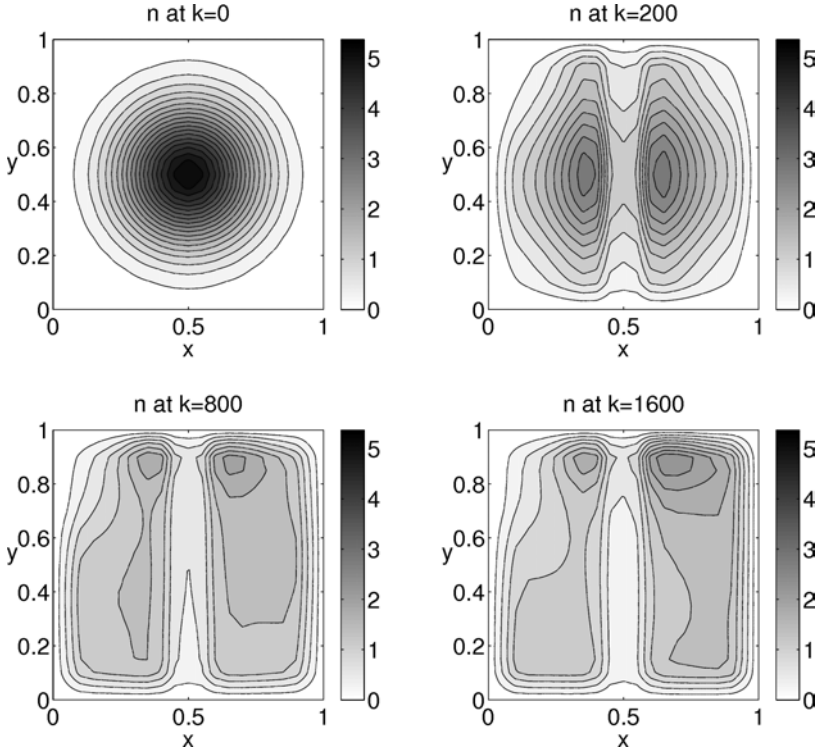
$$\text{scheme 1}: \quad \Delta t = 1.0 \times 10^{-2}\,, \quad \text{scheme 2}: \quad \Delta t = 0.5 \times 10^{-4}\,. \tag{6.5.3}$$

We note that in scheme 2 the CFL condition imposed a rather small increment on the time discretization.

The simulated time evolution of the spin density $n = n_1 + n_2$, the spin polarization $n_{pol} = n_1 - n_2$, the chemical potential A_1 and the chemical potential A_2 are depicted in Figs. 6.2-6.5 (all plotted data was interpolated to a grid of 128×128 points using the MATLAB routine "interp2.m"). In each of these Figures the results obtained from scheme 1 are compared with those obtained from scheme 2 during a time span of 8.0×10^{-2}. Let us briefly explain what is observed, starting with the evolution of the electron density n depicted in Fig. 6.2. At $k = 0$ one identifies the initial Gaussians, given by (6.5.1), which, as time evolves, are gradually split into two parts because of the potential barrier located at the center of the transistor, c.f. Figure 6.1. In the steady-state (at $k = 8$) the electron density has its maximum in the vicinity of the gate and the drain electrode, which is the region where the electron potential $V_{ext} + V_s$ has its lowest value. The region where the potential barrier is located shows a reduced electron density. This can be attributed to the positive barrier height "seen" by spin-up electrons, which impedes these electrons from entering (crossing) this region. Spin-down electrons are much less affected by the barrier, which becomes more transparent when regarding Figs. 6.4 and 6.5 for the respective chemical potentials. One clearly observes the barrier in the chemical potential A_1, while in A_2 it is completely absent. In both cases, however, gradients in A are gradually reduced when approaching the steady-state (at $k = 8$). Last but not least we turn to an interpretation of the obtained spin polarization n_{pol}, depicted in Fig. 6.3. A pattern similar to the one for the spin density evolves; however, n_{pol}

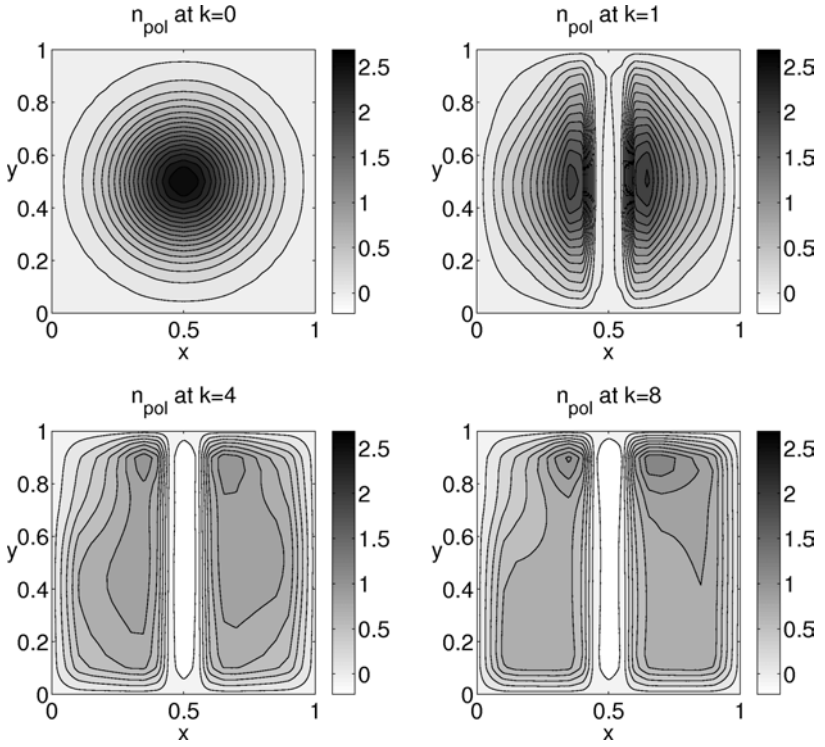


(i) Scheme 1: $\Delta t = 1.0 \times 10^{-2}$.

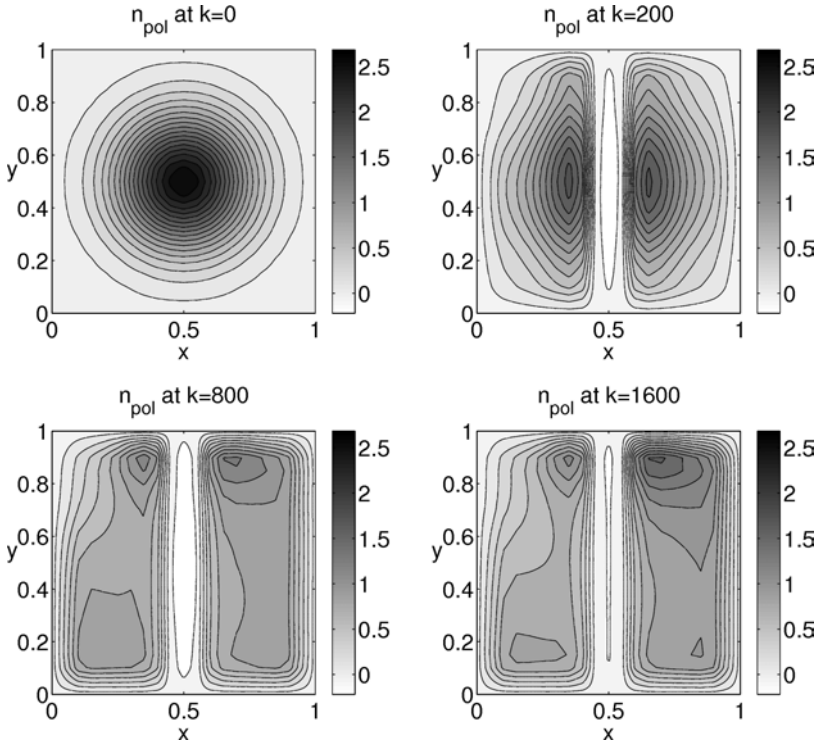


(ii) Scheme 2: $\Delta t = 0.5 \times 10^{-4}$.

Figure 6.2: Time evolution of the electron density $n = n_1 + n_2$ in the transistor geometry depicted in Fig. 6.1.

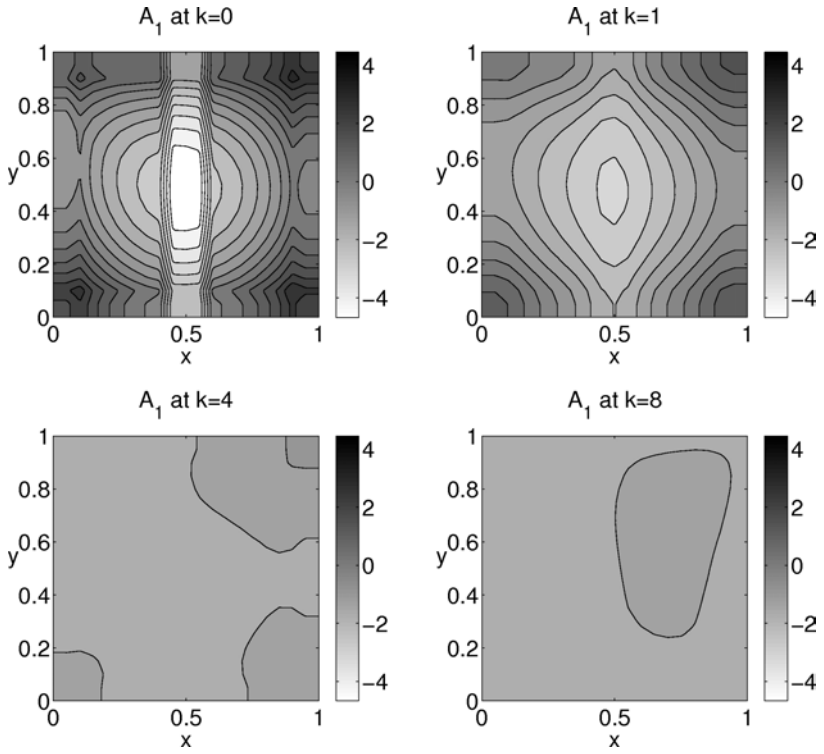


(i) Scheme 1: $\Delta t = 1.0 \times 10^{-2}$.

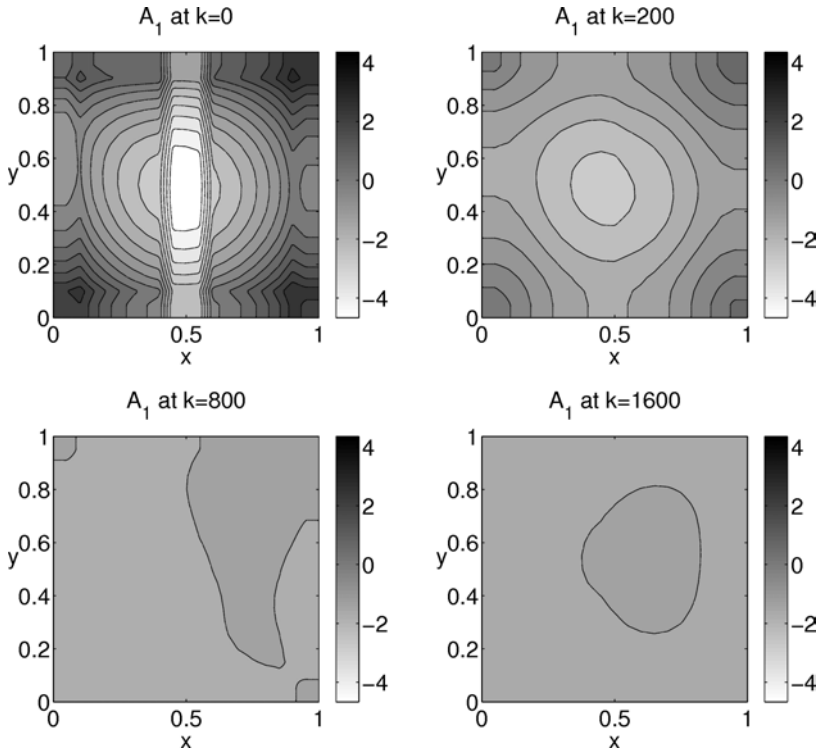


(ii) Scheme 2: $\Delta t = 0.5 \times 10^{-4}$.

Figure 6.3: Time evolution of the spin polarization $n_{pol} = n_1 - n_2$ in the transistor geometry depicted in Fig. 6.1.

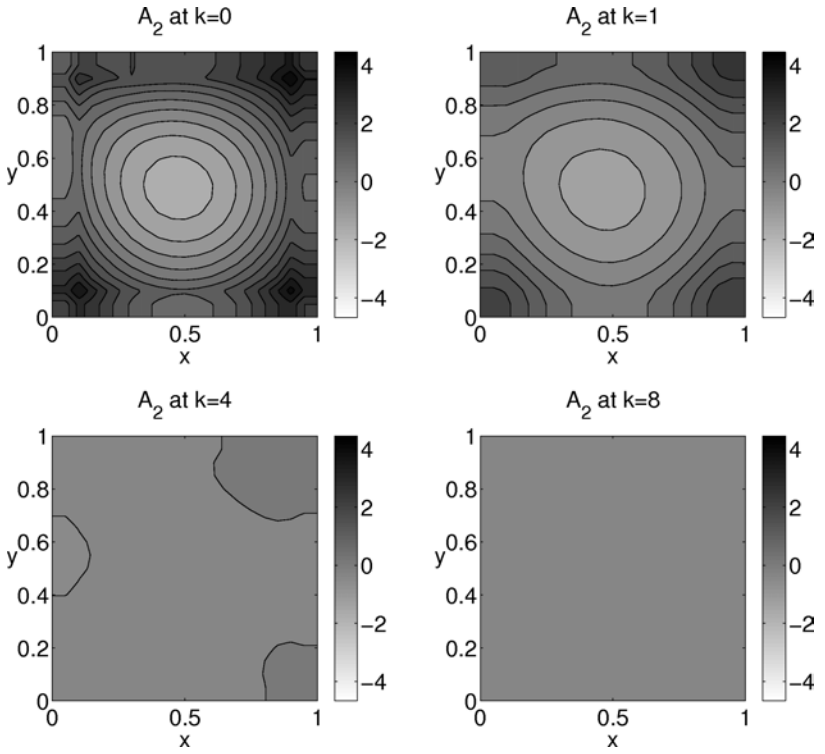


(i) Scheme 1: $\Delta t = 1.0 \times 10^{-2}$.

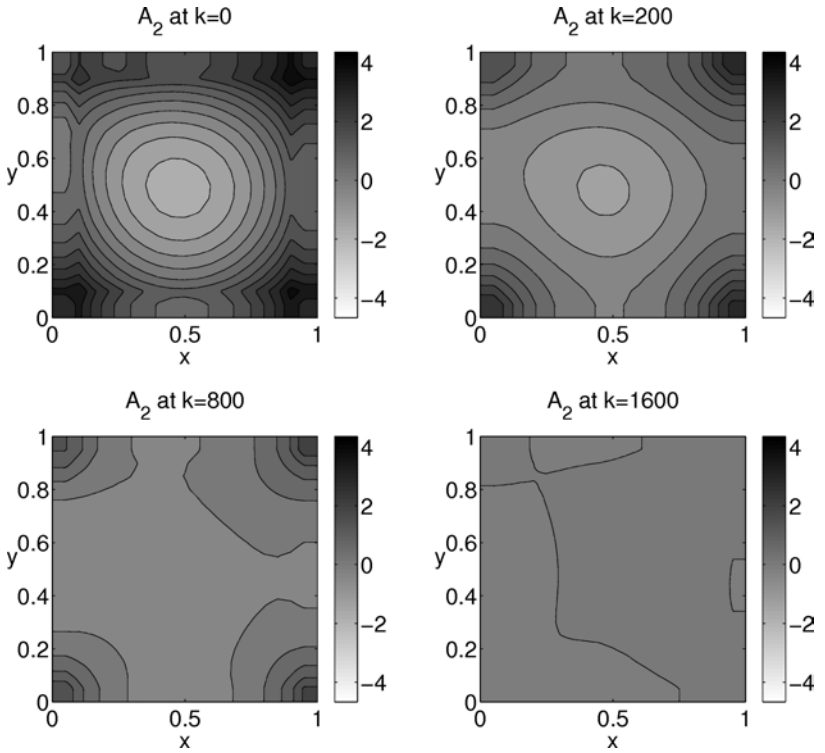


(ii) Scheme 2: $\Delta t = 0.5 \times 10^{-4}$.

Figure 6.4: Time evolution of the chemical potential A_1 in the transistor geometry depicted in Fig. 6.1.



(i) Scheme 1: $\Delta t = 1.0 \times 10^{-2}$.



(ii) Scheme 2: $\Delta t = 0.5 \times 10^{-4}$.

Figure 6.5: Time evolution of the chemical potential A_2 in the transistor geometry depicted in Fig. 6.1.

becomes negative in the region where the potential barrier is located. This is again due to the positive barrier height, which leads to $n_2 > n_1$ in the respective region.

6.6 Conclusion

In this work we proved the existence and uniqueness of a solution of two time discrete versions of the quantum drift-diffusion model (6.2.8)-(6.2.14) on the basis of a functional argument. Furthermore, finite difference approximations of space derivatives resulted in two fully discrete schemes which were later applied to simulate the time evolution of a Rashba electron gas confined to a bounded domain and under the influence of a prescribed external potential. The first scheme is implicit and advances in time the spin chemical potentials, whereas the second scheme is forward Euler and advances in time the spin-up and spin-down densities, respectively. The second scheme is subjected to a CFL stability condition, which results in the use of a considerably smaller time step as compared to the implicit scheme. Our results prove that the quantum drift-diffusion model considered can be applied for the numerical study of spin-polarized effects due to Rashba spin-orbit coupling and, thus, appears to benefit the design of novel spintronics applications.

6.A Perturbed eigenvalue problem

This section is devoted to the computation of the derivatives $\mathrm{d}\lambda_l(A)(\delta A)$ and $\mathrm{d}\psi_l(A)(\delta A)$ of the eigenvalues and eigenfunctions, respectively, of the Hamiltonian (6.2.4), when a small perturbation δA of the chemical potential A is applied. Let us define

$$\delta H = \begin{pmatrix} \delta A_1 & 0 \\ 0 & \delta A_2 \end{pmatrix}, \tag{6.A.1}$$

and start from

$$(H + \delta H)(\psi_l + \mathrm{d}\psi_l) = (\lambda_l + \mathrm{d}\lambda_l)(\psi_l + \mathrm{d}\psi_l) \tag{6.A.2}$$

where H denotes the Hamiltonian (6.2.4). Using $H\psi_l = \lambda_l\psi_l$ one obtains, up to first order in the variations,

$$H\mathrm{d}\psi_l + \delta H\psi_l = \lambda_l\mathrm{d}\psi_l + \mathrm{d}\lambda_l\psi_l\,. \tag{6.A.3}$$

Taking now the scalar product with ψ_k and using the orthonormalitiy of the eigenfunctions,

$$(\psi_k, \psi_l)_{L^2} = \int_\Omega (\psi_k^1\overline{\psi_l^1} + \psi_k^2\overline{\psi_l^2})\,dx = \delta_{kl}\,, \tag{6.A.4}$$

one obtains

$$(\psi_k, H\mathrm{d}\psi_l)_{L^2} + (\psi_k, \delta H\psi_l)_{L^2} = \lambda_l(\psi_k, \mathrm{d}\psi_l)_{L^2} + \mathrm{d}\lambda_l\delta_{kl}\,. \tag{6.A.5}$$

Since H is hermitian we have

$$(\psi_k, H\mathrm{d}\psi_l)_{L^2} = (H\psi_k, \mathrm{d}\psi_l)_{L^2} = \lambda_k(\psi_k, \mathrm{d}\psi_l)_{L^2}\,, \tag{6.A.6}$$

and (6.A.5) can be written as

$$(\psi_k, \delta H\psi_l)_{L^2} = (\lambda_l - \lambda_k)(\psi_k, \mathrm{d}\psi_l)_{L^2} + \mathrm{d}\lambda_l\delta_{kl}\,. \tag{6.A.7}$$

For $l = k$ we obtain

$$\mathrm{d}\lambda_l(A)(\delta A) = (\psi_l, \delta H\psi_l)_{L^2} = \int_\Omega \left(|\psi_l^1(A)|^2\delta A_1 + |\psi_l^2(A)|^2\delta A_2\right)\,dx\,, \tag{6.A.8}$$

and for $l \neq k$, assuming that the spectrum of H is non-degenerate, i.e. $\lambda_l \neq \lambda_k\,\forall l \neq k$, one obtains

$$(\psi_k, \mathrm{d}\psi_l)_{L^2} = \frac{(\psi_k, \delta H\psi_l)_{L^2}}{\lambda_l - \lambda_k}\,. \tag{6.A.9}$$

Since (6.A.9) is the projection of $\mathrm{d}\psi_l$ on the k-th basis vector of the eigenbasis of H we may write

$$\begin{aligned}\mathrm{d}\psi_l(A)(\delta A) &= \sum_{k\neq l}\frac{\psi_k}{\lambda_l - \lambda_k}(\psi_k, \delta H\psi_l)_{L^2} = \\ &= \sum_{k\neq l}\frac{\psi_k(A)}{\lambda_l(A) - \lambda_k(A)}\int_\Omega\left(\psi_k^1(A)\overline{\psi_l^1}(A)\delta A_1 + \psi_k^2(A)\overline{\psi_l^2}(A)\delta A_2\right)\,dx\,.\end{aligned} \tag{6.A.10}$$

6.B The maps $\mathcal{G}(A)$ and $\mathcal{G}_n(A)$

The map $\mathcal{G} : (H^1(\Omega, \mathbb{R}))^2 \to \mathbb{R}$ introduced in (6.3.7) is Gateaux-derivable and its first and second derivative, respectively, in the direction δA read

$$\begin{aligned}\mathrm{d}\mathcal{G}(A)(\delta A) &= -\sum_l e^{-\lambda_l(A)}\int_\Omega\left(|\psi_l^1(A)|^2\delta A_1 + |\psi_l^2(A)|^2\delta A_2\right)\,dx\,,\\ \mathrm{d}^2\mathcal{G}(A)(\delta A) &= -\sum_{l,k}\frac{e^{-\lambda_l} - e^{-\lambda_k}}{\lambda_l - \lambda_k}\left(\int_\Omega \psi_k^1\overline{\psi_l^1}\,\delta A_1\,dx + \int_\Omega \psi_k^2\overline{\psi_l^2}\,\delta A_2\,dx\right)^2\,.\end{aligned}$$

The maps $\mathcal{G}$ and $\mathcal{G}_n$ (which was introduced in (6.3.19)) are thus strictly convex. We obtain formally,

$$\begin{aligned}\mathcal{G}_n(A) &= \sum_l e^{-\lambda_l(A)} + \int_\Omega n_1^k A_1\, dx + \int_\Omega n_2^k A_2\, dx \\ &\geq e^{-\lambda_1(A)} + \int_\Omega n_1^k A_1\, dx + \int_\Omega n_2^k A_2\, dx \xrightarrow[||A_1||_{L^2}+||A_2||_{L^2}\to\infty]{} \infty\,,\end{aligned} \tag{6.B.1}$$

which means that $\mathcal{G}_n$ is even coercive. Here, $\lambda_1(A)$ stands for the smallest eigenvalue of the Hamiltonian $H(A)$,

$$\lambda_1(A) = \min_{\phi\in(H^1(\Omega))^2}(H(A)\phi,\phi)\,, \qquad ||\phi||_{(L^2(\Omega))^2} = 1\,. \tag{6.B.2}$$

In what follows we present a detailed computation of the second derivative of the functional $\mathcal{G}(A)$. We have

$$\begin{aligned}\mathrm{d}^2\mathcal{G}(A)(\delta A) =& -2\sum_l e^{-\lambda_l}\int_\Omega \mathcal{R}\mathrm{e}\left(\overline{\psi_l^1}\mathrm{d}\psi_l^1\delta A_1 + \overline{\psi_l^2}\mathrm{d}\psi_l^2\delta A_2\right)\, dx \\ &+ \sum_l e^{-\lambda_l}\mathrm{d}\lambda_l\int_\Omega \left(|\psi_l^1|^2\delta A_1 + |\psi_l^2|^2\delta A_2\right)\, dx\,.\end{aligned} \tag{6.B.3}$$

Let us define the following integrals,

$$I_1^{kl} := \int_\Omega \psi_k^1\overline{\psi_l^1}\,\delta A_1\, dx \qquad\qquad I_2^{kl} := \int_\Omega \psi_k^2\overline{\psi_l^2}\,\delta A_2\, dx\,. \tag{6.B.4}$$

Remark that from (6.A.8) one deduces

$$\mathrm{d}\lambda_l = I_1^{ll} + I_2^{ll}\,. \tag{6.B.5}$$

Thus, the second line in (6.B.3) can be written as

$$\sum_l e^{-\lambda_l}\left(I_1^{ll} + I_2^{ll}\right)^2\,. \tag{6.B.6}$$

Moreover, from (6.A.10) one obtains

$$\begin{aligned}\mathrm{d}\psi_l^1 &= \sum_{k\neq l}\frac{\psi_k^1}{\lambda_l-\lambda_k}\left(I_1^{kl} + I_2^{kl}\right)\,, \\ \mathrm{d}\psi_l^2 &= \sum_{k\neq l}\frac{\psi_k^2}{\lambda_l-\lambda_k}\left(I_1^{kl} + I_2^{kl}\right)\,,\end{aligned} \tag{6.B.7}$$

and therefore we have

$$\int_\Omega \left(\overline{\psi_l^1} \mathrm{d}\psi_l^1 \delta A_1 + \overline{\psi_l^2} \mathrm{d}\psi_l^2 \delta A_2 \right) dx = \sum_{k \neq l} \frac{1}{\lambda_l - \lambda_k} \left(I_1^{kl} + I_2^{kl} \right)^2 . \tag{6.B.8}$$

The right-hand-side of the first line in (6.B.3) can now be written as

$$-2 \sum_l \sum_{k \neq l} \frac{e^{-\lambda_l}}{\lambda_l - \lambda_k} \left(I_1^{kl} + I_2^{kl} \right)^2 = - \sum_{l,k,l \neq k} \frac{e^{-\lambda_l} - e^{-\lambda_k}}{\lambda_l - \lambda_k} \left(I_1^{kl} + I_2^{kl} \right)^2 . \tag{6.B.9}$$

Adding (6.B.6) and (6.B.9) together and making the convention

$$l = k : \qquad \frac{e^{-\lambda_l} - e^{-\lambda_k}}{\lambda_l - \lambda_k} = -e^{-\lambda_l} , \tag{6.B.10}$$

the second derivative of $\mathcal{G}(A)$ becomes

$$\mathrm{d}^2 \mathcal{G}(A)(\delta A) = - \sum_{l,k} \frac{e^{-\lambda_l} - e^{-\lambda_k}}{\lambda_l - \lambda_k} \left(I_1^{kl} + I_2^{kl} \right)^2 . \tag{6.B.11}$$

6.C Gateaux derivatives of $\mathcal{F}_1$-$\mathcal{F}_4$

The first and second Gateaux derivative, respectively, of the functionals (6.3.9)-(6.3.12) is given by

$$\begin{aligned} \mathrm{d}\mathcal{F}_1(A, V_s)(\delta A, \delta V_s) = &- \Delta t \int_\Omega \nabla \cdot (n_1^k \nabla (A_1 - V_s))(\delta A_1 - \delta V_s)\, dx \\ &- \Delta t \int_\Omega \nabla \cdot (n_2^k \nabla (A_2 - V_s))(\delta A_2 - \delta V_s)\, dx , \end{aligned}$$

$$\begin{aligned} \mathrm{d}^2\mathcal{F}_1(A, V_s)(\delta A, \delta V_s) = &- \Delta t \int_\Omega \nabla \cdot [n_1^k \nabla (\delta A_1 - \delta V_s)](\delta A_1 - \delta V_s)\, dx \\ &- \Delta t \int_\Omega \nabla \cdot [n_2^k \nabla (\delta A_2 - \delta V_s)](\delta A_2 - \delta V_s)\, dx \\ = &\Delta t \int_\Omega n_1^k |\nabla (\delta A_1 - \delta V_s)|^2\, dx + \Delta t \int_\Omega n_2^k |\nabla (\delta A_2 - \delta V_s)|^2\, dx , \end{aligned}$$

$$\begin{aligned} \mathrm{d}\mathcal{F}_2(A, V_s)(\delta A, \delta V_s) = &- \gamma^2 \int_\Omega \Delta V_s \delta V_s\, dx - \int_\Omega (n_1^k + n_2^k) \delta V_s\, dx \\ &+ \int_\Omega n_1^k \delta A_1\, dx + \int_\Omega n_2^k \delta A_2\, dx , \end{aligned}$$

$$\mathrm{d}^2\mathcal{F}_2(A,V_s)(\delta A,\delta V_s) = -\gamma^2\int_\Omega(\Delta\delta V_s)\delta V_s\,dx = \gamma^2\int_\Omega|\nabla\delta V_s|^2\,dx\,,$$

$$\begin{aligned}\mathrm{d}\mathcal{F}_3(A,V_s)(\delta A,\delta V_s) =& -\alpha\Delta t\,\mathcal{R}\mathrm{e}\left\{\int_\Omega\mathcal{D}[n_{21}^k(A_1-A_2)](\delta A_1+\delta A_2-2\delta V_s)\,dx\right\}\\ &+\alpha\Delta t\,\mathcal{R}\mathrm{e}\left\{\int_\Omega n_{21}^k\mathcal{D}(A_1+A_2-2V_s)(\delta A_1-\delta A_2)\,dx\right\},\end{aligned}$$

$$\mathrm{d}^2\mathcal{F}_3(A,V_s)(\delta A,\delta V_s) = 2\alpha\Delta t\,\mathcal{R}\mathrm{e}\left\{\int_\Omega n_{21}^k\mathcal{D}(\delta A_1+\delta A_2-2\delta V_s)(\delta A_1-\delta A_2)\,dx\right\},$$

$$\mathrm{d}\mathcal{F}_4(A)(\delta A) = \frac{2\alpha\Delta t}{\varepsilon}\,\mathcal{I}\mathrm{m}\left\{\int_\Omega(A_1-A_2)(\delta A_1-\delta A_2)(J_x^{21,k}-iJ_y^{21,k})\,dx\right\},$$

$$\mathrm{d}^2\mathcal{F}_4(A)(\delta A) = \frac{2\alpha\Delta t}{\varepsilon}\,\mathcal{I}\mathrm{m}\left\{\int_\Omega(\delta A_1-\delta A_2)^2(J_x^{21,k}-iJ_y^{21,k})\,dx\right\}.$$

To show that $\mathcal{F}$ is strictly convex, it is sufficient to show that

$$\mathrm{d}^2\mathcal{F}(A,V_s)(\delta A,\delta V_s)\geq 0\,,\quad\forall\delta A,\delta V_s\,.$$

One can see immediatly that the terms corresponding to $\mathcal{G}$, $\mathcal{F}_1$ and $\mathcal{F}_2$ are positive. Nevertheless, nothing can be said about the sign of the terms corresponding to $\mathcal{F}_3$ and $\mathcal{F}_4$. Assuming on the other hand that ε is a small parameter, which is a physical hypothesis, one can incorporate these latter terms in the former ones. Inspired by a formal proof in [1], in this work we assume, for some constant $c>0$,

$$n_{21}^k=\mathcal{O}(\varepsilon^2)\,,\quad\mathcal{I}\mathrm{m}(J_x^{21,k}-iJ_y^{21,k}) = 2c\varepsilon\alpha\frac{e^{-A_1^k}-e^{-A_2^k}}{A_2^k-A_1^k}+\mathcal{O}(\varepsilon^3)\,.$$

Remark then that the dominant term in $d^2\mathcal{F}_4$

$$4c\alpha^2\Delta t\left\{\int_\Omega(\delta A_1-\delta A_2)^2\frac{e^{-A_1^k}-e^{-A_2^k}}{A_2^k-A_1^k}\,dx\right\}.$$

is positive.

Concerning the coercivity, it is enough to show that

$$|\mathcal{F}(A,V_s)|\xrightarrow[||A||_{H^1}+||V_s||_{H^1}\to\infty]{}\infty\,.$$

In [7] this property has been shown for the first terms $\mathcal{X} := \mathcal{G} + \mathcal{F}_1 + \mathcal{F}_2$, by proving that if $|\mathcal{X}(A, V_s)| < c_1$ for some constant $c_1 > 0$, than there exists a constant $c_2 > 0$ such that $||A||_{H^1} + ||V_s||_{H^1} < c_2$. We can adapt this result in the present case, by assuming again that ε is a small parameter. Indeed, one can again incorporate the new terms $\mathcal{F}_3 + \mathcal{F}_4$ in $\mathcal{X}$, by proving the existence of some constant $C > 0$, such that

$$C|\mathcal{X}(A, V_s)| \leq |\mathcal{F}(A, V_s)|\,,$$

which finishes the coercivity proof. Thus, the functional $\mathcal{F}$, being strictly convexe and coercive, admits a unique minimum.

6.D Discretization matrices

Let $\mathbb{1}$ stands for the $(N-2)\times(N-2)$ identity matrix. Then we have the following discretization matrices:

$$D_x^+ = \frac{1}{\Delta x}\begin{pmatrix} -\mathbb{1} & \mathbb{1} & & & \\ 0 & -\mathbb{1} & \mathbb{1} & & \\ & \ddots & \ddots & \ddots & \\ & & 0 & -\mathbb{1} & \mathbb{1} \\ & & & 0 & 0 \end{pmatrix} \in \mathbb{R}^{P\times P}\,,$$

$$D_x^- = \frac{1}{\Delta x}\begin{pmatrix} 0 & 0 & & & \\ -\mathbb{1} & \mathbb{1} & 0 & & \\ & \ddots & \ddots & \ddots & \\ & & -\mathbb{1} & \mathbb{1} & 0 \\ & & & -\mathbb{1} & \mathbb{1} \end{pmatrix} \in \mathbb{R}^{P\times P}\,,$$

$$D_y^+ = \frac{1}{\Delta y}\begin{pmatrix} d_y^+ & & \\ & \ddots & \\ & & d_y^+ \end{pmatrix}, \qquad D_y^- = \frac{1}{\Delta y}\begin{pmatrix} d_y^- & & \\ & \ddots & \\ & & d_y^- \end{pmatrix},$$

$$D_y^+, D_y^- \in \mathbb{R}^{P\times P}\,,$$

$$d_y^+ = \begin{pmatrix} -1 & 1 & & & \\ 0 & -1 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 0 & -1 & 1 \\ & & & 0 & 0 \end{pmatrix} \in \mathbb{R}^{(N-2)\times(N-2)},$$

$$d_y^- = \begin{pmatrix} 0 & 0 & & & \\ -1 & 1 & 0 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 1 & 0 \\ & & & -1 & 1 \end{pmatrix} \in \mathbb{R}^{(N-2)\times(N-2)},$$

$$\widetilde{D}_x = \frac{D_x^+ + D_x^-}{2}, \qquad \widetilde{D}_y = \frac{D_y^+ + D_y^-}{2},$$

$$\Delta_{dir} = \Delta^x + \Delta^y \quad \in \mathbb{R}^{P\times P},$$

$$\Delta^x = \frac{1}{(\Delta x)^2} \begin{pmatrix} -2\mathbb{1} & \mathbb{1} & & & \\ \mathbb{1} & -2\mathbb{1} & \mathbb{1} & & \\ & \ddots & \ddots & \ddots & \\ & & \mathbb{1} & -2\mathbb{1} & \mathbb{1} \\ & & & \mathbb{1} & -2\mathbb{1} \end{pmatrix} \in \mathbb{R}^{P\times P}.$$

$$\Delta^y := \frac{1}{(\Delta y)^2} \begin{pmatrix} l_y & & \\ & \ddots & \\ & & l_y \end{pmatrix} \in \mathbb{R}^{P\times P}.$$

$$l_y = \begin{pmatrix} -2 & 1 & & & \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \\ & & & 1 & -2 \end{pmatrix} \in \mathbb{R}^{(N-2)\times(N-2)},$$

$$D_x = \frac{1}{2\Delta x} \begin{pmatrix} 0 & \mathbb{1} & & & \\ -\mathbb{1} & 0 & \mathbb{1} & & \\ & \ddots & \ddots & \ddots & \\ & & -\mathbb{1} & 0 & \mathbb{1} \\ & & & -\mathbb{1} & 0 \end{pmatrix} \in \mathbb{R}^{P\times P},$$

$$D_y := \frac{1}{2\Delta y} \begin{pmatrix} d_y & & \\ & \ddots & \\ & & d_y \end{pmatrix} \in \mathbb{R}^{P\times P}.$$

$$d_y = \begin{pmatrix} 0 & 1 & & & \\ -1 & 0 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 0 & 1 \\ & & & -1 & 0 \end{pmatrix} \in \mathbb{R}^{(N-2)\times(N-2)},$$

$$D_x^b = \frac{1}{\Delta x} \begin{pmatrix} \mathbb{1} & 0 & & & \\ -\mathbb{1} & \mathbb{1} & 0 & & \\ & \ddots & \ddots & \ddots & \\ & & -\mathbb{1} & \mathbb{1} & 0 \\ & & & -\mathbb{1} & \mathbb{1} \end{pmatrix} \in \mathbb{R}^{P\times P}.$$

$$D_y^b = \frac{1}{\Delta y} \begin{pmatrix} d_y^b & & \\ & \ddots & \\ & & d_y^b \end{pmatrix} \in \mathbb{R}^{P\times P}\,,$$

$$d_y^b = \begin{pmatrix} 1 & 0 & & & \\ -1 & 1 & 0 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 1 & 0 \\ & & & -1 & 1 \end{pmatrix} \in \mathbb{R}^{(N-2)\times(N-2)}\,,$$

Acknowledgments. This work has been supported by ...QUATRAIN, GREFIMEFI. S. Possanner acknowledges the support from the Austrian Science Fund, Vienna, under the contract number P21326-N16.

References

[1] L. Barletti and F. Méhats,. *Quantum drift-diffusion modeling of spin transport in nanostructures,* J. Math. Phys. **51**, 053304 (2010).

[2] Y. Bychkov, and E. I. Rashba, *Properties of a 2D electron gas with lifted spectral degeneracy,* JETP Letters **39**, 78 (1984).

[3] P. Degond and C. Ringhofer, *Quantum moment hydrodynamics and the entropy principle,* J. Stat. Phys. **112**(3-4), 587 (2003).

[4] P. Degond, F. Méhats and C. Ringhofer, *Quantum energy-transport and drift-diffusion models,* J Stat. Phys. **118**(3-4), 625-667 (2005).

[5] F. Méhats and O. Pinaud. *A*n inverse problem in quantum statistical physics. J. Stat. Phys. **140**, 565–602 (2010)

[6] F. Méhats and O. Pinaud. *A* problem of moment realizability in quantum statistical physics. Kinet. Relat. Models **4**(4), 1143–1158 (2011)

[7] S. Gallego and F. Méhats, *Entropic discretization of a quantum drift-diffusion model,* SIAM J. Numer. Anal. **43**(5), 1828-1849 (2006).

[8] R. El Hajj, *Diffusion models for spin transport derived from the spinor Boltzmann equation* Preprint.

[9] S. Possanner and C. Negulescu, *Diffusion limit of a generalized matrix Boltzmann equation for spin-polarized transport,* Kinetic and related Models **4**(4), 1159-1191 (2011).

[10] I. $\vec{Z}$utić, J. Fabian and S. Das Sarma, *Spintronics: fundamentals and applications,* Rev. Mod. Phys, **76**, 323 (2002).