universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„The Pauli equation: Modeling and Analytical Aspects "

verfasst von / submitted by

Jakob Möller, MSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2019 / Vienna 2019

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 066 821
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Mathematik
Betreut von / Supervisor:	Univ.-Prof. Dr. Norbert Mauser

Abstract

The Pauli equation was introduced by the Viennese physicist Wolfgang Pauli in 1927. It describes spin-1/2-particles like electrons at velocities which are much slower than the speed of light c, i.e. in the nonrelativistic or Post-Newtonian regime. In this thesis we consider the Pauli equation $i\hbar\partial_t u = H_P u$ where H_P is the Pauli operator, its modeling and analytic aspects. The Pauli equation can be obtained by either adding a spin term to the one-particle magnetic Schrödinger equation and thereby turning it into a two-component spinor equation or as the nonrelativistic limit of the Dirac equation which is a four-component spinor equation. Here we will focus on the first approach. For the Pauli operator some results on self-adjointness exist and we will present important aspects in this thesis. Concerning spectral properties we collect results on so called Lieb-Thirring estimates. Moreover we will briefly present further topics in the theory of the Pauli equation. First we will give a formulation of the Pauli equation in terms of density matrices. Last we will introduce the Pauli-Poiswell system which couples the first order in $1/c$ Pauli equation self-consistently to the first order approximation of the Maxwell system.

Zusammenfassung

Von Wolfgang Pauli im Jahr 1927 entdeckt beschreibt die Pauligleichung Teilchen mit halbzahligem Spin (wie beispielsweise das Elektron) die sich mit Geschwindigkeiten weit langsamer als die Lichtgeschwindigkeit c bewegen. In der vorliegenden Arbeit betrachten wir sowohl verschiedene Herleitungen als auch analytische Aspekte dieser Gleichung. Einerseits kann die Pauligleichung von der magnetischen Schrödingergleichung her konstruiert werden, indem man einen Spinterm $\sigma \cdot B$ addiert. Dadurch erhält man aus der Schrödingergleichung ein gekoppeltes System aus zwei magnetischen Schrödingergleichungen. Andererseits resultiert die Pauligleichung aus dem nichtrelativistischen Limes der Diracgleichung. In der Literatur existieren einige Arbeiten über die Selbstadjungiertheit und Spektrum des Paulioperators, von denen wir eine Auswahl vorstellen. Des Weiteren befassen wir uns mit weiterführenden Themen, darunter eine mögliche Dichtematrixformulierung der Pauligleichung. Zusätzlich wird das Pauli-Poiswell-System vorgestellt, dass die Pauligleichung in konsistenter Weise an die Nährung der Maxwellgleichungen in erster Ordnung in $1/c$ koppelt.

Acknowledgement

First of all I wish to thank my supervisors Norbert Mauser and Gerald Teschl for their excellent support and the time they have spent in supervising this thesis. Also I would like to thank the Wolfgang Pauli Institute for providing me a productive and pleasant work environment. Furthermore I wish to express my gratitude to the SFB project 'Taming Complexity in Partial Differential Equations' (FWF project number **F 6506-N36**) for financial support.

Amongst my colleagues I have to point out Peter Allmer, Arkadij Boijko and Timon Gutleb with whom I shared many lunches, laughs and discussions. They supported me in my studies where they could.

Finally I would like to thank my family, in particular my parents Christine Hidas and Martin Möller and my brother Joseph Möller who provided me a lot of support for this endeavor.

Contents

1 Introduction

Since its discovery Quantum Mechanics has been a greatly successful theory with many applications and a rich mathematical structure as its framework. At the heart of nonrelativistic Quantum Mechanics lies the Schrödinger equation, named after Erwin Schrödinger (1887-1961) who discovered the equation in 1925. Later, one was able to incorporate the theory of Special Relativity into Quantum Mechanics which led to Relativistic Quantum Mechanics and Quantum Field Theory. A central equation of QFT is the Dirac equation, named after Paul Dirac (1902-1984) who discovered this equation only three years later than the Schrödinger equation in 1928 and which describes spin-1/2-particles and predicts antiparticles. For example see his seminal book on Quantum Mechanics [4]. Today one views QFT as the fundamental theory describing three of the four fundamental forces: Electromagnetism, strong force and weak force. A unified description of all four fundamental forces including gravity has not been found to this day.

Albeit not being fundamental in the strict sense Quantum Mechanics in its nonrelativistic form still enjoys much interest since it is applicable in many regimes, most notably at low velocities (for higher velocities corrections are necessary as we shall see). Furthermore the mathematical formulation of Quantum Mechanics and particularly the Schrödinger equation as a second order PDE have been 'hot topics' within mathematical research over the last decades. In the hierarchy between the Schrödinger equation as a fully nonrelativistic and the Dirac equation as a fully relativistic equation the natural question is to ask whether we can make good approximations between the two equations. For example a feature that the Schrödinger equation lacks is the description of spin. The reason for this is that spin is coupled to the magnetic field which is a fundamentally relativistic notion. On the other hand the Dirac equation is rather complicated in structure and is not very suitable for applications in the nonrelativistic regime. It would hence be very useful to find an equation which describes spin but is not fully relativistic and simpler than the Dirac equation. The equation that does the job is the famous Pauli equation which was discovered by Wolfgang Pauli (1900-1958) in 1927. It is a system of two magnetic Schrödinger equations which are coupled via a spin term which involves the magnetic field. The two equations correspond to the two possible spin states of a spin-1/2-particle which are fermions.

In recent years one has proposed a further generalization of the Pauli equation in order to make its coupling to the electromagnetic field more accurate. As said the magnetic field is relativistic in nature resulting from the fact that the speed of propagation of the field is finite. Consider also Einsteins paper [5]. When coupling the Dirac equation to Maxwell's equations (governing the electromagnetic fields) we have two relativistic equations. If we want to switch to a first order in $1/c$ approximation we have to take the electromagnetic coupling into account and approximate it up to first order. This leads to the Pauli-Poiswell system where the name Poiswell stems from the Poisson system (at the level of the Schrödinger equation) and the Maxwell system (at the level of the Dirac equation).

Goal of this thesis is to give an overview on the mathematical aspects of the Pauli equation respectively the Pauli operator. We will introduce the magnetic Schrödinger operator first as it provides the origin of the Pauli equation

in the bottom up approach. Hence we will give some results on self-adjointness here as a warm up. For the full Pauli operator the main focus lies on collecting some crucial theorems on self-adjointness where very different mathematical tools have been applied. For example self-adjointness in [42] is proven via the quadratic form of the Pauli operator and showing its relative form boundedness. On the other hand the proof of self-adjointness in [8] is done by considering the Birman-Schwinger kernel of the Pauli operator and thereby showing boundedness from below, implying the existence of a self-adjoint extension. The other major focus lies on the estimates on the negative eigenvalues which are commonly known as Lieb-Thirring estimates. Here we also consider multiple cases with different conditions on the magnetic field. The theory on spectral theories of Pauli operators is rather well developed with some open questions left. Here we refer to the overview by Erdős in [6].

The structure of the thesis is as follows. In the beginning we will present the most basic facts about the physical background where the principles of Quantum Mechanics, Electromagnetism and the Dirac equation are motivated. Readers who are not familiar with physics will be directed to the suitable literature. The second part will motivate the Pauli equation where we give two approaches to the Pauli equation, one starting with the magnetic Schrödinger equation and adding a spin term, the other starting with the Dirac equation and performing some nonrelativistic limit. The third part is devoted to analytical results of the Pauli equation. We deal with questions of self-adjointness, spectrum and well-posedness as well as certain estimates on the eigenvalues, called Lieb-Thirring estimates. A final part will give an outlook on the already mentioned Pauli-Poiswell system and a possible density matrix formulation. Technical definitions and lemmata can be found in the appendix. A glossary of frequently used symbols can be found at the very end of the thesis.

2 Background from Physics

2.1 Quantum Mechanics

2.1.1 The Principles of Quantum Mechanics

Being a fundamentally different theory than Classical Mechanics, Quantum Mechanics is based on different 'axioms' which set the stage for the mathematical theory behind. Most important is the setting of functional analysis and there most notably the theory of unbounded (self-adjoint) operators. In Quantum Mechanics the framework is no longer $\mathbb{R} \times \mathbb{R}^3$ where physical quantities are described by real- or vector-valued functions but rather a possibly infinite dimensional (complex) Hilbert space $\mathcal{H}$ on which observables are represented by operators acting on vectors which correspond to physical states. The eigenvalues of these operators are considered to be the possible outcomes of a measurement. For example the outcomes of a measurement of spin in z-direction correspond to the eigenvalues $+\hbar/2$ and $-\hbar/2$ of the spin operator $\hat{S}_z$. Hence one demands observables to be self-adjoint operators which is reasonable to assume since the spectrum of self-adjoint operators is contained in the real line. Therefore the question that pops up frequently in Quantum Mechanics and its mathematical formulation is whether a given operator is self-adjoint. For a more in-depth introduction to the physical aspects of Quantum Mechanics see for example [39] or [38]. For mathematical aspects of Quantum Theory there are excellent books by Teschl [44] and Hall [14] which we will refer to frequently throughout the text.

Besides the axiom of (essentially) self-adjoint operators representing observables on a suitable Hilbert space the other principles of Quantum Mechanics are that *a*) physical states are represented by normed vectors $|\psi\rangle \in \mathcal{H}$ on which the observables act, *b*) the expectation value $\mathbb{E}_{|\psi\rangle}(A)$ of an observable $A \in \mathcal{H}$ if the system is in a state $|\psi\rangle$ is real and given by $\langle \psi, A\psi \rangle = \langle A\psi, \psi \rangle$ where $\langle \cdot, \cdot \rangle$ is the inner product on $\mathcal{H}$ and *c*) the time evolution of a state $|\psi(t)\rangle$ is given by a strongly continuous one-parameter unitary group $U(t)$ where the generator of the group corresponds to the energy of the system. In the following section we will discuss the origins of the equation governing the time evolution, the **Schrödinger equation**. Subsequently we will discuss the principles of spin in Quantum Mechanics and define a common notation in physic, the Dirac notation.

2.1.2 The Schrödinger Equation

The Schrödinger equation is the central equation of nonrelativistic Quantum Mechanics. Central in the sense that it is the equation of motion for any quantum system at velocities far slower than the speed of light. However, unlike in Classical Mechanics, the Schrödinger equation does not describe trajectories of particles in the sense that its solutions are real-valued functions of t giving the position of a particle at time t. Instead it deals with **wave functions** which are associated to quantum particles. The wave function ψ is a complex-valued function on $\mathbb{R} \times \mathbb{R}^3$ and it merely gives the **probability** of finding a quantum particle at a given time t at a given spatial point x by **Born's rule**:

$$\mathbb{P}(\text{particle is at}(x,t)) = |\psi(x,t)|^2 = \psi(x,t) \cdot \overline{\psi(x,t)}. \tag{2.1}$$

We can see that the expression $|\psi(x,t)|^2$ should interpreted as a probability density. Therefore we have to ask for some kind of normalization given by

$$\int_{\mathbb{R}^3} |\psi(x,t)|^2 \mathrm{d}x = 1. \tag{2.2}$$

Hence it is reasonable to assume the wave functions to be normalized vectors[1] in $L^2(\mathbb{R}^3)$. This is in fact the most basic example of an explicit realization of the Hilbert space where the physical system is built upon.

Let us now turn to the equation governing wave functions. In Classical Mechanics the energy of a particle is described by

$$E = \frac{p^2}{2m} + V. \tag{2.3}$$

Here p is the momentum, m is the mass and V is some potential. This expression is also the **Hamiltonian function** $H(p,x)$ from which we can deduce the equations of motion. As observable quantities are represented by operators on the Hilbert space $\mathcal{H}$ of the wave functions momentum and energy are found to be represented by the differential operators

$$\hat{p} := -i\hbar\nabla \tag{2.4}$$

$$\hat{E} := i\hbar\partial_t. \tag{2.5}$$

The replacement above is commonly known as **canonical quantization** and for details we refer to standard books, e.g. [38]. From this however it is not hard to get the Schrödinger equation. The idea is to use the energy from equation (2.3), plug in the operators from (2.5) and let them act on a wave function ψ:

$$i\partial_t\psi = -i\frac{\hbar}{2m}\Delta\psi + V\psi. \tag{2.6}$$

It should be noted that we made use of the identity $\nabla\cdot\nabla = \Delta$. Here V is the multiplication operator M_V on $\mathcal{H}$. This equation is the famous **Schrödinger equation**.

2.1.3 Spin

Spin is one of the phenomena in Quantum Mechanics that has no analogue in Classical Mechanics. It is a fundamental property of particles: There are particles with integer spin (Higgs boson, photon, etc.) which are called **bosons** and particles with half-integer spin like the electron and which are called **fermions**. It is to be thought of as some kind of *intrinsic* angular momentum. However thinking in classical terms does not help understanding spin physically so this expression can only be some help in the sense that spin is some kind of angular momentum *in a mathematical sense* as it obeys the angular momentum algebra. Historically spin was discovered by Otto Stern and Walther Gerlach in their famous **Stern-Gerlach experiment**. They set up a strongly inhomogenous magnetic field and sent a collimated beam of neutral silver atoms through.

[1] More correctly, they are considered as rays in a projective Hilbert space due to the fact that wavefunctions differing by a constant phase $e^{i\theta}$ are physically equivalent.

Now neutral particles can only couple to the magnetic field if they have some magnetic moment μ. Classically the force acting on such a particle is given by

$$\vec{F} = \nabla(\mu \cdot \vec{B}). \tag{2.7}$$

where $\vec{B}$ is the magnetic field. If the z-direction is the dominant direction we can approximate this by

$$\vec{F} = \mu_z \frac{\partial B_z}{\partial z} \vec{e}_z \tag{2.8}$$

where the subscript z denotes the z-component of the vector and $\vec{e}_z$ is the unit vector in z-direction. Classically we would expect from a sample of neutral silver atoms that the value of μ_z is randomly distributed and hence we would observe a line in z-direction on a screen. However the outcome observed were two distinct dots. This led to the discovery that the magnetic moment is in fact *not* continuously distributed but *quantized* and the physical observable which is responsible for that behavior is **spin**.

The mathematical description of spin is based on the Lie algebra $\mathfrak{su}(2)$ and its representations. The operators corresponding to the spin in each spatial directions are

$$\hat{S}_i := \frac{\hbar}{2}\sigma_i \tag{2.9}$$

where σ_i is the ith **Pauli matrix** and i ranges from 1 to 3. Explicitly they are given as

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.10}$$

Later we will use an elegant shorthand notation defined as follows:

$$\sigma := \begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3. \end{pmatrix} \tag{2.11}$$

This expression is to be interpreted as a *vector of matrices*.

2.1.4 Bra-Ket notation

In physics one often uses a different notation than in mathematics which is known as **Braket notation** or **Bra-Ket notation** or **Dirac notation**. If $\mathcal{H}$ is the Hilbert space of the quantum system then a vector, i.e. an element of $\mathcal{H}$ is called a **ket** and denoted by $|u\rangle$. On $\mathcal{H}$ we can define continuous (i.e. bounded) linear functionals which form another Hilbert space, the dual $\mathcal{H}^*$. The elements of $\mathcal{H}^*$ are called **bras** and denoted by $\langle v|$. By the Riesz representation theorem we can find a unique $u_v \in \mathcal{H}$ for any functional $v \in \mathcal{H}^*$ such that $v(w) = \langle w, u_v \rangle$ for all $w \in \mathcal{H}$. This justifies the notation

$$v(u) = \langle v, u \rangle =: \langle v|u \rangle \tag{2.12}$$

if we denote the functional v and the element u_v by the same letter. There is a little subtlety in this notation since the inner product is not linear in both slots but sesquilinear. It depends on the convention which of both is chosen to be the linear part. In physics one often chooses linearity in the second rather than

the first slot in order to make sense out of the notation above. Furthermore if A is an operator on $\mathcal{H}$ and $u \in \mathcal{H}$ then we have the functional $\langle v|$ which sends Au to $\langle v|Au\rangle$. In Dirac notation we would have

$$\langle v|Au\rangle = \langle v|\, A\, |u\rangle\,. \tag{2.13}$$

It is then equivalent to think of $\langle v|$ acting on $A\,|u\rangle$ or to think as $\langle v|\, A$ acting on $|u\rangle$. Note that this interpretation is only well-defined if A is self-adjoint. We can also form the expression $|u\rangle\langle v|$ which sends an element $|w\rangle$ to $|u\rangle\,\langle v|w\rangle$. i.e. a map $u\langle v, \cdot\rangle$ from $\mathcal{H}$ to $\mathcal{H}$ such that

$$w \mapsto \langle v, w\rangle u. \tag{2.14}$$

2.2 Electromagnetism

Classical electromagnetism is a classical field theory, the fields being the electric field E and the magnetic field B. They are governed by **Maxwell's equations** which we will present in the following section. Discovered by Maxwell in 1864 the four equations governing the (time dependent) electric and the magnetic field read

$$\nabla \cdot E = \frac{1}{\epsilon_0}\rho \qquad\qquad \nabla \times E + \frac{\partial B}{\partial t} = 0 \tag{2.15}$$

$$\nabla \cdot B = 0 \qquad\qquad \nabla \times B - \mu_0\epsilon_0\frac{\partial E}{\partial t} = \mu_0 J \tag{2.16}$$

where ρ is the charge density, J is the current density, ϵ_0 is the electric field constant (or vacuum permittivity) and μ_0 is the magnetic field constant (or vacuum permeatbility). The speed of light c is related to ϵ_0 and μ_0 by the formula

$$c = \frac{1}{\sqrt{\mu_0\epsilon_0}}. \tag{2.17}$$

Rearranging and using properties of the differential operators rotation, gradient and divergence leads to wave equations for the electromagnetic fields:

$$-\frac{1}{c^2}\frac{\partial^2 E}{\partial t^2} + \nabla^2 E = 0 \tag{2.18}$$

and

$$-\frac{1}{c^2}\frac{\partial^2 B}{\partial t^2} + \nabla^2 B = 0. \tag{2.19}$$

It is possible to move to a potential formalism in the sense that one introduces **potential fields** and defines the electric and the magnetic field as some derivative of them. From the divergence freedom of B we obtain the existence of the **vector potential** A such that B is the curl of A:

$$\nabla \times A = B. \tag{2.20}$$

Rewriting the second equation in (2.15) with the help of (2.20) we obtain

$$\nabla \times \left(E + \frac{\partial A}{\partial t}\right) = 0. \tag{2.21}$$

This means that $E + \partial A/\partial t$ is the gradient of a scalar potential field ϕ and we have

$$E = \nabla\phi - \frac{\partial A}{\partial t}. \tag{2.22}$$

We will make use of those potentials frequently and in section 5.2 we will formulate Maxwell's equations as wave equations for the potentials (choosing Lorentz gauge). In fact the magnetic Schrödinger equation and the Pauli equation are based on the **magnetic Schrödinger operator** which is given by

$$(-i\hbar\nabla - \frac{q}{c}A)^2. \tag{2.23}$$

2.3 Dirac and Klein-Gordon

For this section we heuristically present the Dirac equation. For an extensive mathematical treatment of the Dirac equation see [45]. For a physical introduction and derivation see [12]. In Special Relativity (SR) the relation between energy and momentum is different than in Classical Physics. It is given by

$$E^2 = p^2c^2 + m^2c^4. \tag{2.24}$$

Expanding the square root of the RHS in powers of c^2 gives the famous expression for the rest mass as the leading term:

$$E = mc^2 + T + \cdots \tag{2.25}$$

where T is the relativistic kinetic energy. We will take (2.24) as starting point for the Dirac equation similarly to the Schrödinger equation above.

In SR space and time are considered to be part of one mathematical object called **spacetime** which is a four dimensional Minkowski space $\mathbb{M}^4$, i.e. a flat Lorentzian manifold of dimension four. Hence the central objects are **four-vectors** whose components will be denoted by a greek superscript ranging from 0 to 3:

$$x^\mu,\ \mu = 0, 1, 2, 3. \tag{2.26}$$

The **four-momentum** p^μ is given by

$$p = (p^0, \mathbf{p}) \tag{2.27}$$

where $p^0 = E$ is the relativistic energy and $\mathbf{p}$ denotes the usual three-momentum. When transitioning to Quantum Mechanics we replace p^μ by the derivative $i\hbar\eta_{\mu\nu}\partial^\nu$. Similarly to how we obtained the Schrödinger equation we will use the energy-momentum relation (2.24) and the operator for p^μ. We then get

$$p^\mu p_\mu - m^2c^2 = 0 \tag{2.28}$$

where the Einstein summation convention is understood, i.e. $p^\mu p_\mu = \sum_{\mu=0}^{3} p^\mu p_\mu$ and by replacing p^μ we obtain

$$(-\hbar^2\partial^\mu\partial_\mu - m^2c^2)\psi = 0. \tag{2.29}$$

This equation is called **Klein-Gordon equation**. It his however not the right equation for spin-1/2-particles for the simple reason that it fails to properly

describe experiments, e.g. it does not predict the energy levels of the hydrogen atom correctly. This is because electrons are spin-$\frac{1}{2}$-particles and the Klein-Gordon equation in fact only describes particles with spin 0 or otherwise put does not include spin at all. Therefore we have to stretch out a little bit in order to find the right equation. The trick is to factorize the energy-momentum relation (2.28):

$$p_\mu p^\mu - m^2 = (\beta^\lambda p_\lambda + m)(\gamma^\sigma p_\sigma - m) = 0. \tag{2.30}$$

Multiplying out the RHS gives

$$p_\mu p^\mu - m^2 = \beta^\lambda \gamma^\sigma p_\lambda p_\sigma - m(\beta^\lambda - \gamma^\lambda) p_\lambda - m^2. \tag{2.31}$$

Comparison of the coefficients yields the following conditions:

$$\beta^\lambda = \gamma^\lambda \tag{2.32}$$

$$p_\mu p^\mu = \gamma^\lambda \gamma^\sigma p_\lambda p_\sigma. \tag{2.33}$$

These requirements can however not be fulfilled by numbers. It was Dirac's idea to postulate the objects γ^μ to be 4×4 matrices. In fact the conditions define a Clifford algebra for which the minimal dimension is four. There are higher dimensions for which the Dirac equation would be realizable. Then we can express the conditions above in the anticommutation relation

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu} \mathbb{1}. \tag{2.34}$$

In standard notation the **Dirac matrices** are given by

$$\gamma^0 = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \qquad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}. \tag{2.35}$$

where σ^i, $i = 1, 2, 3$ denote the *Pauli matrices*. From (2.30) we deduce the equation

$$\gamma^\mu p_\mu - mc = 0 \tag{2.36}$$

and by substituting p_μ by $i\hbar\partial_\mu$ we obtain

$$(i\hbar\gamma^\mu \partial_\mu - mc)\Psi = 0. \tag{2.37}$$

This equation is called **Dirac equation**. Here, Ψ does not denote a complex-valued wave function as above but an object called **Dirac spinor**. It is a four dimensional vector consisting of four individual wave functions:

$$\Psi := \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4. \end{pmatrix} \tag{2.38}$$

In fact it can be considered as an element of $\mathbb{C}^4 \otimes L^2(\mathbb{R}^3)$ or simply $(L^2(\mathbb{R}^3)^4$.

Let us explain how the Dirac equation is to be interpreted physically. The first two components of the Dirac spinor Ψ can easily seen to represent the wave functions for an electron with spin up ($+\frac{1}{2}$) and for an electron with spin down ($-\frac{1}{2}$). The latter two components however pose a problem as it is not *a priori*

clear what kind of physical situation they should represent since the energy of those states is *negative*. Dirac's original idea was to postulate an infinite 'sea' of electrons which fill up all the states of negative energy represented by the latter two components. Due to various metaphysical reasons this interpretation is far from satisfactory. Feynman and Stückelberg (their original papers on the subject are [10] and [43]) proposed a fundamentally different interpretation which is accepted today saying that the two lower components represent a different kind of particle, the **antiparticle** of the electron, called **positron**.

3 The Pauli equation: Derivation

The Schrödinger equation is not a relativistic equation as we have seen above and it fails to include spin. Moreover, magnetic fields have not been incorporated to the Schrödinger equation either. In other words the regime within which the non-relativistic Schrödinger equation works is that of velocities far below c whereas the Dirac equation is a 'fully' relativistic equation and includes spin. This points at a certain hierarchy between the equations and the Pauli equation is somewhat *in the middle* between Dirac and Schrödinger. In fact it is a $O(1/c)$-correction to the Schrödinger equation. Higher order corrections, e.g $O(1/c^2)$ include the *Darwin term* and the *Zitterbewegung*. For details consider the paper by Mauser [32]. More precisely we search for an equation of Schrödinger type which includes the description of spin *but* works in the nonrelativistic regime. There are many ways to model the Pauli equation and we will present two rather heuristic approaches before delving deeper into the mathematical questions.

3.1 Bottom-up: From Schrödinger to Pauli

If one wants to incorporate the interaction of quantum particles with an electromagnetic field it seems reasonable to take the Schrödinger equation for a free particle and add some suitable interaction term with the electromagnetic field in question. The resulting equation is called **magnetic Schrödinger equation**. Let us state again the Schrödinger equation with general Hamiltonian H:

$$i\hbar\partial_t u = Hu. \tag{3.1}$$

The classical Hamiltonian for a point particle with charge q is given by

$$H(x,p,t) = \frac{1}{2m}\left(p - \frac{q}{c}A(x,t)\right)^2 + q\phi(x,t). \tag{3.2}$$

As before, switching to quantum theory is done by replacing $p \to -i\hbar\nabla$, $A \to \hat{A}$ and $\phi \to \hat{\phi}$. We make the assumption that $\hat{A}$ and $\hat{\phi}$ are just multiplication operators. This gives

$$H = \frac{1}{2m}\left(-i\hbar\nabla - \frac{q}{c}A(x,t)\right)^2 + q\phi(x,t) \tag{3.3}$$

and hence the **magnetic Schrödinger equation** reads

$$i\hbar\partial_t u = \frac{1}{2m}\left(-i\hbar\nabla - \frac{q}{c}A(x,t)\right)^2 u + q\phi(x,t)u. \tag{3.4}$$

It is important to emphasize at this point that this is a *scalar* equation in the sense that its solutions are scalar wave functions. The addition of spin will naturally lead to a redefinition of the wave function u which will be done below. Of course if we turn off any electric or magnetic fields (i.e. setting $A = 0$ and $\phi = 0$) we obtain the ordinary Schrödinger equation for a free particle. In the literature one often refers to the term

$$\left(-i\hbar\nabla - \frac{q}{c}A(x,t)\right)^2 \tag{3.5}$$

as the **magnetic Schrödinger operator**. Now as pointed out before spin is not incorporated in this description. Extending the Schrödinger equation to a

system sensible to spin comes down to adding some sort of coupling term of the magnetic field B with the spin operators σ_i. As spin-$\frac{1}{2}$ yields two states (up and down) we expect to obtain to equations coupled by σ_i. In fact we will add the famous **Stern-Gerlach term** which is defined as

$$-\frac{q\hbar}{2mc}\sigma \cdot B. \tag{3.6}$$

where σ is the vector of Pauli matrices from above. The term $\sigma \cdot B$ explicitly reads:

$$\sigma \cdot B = B_1\sigma_1 + B_2\sigma_2 + B_3\sigma_3. \tag{3.7}$$

Adding (3.6) to (3.4) is only possibly if u is a 2-spinor

$$u(x,t) = \begin{pmatrix} u_1(x,t) \\ u_2(x,t) \end{pmatrix} \tag{3.8}$$

we obtain

$$i\hbar\partial_t u = \frac{1}{2m}\left(-i\hbar\nabla - \frac{q}{c}A(x,t)\right)^2 u + q\phi(x,t)u - \left(\frac{q\hbar}{2mc}\sigma \cdot B\right)u. \tag{3.9}$$

which is the **Pauli equation**. Again, we make clear that the equation is no longer a scalar equation but a vector equation, i.e. a coupled system. As we can see this equation is closely related to the magnetic Schrödinger equation for two reasons: If there was no magnetic field involved we would have two decoupled magnetic Schrödinger equations. In the case where $B \neq 0$ we have additional terms involving the components of B, i.e. B_i which can be viewed as a perturbation of the system. Furthermore the two equations for u_1 and u_2 are coupled if $B_1, B_2 \neq 0$. In the case $B_1 = B_2 = 0$ and $B_3 \neq 0$ the system is however decoupled since σ_3 is diagonal. If we write equation (3.9) explicitly we have

$$i\hbar\partial_t u_1 = \frac{1}{2m}\left(-i\hbar\nabla - \frac{q}{c}A\right)^2 u_1 + q\phi u_1 - \frac{q\hbar}{2mc}\left(B_3u_1 + (B_1 - iB_2)u_2\right) \tag{3.10}$$

$$i\hbar\partial_t u_2 = \frac{1}{2m}\left(-i\hbar\nabla - \frac{q}{c}A\right)^2 u_2 + q\phi u_2 - \frac{q\hbar}{2mc}\left(-B_3u_2 + (B_1 + iB_2)u_1\right). \tag{3.11}$$

3.2 Top-down: From Dirac to Pauli

Alternatively one obtains the Pauli equation as the first order in $1/c$ approximation to the Dirac equation which we already discussed. The rigorous nonrelativistic limits of the Dirac equation has been extensively studied in the literature. A first rather heuristic approach was done by Foldy and Wouthuysen in [11]. Subsequent works include the papers by Hunziker [16], White [47], Mauser [31] and [32], Bechouche et al. [1], Masmoudi and Mauser [30] and Mauser and Selberg [33]. The idea is to regard the inverse of the speed of light c^{-1} as a parameter which one can send to zero and do asymptotic analysis of the Dirac equation with, hence the name *nonrelativistic*. However we are not interested in the rigorous justification of the limit wherefore we will follow the heuristic explanation in [17] and [20]. The method employed in order to couple the Dirac equation to an electromagnetic field is commonly called **minimal coupling** in

the physics literature. It boils down to replacing the momentum operator $i\partial_\mu$ by $i\partial_\mu - qA_\mu$. The advantage of this coupling is that it leaves the modified Dirac equation gauge invariant under the common gauge transformations of Ψ and A_μ. The electromagnetic Dirac equation reads

$$(i\gamma^\mu \partial_\mu - q\gamma^\mu A_\mu - m)\Psi = 0. \tag{3.12}$$

For the moment we will rewrite equation (3.12) in a more explicit way:

$$i\partial_t \Psi = (\alpha \cdot (-i\nabla - qA) + \beta m + q\phi) \tag{3.13}$$

where $\beta = \gamma^0$ and

$$\alpha_k = \begin{pmatrix} 0 & \sigma_k \\ \sigma_k & 0 \end{pmatrix}. \tag{3.14}$$

Now we split the four spinor Ψ into two two spinors φ and χ such that

$$\Psi = \begin{pmatrix} \varphi \\ \chi \end{pmatrix}. \tag{3.15}$$

Together with this definition we obtain two equations:

$$i\partial_t \varphi = \sigma \cdot (-i\nabla - qA)\,\chi + q\phi\varphi + m\varphi \tag{3.16}$$
$$i\partial_t \chi = \sigma \cdot (-i\nabla - qA)\,\varphi + q\phi\chi - m\chi. \tag{3.17}$$

Introducing the new spinors $\tilde{\varphi}\exp(-imt) := \varphi$ and $\tilde{\chi}\exp(-imt) := \chi$ we arrive at

$$i\partial_t \tilde{\varphi} = \sigma \cdot (-i\nabla - qA)\,\tilde{\chi} + q\phi\tilde{\varphi} \tag{3.18}$$
$$i\partial_t \tilde{\chi} = \sigma \cdot (-i\nabla - qA)\,\tilde{\varphi} + q\phi\tilde{\chi} - 2m\tilde{\chi}. \tag{3.19}$$

Now the terms $q\phi\tilde{\chi}$ and $i\partial_t\tilde{\chi}$ are of second order in the parameter $\epsilon = c^{-1}$ and we can neglect them in a first order approximation. Whence

$$\tilde{\chi} \approx \frac{1}{2m}\sigma \cdot (-i\nabla - qA)\,\tilde{\varphi} \ll \tilde{\varphi} \tag{3.20}$$

That is in the nonrelativistic limit the lower antiparticle spinor components become very small with respect to the upper particle components. We see that the particle-antiparticle pair is a relativistic effect. If we use this approximation in the equation for the upper component we obtain

$$i\partial_t \tilde{\varphi} = \left(\frac{(\sigma \cdot (-i\nabla - qA))^2}{2m} + q\phi\right)\tilde{\varphi}. \tag{3.21}$$

This can be rewritten in an elegant way using the identity

$$(\sigma \cdot (-i\nabla - qA))^2 = (-i\nabla - qA)^2 - q\sigma \cdot B \tag{3.22}$$

which can be verified directly and uses the so called *Pauli vector identity*. We then have

$$i\partial_t \tilde{\varphi} = \frac{1}{2m}(-i\nabla - qA)^2\tilde{\varphi} + q\phi\tilde{\varphi} - \frac{q}{2m}\sigma \cdot B\tilde{\varphi} \tag{3.23}$$

which we recognize as the Pauli equation.

4 The Pauli equation: Analysis

We have seen in the previous chapter that the Pauli equation arises from the magnetic Schrödinger equation by adding some spin coupling term. This also means that the Pauli equation is closely related to the magnetic Schrödinger equation. In fact they have a very similar structure. The Hamiltonian for the magnetic Schrödinger equation H_{ms} is of the form

$$H_{\text{ms}} = (-i\nabla - A)^2 + \phi \tag{4.1}$$

for the vector potential A and the scalar potential ϕ. The Pauli Hamiltonian H_P is defined by

$$H_P := H_{\text{ms}}\mathbb{1} + \sigma \cdot B. \tag{4.2}$$

Here we see that the main difference is the additional coupling term to the magnetic field B where B is defined as the curl of A, i.e. $B = \text{curl}A = \nabla \times A$ and without the magnetic field we would have two decoupled magnetic Schrödinger equations. Therefore we will first try to understand self-adjointness of the magnetic Schrödinger operator before turning to the Pauli equation.

4.1 Self-adjointness and spectral properties of the magnetic Schrödinger operator

4.1.1 Gauge freedom

Suppose that A is a smooth potential, i.e. $A \in (C^\infty(\mathbb{R}^d))^d$. We will discuss how the gauge freedom of A affects the magnetic Schrödinger equation. To this end take two vector potentials $A_1, A_2 \in (C^\infty(\mathbb{R}^d)^d$ which have the same curl, i.e.

$$\nabla \times A_1 = \nabla \times A_2 = B. \tag{4.3}$$

Then of course the curl of their difference vanishes:

$$\nabla \times (A_1 - A_2) = 0 \tag{4.4}$$

which means that the function $A_1 - A_2$ is proportional to the gradient of a smooth function $\lambda \in C^\infty(\mathbb{R}^d)$:

$$A_1 - A_2 = \nabla\lambda. \tag{4.5}$$

Then the expressions $(-i\nabla - A_j)$, $j = 1, 2$ are unitarily equivalent to each other with the unitary operator being $\exp\{i\lambda\}$ such that

$$e^{i\lambda}(-i\nabla - A_1)e^{-i\lambda} = (-i\nabla - A_2). \tag{4.6}$$

To see this apply $(-i\nabla - A_2)$ to $e^{-i\lambda}u$ and use (4.5). Moreover,

$$e^{i\lambda}H_{\text{ms}}(A_1)e^{-i\lambda} = H_{\text{ms}}(A_2). \tag{4.7}$$

This gives us the important insight that the equation only depends on the magnetic field (which is the actual physical observable) and not on the particular choice of the vector potential as long as it satisfies (4.3). As long as this holds the magnetic Schrödinger operators are unitary equivalent under a multiplication operator (and hence have the same spectral properties). We will return to the question of the spectrum later.

4.1.2 Self-adjointness

We will look at expressions of the form

$$H_{\mathrm{ms}}u = \left((-i\nabla - A)^2 + \phi\right) u \tag{4.8}$$

where ϕ is considered as a multiplication operator M_ϕ. Let us rewrite this equation:

$$H_{\mathrm{ms}}u = \left(-\Delta + i\nabla \cdot A + 2iA \cdot \nabla + A^2 + \phi\right) u \tag{4.9}$$

where of course the expression $A \cdot \nabla$ stands for $\sum_{j=0}^{d} A_j(x)\partial_j$. The factor 2 in front of $A \cdot \nabla$ shows up because the product rule has to be applied to the term Au. Thus we can view the term

$$W := i\nabla \cdot A + 2iA \cdot \nabla + A^2 + \phi \tag{4.10}$$

as some sort of perturbation of the Laplacian $-\Delta$ and try to apply the standard tools like Kato-Rellich. For instance let us suppose that V, A^2 and $(\nabla \cdot A)$ are bounded operators which can be ensured by demanding them to be bounded functions in L^2. We then have

$$\|Wu\|^2 \leq C\left(\|u\|^2 + \|\nabla u\|^2\right) = C\left(\|u\|^2 + \langle u, -\Delta u\rangle\right) \tag{4.11}$$

$$\leq C\left(\|u\|^2 + \|u\|\|\Delta u\|\right) \leq \epsilon\|\Delta u\|^2 + C_\epsilon\|f\|^2. \tag{4.12}$$

for any ϵ with suitable $C_\epsilon \geq 0$ and for all $u \in \mathfrak{D}(-\Delta) = H^2(\mathbb{R}^d)$. The last step is due to the fact that $2ab \leq \epsilon a^2 + \frac{1}{\epsilon}b^2$ for $a, b \geq 0$.. Then W is relatively bounded with respect to $-\Delta$ and by Kato-Rellich $H_{\mathrm{ms}} = -\Delta + W$ is self-adjoint on $\mathfrak{D}(-\Delta)$. Of course those conditions are rather strict as they exclude e.g. Coulomb-like potentials for ϕ. In order to include Coulomb-like potentials we make the following assumption for $d \geq 3$:

$$\phi(x) \leq C\left(\frac{1}{|x|+1}\right), \quad C \in \mathbb{R}. \tag{4.13}$$

Then we have the following theorem, see [46]:

Theorem 4.1. *Let $A\colon \mathbb{R}^d \to \mathbb{R}^d$ be measurable, A^2 and $\nabla \cdot A$ bounded and $\phi\colon \mathbb{R}^d \to \mathbb{R}$ be measurable such that (4.13) holds. Then the operator W, defined in (4.10) is relatively bounded with respect to $-\Delta$ on $H^2(\mathbb{R}^d)$ with relative bound 0. In particular $H_{\mathrm{ms}} = -\Delta + W$ is self adjoint on $H^2(\mathbb{R}^d)$ (and essentially self-adjoint on $C_0^\infty(\mathbb{R}^d)$).*

Proof. The proof is similar to the considerations above. The only problem is the potential term. Let $u \in \mathfrak{D}(-\Delta) = H^2(\mathbb{R}^d)$. If we have

$$\int |u(x)\phi(x)|^2 \mathrm{d}x = C\int \left|\frac{u(x)}{|x|} + u(x)\right| \mathrm{d}x \leq C\left(\int \frac{|u(x)|^2}{|x|^2} + \int |u(x)|^2 \mathrm{d}x\right) \tag{4.14}$$

the term we have to worry about is the term involving $|x|^{-2}$. By Hardy's inequality for $u \in C_0^\infty(\mathbb{R}^d)$ it holds that for $d \geq 3$

$$\int \frac{|u(x)|^2}{|x|^2}\mathrm{d}x \leq \frac{4}{(d-2)^2}\int \sum_{j=1}^{d} |\partial_j u(x)|^2 \mathrm{d}x. \tag{4.15}$$

Hence we may estimate the $|x|^{-2}$-term by the first derivative of u which enters the $\|\nabla u\|$ term. Note that albeit holding only for functions in $C_0^\infty(\mathbb{R}^d)$ the reasoning works fine since $C_0^\infty(\mathbb{R}^d)$ is a dense subset of $H^2(\mathbb{R})$. By the same arguments as above we have

$$\|Wu\|^2 \leq \epsilon\| - \Delta u\|^2 + C_\epsilon \|u\|^2 \tag{4.16}$$

and hence have established the right $-\Delta$-bound for W and can conclude by Kato-Rellich. □

We can improve this result further by just demanding decay at infinity for the terms in W. The crucial point of the theorem is that decay at infinity implies relative compactness of W with respect to $-\Delta$.

Theorem 4.2. *[46]. Let $d \leq 3$. Let A be bounded and $\nabla \cdot A, \phi \in L^2_{\text{loc}}(\mathbb{R}^d)$. If*

$$|A(x)| + |(\nabla \cdot A)(x)| + |\phi(x)| \to 0 \quad \text{as } |x| \to \infty \tag{4.17}$$

then W is relatively compact with respect to $-\Delta$.

In particular $H_{\text{ms}} = -\Delta + W$ is essentially self-adjoint on $C_0^\infty(\mathbb{R}^d)$ and self adjoint on $H^2(\mathbb{R}^d)$. Moreover $\sigma_{\text{ess}}(H_{\text{ms}}) = [0, \infty)$ and there are at most countably many eigenvalues of finite multiplicity below zero which can only accumulate at zero.

Proof. Let χ_r be the characteristic function of the ball of radius r centered at zero, i.e. $B_0(r)$. Recall that the operators $I\colon H^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$, $If = f$ and $\partial_j\colon H^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ are bounded. Whence

$$\|(I - \chi_r)W\|_{\mathcal{L}(H^2(\mathbb{R}^d), L^2(\mathbb{R}^d)} \to 0 \quad \text{as } r \to \infty. \tag{4.18}$$

Hence $\chi_r W$ converges to W in the operator norm topology and we can prove our claim by showing that $\chi_r W$ is $-\Delta$-compact. In $\chi_r W$ there are two types of terms: The first type are those of a multiplication by a function $q \in L^2(\mathbb{R}^d)$. Recall that the operator q is relatively compact with respect to $-\Delta$ if $qR_{-\Delta}(z) = q(-\Delta + zI)^{-1}$ is compact for some $z \in \rho(-\Delta)$. This is the case since for $d \leq 3$ the operator $q(-\Delta + I)^{-1}$ is Hilbert-Schmidt and hence compact.

The second type of operator is of the form $u \mapsto q\nabla u$ where q is bounded (recall that we are assuming that A is bounded). Here we have

$$\|q\nabla u\| \leq C\|\nabla u\| \leq \epsilon\|\Delta u\| + C_\epsilon\|u\| \tag{4.19}$$

whence this operator has $-\Delta$-bound zero and thus already $-\Delta$-compact. This finishes the proof. □

4.1.3 Spectrum

In section 4.1.1 we have already mentioned spectral properties. The following theorem [3] gives conditions on when the magnetic Schrödinger operator and the Pauli operator have the same essential spectrum.

Theorem 4.3. *Let $A \in (C^\infty(\mathbb{R}^3)^3$ and let $|B(x)| \to 0$ as $|x| \to \infty$. Suppose ϕ is $H_0 := -\Delta$-compact where $-\Delta$ is self-adjoint in $L^2(\mathbb{R}^3)$. Then*

$$\sigma_{ess}(H_{\text{ms}}) = \sigma_{ess}(H_P) = [0, \infty). \tag{4.20}$$

Proof. The proof, as presented in [3] consists of five steps. First one reduces the case to $\phi = 0$ by the use of the diamagnetic inequality B.1. Then one shows that the spectrum is contained in $[0, \infty)$. In steps 3 to 5 one uses Weyl's criterion to determine the essential spectrum. The idea is to use the gauge freedom of A to choose a suitable Weyl sequence. This will also make use of the fact that $B \to 0$ at infinity. In the proof we will use the notation $H(A, \phi)$ to denote the dependence of on A and ϕ of the Hamiltonian.

Step 1: ϕ and $\sigma \cdot B$ are $H(A, 0)$-compact perturbations. To see this use the diamagnetic inequality

$$\left| e^{-tH(A,0)} u \right| \leq e^{-tH_0} |u|, \quad t \in \mathbb{R}, u \in \mathcal{H} \tag{4.21}$$

to show

$$|\phi \left(H(A, 0) + I\right)^{-1} u| \leq |\phi| \left(H_0 + I\right)^{-1} |u| \tag{4.22}$$

which ensures that ϕ must also be $H(A, 0)$-compact (since it is H_0-compact by assumption). This follows from the definition of relative compactness A.5 and A.11. By Weyl's theorem A.6 if ϕ is a $H(A, 0)$-compact operator the perturbed operator $H(A, 0) + \phi = H(A, \phi)$ has the same essential spectrum as the unperturbed operator $H(A, 0)$. Hence

$$\sigma_{\text{ess}}(H(A, 0)) = \sigma_{\text{ess}}(H(A, \phi)). \tag{4.23}$$

The case for the Pauli operator follows easily since $|\sigma \cdot B| \to 0$ as $|x| \to \infty$ and hence is also a $H(A, 0)$-compact perturbation.

Step 2: Evidently $H(A, 0)$ is the square of $-i\nabla - A$ and hence a positive operator from which we deduce that the spectrum must be contained in the nonnegative reals, i.e.

$$\sigma(H(A, 0)) \subseteq [0, \infty). \tag{4.24}$$

In the next three steps we will show the other inclusion, i.e. that $[0, \infty) \subseteq \sigma(H(A, 0))$ by using Weyl's theorem.

Step 3: (Weyl's theorem) The proof uses a slight variant of Weyl's theorem A.6. Let A be self-adjoint and nonnegative, i.e. $A \geq 0$. Then $z \in \sigma_{\text{ess}}$ if there is an orthonormal sequence u_n converging weakly to zero such that $\|(A+I)^{-1}(A - zI)u_n\| \to 0$.

Step 4: In fact this step is a lemma in its own right:

Lemma 4.4. *There is a sequence $\{x_n\} \in \mathbb{R}^3$ with $|x_n| \to \infty$ as $n \to \infty$ and a sequence $\{A_n\}$ of vector potentials such that*

$$\nabla \times A_n = B \quad \forall n \in \mathbb{N} \tag{4.25}$$

and there is a $C > 0$ such that

$$\sup_{|x - x_n| \leq n} |A_n(x)| \leq \frac{C}{n}. \tag{4.26}$$

Proof. Since B decays at infinity by assumption we have the following estimate for suitable C:

$$\sup_{|x - x_n| \leq n} |B(x)| \leq \frac{C}{n^2}. \tag{4.27}$$

Now define a sequence of vector potentials as follows: Let $A_n(x) = (t_n, s_n, 0)$ with

$$t_n(x) := -\int_{y_n}^{y} B_3(x,t,0)\mathrm{d}t + \int_{z_n}^{z} B_2(x,y,s)\mathrm{d}s \tag{4.28}$$

and

$$s_n(x) := -\int_{z_n}^{z} B_1(x,y,t)\mathrm{d}t. \tag{4.29}$$

Then it is easily checked that $\nabla \times A_n = B$ using $\nabla \cdot B = 0$ (by Maxwell's equations). This proves the claim by the estimate on B in equation (4.27). □

Step 5: Now we have all the ingredients to construe the Weyl sequence which by step 3 proves the claim. Recall that we want a sequence $\{u_n\} \subset \mathcal{H}$ such that

$$\|(H(A)+I)^{-1}(H(A)-k^2)u_n\| \underset{n\to\infty}{\to} 0. \tag{4.30}$$

Since $\nabla \times A_n = B$ we have

$$\nabla \times (A - A_n) = 0 \tag{4.31}$$

and hence there is a gauge function λ_n for each n such that $H(A,0)$ and $H(A_n,0)$ are unitarily equivalent:

$$H(A,0) = e^{i\lambda_n} H(A_n,0) e^{-i\lambda_n}. \tag{4.32}$$

The sequence $\{u_n\}$ is now construed as follows. Extract a subsequence $\{x_n\}$ with $|x_n - x_{n-1}| > 2n$. Furthermore we will use some bump function g which is zero outside the unit ball and one inside $B_{\frac{1}{2}}(0)$. Then we have a sequence of bump functions $\{g_n\}$ defined as $g_n = \alpha_n g(n^{-1}(x - x_n))$ where α_n is just a normalization factor such that $\|g_n\|_2 = 1$ for all n. Then define

$$u_n := e^{i\lambda_n(x)} e^{ik\cdot x} g_n(x). \tag{4.33}$$

Then $\{u_n\}$ is orthonormal in $L^2(\mathbb{R}^3)$ and $u_n \rightharpoonup 0$. The only thing that remains to show is the convergence in equation (4.30). By unitary equivalence we may substitute the vector potentials to obtain

$$\|(H(A_n)+I))^{-1}(H(A_n)-k^2)\phi_n\| \tag{4.34}$$

where $\phi_n(x) := e^{ik\cdot x} g_n(x)$. Explicitly $H(A_n)$ can be written as $H(A_n) = H_0 + (i\nabla + A_n)A_n + iA_n \cdot \nabla$. Thus,

$$\begin{aligned}
&\|(H(A_n)+I))^{-1}(H(A_n)-k^2)\phi_n\| = && (4.35)\\
=&\|(H(A_n)+I))^{-1}(H_0 - k^2 + (i\nabla + A_n) + iA_n\cdot\nabla)\phi_n\| = && (4.36)\\
=&\|(H(A_n)+I))^{-1}(e^{ik\cdot x}H_0 g_n + 2ie^{ik\cdot x}k\cdot\nabla g_n && (4.37)\\
&+ (i\nabla + A_n)A_n + iA_n\cdot\nabla)\phi_n\| && (4.38)\\
\leq&\|H_0 g_n\| + 2|k|\|\nabla g_n\| + \|(H(A_n)+I))^{-1}(i\nabla + A_n)\|\|A_n g_n\| && (4.39)\\
&+ \|A_n\cdot\nabla g_n\| \to 0 && (4.40)
\end{aligned}$$

as n goes to infinity. Hence for any k^2 we have $k^2 \in \sigma_{\mathrm{ess}}(H(A,0))$ and hence

$$[0,\infty) \subseteq \sigma_{\mathrm{ess}}(H(A,0)) \tag{4.41}$$

which proves the theorem. □

4.2 Self-adjointness of the Pauli operator

It is easily seen that the conditions for which the Pauli operator H_P is self-adjoint are strongly dependent on the magnetic field. However there are some special results and we will present excerpts of them here. The first result which we will discuss now is the 1997 paper by Sobolev [42] where the author considers the three dimensional case. The two dimensional case is covered in the preceding paper [41] which we will leave out. Note that in the same paper Lieb-Thirring estimates are covered and will be mentioned later in this section. Sobolev distinguishes two cases: Arbitrary fields and fields with constant direction. The result on fields with constant direction is stronger in the sense that the regularity assumptions for the magnetic field are weaker. However we will only consider the case for arbitrary directions here. The conditions on the magnetic field go as follows: Suppose there is a function $\ell \in C(\mathbb{R}^3)$ such that $\ell(x) - \ell(y)| \leq \varrho|x-y|$ for $\varrho \in [0,1)$ for all $x, y \in \mathbb{R}^3$. Moreover define the set $K(x)$ by

$$K(x) = \{y \in \mathbb{R}^3 | |x-y| < \ell(x)\}. \tag{4.42}$$

Assume there is a positive real-valued function $b \in L^\infty_{\text{loc}}(\mathbb{R}^3)$ and constants C_1,C_2 and c such that

$$|B(x)| \leq b(x) \tag{4.43}$$

$$C_1 b(x) \leq b(y) \leq C_2 b(x), \quad y \in K(x) \tag{4.44}$$

$$b(x)\ell(x)^2 \geq c \tag{4.45}$$

for almost all $x \in \mathbb{R}^3$. Then we have the following theorem:

Theorem 4.5. *Let B fulfill the conditions above for given $b(x)$ and $\ell(x)$. Let ϕ be a real valued potential and let $p > 3/2$ such that*

$$\sup_{x \in \mathbb{R}^3} \int\limits_{K(x)} |\phi(y)|^p (b(y)^{\frac{3}{2}+1} \mathrm{d}y < \infty. \tag{4.46}$$

Furthermore let H_P be the Pauli operator and H_P^0 the Pauli operator with $\phi = 0$. Then the form $H_P[\cdot,\cdot] = H_P^0[\cdot,\cdot] + \Phi[\cdot,\cdot]$ is closed on the form domain $\mathfrak{Q}[H_P] = \mathfrak{Q}[H_P^0]$.

The form domain of H_P^0 is defined via the symmetric operator T defined as

$$T := \sum_k \sigma_k \Pi_k = \begin{pmatrix} \Pi_3 & \Pi_1 - i\Pi_2 \\ \Pi_1 + i\Pi_2 & -\Pi_3 \end{pmatrix} \tag{4.47}$$

where $\Pi_k := -i\partial_k - A_k$. Then $\mathfrak{Q}[H_P^0] = \mathfrak{D}(T)$. The quadratic form H_P^0 is defined by $H_P^0[u,u] := \|Tu\|^2$. The Pauli operator corresponds to this quadratic form by $H_P^0 = T^*T$. Note that the form H_P^0 is semi-bounded since the Pauli operator H_P^0 is nonnegative. Hence showing that the corresponding form is closed shows that H_P^0 is self-adjoint. The operator T is however easily seen to be closed since the operators Π_i are symmetric and hence closable and we only consider their closures. The form defined by the closed operator T above is hence also closed. Moreover if the magnetic field B is bounded then so is b and

the condition on ϕ reduces to

$$\sup_{x\in\mathbb{R}^3}\int\limits_{|x-y|\leq 1}|\phi(y)|^p\mathrm{d}y<\infty. \tag{4.48}$$

Let us give two examples for the magnetic field.

Examples 4.6. (i) Let $B\in C^1(\mathbb{R}^3)$ and $B>0$ such that

$$|\nabla B(x)|\leq CB(x)^{\frac{3}{2}} \tag{4.49}$$

for all $x\in\mathbb{R}^3$. Moreover choose $\ell(x):=\epsilon B(x)^{-1/2}$ and $b(x)=B(x)$. Then (4.43) is immediately fulfilled. Furthermore $b(x)\ell(x)=B(x)\epsilon^2B(x)^{-1}=\epsilon^2$ which implies (4.45). The condition on $\ell(x)$ follows from (4.49) since

$$|\nabla\ell(x)|\leq\frac{\epsilon}{2}B(x)^{-\frac{3}{2}}|\nabla B(x)|\leq\frac{C\epsilon}{2} \tag{4.50}$$

from which we deduce that ℓ is Lipschitz continuous. Condition (4.44) is ensured by the bound

$$\frac{1}{(1+\varrho)^2}B(x)\leq B(y)\leq\frac{1}{(1-\varrho)^2}B(x) \tag{4.51}$$

for all $y\in K(x)$. As a remark note that $B(x)$ can be chosen to be $\exp(-|x|^m)$ which diverges at infinity. On the other hand if $B=|x|^{-\alpha}$ with $|\alpha|>2$ then the (4.49) is not fulfilled.

(ii) (Compactly supported magnetic field). Let B be defined as follows:

$$B(x):=\begin{cases}B_\circ & |x|\leq R\\ 0 & |x|>R\end{cases}. \tag{4.52}$$

Moreover let $\ell(x):=\varrho\sqrt{1+|x|^2}$ for $\varrho\in(0,1)$ and

$$b(x):=\begin{cases}B_\circ & |x|\leq R\\ \frac{2B_\circ}{1+\frac{|x|^2}{R^2}} & |x|>R.\end{cases} \tag{4.53}$$

Then clearly $B(x)\leq b(x)$ for all x. We have

$$b(x)\ell(x)^2=\varrho^2B_\circ(1+|x|^2)\geq\varrho^2B_\circ \tag{4.54}$$

for $|x|^2\leq R$. On the other hand for $|x|>R$

$$b(x)\ell(x)^2=2\varrho^2B_\circ\frac{1+|x|^2}{1+\frac{|x|^2}{R^2}}\geq\frac{2\varrho^2B_\circ}{2}=\rho^2B_\circ. \tag{4.55}$$

since $(1+R^2)/2\geq 1/2$ and $\lim_{x\to\infty}(1+x^2)(1+x^2/R^2)^{-1}=1$. Hence $b(x)\ell(x)^2\geq 2\varrho^2B_\circ$. To see that $\ell(x)$ has a bounded derivative consider

$$|\nabla\ell(x)|\leq\varrho\frac{r}{\sqrt{1+r^2}}=\varrho\sqrt{\frac{r^2}{1+r^2}}\leq\varrho \tag{4.56}$$

where $r\geq 0$.

Let us look at the proof in [42]. We will merely present the idea and leave out the details and refer to the paper where they are worked out. The goal of course is to show closedness of the form

$$H_P[\cdot,\cdot] = H_P^0[\cdot,\cdot] + \Phi[\cdot,\cdot] \tag{4.57}$$

defined above on the form domain $\mathfrak{Q}[H_P^0] = \mathfrak{Q}[H_P]$. This is done by showing relative form boundedness of $\Phi[\cdot,\cdot]$ with respect to $H_P^0[\cdot,\cdot]$, i.e. show that the inequality

$$\||\phi|^{\frac{1}{2}}u\|^2 \leq \epsilon H_P^0[u,u] + C\|u\|^2 \tag{4.58}$$

holds for all u in $\mathfrak{Q}[H_P^0]$. The first step is to show the following inequality

Lemma 4.7. *Let $p \geq 3/2$, $W \in L^p(\mathbb{R}^3)$ and $u \in (C_0^\infty(\mathbb{R}^3)^2$. Then*

$$\||W|^{\frac{1}{2}}u\|_2^2 \leq CH_\omega(u,W) \tag{4.59}$$

for all $\omega > 0$ where

$$H_\omega(u,W) := \ \omega\|W\|_p^{\frac{2p}{3}} H(A)[u,u] + \omega^{\frac{\beta-1}{\beta}}\|u\|_2^2 \tag{4.60}$$

where $\beta = 1 - \frac{3}{2p}$.

Proof. First of all note that due to the condition $p > 3/2$ we have $\beta > 0$. Now by Hölder's inequality and the fact that for the conjugate exponent q we have $q = 2p(p-1)^{-1}$ it holds that for all $u \in C_0^\infty(\mathbb{R}^3)$

$$\||W|^{\frac{1}{2}}u\|_2^2 \leq \|W\|_p\|u\|_q^2 \leq C\|W\|_p\|\partial u\|_2^{2(1-\beta)}\|u\|_2^{2\beta}. \tag{4.61}$$

The last inequality is due to the inequality

$$\|u\|_q \leq C_q\|\nabla u\|_2^{1-\beta}\|u\|_2^\beta \tag{4.62}$$

for $\beta = \frac{3}{q} - \frac{1}{2}$. For a more precise statement of this inequality and references see A.14. By Young's inequality for numbers with the choice $a = \|W\|_p\|\partial u\|_2^{2(1-\beta)}$ and $b = \|u\|^{2\beta}$ as well as the exponents $p = (1-\beta)^{-1}$ and $q = \beta^{-1}$ we obtain the bound

$$\||W|^{\frac{1}{2}}u\|_2^2 \leq \omega\|W\|_p^{\frac{1}{1-\beta}}\|\partial u|_2^2 + \omega^{\frac{\beta-1}{\beta}}\|u\|^2 \tag{4.63}$$

for all $\omega > 0$. The last step is to use Corollary 3.1 in [42] which tells that whenever a multiplication operator M_f by a measurable function f has a bound of the form

$$\|M_f u\|_2^2 \leq \sum_{k=1}^3 \gamma_k^2\|\partial_k u\|_2^2 + M\|u\|_2^2 \tag{4.64}$$

for all $u \in C_0^\infty(\mathbb{R}^3)$ for positive constants γ_k and M then the bound holds with the partial derivative ∂_k replaced by the magnetic derivative $\Pi_k := \partial_k - iA_k$. This is a consequence of the diamagnetic inequality. Together with this bound the proof is completed. □

With this lemma at hand we are able to prove the main argument. Let K_j be open balls covering $\mathbb{R}^n$ with no ball intersecting more than $N = N(\varrho) < \infty$ many other balls (this is also known as the finite intersection property) and let φ_j be the corresponding partition of unity $\{\varphi_j\}_{j\in\mathbb{N}}$ with $\varphi_j \in C_0^\infty(K_j)$ such that $\sum_j \varphi_j^2 = 1$ and $|\partial_m \varphi_j| \leq C_m \ell^{-|m|}$ for all m. This covering and partition exist due to Lemma 3.5 in [42] which it itself a consequence of Theorem 1.4.10 in [15]. Denote by $\chi_j := \chi_{K_j}$ the characteristic function of K_j. Then

$$\||\phi|^{\frac{1}{2}} u\|^2 = \int \sum_j \chi_j |\phi| \varphi_j^2 |u|^2 \mathrm{d}x = \sum_j \langle \chi_j |\phi| u_j, u_j \rangle \tag{4.65}$$

where in the last step we have set $u_j := \varphi_j u$. Now define

$$I_j := I_j(V) := \left(\int_{K_j} |\phi(x)|^p \mathrm{d}x \right)^{\frac{2}{3}}. \tag{4.66}$$

By Lemma 4.7 we have with $W = \chi_j V$ and $u = u_j$

$$\||\phi|^{\frac{1}{2}} u_j\|^2 \leq C H_\omega(u_j, \chi_j V) \tag{4.67}$$

for all ω. Now we rewrite the RHS of the above inequality:

$$\begin{aligned}
H_\omega(u_j, \chi_j \phi) &= \omega \|\chi_j \phi\|_p^{\frac{2p}{3}} H(A)[u_j, u_j] + \omega^{\frac{\beta-1}{\beta}} \|u_j\|_2^2 && (4.68)\\
&= \omega \left(\int_{K_j} |\phi|^p \mathrm{d}x \right)^{\frac{2}{3}} H(A)[u_j, u_j] + \omega^{\frac{\beta-1}{\beta}} \|u_j\|_2^2 && (4.69)\\
&= \omega I_j H(A)[u_j, u_j] + \omega I_j \langle \sigma B u_j, u_j \rangle - \omega I_j \langle \sigma B u_j, u_j \rangle + \omega^{\frac{\beta-1}{\beta}} \|u_j\|_2^2 && (4.70)\\
&= \omega I_j H_P^0[u_j, u_j] + \omega I_j \langle \sigma B u_j, u_j \rangle + \omega^{\frac{\beta-1}{\beta}} \|u_j\|_2^2. && (4.71)
\end{aligned}$$

Let us estimate the term $H_P^0[u_j, u_j]$. By setting $b_j := b(x_j)$ and $\ell_j := \ell(x_j)$ with $x_j \in K_j$ we have

$$\begin{aligned}
H_P^0[u_j, u_j] = \|T u_j\|^2 &= \|T \varphi_j u + [T, \varphi_j] u\| && (4.72)\\
&\leq 2\|T \varphi_j u\|^2 + 2\|[T, \varphi_j] u\|^2 && (4.73)\\
&= 2\|T \varphi_j u\|^2 + C \ell_j^{-2} \|\chi_j u\|^2 && (4.74)
\end{aligned}$$

where in the last step we used the property of the partition of unity and since the term $[T, \varphi_j]$ contains derivatives only up to order 1. Now by noting that $\langle \sigma B u_j, u_j \rangle \leq \langle B u_j, u_j \rangle$, that $|B|$ is bounded from above by b and that $b(x)\ell(x)^2 \geq c$ we deduce the following inequality for $H_\omega(u_j, \phi\chi_j)$:

$$H_\omega(u_j, \phi\chi_j) \leq 2\omega I_j \|\varphi_j T u\|^2 + C_1 \omega I_j b_j \|\chi_j u\|^2 + C_2 \omega^{\frac{\beta-1}{\beta}} \|u_j\|^2. \tag{4.75}$$

Condition (4.46) from the theorem tells us that the expression

$$I(\phi_j) b_j = b_j \left(\int_{K_j} |\phi|^p \mathrm{d}x \right)^{\frac{2}{3}} \leq C \left(\int_{K_j} |\phi|^p b(x)^{\frac{3}{2}} \mathrm{d}x \right)^{\frac{2}{3}} \tag{4.76}$$

is uniformly bounded in j. Putting all ingredients together we obtain

$$\||\phi|^{\frac{1}{2}}u_j\|^2 \leq C_1\omega\|\phi_j Tu\|^2 + C_2\left(\omega + \omega^{\frac{\beta-1}{\beta}}\right)\|\chi_j u\|^2. \tag{4.77}$$

Note that the constants C_i are independent of j. By summing up the balls K_j we arrive at

$$\||\phi|^{\frac{1}{2}}u\|^2 \leq C_1\omega H_P^0[u,u] + C_2(\omega)\sum_j \|\chi_j u\|^2. \tag{4.78}$$

The last thing to be noted is that due to the finite intersection property the last term is smaller than $C_2'(\omega)\|u\|^2$. Picking any suitable ω will lead to the desired inequality and hence prove relative form boundedness and ultimately self-adjointness of the Pauli operator. Note that the conditions on B and the potential ϕ were crucial throughout the proof, guaranteeing the uniform boundedness of the integrals I_j in j.

The next theorem relies on the Birman-Schwinger kernel for the Pauli operator. For details on the Birman-Schwinger kernel we refer to the following section. The details are somewhat left out but they can be found in [8]. Theorem 2.4 in this paper gives Lieb-Thirring type estimates for a magnetic field with arbitrary direction such that B is of the form $B = B^< + B^>$ such that $B^< \in L^\infty(\mathbb{R}^3)$ and $B^> \in L^2(\mathbb{R}^3)$. As a result of those estimates one also obtains self-adjointness or rather existence of a self-adjoint extension. The idea is to show that H_P is bounded from below by showing compactness of the corresponding Birman-Schwinger kernel. Then one can use theorem A.8 (Friedrich extension).

Theorem 4.8. *Let $B = B^< + B^>$ be such that $B^< \in L^\infty(\mathbb{R}^3)$ and $B^> \in L^2(\mathbb{R}^3)$. Moreover let ϕ be an external potential such that $|\phi|_- \in L^\gamma(\mathbb{R}^3) \cap L^{\gamma+3/2}(\mathbb{R}^3)$. Then H_P is bounded from below.*

Proof. First we need to define the Birman-Schwinger kernel K_e of the full Pauli operator $H_P = (-i\nabla - A)^2 + \phi - \sigma \cdot B$. Let H_P^0 be the Pauli operator with $\phi = 0$. Then K_e is defined as

$$K_e := \left|\phi + \frac{e}{2}\right|_-^{\frac{1}{2}} \left(H_P^0 + \frac{e}{2}\right)^{-1} \left|\phi + \frac{e}{2}\right|_-^{\frac{1}{2}}. \tag{4.79}$$

Let P_L be the spectral projection of H_P^0 on $[0, L]$ where $L \geq 0$ acting on $L^2(\mathbb{R}^3)$. Then $K_e = K_e^< + K_e^>$ with

$$K_e^< = \left|\phi + \frac{e}{2}\right|_-^{\frac{1}{2}} P_L \left(H_P^0 + \frac{e}{2}\right)^{-1} P_L \left|\phi + \frac{e}{2}\right|_-^{\frac{1}{2}} \tag{4.80}$$

$$K_e^> := \left|\phi + \frac{e}{2}\right|_-^{\frac{1}{2}} (I - P_L) \left(H_P^0 + \frac{e}{2}\right)^{-1} (I - P_L) \left|\phi + \frac{e}{2}\right|_-^{\frac{1}{2}}. \tag{4.81}$$

In [8] the corresponding Lieb-Thirring estimates are proven. In the proof one shows that $\mathrm{tr}\big((K_e^>)^2\big)$ is finite, hence $K_e^>$ is Hilbert-Schmidt and therefore compact. The claim that $K_e^<$ is compact can be established as follows. It is proven in equation (49) in [8] that for all $\rho > 0$ the number of eigenvalues above ρ is finite. From this we deduce that we can split $K_e^<$ into a part which can be

estimated by ρ (for the eigenvalues smaller than ρ) and into a compact part depending on ρ (for the eigenvalues above ρ of which there are only finitely many). Then by letting $\rho \to 0$ the first part converges to zero whereas the second part converges to a compact operator (the limit of compact operators is compact). This proves the claim. Hence K_e must also be compact since the sum of compact operators is compact.

Since K_e is compact it follows that $|\phi + e/2|_-$ is a relatively compact perturbation of $H_P^0 + e/2$.[2] Thus, if we set $U_e := H_P^0 + e/2 - |\phi - e/2|_-$ we have that U_e is bounded from below since it is the relatively compact perturbation of a positive operator. On the other hand H_P^0 can be rewritten as

$$H_P^0 = U_e + \left|\phi + \frac{e}{2}\right|_+ - e \tag{4.82}$$

which is bigger that $U_e - e$ which is bounded from below. Hence H_P^0 is bounded from below. Consequently H_P is bounded from below. □

4.3 Spectrum of the Pauli operator: Lieb-Thirring estimates

Roughly speaking **Lieb-Thirring estimates** are estimates on the moments of the negative eigenvalues of a Schrödinger operator. Originally they were introduced by Lieb and Thirring in [29] where they considered an operator of the form $H = -\Delta + V(x)$. The results have been extended to magnetic Schrödinger operators by use of the diamagnetic inequality and in recent developments to the Pauli operator. We give a short overview on the results.

4.3.1 Lieb-Thirring estimates for the (magnetic) Schrödinger operator with external potential

Let $H = -\Delta + V(x)$ on $L^2(\mathbb{R}^d)$ and denote by $\lambda_1 \leq \lambda_2 \leq \cdots < 0$ the negative eigenvalues of H if they exist. Then denote by $\mathfrak{M}_\gamma(V)$ the sum

$$\mathfrak{M}_\gamma(V) := \sum_j |\lambda_j|^\gamma \tag{4.83}$$

for $\gamma \geq 0$. Lieb and Thirring derived the following bound for $\mathfrak{M}_\gamma$ in [29]. If $\gamma > \max(0, 1 - \frac{d}{2})$ then

$$\mathfrak{M}_\gamma(V) = \sum_j |\lambda_j|^\gamma \leq L_{d,\gamma} \int_{\mathbb{R}^d} V_-(x)^{\gamma + \frac{d}{2}} \mathrm{d}x \tag{4.84}$$

where V_- denotes the negative part of V, i.e. $V_-(x) := \max(-V(x), 0)$. The case where $\gamma = 0$ and $d \geq 3$ which corresponds to estimating the number of eigenvalues has been covered by Cwikel, Lieb and Rosenbljum independently. As a side remark note that the sum in (4.84) makes sense since we can assume that the Schrödinger operator is semibounded from below and that the eigenvalues do not accumulate at negative infinity. In order to prove this inequality we will follow the lecture notes by Erdős [7] who follows the book by Lieb [25].

[2] Recall that for an operator K to be relatively compact with respect to A it has to hold that $KR_A(z)$ is compact for some z in the resolvent of A.

Another presentation of the proof can be found in chapter 12 in [23]. First we need a preparatory result by Birman and Schwinger, the **Birman-Schwinger principle**. It reduces the problem of negative eigenvalues of the Schrödinger operator to the eigenvalue of a compact operator, the **Birman-Schwinger operator** K_e.

Theorem 4.9. (Birman-Schwinger) *Let $V \in L^{d/2} + L^\infty$. Let $e > 0$ and define*

$$K_e := \sqrt{V_-}\,(-\Delta + e)^{-1}\,\sqrt{V_-} \tag{4.85}$$

where V_- denotes the negative part of V. Denote by N_e the number of eigenvalues of $H = -\Delta + V$ less than or equal to $-e$ and by B_e the number of eigenvalues of K_e bigger than or equal to 1. Then

$$N_e = B_e. \tag{4.86}$$

The operator K_e is motivated as follows. If $-e$ is a negative eigenvalue of $H = -\Delta - V_-$, i.e. $e > 0$ and $(-\Delta + e)\tilde{u} = V_-\tilde{u}$ where $\tilde{u}$ is a normalized eigenfunction then define

$$u(x) := \sqrt{V_-(x)}\tilde{u}(x) \tag{4.87}$$

such that

$$(-\Delta + e)\tilde{u} = \sqrt{V_-}u. \tag{4.88}$$

If $(-\Delta + e)$ has an inverse which is true if $e \in \rho(-\Delta)$ then we have

$$\tilde{u} = (-\Delta + e)^{-1}\sqrt{V_-}u \tag{4.89}$$

such that

$$u = K_e u \tag{4.90}$$

with K_e as defined in 4.9. Hence we are let to believe that whenever $-e$ is a negative eigenvalue of H then 1 is an eigenvalue of K_e. Let us look at this more thoroughly. First let us show that $u \in L^2$ if $\tilde{u}$ is an eigenfunction of H. Then $\tilde{u} \in H^1$ and we have the following chain of inequalities where we use Hölder's inequality with $p = \frac{d}{2}$ and $p^* = \frac{d}{d-2}$, Sobolev's inequality and the fact that by assumption $V_- = V_1 + V_2$ with $V_1 \in L^{\frac{d}{2}}$ and $V_2 \in L^\infty$:

$$\|u\|_2^2 = \int \left|\sqrt{V_-}\tilde{u}\right|^2 = \int V_-|\tilde{u}|^2 = \int (V_1 + V_2)|\tilde{u}|^2 \tag{4.91}$$

$$\leq \|V_1\|_{\frac{d}{2}}\|\tilde{u}\|^2_{\frac{2d}{d-2}} + \|V_2\|_\infty\|\tilde{u}\|_2^2 \tag{4.92}$$

$$\leq C\|V_-\|_{\frac{d}{2}}\|\nabla\tilde{u}\|_2^2 + \|V_2\|_\infty < \infty. \tag{4.93}$$

Also it immediately follows that $M_{\sqrt{V_-}}\colon H^1 \to L^2$ is bounded since

$$\int \left|\sqrt{V_-}\tilde{u}\right|^2 \leq C_V\|\tilde{u}\|^2_{H^1}. \tag{4.94}$$

Next we show that K_e is a bounded operator on L^2.

Lemma 4.10. *The operator K_e defined in equation (4.85) is bounded on L^2.*

Proof. In order to see this define

$$B := \sqrt{V_-}(-\Delta + e)^{-1/2}. \tag{4.95}$$

The part $(-\Delta + e)^{-1/2}$ maps L^2 to H^1 which we can see by the estimate

$$\|(-\Delta + e)^{\frac{1}{2}} f\|_{H^1}^2 = \int (1 + (2\pi k)^2) \left|((2\pi k)^2 + e)^{-\frac{1}{2}} \hat{f}(k)\right|^2 \mathrm{d}k \tag{4.96}$$

$$\leq \max(1, e^{-1}) \|f\|_2^2 \tag{4.97}$$

since $(1 + (2\pi k)^2(e + 2\pi k)^{-2} \leq \max(1, e^{-1})$. Now $\sqrt{V_-}$ is a multiplication operator from H^1 to L^2 so indeed B is bounded on L^2. The adjoint of B is given by

$$B^* = (-\Delta + e)^{-\frac{1}{2}}(\sqrt{V_-}\cdot) \tag{4.98}$$

so $K_e = BB^*$ and hence bounded on L^2. □

Conversely let $u \in L^2$ be an eigenfunction of K_e with eigenvalue 1 and define $\tilde{u} := (-\Delta + e)^{-1}\sqrt{V_-}u$. Then it is easy to see that $\tilde{u}$ is an eigenfunction of H with eigenvalue $-e$. It only remains to show that $\tilde{u} \in L^2$. Hence consider

$$\|\tilde{u}\|_2^2 = \langle \sqrt{V_-}u, (-\Delta + e)^{-2}\sqrt{V_-}u\rangle \tag{4.99}$$

$$\leq e^{-1}\langle \sqrt{V_-}u, (-\Delta + e)^{-1}\sqrt{V_-}u\rangle \tag{4.100}$$

$$= e^{-1}\langle u, K_e u\rangle \tag{4.101}$$

$$= e^{-1}\|u\|_2^2. \tag{4.102}$$

The remaining ingredient for the proof of 4.9 is the fact that K_e is compact. In fact we will show that K_e is Hilbert-Schmidt which implies compactness.

Lemma 4.11. *K_e is a Hilbert-Schmidt operator if $V_- \in L^2$.*

Proof. We show the claim for $d = 3$ which is the interesting case. Recall that an operator is Hilbert-Schmidt if its Hilbert-Schmidt norm is finite, i.e. $\mathrm{tr}(A^*A) < \infty$. In this case $A = K_e$ and hence self-adjoint so we compute

$$\mathrm{tr}(K_e^2) = \mathrm{tr}(\sqrt{V_-}(\Delta + e)^{-1}V_-(\Delta + e)^{-1}\sqrt{V_-}) = \tag{4.103}$$

$$= \iint V_-(x)|(\Delta + e)^{-1}(x, y)|^2 V_-(y)\mathrm{d}x\mathrm{d}y. \tag{4.104}$$

The operator kernel of the integral operator $(\Delta + e)^{-1}$ is known and given by

$$\frac{1}{4\pi}\frac{e^{-\sqrt{e}|x-y|}}{|x-y|}. \tag{4.105}$$

We then have

$$\mathrm{tr}(K_e^2) = \frac{1}{16\pi^2}\iint V_-(x)\frac{e^{-2\sqrt{e}|x-y|}}{|x-y|^2}V_-(y)\mathrm{d}x\mathrm{d}y. \tag{4.106}$$

and eventually

$$\mathrm{tr}(K_e^2) \leq \frac{1}{16\pi^2}\iint V_-(x)^2\frac{e^{-2\sqrt{e}|x-y|}}{|x-y|^2}\mathrm{d}x\mathrm{d}y = \frac{C}{e^2}\int V_-^2. \tag{4.107}$$

By assumption $V_- \in L^2$ and hence $\mathrm{tr}(K_e^2) < \infty$. □

Proof. (of Theorem 4.9). As K_e is a positive compact self-adjoint operator its eigenvalues (depending on e) can be written in decreasing order

$$\mu_1(e) \geq \mu_2(e) \geq \cdots \geq 0. \tag{4.108}$$

The functions $\mu_j(e)$ are continuous functions of e. To see this observe that for $0 < e \leq e'$ we have that $0 \leq K_e - K_{e'} \leq ((e'-e)/e)K_e$ (which can be calculated directly). Then by the min-max-principle (see [23]) one has that

$$|\mu_j(e) - \mu_j(e')| \leq \frac{e'-e}{e'}\|K_e\| \tag{4.109}$$

This proves the claim that $\mu_j(e)$ is continuous in e. Hence if we fix some e and let e' increase monotonically from e to infinity all $\mu_j(e)$ decrease (since K_e is monotonically decreasing in e) and by reaching one the number $-e'$ is an eigenvalue of H. Since $K_e \to 0$ as $e \to \infty$ for sufficiently large e' all eigenvalues of K_e are below one and hence the number of eigenvalues of K_e which are above or equal to one must equal the number of eigenvalues of H below or equal to $-e$. □

Now we are ready to prove the Lieb-Thirring estimates. We leave out the case $d = 3, \gamma = 0$ which counts the eigenvalues. The proof uses different machinery and the reader is referred to the papers by Cwickel [2], Lieb [22] and Rosenbljum [37].

Theorem 4.12. (Lieb-Thirring) *Let $\gamma \geq 0$ and let $V_- \in L^{\gamma+d/2}(\mathbb{R}^d)$. Then there is a constant $L_{\gamma,d} < \infty$ such that the sum of the γ-th moments of the eigenvalues of the Schrödinger operator $H = -\Delta + V$ can be estimated by*

$$\mathfrak{M}(V) = \sum_j |\lambda_j|^\gamma \leq L_{\gamma,d} \int V_-^{\gamma+\frac{d}{2}}. \tag{4.110}$$

Here γ can have the following ranges:

$$\gamma \geq \frac{1}{2} \qquad \text{if } d = 1 \tag{4.111}$$

$$\gamma > 0 \qquad \text{if } d = 2 \tag{4.112}$$

$$\gamma \geq 0 \qquad \text{if } d = 3. \tag{4.113}$$

Proof. As said we will not prove the case $\gamma = 0$. Suppose $\gamma > 0$. If we interpret N_e as a sum of characteristic functions for the set where e is smaller than λ_j we have

$$N_e = \sum_j \chi_{\{e\in\mathbb{R}:e<|\lambda_j|\}} \tag{4.114}$$

and we can write the sum of the eigenvalues $|\lambda_j|$ as

$$\sum_j |\lambda_j| = \int_0^\infty \sum_j \chi_{\{e\in\mathbb{R}:e<|\lambda_j|\}} \mathrm{d}e = \int_0^\infty N_e \mathrm{d}e. \tag{4.115}$$

Here we may interchange integration and summation since the integrand is positive. By a change of variables ($\mathrm{d}e \to \gamma e^{\gamma-1}\mathrm{d}e$) and by the formula

$$|\lambda_j|^\gamma = \int_0^\infty \chi_{\{e\in\mathbb{R}:e<|\lambda_j|^\gamma\}} \mathrm{d}e \tag{4.116}$$

we can write

$$\sum_j |\lambda_j|^\gamma = \gamma \int_0^\infty e^{\gamma-1} N_e \mathrm{d}e. \tag{4.117}$$

For B_e we have the estimate

$$B_e \leq \sum_{\mu_j \geq 1} \mu_j^m \leq \mathrm{tr}(K_e^m) \tag{4.118}$$

and hence

$$\mathfrak{M}(V) \leq \gamma \int_0^\infty e^{\gamma-1} \mathrm{tr}(K_e^m) \mathrm{d}e \tag{4.119}$$

by the Birman-Schwinger principle. By theorem B.2 we have then

$$\begin{aligned}
\mathrm{tr}(K_e^m) &\leq \mathrm{tr}(V_-^{\frac{m}{2}} (\Delta + e)^{-m} V_-^{\frac{m}{2}}) && (4.120)\\
&= \int V_-^m(x_1)(\Delta+e)^{-1}(x_1,x_2)\cdots(\Delta+e)^{-1}(x_m,x_1)\mathrm{d}x_1\cdots\mathrm{d}x_m && (4.121)\\
&= \int V_-(x)^{\frac{m}{2}} (\Delta+e)^{-m}(0,0)\mathrm{d}x && (4.122)\\
&= \int \frac{1}{((2\pi k)^2+e)^m}\mathrm{d}k \int V_(x)^m \mathrm{d}x. && (4.123)
\end{aligned}$$

The expression $\int((2\pi k)^2+e)^{-m}$ is finite iff $2m > d$ and given by

$$C_{m,d} e^{-m+\frac{d}{2}} \tag{4.124}$$

where

$$C_{m,d} = (4\pi)^{(-d/2)} \Gamma(m - d/2)/\Gamma(m). \tag{4.125}$$

It is now necessary to use a little trick to circumvent the divergence of the integral (4.119) when inserting the expression $e^{-m+\frac{d}{2}}$. Note that this would yield an integrand of the from $e^{\gamma-1-m+d/2}$ which diverges regardless of the choice of exponents. Therefore we introduce a "shifted" potential $\tilde{V}_e$ depending on e given by

$$\tilde{V}_e := |V(x) + \frac{e}{2}|_- = \max\{-V(x) - \frac{e}{2}, 0\}. \tag{4.126}$$

Clearly the number of eigenvalues of the unshifted Hamiltonian $-\Delta - V_-$ below $-e$ is equal to the number of eigenvalues of the shifted Hamiltonian $-\Delta - V_- + e/2$ which lie below $-e/2$ and hence

$$N_e(-V_-) = N_{e/2}(-V_- + \frac{e}{2}) \leq N_{e/2}(-\tilde{V}_e). \tag{4.127}$$

This is because $-\tilde{V}_e \leq -V_- + e/2$ which relies on the fact that $V_- \leq a + |V + a|_-$

for any $a > 0$. Repeating the argument above with $\tilde{V}_e$ and $e/2$ we obtain

$$\mathfrak{M}(V) \leq C_{d,m,\gamma} \int_0^\infty e^{\gamma-1} \left(\frac{e}{2}\right)^{-m+d/2} \int_{\mathbb{R}^n} \tilde{V}_e(x)^m \mathrm{d}x\mathrm{d}e \tag{4.128}$$

$$\lesssim C_{d,m,\gamma} \int_{\mathbb{R}^n} \int_0^\infty e^{\gamma-1-m+d/2} \left|V(x) + \frac{e}{2}\right|_-^m \mathrm{d}e\mathrm{d}x \tag{4.129}$$

$$= C_{d,m,\gamma} \int_{\mathbb{R}^n} \int_0^{2V_-(x)} e^{\gamma-1-m+d/2} \left(V_-(x) + \frac{e}{2}\right)^m \mathrm{d}e\mathrm{d}x \tag{4.130}$$

$$\lesssim C_{d,m,\gamma} \int_{\mathbb{R}^n} V_-^{\gamma+d/2} \mathrm{d}x \int_0^1 u^{\gamma-1-m+d/2}(1-u)^m \mathrm{d}u. \tag{4.131}$$

The last step is due to the change of variables $e = 2V_-(x)u$ and the second to last step to the fact that for $e \geq 2V_-(x)$ the expression $|V(x) + e/2|_-$ is zero. Now note that for $m \geq \gamma + d/2$ the integral diverges whence we choose $m < \gamma + d/2$. As we have seen before $2m > d$ has to be valid, too. Combining these conditions yields the Lieb-Thirring bound (4.84) for $\gamma > 0$ and $d \geq 2$ and for $\gamma > 1/2$ and $d = 1$. □

4.3.2 Lieb-Thirring estimates for the Pauli operator

In this section we will briefly mention and discuss results on Lieb-Thirring estimates for the Pauli operator without going to much into details. Main works include the series of papers by Lieb, Solovej and Yngvason: [26], [27] and [28], the works by Erdős [9], [8] and by Sobolev [41] (covering the two-dimensional case), [42] (covering the three-dimensional case). Lieb, Solovej and Yngvason established Lieb-Thirring bounds for the three dimensional Pauli operator $H_P = (-i\nabla - A)^2 + \phi - \sigma \cdot B$ where the magnetic field is constant. For the bound of the sum of the negative eigenvalues of H_P they established the following result:

$$\sum_j |\lambda_j| \leq C' \int \phi_-^{\frac{5}{2}} + C'' \int |B| \phi_-^{\frac{3}{2}}, \quad \gamma > \frac{1}{2}. \tag{4.132}$$

Clearly the first term is the Lieb-Thirring estimate for a Schrödinger operator without magnetic field. Sobolev in [42] calls (4.132) the **strong** version. The **weak** version is given by

$$\sum_j |\lambda_j| \leq C' \int \phi_-^{\frac{5}{2}} + C'' \int |B|^{\frac{3}{2}} \phi_-, \quad \gamma \geq 1. \tag{4.133}$$

The result for arbitrary fields in [42] depends on the conditions already mentioned above in the statement about self-adjointness of the Pauli operator. Under those conditions we have the following theorem.

Theorem 4.13. (Theorem 2.2 in [42]) *Let the hypotheses of Theorem 4.5 hold. Moreover suppose that* $\phi_- \in L^{\gamma+3/2}(\mathbb{R}^3)$ *and* $\phi_-^\gamma b^{3/2} \in L^1(\mathbb{R}^3)$ *for* $\gamma \geq 1$. *Then*

the Lieb-Thirring bound for H_P is given by

$$\mathfrak{M}_\gamma(\phi) \leq C'_\gamma \int \phi_-(x)^{\gamma+3/2}\mathrm{d}x + C''_\gamma \int \phi_-(x)^\gamma b(x)^{3/2}\mathrm{d}x \tag{4.134}$$

where the constants depend only on γ and the constants in the estimates on $b(x)$ and $\ell(x)$.

Remark 4.14. One can easily rescale the magnetic field without touching the validity of 4.13. In fact let $\mu > 0$ and define $\tilde{A} := \mu A$. Then $\tilde{B} = \nabla \times (\mu A) = \mu(\nabla \times A) = \mu B$. The estimates for B hold for $\tilde{b} = \mu b$ and $\tilde{\ell} = \ell$ if $\mu \geq c$. An interesting example of μ involves the reduced Planck constant $\hbar$. Denote by

$$H^0_{P,\hbar} := \begin{pmatrix} (-i\hbar\nabla - A)^2 & 0 \\ 0 & (-i\hbar\nabla - A)^2 \end{pmatrix} - \hbar(\sigma \cdot B) \tag{4.135}$$

the Pauli operator with $\hbar$. We can also write

$$H^0_{P,\hbar} := \hbar^2 \left(\begin{pmatrix} (-i\nabla - \frac{A}{\hbar})^2 & 0 \\ 0 & (-i\nabla - \frac{A}{\hbar})^2 \end{pmatrix} - (\sigma \cdot \frac{B}{\hbar}) \right) = \hbar^2 H^0_P(A/\hbar) \tag{4.136}$$

By introducing the potential V we have

$$H_{P,\hbar} = \hbar^2 H^0_P(A/\hbar) + V = \hbar^2 \left(H^0_P(A/\hbar) + \frac{V}{\hbar^2} \right) = \hbar^2 H_P(A/\hbar, V/\hbar^2). \tag{4.137}$$

Now let us look at $\mathfrak{M}_\gamma$:

$$\mathfrak{M}_\gamma(H_{P,\hbar}(A,V)) = \mathfrak{M}_\gamma(\hbar^2 H_{P,\hbar}(A/\hbar, V/\hbar^2)) \tag{4.138}$$

$$= \sum_j |\hbar^{-2}\lambda_j|^\gamma \tag{4.139}$$

$$= \hbar^{-2\gamma} \sum_j |\lambda_j|^\gamma \tag{4.140}$$

On the other hand

$$\mathfrak{M}_\gamma(H_{P,\hbar}(A,V)) \leq C'_\gamma \int (\frac{V_-}{\hbar^2})(x)^{\gamma+3/2}\mathrm{d}x + C''_\gamma \int (\frac{V_-}{\hbar^2})(x)^\gamma (\frac{b(x)}{\hbar})^{3/2}\mathrm{d}x \tag{4.141}$$

$$= \frac{1}{\hbar^{2\gamma}} \left(C'_\gamma \frac{1}{\hbar^3} \int V_-(x)^{\gamma+3/2}\mathrm{d}x + C''_\gamma \frac{1}{\hbar^{3/2}} \int V_-(x)^\gamma b(x)^{3/2}\mathrm{d}x \right) \tag{4.142}$$

The factor $\hbar^{-2\gamma}$ is eliminated by the same factor in $\mathfrak{M}_\gamma(H_{P,\hbar}(A,V))$ so finally

$$\sum_j |\lambda_j|^\gamma \leq C'_\gamma \frac{1}{\hbar^3} \int V_-(x)^{\gamma+3/2}\mathrm{d}x + C''_\gamma \frac{1}{\hbar^{3/2}} \int V_-(x)^\gamma b(x)^{3/2}\mathrm{d}x \tag{4.143}$$

An improvement of the foregoing results for the case $\gamma = 1$, i.e. the case of summing up the eigenvalues is Theorem 3.1 in [9]. The conditions on B are much more gentle and require only some degree of differentiability. In order to phrase this theorem we need to collect some definitions first.

Definition 4.15. (Lengthscales of the magnetic field) *Let $B \in C^4(\mathbb{R}^3)$. Let $L \geq 0$ and $x \in \mathbb{R}^3$ and define*

$$B_L(x) := \sup_{y \in \mathbb{R}^3} \{|B(y) \colon |x - y| \leq L\} \tag{4.144}$$

$$b_L(x) := \inf_{y \in \mathbb{R}^3} \{|B(y) \colon |x - y| \leq L\}. \tag{4.145}$$

In other words B_L and b_L are the supremum and infimum of the magnetic field strength on the ball of radius L about the point x. The **magnetic lengthscale of** *B, $L_m(x)$, is defined as*

$$L_m(x) := \sup\{L > 0 \colon B_L(x) \leq L^{-2}\}. \tag{4.146}$$

The **variation lengthscale of** *B $L_v(x)$ is given by*

$$L_v(x) := \sup\{L \geq 0 \colon L^\gamma \cdot \sup_{\gamma=1,2,3,4} \{|\nabla^\gamma B(y)| \colon |x - y| \leq L\} \leq b_L(x)\}. \tag{4.147}$$

Moreover the **combined lengthscale** *$L_c(x)$ is defined as*

$$L_c(x) := \max\{L_m(x), L_v(x)\}. \tag{4.148}$$

The definition of the lengthscales can be motivated as follows. The magnetic lengthscale L_m is the upper bound on the radius of the ball for which the supremum of the magnetic field strength on the ball is bounded by the inverse square of said radius. The variational lengthscale somehow determines the scale on which B varies.

Theorem 4.16. (Uniform Lieb-Thirring estimates) *Let $A, B \in C^4(\mathbb{R}^3)$ such that $\nabla \times A = B$, i.e. A is the vector potential and B is the magnetic field. Denote by H_P the full Pauli operator with electric potential V. Then the sum of the negative eigenvalues $\mathfrak{M}_1(V) = \sum_j |\lambda_j|$ has the following Lieb-Thirring bound*

$$\mathfrak{M}(V) \leq C_1 \int |V|_-^{5/2} + C_2 \int |B||V|_-^{3/2} + C_3 \int \left(\frac{1}{L_c} |B||V|_- + \frac{1}{L_c^3} |V|_- \right) \tag{4.149}$$

whenever the integrals exist.

The proof of 4.16 again relies on the Birman-Schwinger principle. Let us sketch the proof. Define the **tempered lengthscale** $L(x)$ as

$$L(x) := \frac{1}{2} L_c(x). \tag{4.150}$$

Furthermore define the function $\ell(x)$:

$$\ell(x) := \epsilon L(x) \tag{4.151}$$

and let

$$P(x) := \epsilon^{-5} \ell(x)^{-2}. \tag{4.152}$$

In order to stick to the notation of [9] denote by R_f the resolvent of $H_P(f) = (\sigma \cdot (-i\nabla - A))^2 + f$ where $f > 0$ is a positive function (note that the operator

$H_P(f)$ is positive). Now by the Birman-Schwinger principle the sum of the eigenvalues of the negative part of H_P can be written as

$$|\operatorname{tr}(H_{P,-})| = \int_0^\infty N_1(V^{1/2}R_eV^{1/2})\mathrm{d}e \tag{4.153}$$

where $N_\mu(A)$ denotes the number of eigenvalues of the operator A greater than or equal to μ. By the second resolvent identity $R_e - R_{P+e} = R_{P+e}PR_{P+e}$ and hence

$$R_e = R_{P+e} + R_{P+e}PR_{P+e} + R_{P+e}PR_ePR_{P+e}. \tag{4.154}$$

Moreover $P \leq (\sigma\cdot(-i\nabla - A))^2 + e + P$ and $R_e \leq e^{-1}$ by positivity of H_P. From this and from the resolvent identity above we immediately deduce

$$R_e \leq 2R_{P+e} + e^{-1}R_{P+e}P^2R_{P+e}. \tag{4.155}$$

It is easy to see that for positive operators A_1 and A_2 we have the following identity:

$$N_{\mu_1+\mu_2}(A_1 + A_2) \leq N_(\mu_1)(A_1) + N_{\mu_2}(A_2). \tag{4.156}$$

This allows to split the integral (4.153) into two parts. By the two last equation we then obtain

$$\begin{aligned}|\operatorname{tr}(H_{P,-})| \leq &\int_0^\infty N_{\frac{1}{4}}(V^{1/2}R_{P+e}V^{1/2})\mathrm{d}e+\\ &\int_0^\infty N_e(2V^{1/2}R_{P+e}P^2R_{P+e}V^{1/2})\mathrm{d}e.\end{aligned} \tag{4.157}$$

Theorem 4.3 in [9] gives the needed estimates for the two integrals but we will not try to retrace the proof here. Let us simply give a short physical interpretation of the terms occurring. The first term corresponds to the number of eigenvalues of the shifted operator $|(\sigma\cdot(-i\nabla - A))^2 + P - 4V|_-$. That is it contains the contribution from eigenfunctions with energy at least $O(P)$. It is also called **positive energy regime**, cf. [9]. The second term contains the contribution of eigenfunctions with energy near zero and hence is called **zero mode regime**.

Moreover let us quote another result due to Lieb-Loss and Solovej in [24] where the only condition on B is that it is in L^2.

Theorem 4.17. [Theorem 2 in [24]]. *Let $B \in L^2(\mathbb{R}^3)$. If $V \in L^4(\mathbb{R}^3)\cap L^{5/2}(\mathbb{R}^3)$ then the sum of the negative eigenvalues $\mathfrak{M}_1(V) = \sum_j |\lambda_j|$ has the following Lieb-Thirring bound*

$$\mathfrak{M}(V) \leq C_1\int |V|_-^{5/2} + C^2\left(\int |B|^2\right)^{\frac{3}{4}} + \left(\int |V|_-^4\right)^{\frac{1}{4}} \tag{4.158}$$

$$= C_1\int |V|_-^{5/2} + C^2\|B\|_2^{3/2}\|V\|_4. \tag{4.159}$$

5 Further topics

5.1 Quantum Statistical Mechanics: Density matrix formulation

Quantum Mechanics is commonly interpreted as a *probabilistic* theory. Outcomes of measurements are described by the spectrum of self-adjoint operators and occur with certain probability as for example in a spin measurement in z-direction. However, the somewhat *intrinsic* probabilistic nature aside, Quantum Mechanics can be viewed in a *statistical* way very similar to Classical Mechanics. In a setting where a great number of particles is observed (e.g. a quantum gas) it is neither efficient nor realistic to be able to describe every particle by its wave function and corresponding Schrödinger equation since the number of equations to solve grows with the size of the system. Hence one has to resort to a statistical theory which is known as *Statistical Quantum Mechanics*. The density matrix formulation is central to Statistical Quantum Mechanics. Here, not L^2-wave functions are viewed as the fundamental objects but a bounded operator, called **density operator** which describes all the states and whose time evolution is given by the **von Neumann-Heisenberg equation**. We will give a mathematical foundation of the density operator based on chapter 19 in [14] and try to apply it to the Pauli Hamiltonian.

What one observes in experiment are of course not the observables or the states themselves but rather the expectation values of them. Hence we wish to obtain an understanding of those expectation values. We can for the moment restrict ourselves to bounded operators since for an unbounded self-adjoint operator A we can look at the expectation value of the operator $\chi_E(A)$ where $E \subset \mathbb{R}$ is a Borel set $\chi_E(A)$ is defined by functional calculus. If $\mathbb{E}(1_E(A))$ is known for every E then we can reconstruct the expectation value for A. In the end we want to express the system by its density operator and express expectation values via the density operator. Therefore we need some theoretical considerations.

Definition 5.1. *Let $\hat{\rho} \in \mathcal{B}(\mathcal{H})$ be a bounded operator. Then $\hat{\rho}$ is called a* **density operator** *if $\hat{\rho}$ is self-adjoint, nonnegative and trace class with $\operatorname{tr}(\hat{\rho}) = 1$.*

Definition 5.2. *A map $\Phi\colon \mathcal{B}(\mathcal{H}) \to \mathbb{C}$ is called a* **family of expectation values** *if the following hold*

(i) $\Phi(I) = 1$.

(ii) If A is self-adjoint then $\Phi(A) \in \mathbb{R}$. If A is also nonnegative then $\Phi(A) \geq 0$.

(iii) (Continuity) *Let $\{A_n\}$ be a sequence in $\mathcal{B}(\mathcal{H})$. If $\|A_n u - Au\| \to 0$ for all $u \in \mathcal{H}$ then $\Phi(A_n) \to \Phi(A)$ in $\mathbb{C}$.*

Theorem 5.3. *Let $\hat{\rho}$ be a density operator. Then the map $\Phi_{\hat{\rho}}\colon \mathcal{B}(\mathcal{H}) \to \mathbb{C}$ defined by*

$$\Phi_{\hat{\rho}}(A) := \operatorname{tr}(\hat{\rho}A) = \operatorname{tr}(A\hat{\rho}) \tag{5.1}$$

is a family of expectation values.

Proof. Item (i) is clear since $\mathrm{tr}(\hat{\rho}) = 1$ by definition. For item (ii) consider a self-adjoint operator A. Then $\mathrm{tr}(\hat{\rho}A^*) = \mathrm{tr}(A^*\hat{\rho}) = \mathrm{tr}((\hat{\rho}A)^*) = \overline{\mathrm{tr}(\hat{\rho}A)}$ from which the first part of (ii) follows. For the second part consider the self-adjoint square root $\hat{\rho}^{1/2}$ which exists since by assumption ρ is nonnegative. The operator $\hat{\rho}^{1/2}$ is clearly Hilbert-Schmidt and so is $A\hat{\rho}^{1/2}$ since the product of a trace class operator with a bounded operator is trace class. Since $\mathrm{tr}(AB) = \mathrm{tr}(BA) < \infty$ for two bounded Hilbert-Schmidt operators we have

$$\mathrm{tr}(\hat{\rho}A) = \mathrm{tr}\Big(\hat{\rho}^{1/2}\hat{\rho}^{1/2}A\Big) = \mathrm{tr}\Big(\hat{\rho}^{1/2}A\hat{\rho}^{1/2}\Big) \geq 0 \tag{5.2}$$

since $\hat{\rho}^{1/2}A\hat{\rho}^{1/2}$ is non-negative[3]. To see (iii) consider a sequence $A_n u$ converging to Au. By the uniform boundedness principle there is a constant C such that $\|A_n\| \leq C$ for all n. Now choose an orthonormal basis $\{e_j\}$ in $\mathcal{H}$ such that

$$|\langle e_j, \hat{\rho}^{1/2}A_n\hat{\rho}^{1/2}e_j\rangle| = |\langle \hat{\rho}^{1/2}e_j, A_n\hat{\rho}^{1/2}e_j\rangle| \leq C\|\hat{\rho}^{1/2}e_j\|^2 < \infty. \tag{5.3}$$

Whence by dominated convergence

$$\mathrm{tr}\Big(\hat{\rho}^{1/2}A\hat{\rho}^{1/2}\Big) = \lim_{n\to\infty} \mathrm{tr}\Big(\hat{\rho}^{1/2}A_n\hat{\rho}^{1/2}\Big) \tag{5.4}$$

thereby proving the claim. □

The next result ensures that given a family of expectation values and given a bounded operator $A \in \mathcal{B}(\mathcal{H})$ we can always find a unique density matrix such that $\Phi(A) = \mathrm{tr}(\rho A)$.

Theorem 5.4. *Let Φ be a family of expectation values. Then there is a unique density operator $\hat{\rho}$ such that*

$$\Phi(A) = \mathrm{tr}(\hat{\rho}A) \tag{5.5}$$

for all $A \in \mathcal{B}(\mathcal{H})$.

Proof. See Theorem 19.9 in [14]. □

To see how the notion of a state in the usual sense is included in the density matrix formulation consider a ket $|u\rangle$ and denote by $|u\rangle\langle u|$ the orthogonal projection on the span of the vector $u \in \mathcal{H}$. An orthogonal projection is by definition bounded, self-adjoint and nonnegative so let us compute its trace. In a suitable orthonormal basis such that $u = e_1$ we immediately see that $\mathrm{tr}(|u\rangle\langle u|) = 1$. Whence $|u\rangle\langle u|$ is a density matrix. The expression $\mathrm{tr}(|u\rangle\langle u|\, A)$ is easily calculated:

$$\mathrm{tr}(|u\rangle\langle u|\, A) = \sum_j \langle e_j, u\rangle\langle u, Ae_j\rangle = \langle u, Au\rangle \tag{5.6}$$

which we recognize as the usual expectation value of A. Furthermore we will call a density operator **pure** or a **pure state** if $\hat{\rho}$ is the orthogonal projection on the span of a unit vector $u \in \mathcal{H}$. If this is not the case then $\hat{\rho}$ is called **mixed** or a **mixed state**. Clearly a density operator is pure if and only if $\mathrm{tr}\big(\hat{\rho}^2\big) = 1$. A generic density operator for a mixed state can be described as follows. Let $\{e_j\}$

[3]This is because $\langle u, \hat{\rho}^{1/2}A\hat{\rho}^{1/2}u\rangle = \langle \hat{\rho}^{1/2}u, A\hat{\rho}^{1/2}u\rangle \geq 0$ by nonnegativity of A.

be an orthonormal basis for $\mathcal{H}$ corresponding to an observable A. Let $w_j \in [0,1]$ such that $\sum_j w_j = 1$. Then

$$\rho = \sum_j w_j \, |e_j\rangle\langle e_j| = \sum_j w_j \langle e_j, \cdot\rangle e_j \tag{5.7}$$

is clearly a density operator. It can be interpreted as the weighted projections on the eigenspaces. For example take the S_z eigenstate $|\uparrow\rangle$. If we take $\hat{\rho} = |\uparrow\rangle\langle\uparrow|$ then $\hat{\rho}$ is pure and the expectation value of the spin observable S_z in z-direction is $\operatorname{tr}(S_z \, |\uparrow\rangle\langle\uparrow|) = \hbar/2$. If $\rho = \frac{1}{2} \, |\uparrow\rangle\langle\uparrow| + \frac{1}{2} \, |\downarrow\rangle\langle\downarrow|$ then ρ is mixed and

$$\operatorname{tr}\left(S_z(\frac{1}{2} \, |\uparrow\rangle\langle\uparrow| + \frac{1}{2} \, |\downarrow\rangle\langle\downarrow|) \right) = \frac{\hbar}{4} \operatorname{tr}(|\uparrow\rangle\langle\uparrow| - |\downarrow\rangle\langle\downarrow|) = 0 \tag{5.8}$$

as expected. Hence we see that the density operator fully describes the state of a quantum system. If A is a bounded operator on $\mathcal{H}$ then the expectation value of A in the state $\hat{\rho}$ is given by $\operatorname{tr}(\hat{\rho}A) = \operatorname{tr}(A\hat{\rho})$. In the case where A is unbounded (but self-adjoint) the expression $\operatorname{tr}(\hat{\rho}A)$ and $\operatorname{tr}(A\hat{\rho})$ need not be equal. It is however possible to construct a measure $\mu^A_{\hat{\rho}}(E) := \operatorname{tr}(\hat{\rho}\chi_E(A))$ and define the expectation value of A to be

$$\mathbb{E}(A) := \int_{\mathbb{R}} \lambda \mathrm{d}\mu^A_{\hat{\rho}}(\lambda) \tag{5.9}$$

This can be made formal in some sense. The following discussion can be retracted in chapter 4.1 in [21]. A **C^*-algebra** X is an associative algebra over $\mathbb{C}$ equipped with an involution, i.e a real linear map $A \mapsto A^*$ for $A, A^* \in X$ such that $A^{**} = A$, $(AB)^* = B^*A^*$ and $(\lambda A)^* = \overline{\lambda}A^*$ for $\lambda \in \mathbb{C}$ and with a norm $\|\cdot\|_X$ such that $\|AB\| \leq \|A\|\|B\|$ and $\|A^*A\| = \|A\|^2$. In this setting a **state** ω on A C^*-algebra X is a complex linear map on X mapping to $\mathbb{C}$ such that $\omega(A^*A) \geq 0$ (**positivity**) and $\|\omega\| = 1$ (**normalization**).

Theorem 5.5. *Let A be an observable and let $\rho \in \mathcal{B}(\mathcal{H})$ be a density operator. Then ρ induces a unique probability measure $\mu_{\hat{\rho}}$ on the spectrum $\sigma(A)$ such that*

$$\operatorname{tr}(\hat{\rho}f(A)) = \int_{\sigma(A)} f \mathrm{d}\mu^A_{\hat{\rho}} \tag{5.10}$$

for $f \in C(\sigma(A))$ and

$$\mu^A_{\hat{\rho}}(E) := \operatorname{tr}(\hat{\rho}\chi_E(A)). \tag{5.11}$$

The expression $\chi_E(A)$ is defined by functional calculus. The measure $\mu^A_{\hat{\rho}}$ is also called **Born measure**.

The time evolution of a density operator depending on a real parameter t is given by the **von Neumann-Heisenberg equation**

$$i\partial_t \hat{\rho}(t) = -[\hat{\rho}, H] \tag{5.12}$$

where H is the Hamiltonian of the system. The **commutator** $[\cdot,\cdot]$ is defined for two operators $A, B \colon \mathcal{H} \to \mathcal{H}$ where A is assumed to be bounded as

$$[A, B] := AB - BA \tag{5.13}$$

such that whenever $u \in \mathfrak{D}(B)$ then $Au \in \mathfrak{D}(B)$ (otherwise the commutator would not be well-defined). This definition is necessary since the Hamiltonian might be unbounded. The von Neumann-Heisenberg equation can be easily derived from the Schrödinger equation. Let $\hat{\rho}(u) = \sum_j w_j \langle \psi_j, u \rangle \psi_j$. Then

$$
\begin{aligned}
i(\partial_t \hat{\rho}) &= \sum_j w_j \partial_t \left(\langle \psi_j, \cdot \rangle \psi_j \right) && (5.14)\\
&= i \sum_j w_j \left(\langle \dot{\psi_j}, \cdot \rangle \psi_j + \langle \psi_j, \cdot \rangle \dot{\psi_j} \right) && (5.15)\\
&= \sum_j w_j \left(\langle H\psi_j, \cdot \rangle \psi_j + \langle \psi_j, \cdot \rangle H\psi_j \right) && (5.16)\\
&= \sum_k w_j \left(\langle H\psi_j, \cdot \rangle \psi_j + \langle \psi_j, H\cdot \rangle \psi_j - \langle \psi_j, H\cdot \rangle \psi_j + \langle \psi_j, \cdot \rangle H\psi_j \right) && (5.17)\\
&= \sum_j w_j \left(-\langle \psi_j, H\cdot \rangle \psi_j + H(\langle \psi_j, \cdot \rangle \psi_j) \right) && (5.18)\\
&= -[\hat{\rho}, H]. && (5.19)
\end{aligned}
$$

The formal solution of the von Neumann-Heisenberg equation is given by

$$\hat{\rho}(t) = e^{-itH} \hat{\rho}_0 e^{itH} \tag{5.20}$$

where $\hat{\rho}_0 := \hat{\rho}(0)$. Alternatively one can view the density operator as an integral operator acting on L^2. The action of $\hat{\rho}$ on a function $u \in L^2$ can then be defined as

$$\hat{\rho}(u) = \int \rho(x, y, t) u(y, t) \mathrm{d}y. \tag{5.21}$$

The operator kernel $\rho(x, y, t)$ is often called **density matrix** (as opposed to the density operator $\hat{\rho}$). In the pure case the density matrix is simply given by

$$\rho(x, y, t) := \psi(x, t) \bar{\psi}(y, t) \tag{5.22}$$

whereas in the mixed case we have

$$\rho(x, y, t) := \sum_j w_j \psi_j(x, t) \bar{\psi}_j(y, t). \tag{5.23}$$

Let us calculate the time derivative of $\rho(x, y, t)$. For simplicity we will consider the pure case.

$$
\begin{aligned}
\partial_t \rho(x, y, t) &= \partial_t (\psi(x, t) \bar{\psi}(y, t)) && (5.24)\\
&= \dot{\psi}(x, t) \bar{\psi}(y, t) + \psi(x, t) \dot{\bar{\psi}}(y, t). && (5.25)
\end{aligned}
$$

Now ψ has to satisfy the Pauli equation whereas $\bar{\psi}$ has to satisfy the equation

$$-i\partial_t \bar{\psi} = H_P \bar{\psi}. \tag{5.26}$$

This yields

$$\partial_t \rho(x, y, t) = (-iH_{P,x}) \psi(x, t) \bar{\psi}(y, t) + \psi(x, t) (iH_{P,y}) \bar{\psi}(y, t) \tag{5.27}$$

The subscript emphasizes that the Hamiltonian is acting on the respective variable (x or y) only. If we multiply the left hand side with i we obtain

$$i\partial_t \rho(x,y,t) = H_{P,x}\rho(x,y,t) - H_{P,y}\rho(x,y,t). \tag{5.28}$$

This is consistent with the von Neumann-Heisenberg equation since the term with the positive sign corresponds to $H_P\hat{\rho}$ for the reason that we integrate with respect to y and so the Hamiltonian with respect to x acts on the density operator. By a similar argument the term with the negative sign corresponds to $\hat{\rho}H_P$. Now recall the Pauli Hamiltonian in its explicit form and subjected to the Coulomb gauge $\nabla \cdot A = 0$:

$$H_{P,i} = -\Delta_i + 2iA\nabla_i + A^2 - \sigma \cdot B \tag{5.29}$$

where $i = x, y$. This equation can be considered as the density matrix analogue of the von Neumann-Heisenberg equation. Then we have explicitly for the density matrix:

$$i\partial_t \rho(x,y,t) = \left(-\Delta_x + 2iA\nabla_x + A^2 - \sigma \cdot B\right)\rho(x,y,t) \tag{5.30}$$

$$+ \left(\Delta_y - 2iA\nabla_y - A^2 + \sigma \cdot B\right)\rho(x,y,t). \tag{5.31}$$

5.2 The Pauli-Poiswell system

When coupling the magnetic Schrödinger equation to spin to obtain the Pauli equation one is left with an important inconsistency. Namely the Pauli equation is a $O(\epsilon)$ approximation of the Dirac equation where $\epsilon := 1/c$. The magnetic field however is actually a relativistic notion and hence when coupling Pauli to an electromagnetic field one has to take into account that it is only consistent if all terms in the system are approximations of the same order. More precisely, the electromagnetic fields are described by Maxwell's equations as introduced above. In the case where no magnetic field is present, i.e. in the Newtonian limit, the electric field is described by a Poisson equation. Therefore we need to find the correct first order in $1/c$ approximation of Maxwell's equations which is where the expression *Poiswell* comes from. It has been proposed by Masmoudi and Mauser in [30]. We will present the results in this section. For convenience and readability we will switch to index notation where we agree on the Einstein summation convention that repeating indices are summed over from 1 to 3. Consider the Dirac equation in a rescaled form as being presented in [1]:

$$i\hbar\partial\Psi = -\frac{i\hbar}{\epsilon}\gamma^0\gamma^k\partial_k\Psi + \frac{1}{\epsilon^2}\gamma^0\Psi - A_k\gamma^0\gamma^k\Psi - \phi\Psi \tag{5.32}$$

$$\Psi(t=0,x) = \Psi_I(x). \tag{5.33}$$

Furthermore we have the position density $n(t,x) := J_0(t,x) = \Psi(t,x)\overline{\Psi}(t,x)$ and the current density $J_k(t,x) = \epsilon^{-1}\gamma^0\gamma^k\Psi(t,x)\overline{\Psi}(t,x)$. The **Dirac-Maxwell system** is then obtained by coupling to the Maxwell equations in the same scaling:

$$\epsilon\partial_t E_k - e_{kij}\partial_i B_j = -\epsilon J_k \qquad \partial_k E_k = n \tag{5.34}$$

$$\epsilon\partial_t B_k + e_{kij}\partial_i E_j = 0 \qquad \partial_k B_k = 0. \tag{5.35}$$

Together with the Lorentz gauge $-\epsilon\partial_t\phi = \partial_k A_k$ we can rewrite Maxwell's equations in the potential formalism

$$-\epsilon^2\partial_{tt}^2\phi + \Delta\phi = -n \tag{5.36}$$
$$-\epsilon^2\partial_{tt}^2 A_k + \Delta A_k = -\epsilon J_k. \tag{5.37}$$

We will now perform a somewhat heuristic argument for the modeling of the Pauli equation coupled to the electromagnetic field. Denote by a superscript the explicit ϵ-dependence of the fields. Then we formally have

$$E_k^\epsilon = E_k^0 + \epsilon E_k^1 + O(\epsilon^2) \qquad B_k^\epsilon = B_k^0 + \epsilon B_k^1 + O(\epsilon^2) \tag{5.38}$$
$$\phi^\epsilon = \phi^0 + \epsilon\phi^1 + O(\epsilon^2) \qquad A_k^\epsilon = A_k^0 + \epsilon A_k^1 + O(\epsilon^2). \tag{5.39}$$

Plugging these equations into the Dirac-Maxwell system and comparing powers we obtain the following conditions for the leading terms of order $O(1)$:

$$e_{kij}\partial_i E_j^0 = 0 \qquad \partial_k E_k^0 = n_k \tag{5.40}$$
$$e_{kij}\partial_i B_j^0 = 0 \qquad \partial_k B_k^0 = n_k \tag{5.41}$$
$$E_k^0 = -\partial_k\phi^0 \tag{5.42}$$

from which we deduce $B_k^0 = 0$ which is consistent since the magnetic field is a relativistic effect and hence the $O(1)$-term should vanish. Moreover $\Delta\phi^0 = -n$. For terms of order $O(\epsilon)$ we obtain

$$\partial_t E_k^0 - e_{kij}\partial_i B_j^1 = -J_k \qquad \partial_k B_k^1 = 0 \tag{5.43}$$
$$e_{kij}\partial_i E_j^1 = 0 \qquad \partial_k E_k^1 = 0 \tag{5.44}$$
$$-\partial_t\phi^0 = \partial_k A_k^1 \tag{5.45}$$

from which we immediately deduce $E_k^1 = 0$. Moreover by the formula $B_k^1 = e_{kij}\partial_i A_j$ and $e_{kij}e_{jlm}\partial_i\partial_l A_m^1 = \partial_k\partial_i A_i^1 - \partial_{jj}^2 A_k^1$ one deduces that $\partial_{jj}^2 A_k^1 = -J_k$. Finally we obtain the **Poiswell system** as the $O(\epsilon)$ approximation of Maxwell's equation in the scaling above:

$$\Delta\phi^\epsilon = -n \tag{5.46}$$
$$\Delta A^\epsilon = -\epsilon J. \tag{5.47}$$

Hence we eventually arrive at the **Pauli-Poiswell system** if we couple the scaled Pauli equation to the scaled Poiswell system:

$$i\hbar\partial_t u^{\epsilon,\hbar} = \frac{1}{2}(i\hbar\nabla - \epsilon A^{\epsilon,\hbar})u^{\epsilon,\hbar} - \phi^\epsilon u^{\epsilon,\hbar} - \frac{\epsilon\hbar}{2}\sigma^k e_{kij}\partial_i A_j^{\epsilon,\hbar} u^{\epsilon,\hbar} \tag{5.48}$$
$$u^{\epsilon,\hbar}(t=0,x) = u_I^{\epsilon,\hbar}(x) \tag{5.49}$$
$$n^{\epsilon,\hbar}(t,x) = |u^{\epsilon,\hbar}(t,x)|^2 \tag{5.50}$$
$$J_k^{\epsilon,\hbar} = \mathfrak{Im}\left(\overline{u^{\epsilon,\hbar}}\cdot(\hbar\partial_k + i\epsilon A_k)u^{\epsilon,\hbar}\right) + e_{kij}\partial_i\left(\left(\overline{u^{\epsilon,\hbar}}\cdot\sigma\right)u^{\epsilon,\hbar}\right)_j \tag{5.51}$$
$$\Delta\phi^{\epsilon,\hbar}(t,x) = -n^{\epsilon,\hbar}(t,x) \tag{5.52}$$
$$\Delta A^{\epsilon,\hbar}(t,x) = -\epsilon J^{\epsilon,\hbar}(t,x). \tag{5.53}$$

Of course we can rewrite this a little. Using the definition of n we can rewrite $\Delta\phi^{\epsilon,\hbar} = -n^{\epsilon,\hbar}$ as

$$\Delta\phi^{\epsilon,\hbar}(t,x) = -|u^{\epsilon,\hbar}|^2. \tag{5.54}$$

Moreover the equation for A can be rewritten as

$$\Delta A_k^{\epsilon,\hbar} = -\epsilon^2 A_k |u^{\epsilon,\hbar}|^2 - \epsilon\hbar\mathfrak{Im}\left(\overline{u^{\epsilon\hbar}}\partial_k u^{\epsilon\hbar}\right) + e_{kij}\partial_i\left(\left(\overline{u^{\epsilon,\hbar}}\cdot\sigma\right)u^{\epsilon,\hbar}\right)_j \tag{5.55}$$

and therefore removing the equations for J and n. The Pauli-Poiswell system has the interesting property to have a much simpler structure than the Dirac-Maxwell system by avoiding the notion of antiparticles which is not very useful for the range of applications one usually considers. On the contrary it is a first order correction to the Schrödinger-Poisson system and hence one hopes that one can improve experimental accuracy by using Pauli-Poiswell. One could ask whether higher order corrections would be helpful but this is unfortunately not the case since Maxwell's equations depend singularly on ϵ in the transition from the hyperbolic wave equations (5.37) to the elliptic Poisson-like equations above.

The Pauli-Poiswell system is not well-understood until now. Questions like well-posedness, dependence on the data and other analytical questions are almost all open. On can hope to use results from the Pauli operator but obviously the system is much more delicate than the Pauli operator alone. Numerically we can make some suggestions on how to proceed with the Pauli-Poiswell system. As mentioned earlier the Pauli equation alone can be solved by a four term operator splitting approach. This has been done for example by Gutleb in [13]. Since the Pauli-Poiswell system is coupled we need a more sophisticated approach for solving it. The usual way is to solve the potential equation for the electric potential ϕ, i.e 5.52 for initial density n^0 which can be immediately calculated from the inital condition u_I. Additionally the initial current density J^0 is needed which can only partly be derived from the initial condition u_I as it also depends on the vector potential A. Hence we have to provide an initial vector potential A^0 as well. With the initial current density J^0 and the initial particle density n^0 at hand we can solve the original Pauli equation (e.g. by the established splitting approach) and reiterate this procedure. The solvers for the potentials shouldn't be too difficult since they are described a linear Poisson equation. This question will be adressed in future works.

A Functional Analysis

A.1 Self-adjoint operators

Let $\mathcal{H}$ be a Hilbert space and A be a densely defined linear operator with domain $\mathfrak{D}(A)$. Then A is called **symmetric** if

$$\langle \varphi, A\psi \rangle = \langle A\varphi, \psi \rangle, \quad \psi, \phi \in \mathfrak{D}(A). \tag{A.1}$$

The **adjoint operator** A^* of a densely defined linear operator A is then defined by the domain

$$\mathfrak{D}(A^*) := \{\psi \in \mathcal{H} | \exists \psi' \in \mathcal{H} \,:\, \langle \psi, A\varphi \rangle = \langle \psi', \varphi \rangle, \forall \varphi \in \mathfrak{D}(A)\} \tag{A.2}$$

and the action

$$A^*\psi = \psi'. \tag{A.3}$$

The notion of self-adjointness for unbounded operators is cleary much more subtle than in the bounded case since the question of the domain is omnipresent. It might happen that $\mathfrak{D}(A^*)$ isn't dense or that even $\mathfrak{D}(A^*) = 0$. We will call A **self-adjoint** if $A = A^*$ and $\mathfrak{D}(A) = \mathfrak{D}(A^*)$.

Examples A.1.

- As an important example take the **multiplication operator** M_A on $L^2(\mathbb{R}^d, d\mu)$. It is defined by

$$\begin{cases} (M_A u)(x) := A(x)u(x) \\ \mathfrak{D}(A) := \left\{u \in L^2(\mathbb{R}^d, d\mu) \colon M_A u \in L^2(\mathbb{R}^d, d\mu)\right\} \end{cases} \tag{A.4}$$

  where $A \colon \mathbb{R}^d \to \mathbb{C}$ is a measurable function. Then A is self-adjoint. For a proof see [44].

- Define A as

$$Au := -i\partial_x u, \quad \mathfrak{D}(A) := H^2(a,b) \cap \{u | u(a) = u(b)\} \tag{A.5}$$

  Then A is self-adjoint.

We say that an operator is **essentially self-adjoint** if A is symmetric and its closure is self-adjoint (Symmetric operators are always closable). Moreover, if A is essentially self-adjoint then the closure $\overline{A}$ is the unique self-adjoint extension of A.

Theorem A.2. *Let A be a self adjoint operator on a Hilbert space $\mathcal{H}$. Then its spectrum $\sigma(A)$ is contained in the real line $\mathbb{R}$. The converse also holds (if A is not self-adjoint then $\sigma(A) \backslash \mathbb{R} \neq \emptyset$).*

Proof. □

The next theorem is due to Kato and Rellich. It ensures self-adjointness of an operator of the form $A + B$ where A is known to be self-adjoint and B is somewhat controlled by A. Before we need to define **relative boundedness** of an operator B:

Definition A.3. *Let A,B be operators. B is called* A**-bounded** *or* **relatively bounded** *with respect to A if* $\mathfrak{D}(A) \subset \mathfrak{D}(B)$ *and there exist constants* α, β *such that*

$$\|Bu\| \leq \alpha \|Au\| + \beta \ \|u\|, \quad u \in \mathfrak{D}(A). \tag{A.6}$$

Now the **Kato-Rellich theorem** goes as follows.

Theorem A.4. (Kato-Rellich) *Let A be (essentially) self-adjoint. Let B be symmetric and relatively bounded with respect to A and with relative bound* < 1*. Then* $A + B$ *with* $\mathfrak{D}(A + B) = \mathfrak{D}(A)$ *is (essentially) self-adjoint.*

Proof. □

Another important theorem is **Weyl's theorem**. It states that the essential spectrum is stable under compact perturbations.

Definition A.5. *[44], p.128. Let A be an operator with resolvent* $R_A(z)$*,* $z \in \rho(A)$*. The operator K is called* **relatively compact** *with respect to A or* A***-compact*** *if*

$$K R_A(z) \in \mathfrak{C}(\mathcal{H}) \tag{A.7}$$

where $\mathfrak{C}(\mathcal{H})$ *denotes the set of compact operators on* $\mathcal{H}$*. This only needs to be true for one* $z \in \rho(A)$ *since it follows for all* $z \in \rho(A)$ *by the first resolvent identity.*

Theorem A.6. (Weyl) *Let A be self-adjoint and K be symmetric. If K is* $A-$*compact then* $A+K$ *is self-adjoint with domain* $\mathfrak{D}(A)$*. Moreover,* $\sigma_{\text{ess}}(A + K) = \sigma_{\text{ess}}(A)$*.*

Proof. The first part follows from the fact that an operator which is relatively compact with respect to a self-adjoint operator A is already relatively bounded with respect to with relative bound zero. The detailed reason can be found in Lemma 6.22 in [44]. The second part is actually more general since the claim holds for any two self adjoint operators A and B for which compactness of $R_A(z) - R_b(z) \in \mathfrak{C}$ for some $z \in \rho(A) \cap \rho(B)$ holds. In particular this is the case for A and $B := A + K$ where K is relatively compact. This is proven in Theorem 6.19 in [44]. □

An important criterion for a point $\lambda \in \mathbb{C}$ to be in the essential spectrum is **Weyl's criterion**.

Theorem A.7. (Weyl's criterion) *Let* $\lambda \in \mathbb{C}$ *and A be self-adjoint. Then* $\lambda \in \sigma_{\text{ess}}(A)$ *iff there is a normalized sequence* $\{u_n\}$ *converging weakly to zero such that* $\|(A - \lambda)u_n\| \to 0$*. The sequence can be chosen orthonormal.*

Proof. See Lemma 6.17 in [44]. □

Next we look at quadratic forms and how they can be used to show self-adjointness. Recall that an operator is called **nonnegative** if $\langle u, Au \rangle \geq 0$ for all $u \neq 0$ in $\mathfrak{D}(A)$. On $\mathfrak{D}(A)$ define the scalar product

$$\langle u, v \rangle_A := \langle u, (A + 1)v \rangle. \tag{A.8}$$

Define $\mathcal{H}_A$ to be the completion of $\mathfrak{D}(A)$ with respect to $\langle\cdot,\cdot\rangle_A$. Then $\mathfrak{D}(A) \subset \mathcal{H}_A \subset \mathcal{H}$. Define the quadratic form on $\mathcal{H}_A$ to be

$$q_A(u) := \langle u, u\rangle_A - \|u\|^2 \tag{A.9}$$

where $\mathfrak{Q}(A) := \mathcal{H}_A$ is called the **form domain** of A. A symmetric operator A is called **bounded from below** or **semi-bounded** if

$$q_A(u) = \langle u, Au\rangle \geq \gamma\|u\|^2 \tag{A.10}$$

for $\gamma \in \mathbb{R}$. The following theorem gives the remarkable result that a symmetric operator which is bounded from below has a unique self-adjoint extension which is sometimes called **Friedrich extension**.

Theorem A.8. (Friedrich extension). *Let A be a symmetric operator. If A is bounded from below by γ then A has a unique self-adjoint extension $\tilde{A}$ which is bounded from below by γ and for which $\mathfrak{D}(\tilde{A}) \subset \mathcal{H}_{A-\gamma}$. It is also the only self-adjoint extension for which this holds.*

Proof. See Theorem 2.12 in [44]. □

Now let $s\colon \mathfrak{Q} \times \mathfrak{Q} \to \mathbb{C}$ be a sesquilinear form and let q be the associated quadratic form, i.e. $q(u) := s(u, u)$. q is called **hermitian** if it is real-valued. A hermitian form is called, as before, **semi-bounded** if $q(u) \geq \gamma\|u\|^2$. The norm associated with (semi-bounded) q is defined by $\|u\|_q := q(u) + (1-\gamma)\|u\|^2$. The completion of $\mathfrak{Q}$ with respect to this norm is denoted by $\mathcal{H}_q$. q is called **closable** if every for every Cauchy sequence which converges to zero in the $\|\cdot\|$-norm it follows that it also converges to zero in the $\|\cdot\|_q$-norm. Then $\mathcal{H}_q \subset \mathcal{H}$ and the extension of q to $\mathcal{H}_q$ is called the **closure** of q. Moreover if $\mathfrak{Q} = \mathcal{H}_q$ then q is **closed**. We then have the important theorem that ensures that for a given quadratic form we obtain a self-adjoint operator:

Theorem A.9. *Let q be a closed semi-bounded quadratic form. Then there is a unique self-adjoint operator A such that $\mathfrak{Q}$ is the form domain $\mathfrak{Q}_A$ of A and $q = q_A$. Explicitly A is given by*

$$\mathfrak{D}(A) := \{u \in \mathcal{H}_q | \exists \tilde{u} \in \mathcal{H}\colon s(w, u) = \langle w, \tilde{u}\rangle \forall w \in \mathcal{H}_q\} \tag{A.11}$$

$$Au := \tilde{u} - (1-\gamma)u \tag{A.12}$$

where s is the sesquilinear form corresponding to q.

Proof. Theorem 2.13 in [44]. □

After having seen that every closed semi-bounded form corresponds to a self-adjoint operator let us look at perturbations of forms. First we need relative form boundedness. Let $A \geq \gamma$ be a self-adjoint operator which is bounded from below. Let $q\colon \mathfrak{Q} \to \mathbb{R}$ be a hermitian form such that the form domain of A is contained in $\mathfrak{Q}$, i.e. $\mathfrak{Q}(A) \subseteq \mathfrak{Q}$. Then q is called **relatively form bounded** with respect to q_A if there are constants $a, b \geq 0$ such that

$$|q(u)| \leq a q_{A-\gamma}(u) + b\|u\|^2 \tag{A.13}$$

for all $u \in \mathfrak{Q}(A)$. The infimum over all a then called the **form bound** of q with respect to q_A. The most important application of relative form boundedness is the form analogue of the Kato-Rellich theorem and is due to Kato, Lions, Lax, Milgram and Nelson whence it is called **KLMN theorem**

Theorem A.10. [KLMN]. *Let $q_A\colon \mathfrak{Q}(A) \to \mathbb{R}$ be a closed semi-bounded hermitian form and q relatively bounded with respect to q_A with $a < 1$. Then $q_A + q$ is closed with $\mathfrak{Q} = \mathfrak{Q}(A)$. In particular there exists a semi-bounded self-adjoint operator corresponding to $q + q_A$.*

Proof. See Theorem 6.24 in [44]. □

The next result on compactness is needed the proof of Theorem 4.3. For its proof we refer to [35].

Theorem A.11. *Let X, Y be Banach function spaces and $A \in \mathfrak{C}(X, Y)$ such that A is positivity preserving (see [36]). Then $B\colon X \to Y$ is compact if*

$$|Bx(t)| \leq A|x|(t). \tag{A.14}$$

A.2 Lebesgue and Sobolev spaces

A.2.1 Lebesgue spaces

This section is based on section 0.6. in [44]. Let (X, E, μ) be some σ-finite measure space and let $\mathcal{L}^p(X, \mathrm{d}\mu)$ where $1 \leq p$ be the set of all complex-valued measurable functions such that

$$\|f\|_p := \left(\int_X |f|^p \mathrm{d}\mu \right)^{\frac{1}{p}} < \infty. \tag{A.15}$$

This space is linear. In order to turn it into a Banach space denote by

$$\mathcal{N}(X, \mathrm{d}\mu) := \{f | f(x) = 0 \ \mu - \text{almost everywhere}\}. \tag{A.16}$$

the set of functions vanishing almost everywhere. $\mathcal{N}(X, \mathrm{d}\mu)$ is a linear subspace of $\mathcal{L}^p(X, \mathrm{d}\mu)$. The **Lebesgue L^p space** is then defined as

$$L^p(X, \mathrm{d}\mu) := \mathcal{L}^p(X, \mathrm{d}\mu)/\mathcal{N}(X, \mathrm{d}\mu). \tag{A.17}$$

with norm $\|\cdot\|_p$ defined as above. Hence the elements of L^p are in fact equivalence classes of functions. Let

$$\|f\|_\infty := \inf\{C | \mu(\{c | |f(x)| > C\}) = 0\} \tag{A.18}$$

be the infimum over the constants for which the measure of the set of points x for which $f(x)$ is greater than C is zero, i.e. the **essential supremum**. Then the space $L^\infty(X, \mathrm{d}\mu)$ of all measurable functions having an essential supremum is a Banach space. The next theorem is frequently used. Note that two indices p, q are called **dual** if $1/p + 1/q = 1$.

Theorem A.12. (Hölder's inequality) *Let $p, q \in [1, \infty]$ be dual indices. If $f \in L^p(X, \mathrm{d}\mu)$ and $g \in L^q(X, \mathrm{d}\mu)$ then fg is in $L^1(X, \mathrm{d}\mu)$ with the estimate*

$$\|fg\|_1 \leq \|f\|_p \|g\|_q. \tag{A.19}$$

Proof. Without loss of generality assume that f and g are normalized, i.e. $\|f\|_p = \|g\|_q = 1$. Recall that for arbitrary numberts $a, b \geq 0$ we have

$$a^{\frac{1}{p}} b^{\frac{1}{q}} \leq \frac{a}{p} + \frac{b}{q}. \tag{A.20}$$

Now set $|f|^p := a$ and $|g|^q := b$. Then integrate over X to obtain

$$\int |fg| \leq \frac{1}{p} \int |f|^p + \frac{1}{q} \int |g|^q = 1. \tag{A.21}$$

□

Alternatively L^p can be viewed as the completion of the smooth functions with compact support on X with respect to the L^p-norm if X is a subset of $\mathbb{R}^n$.

Theorem A.13. *Let $X \subset \mathbb{R}^n$ and let μ be a regular Borel measure. Then the set $C_c^\infty(X)$ of all smooth functions with compact support is dense in $L^p(X, \mathrm{d}\mu)$ for $p \in [1, \infty)$.*

Proof. See Theorem 0.36. in [44]. □

Note that this fails in the case $p = \infty$ since the uniform limit of continuous functions is again continuous whereas a function in $L^\infty(X, \mathrm{d}\mu)$ need not be continuous.

A.2.2 Sobolev spaces

The **Sobolev space** $W^{k,p}(U)$ where U is some nonempty open set in $\mathbb{R}^n$ is the set of all functions in $L^p(U)$ which have weak derivatives up to order k in $L^p(U)$. The norm on $W^{k,p}(U)$ is defined as

$$\|f\|_{k,p} := \begin{cases} \left(\sum\limits_{|\alpha| \leq k} \|\partial_\alpha f\|_p^p \right)^{\frac{1}{p}}, & 1 \leq p < \infty \\ \max\limits_{|\alpha| \leq k} \|\partial f\|_\infty & p = \infty. \end{cases} \tag{A.22}$$

In the case where $p = 2$, $W^{k,2}(U)$ is a Hilbert space and denoted by $H^k(U)$. Moreover, denote by $W_0^{1,p}(U)$ the closure of $C_c^\infty(U)$ (i.e. the C^∞-functions with compact support in U) in the norm of $W^{1,p}(U)$.

In the proof of 4.5 we need the following theorem which can be found in [19] as Theorem 2.2. It is also known as **Gagliardo-Nirenberg interpolation inequality**.

Theorem A.14. *Let $u \in W_0^{1,m}(U)$. For any $r \geq 1$ the inequality*

$$\|u\|_q \leq C \|\nabla u\|_m^{1-\beta} \|u\|_r^\beta \tag{A.23}$$

holds with

$$1 - \beta = \left(\frac{1}{r} - \frac{1}{q} \right) \left(\frac{1}{n} - \frac{1}{m} + \frac{1}{r} \right)^{-1}. \tag{A.24}$$

The range of q is $[r, nm(n-m)]$ in the case $n > 1$, $n > m$ and $r \leq nm/(n-m)$. For the other cases see Theorem 2.2 in [19].

For 4.5 we have the situation $n = 3$, $m = r = 2$, and thus $\beta = 3/q - 1/2$. Moreover $q \in [2, 6]$.

B Technical lemmata

B.1 Hardy's inequality

Let $u \in C_0^\infty(\mathbb{R}^d)$ and $d \geq 3$. Then

$$\int \frac{|u(x)|^2}{|x|^2} \leq \frac{4}{(d-2)^2} \int \sum_{j=1}^{d} |\partial_j u(x)|^2 \mathrm{d}x. \tag{B.1}$$

Proof. See Theorem 17.1 in [46]. □

B.2 Diamagnetic inequality

The diamagnetic inequality is associated with Kato's inequality. In fact it can be considered as the semigroup version of Kato's inequailty. Consider the operator $H = (-i\nabla - A)^2 + \phi$ with domain $\mathfrak{D}(H)$. Denote $-\Delta$ by H_0.

Theorem B.1. (Diamagnetic inequality) *Let $f \in L^2(\mathbb{R}^d)$ and H as above. Then*

$$\left|e^{-tH} f\right| \leq e^{-tH_0}|f|, \quad t \in \mathbb{R}. \tag{B.2}$$

Proof. By Trotter's product formula (see [18]) we have

$$e^{-tH} = \lim_{N\to\infty} \left[e^{\frac{t}{N}D_1^2} \cdots e^{\frac{t}{N}D_d^2} e^{-\frac{t}{N}\phi}\right]^N \tag{B.3}$$

where $D_j := \partial_j - iA_j$. Define

$$\lambda_j(x) := \int_0^{x_j} A(x_1, ..., x_{j-1}, y, x_{j+1}, ..., x_d)\mathrm{d}y \tag{B.4}$$

for $A \in L^2_{\mathrm{loc}}(\mathbb{R}^d)$. Then by Lemma 2.5 in [40] we have unitary equivalence of $-iD_j$ and $-i\partial_j$ by multiplication with $\exp\{i\lambda_j\}$:

$$-iD_j = e^{i\lambda_j}(-i\partial_j)e^{-i\lambda_j}. \tag{B.5}$$

Whence

$$e^{\frac{t}{N}D_j^2} = e^{i\lambda_j} e^{\frac{t}{N}\partial_j} e^{-i\lambda_j}. \tag{B.6}$$

Together with the resulting estimate $|\exp\big(tD_j^2\big)f| \leq \exp(t\partial_j)|f|$ for $f \in L^2(\mathbb{R}^d)$ the claim follows from the Trotter formula (B.3) and from the fact that $|\exp(-tV/N)| \leq 1$. □

B.3 Traces of Powers

This theorem is needed in the proof of the Lieb-Thirring bound. Its proof can be found in Theorem 4.5 in [25].

Theorem B.2. (Traces of Powers) *Let A,B be positive self-adjoint operators. Then for all $m \geq 1$*

$$\mathrm{tr}(B^{\frac{1}{2}} A B^{\frac{1}{2}}) \leq \mathrm{tr}(B^{\frac{m}{2}} A^m B^{\frac{m}{2}}). \tag{B.7}$$

C Asymptotic analysis

For an overview on asymptotic analysis see [34].

Definition C.1. *Let f and g be some complex-valued functions defined on some set U in $\mathbb{R}^n$. Then one writes*

$$f(x) = O(g(x)), \quad x \in U \tag{C.1}$$

if there is a constant C such that

$$|f(x)| \leq C|g(x)| \tag{C.2}$$

for all $x \in U$.

A stronger version of the above is the following

Definition C.2. *Let f and g be as above and let $x_0 \in \overline{D}$. Then one writes*

$$f(x) = o(g(x)) \tag{C.3}$$

as $x \to x_0$ for $x \in D$ if for any $\epsilon > 0$ there is a $\delta > 0$ such that

$$|f(x)| \leq \epsilon|g(x)| \tag{C.4}$$

whenever $|x - x_0| < \delta$. In other words, $f(x) = o(g(x))$ if

$$\lim_{\epsilon \to 0} \frac{|f(\epsilon)|}{|g(\epsilon)|} = 0. \tag{C.5}$$

The next thing we are going to need are asymptotic expansions.

Definition C.3. *Let X be a normed space and let $u_\epsilon, v_\epsilon \in X$ be elements depending on ϵ. Then u_ϵ and v_ϵ are called* **asymptotically equivalent** *if*

$$\|u_\epsilon - v_\epsilon\|_X = o(\|u_\epsilon\|) \tag{C.6}$$

as $\epsilon \to 0$. Note that this definition is symmetric with respect to u_ϵ and v_ϵ. One also calls v_ϵ an **asymptotic approximation** *for u_ϵ. Now let $\{u_k\}$ for $k \in \mathbb{N}_0$ be a sequence. Then the formal sum*

$$\sum_{k=0}^{\infty} u_k \epsilon^k \tag{C.7}$$

is called **asymptotic expansion** *(of order N) if*

$$\|u_\epsilon - \sum_{k=0}^{n} u_k \epsilon^k\| = o(\epsilon^n) \tag{C.8}$$

for all n (up to N).

There is a more or less informal classification of perturbed problems. Namely we call problems **regularly perturbed** if the perturbed problem for $\epsilon \neq 0$ is *of the same type*, i.e if for example a second order homogeneous ODE remains a second order homogeneous ODE after perturbation. All other problems are called **singularly perturbed**.

References

[1] Bechouche, Philippe, M. N. J., and Poupaud, F. (Semi)-nonrelativistic limits of the Dirac equation with external time-dependent electromagnetic field. *Communications in mathematical physics 197*, 2 (1998), 405–425.

[2] Cwikel, M. Weak type estimates for singular values and the number of bound states of Schrödinger operators. *Annals of Mathematics* (1977), 93–100.

[3] Cycon, H. L., Froese, R. G., Kirsch, W., and Simon, B. *Schrödinger operators: With application to quantum mechanics and global geometry*. Springer, 2009.

[4] Dirac, P. A. M. *The principles of quantum mechanics*. No. 27. Oxford university press, 1981.

[5] Einstein, A. Zur elektrodynamik bewegter körper. In *Das Relativitätsprinzip*. Springer, 1923, pp. 26–50.

[6] Erdős, L. Recent developments in quantum mechanics with magnetic fields. *arXiv preprint math-ph/0510055* (2005).

[7] Erdős, L. Lieb-Thirring inequalities: lecture notes. *Mathematisches Institut der LMU München* (2013).

[8] Erdős, L. Magnetic Lieb-Thirring inequalities. *Communications in mathematical physics 170*, 3 (1995), 629–668.

[9] Erdős, L., and Solovej, J. P. Uniform Lieb-Thirring inequality for the three-dimensional Pauli operator with a strong non-homogeneous magnetic field. In *Annales Henri Poincaré* (2004), vol. 5, Springer, pp. 671–741.

[10] Feynman, R. P. Space-time approach to non-relativistic quantum mechanics. In *Feynman's Thesis—A New Approach To Quantum Theory*. World Scientific, 2005, pp. 71–109.

[11] Foldy, L. L., and Wouthuysen, S. A. On the Dirac theory of spin-1/2-particles and its non-relativistic limit. *Physical Review 78*, 1 (1950), 29.

[12] Griffiths, D. *Introduction to elementary particles*. John Wiley & Sons, 2008.

[13] Gutleb, T. An operator splitting approach for the Pauli equation. Master thesis, University of Vienna, 2018.

[14] Hall, B. C. *Quantum theory for mathematicians*, vol. 267. Springer, 2013.

[15] Hörmander, L. *Linear partial differential operators*, vol. 116. Springer, 2013.

[16] Hunziker, W. On the nonrelativistic limit of the Dirac theory. *Communications in Mathematical Physics 40*, 3 (1975), 215–222.

[17] Itzykson, C., and Zuber, J.-B. *Quantum field theory*. Courier Corporation, 2012.

[18] Kato, T., and Masuda, K. Trotter's product formula for nonlinear semigroups generated by the subdifferentials of convex functionals. *Journal of the Mathematical Society of Japan 30*, 1 (1978), 169–178.

[19] Ladyzhenskaia, O. A., Solonnikov, V. A., and Ural'ceva, N. N. *Linear and quasi-linear equations of parabolic type*, vol. 23. American Mathematical Soc., 1968.

[20] Landau, L. D., Pitaevskii, L. P., Lifshitz, E. M., and Berestetskii, V. B. Relativistic quantum theory.

[21] Landsman, K. *Foundations of Quantum Theory*, vol. 188. 01 2017.

[22] Lieb, E. H. The number of bound states of one-body Schrödinger operators and the Weyl problem. In *The Stability of Matter: From Atoms to Stars*. Springer, 1997, pp. 241–252.

[23] Lieb, E. H., and Loss, M. Analysis, volume 14 of graduate studies in mathematics. *American Mathematical Society, Providence, RI, 4* (2001).

[24] Lieb, E. H., Loss, M., and Solovej, J. P. Stability of matter in magnetic fields. In *The Stability of Matter: From Atoms to Stars*. Springer, 1997, pp. 432–436.

[25] Lieb, E. H., and Seiringer, R. *The stability of matter in quantum mechanics*. Cambridge University Press, 2010.

[26] Lieb, E. H., Solovej, J. P., and Yngvason, J. Asymptotics of heavy atoms in high magnetic fields: I. Lowest Landau band regions. *Communications on Pure and Applied Mathematics 47*, 4 (1994), 513–591.

[27] Lieb, E. H., Solovej, J. P., and Yngvason, J. Asymptotics of heavy atoms in high magnetic fields: Ii. Semiclassical regions. *Communications in mathematical physics 161*, 1 (1994), 77–124.

[28] Lieb, E. H., Solovej, J. P., and Yngvason, J. Ground states of large quantum dots in magnetic fields. In *The Stability of Matter: From Atoms to Stars*. Springer, 2001, pp. 171–190.

[29] Lieb, E. H., and Thirring, W. E. Inequalities for the moments of the eigenvalues of the Schrödinger Hamiltonian and their relation to Sobolev inequalities. In *The Stability of Matter: From Atoms to Stars*. Springer, 1991, pp. 135–169.

[30] Masmoudi, N., and Mauser, N. J. The Selfconsistent Pauli Equation. *Monatshefte für Mathematik 132*, 1 (Apr 2001), 19–24.

[31] Mauser, N. J. Rigorous derivation of the Pauli equation with time-dependent electromagnetic field. *VLSI Design 9*, 4 (1999), 415–426.

[32] Mauser, N. J. Semi-relativistic approximations of the dirac equation: first and second order corrections. *Transport Theory and Statistical Physics 29*, 3-5 (2000), 449–464.

[33] Mauser, N. J., and Selberg, S. Convergence of the dirac–maxwell system to the vlasov–poisson system. *Communications in Partial Differential Equations 32*, 3 (2007), 503–524.

[34] Miller, P. D. *Applied asymptotic analysis*, vol. 75. American Mathematical Soc., 2006.

[35] Pitt, L. D. A compactness condition for linear operators on function spaces. *Journal of Operator Theory* (1979), 49–54.

[36] Reed, M., and Simon, B. *II: Fourier Analysis, Self-Adjointness*, vol. 2. Elsevier, 1975.

[37] Rosenbljum, G. Distribution of the discrete spectrum of singular differential operators. *Soviet Math. Doki. 202* (1972), 1012–1015.

[38] Sakurai, J. J., and Commins, E. D. Modern quantum mechanics, revised edition, 1995.

[39] Schwabl, F. *Advanced quantum mechanics*. Springer Science & Business Media, 2005.

[40] Simon, B. Maximal and minimal Schrödinger forms. *Journal of Operator Theory* (1979), 37–47.

[41] Sobolev, A. V. On the Lieb-Thirring estimates for the Pauli operator. *Duke Mathematical Journal 82*, 3 (1996), 607–636.

[42] Sobolev, A. V. Lieb-Thirring inequalities for the Pauli operator in three dimensions. In *Quasiclassical Methods*. Springer, 1997, pp. 155–188.

[43] Stueckelberg, E. La signification du temps propre en mécanique ondulatoire. *Helv. Phys. Acta 14* (1941), 321–322.

[44] Teschl, G. Mathematical methods in quantum mechanics. *Graduate Studies in Mathematics 99* (2009).

[45] Thaller, B. *The Dirac equation*. Springer Science & Business Media, 2013.

[46] Weidmann, J. *Linear operators in Hilbert spaces*, vol. 68. Springer Science & Business Media, 2012.

[47] White, G. B. Splitting of the Dirac operator in the nonrelativistic limit. *Ann. Inst. Henri Poincaré 53*, 1 (1990), 109–121.

Nomenclature

χ_U	characteristic function of the set U
γ^μ	μth Dirac matrix
$\mathcal{H}$	Hilbert space
$\mathfrak{C}$	set of compact operators
$\mathfrak{D}(A)$	domain of the operator A
$\mathfrak{M}_\gamma$	sum of the moments of the eigenvalues of H
$\mathfrak{Q}$	form domain of a quadratic form
∇	Nabla operator
ϕ	scalar potential
Ψ	Dirac spinor
$\rho(A)$	resolvent set of A
$\sigma(A)$	spectrum of A
$\sigma \cdot B$	stands for $\sum_{j=1}^{3} \sigma_j B_j$
σ_i	ith Pauli matrix
σ_{ess}	essential spectrum of A
A	vector potential
B	magnetic field
$C^\infty(U)$	class of infinitely differentiable functions on the set U
$C_0^\infty(U)$	class of C^∞-functions vanishing at the boundary
$C_c^\infty(U)$	C^∞-functions having compact support in U
E	electric field
e_{kij}	Levi-Civita symbol
$H^2(U)$	the Sobolev space $W^{2,2}(U)$ on the set U
H_0	free Schrödinger Hamiltonian
H_P	Pauli operator
H_P^0	Pauli operator with $\phi = 0$
H_{ms}	magnetic Schrödinger operator
J_k	components of the current density
$L^p(U)$	space of Lebesgue p integrable functions on U

M_f multiplication operator with respect to $f(x)$.

$R_A(z)$ resolvent of A

$W_0^{1,p}(U)$ closure of $C_c^\infty(U)$ in $W^{1,p}(U)$