The Aharonov-Bohm effect

Proseminar on Algebra, Topology & Group theory in physics
ETH Zürich

Simon Wächter

March 30, 2018

Contents

1 Introduction

The aim of this report is to study the influence of the *gauge potentials* of electrodynamics in non-relativistic *quantum mechanics*. In contrast to *classical mechanics*, where the equation(s) of motion contain only the electric and magnetic field, in quantum mechanics the Schrödinger equation explicitly contains the electromagnetic potentials. This fact was known since the beginning of quantum mechanics, but it wasn't until the publishing of the paper *Significance of Electromagnetic Potentials in the Quantum Theory* [7] by Y. Aharonov and D. Bohm in 1959, that the consequences of this fact received serious attention.

This report is divided in two bigger parts. In section 2 the Schrödinger equation of a charged particle in an electromagnetic field is recapitulated. In particular the behavior of an eigenfunction to the Schrödigner equation under gauge transformations is considered. We will find that in order to keep physical observables independent of gauge the wave functions must change with the gauge used. This will be illustrated in the example of a charged particle in a constant magnetic field. In the sections 3, 4 & 5 the "physicality" of the gauge potentials is exemplarily presented in the case of a solenoid with a magnetic field restricted to the interior of the solenoid. In particular in section 3 we look at effects of the gauge potentials on the energy spectrum, in section 4 we look at the Aharonov-Bohm effect, which occurs e.g. in a modified double slit experiment and in section 5 we look at the connection to topology.

This text is part of the proseminar *Algebra, Topology and Group theory in physics* organized by Prof. Matthias Gaberdiel, taking place in the spring semester 2018 at ETH Zürich. It is the report associated with a talk held by myself in March 2018. Most topics of this text are based on chapters 1.4 and 1.5 of David Tong's lecture script on the Quantum Hall effect [1].

2 Schrödinger equation for charged particle

In this section we recapitulate the classical equation of motion of a charged particle in an electromagnetic field and derive the transformation of the Schrödinger equation and the corresponding wave function of such a particle. We finish with an example, where we explicitly solve the Schrödinger equation of a charged particle in a constant magnetic field. This will yield a discrete energy spectrum called *Landau levels* named after the soviet physicist Lew Landau. Landau levels play a role in the explanation of the integer and fractional quantum hall effect [1], which is treated in another proseminar.

2.1 Lorentz force, Lagrangian and Hamiltonian

Let's consider the motion of a particle of mass m and charge q in a generic electromagnetic field. Classically the motion is given by the Lorentz force (in cgs units):

$$m\frac{d\vec{v}}{dt} = q\vec{E} + \frac{q}{c}\vec{v}\times\vec{B} \tag{1}$$

The E- and B-field can be expressed in terms of the scalar and the vector potential.

$$\vec{E} = -\vec{\nabla}V - \frac{1}{c}\frac{d}{dt}\vec{A} \qquad\qquad \vec{B} = \vec{\nabla}\times\vec{A} \tag{2}$$

The equation of motion can be derived from the Lagrangian or the Hamiltonian, which are

$$\mathcal{L} = \frac{m}{2}\vec{v}^2 + \frac{q}{c}\vec{v}\cdot\vec{A} - qV \qquad\qquad \mathcal{H} = \vec{p}\cdot\vec{v} - \mathcal{L} = \frac{1}{2m}\left(\vec{p} - \frac{q}{c}\vec{A}\right)^2 + qV \tag{3}$$

where $\vec{p} = m\vec{v} + \frac{q}{c}\vec{A}$ is the *canonical momentum* and the term $\Pi = m\vec{v}$ is the *mechanical momentum*.

2.2 Schrödinger equation and gauge transformations

We can now formulate the Schrödinger equation by using the correspondence principle where we replace $\vec{p} \to -i\hbar\vec{\nabla}$ in the Hamiltonian of eqn. (3).

$$i\hbar\frac{\partial}{\partial t}\Psi = \mathcal{H}\Psi = \left[\frac{1}{2m}\left(-i\hbar\vec{\nabla} - \frac{q}{c}\vec{A}\right)^2 + qV\right]\Psi \tag{4}$$

From electrodynamics we remember that both E- and B-Field and in general all physically observable quantities are invariant under gauge transformations of the scalar and vector potential, that are:

$$V \to \widetilde{V} = V - \frac{1}{c}\frac{\partial}{\partial t}\Lambda \qquad\qquad \vec{A} \to \widetilde{\vec{A}} = \vec{A} + \vec{\nabla}\Lambda \tag{5}$$

with Λ some scalar function.
Since the classical Lorentz force only depends on the EM-field $\vec{E}$ and $\vec{B}$ but the Schrödinger equation depends on the scalar and vector potential, we might wonder whether the wave function Ψ depends on the gauge and whether charged particles respond to the EM-field or the EM-potentials. We will now show that the wave function for the Schrödinger equation with the new potentials is [5]

$$\widetilde{\Psi}(\vec{x},t) = \exp\left(\frac{iq\Lambda(\vec{x},t)}{\hbar c}\right)\Psi(\vec{x},t) \tag{6}$$

Proof. We multiply the Schrödinger eqn. from the right with $\exp\left(\frac{iq}{\hbar c}\Lambda(\vec{x},t)\right)$ and use the identity

$$\exp\left(f(x,y)\right)\partial_x = \left(\partial_x - \partial_x f(x,y)\right)\exp\left(f(x,y)\right)$$

which yields

$$\left[\frac{1}{2m}\left(\frac{\hbar}{i}\vec{\nabla} - \frac{q}{c}\vec{A} - \frac{\hbar}{i}\frac{iq}{\hbar c}\vec{\nabla}\Lambda\right)^2 + qV\right]\exp\left(\frac{iq\Lambda}{\hbar c}\right)\Psi = i\hbar\left(\partial_t - \frac{iq}{\hbar c}\partial_t\Lambda\right)\exp\left(\frac{iq\Lambda}{\hbar c}\right)\Psi$$

which is identical to

$$\left[\frac{1}{2m}\left(\frac{\hbar}{i}\vec{\nabla} - \frac{q}{c}\widetilde{\vec{A}}\right)^2 + q\widetilde{V}\right]\widetilde{\Psi} = i\hbar\partial_t\widetilde{\Psi}.$$

With this change of the wave function the Schrödinger equation stays invariant. □

We obtain the same result when considering the correspondence between a classical observable and the expectation value of a quantum mechanical observable. Classically we find that the trajectory $\vec{x}$ and the mechanical momentum Π are invariant whereas the canonical momentum $\vec{p}$ is not invariant under gauge transformations. From the requirement that the same should hold for the expectation values of the respective quantum mechanical operators we find again (6). Luckily with this additional phase factor of the wave function, the probability density $\rho = \Psi^*\Psi$ and the probability flux $\vec{j} = \frac{\hbar}{m}\Im(\Psi^*\vec{\nabla}'\Psi) - \frac{q}{mc}\vec{A}\rho$ stay invariant under gauge transformation. This can best be seen when writing Ψ as $\sqrt{\rho}\exp(iS(\vec{x})/\hbar)$ and therefore $\vec{j} = \frac{\rho}{m}(\vec{\nabla}S - \frac{q\vec{A}}{c})$, where S is some real scalar function. Since $S \to S + \frac{q\Lambda}{c}$ the terms with Λ just cancel each other. [3]

2.3 Free charge in constant magnetic field - Landau levels

In this subsection we treat the special case of a free particle of charge e in a constant homogeneous magnetic field with no other potential. We choose the magnetic field to point in positive z-direction

$\vec{B} = B\hat{z}$ and for the vector potential we choose the so-called symmetric gauge. Since $\vec{B}$ is constant we can write

$$\vec{A} = -\frac{1}{2}\vec{r} \times \vec{B} \qquad A_x = -\frac{1}{2}yB \qquad A_y = \frac{1}{2}xB \tag{7}$$

Plugging these expressions into our Hamiltonian from (3) with $V = 0$ we find

$$\mathcal{H} = \frac{1}{2m}(-\hbar^2\vec{\nabla}^2 - \frac{eB}{c}L_z + \frac{e^2B^2}{4c^2}(x^2 + y^2)) \tag{8}$$

where $L_z = -i\hbar(x\partial_y - y\partial_x)$ is the z-component of the angular momentum operator. In order to make calculations easier we rewrite this equation using magnetic units

$$l_B = \sqrt{\frac{\hbar}{m\omega_c}} \qquad \omega_c = \frac{eB}{mc} \qquad l_B = a_0\frac{4.83 \times 10^4}{\sqrt{B(gauss)}} \tag{9}$$

where l_B is called magnetic length, ω_c is the cyclotron frequency and a_0 is the Bohr radius.
Using l_B as the unit of length, $\hbar\omega_c$ as the unit of energy and measuring L_z in units of $\hbar$ the Hamiltonian now reads

$$\mathcal{H} = -\frac{1}{2}(\vec{\nabla}^2 + L_z) + \frac{x^2 + y^2}{8} \tag{10}$$

2.3.1 Derivation of the energy spectrum using creation and destruction operators

First we rewrite the Hamiltonian (10) in terms of the Hamiltonian for zero angular momentum and a part containing only the angular momentum part.

$$\mathcal{H} = \mathcal{H}_0 - \omega L_z \tag{11}$$

$$\mathcal{H}_0 = \frac{1}{2}(p_x^2 + \omega^2x^2 + p_y^2 + \omega^2y^2) \tag{12}$$

where we set $\omega = 1/2 \cdot \omega_c = 1/2$ to have $\mathcal{H}_0$ in usual harmonic oscillator form. We now introduce the well known creation and destruction operators of the harmonic oscillator, that satisfy the canonical commutation relations $\left[a_j, a_j^\dagger\right] = 1$ for $j \in \{x, y\}$. Explicitly we have

$$a_j = \frac{1}{\sqrt{2\omega}}(\omega \cdot j + ip_j) \qquad j = \frac{a_j + a_j^\dagger}{\sqrt{2\omega}} \qquad p_j = -i\sqrt{\frac{\omega}{2}}(a_j - a_j^\dagger) \tag{13}$$

Rewriting the Hamiltonian $\mathcal{H}_0$ and the angular momentum L_z in terms of these operators we get

$$\mathcal{H}_0 = \omega(a_x^\dagger a_x + \frac{1}{2} + a_y^\dagger a_y + \frac{1}{2}) \tag{14}$$

$$L_z = -i(a_x^\dagger a_y - a_y^\dagger a_x) \tag{15}$$

Now because we work in two dimensions it's best to work in complex components rather than in cartesian. We therefore define the complex numbers $z = x + iy$ and $\bar{z} = x - iy$ and express them in terms of the above operators:

$$z = \frac{1}{\sqrt{2\omega}}((a_x + ia_y) + (a_x^\dagger + ia_y^\dagger)) \qquad \bar{z} = \frac{1}{\sqrt{2\omega}}((a_x - ia_y) + (a_x^\dagger - ia_y^\dagger)) \tag{16}$$

But this is still in terms of Cartesian operators. Therefore we define new, so-called "charged" operators a and b, which also satisfy the canonical commutation relations $\left[a, a^\dagger\right] = \left[b, b^\dagger\right] = 1$ and $[a, b] = \left[a, b^\dagger\right] = 0$, i.e. the a's and b's form a set of independent creation and destruction operators.

$$a = \frac{1}{\sqrt{2}}(a_x + ia_y) \qquad a^\dagger = \frac{1}{\sqrt{2}}(a_x^\dagger - ia_y^\dagger) \tag{17}$$

$$b = \frac{1}{\sqrt{2}}(b_x - ib_y) \qquad b^\dagger = \frac{1}{\sqrt{2}}(b_x^\dagger + ib_y^\dagger) \tag{18}$$

z reads then

$$z = \frac{1}{\sqrt{2\omega}}(a + b^\dagger) \tag{19}$$

We now want to express our Hamiltonian in terms of these new operators. We have

$$a_x = \frac{1}{\sqrt{2}}(a + b) \qquad\qquad a_y = -\frac{i}{\sqrt{2}}(a - b) \tag{20}$$

and therefore

$$\mathcal{H}_0 = \omega(a^\dagger a + \frac{1}{2} + b^\dagger b + \frac{1}{2}) \tag{21}$$

$$L_z = (b^\dagger b - a^\dagger a) \tag{22}$$

and with using $\omega = 1/2$ we obtain [2]

$$\mathcal{H} = a^\dagger a + \frac{1}{2} \tag{23}$$

This is an interesting result. On the one hand we see that Hamiltonian is of the form of a 1-D harmonic oscillator and therefore has discrete energy levels $E_n = n + 1/2$ with $n \in \mathbb{N}_0$. These energy levels are called Landau levels. On the other hand we see that the Hamiltonian is independent of the creation and destruction operators b and $b^\dagger$. From the theory of the harmonic oscillator we know that by acting with $b^\dagger$ on an arbitrary eigenstate $|n, m\rangle$ we obtain an orthogonal eigenstate $|n, m+1\rangle$, but with the same energy $(n+1/2)$. For this reason we conclude that the Landau levels must be hugely degenerate. There are infinitely many (orthogonal) eigenstates with the same energy.

2.3.2 Derivation of the ground state wave functions

To find the ground state wave functions we express the operators a and b in terms of differential operators:

$$\partial_z = \frac{1}{2}(\partial_x - i\partial_y) \qquad\qquad \partial_{\bar{z}} = \frac{1}{2}(\partial_x + i\partial_y) \tag{24}$$

$$a = \frac{1}{\sqrt{2}}(\frac{z}{2} + 2\partial_{\bar{z}}) \qquad\qquad a^\dagger = \frac{1}{\sqrt{2}}(\frac{\bar{z}}{2} - 2\partial_z) \tag{25}$$

$$b = \frac{1}{\sqrt{2}}(\frac{\bar{z}}{2} + 2\partial_z) \qquad\qquad b^\dagger = \frac{1}{\sqrt{2}}(\frac{z}{2} - 2\partial_{\bar{z}}) \tag{26}$$

We want to show that a general ground state wave function $\langle (z, \bar{z}) \mid n = 0\rangle = \sum_{m=0}^{\infty} a_m \langle (z, \bar{z}) \mid 0, m\rangle$ can be expressed by an arbitrary holomorphic function $f(z)$ multiplied by a Gaussian $\exp(-\frac{z\bar{z}}{4})$ (up to normalization). Using that a general ground state wave function satisfies $\langle (z, \bar{z}) \mid a \mid n = 0\rangle = 0$, we have

$$\frac{1}{\sqrt{2}}(\frac{z}{2} + 2\partial_{\bar{z}})\Psi_0(z, \bar{z}) = 0 \tag{27}$$

$$\frac{1}{\sqrt{2}}(\frac{z}{2} + 2\partial_{\bar{z}})f(z)\exp(-\frac{z\bar{z}}{4}) = 0 \tag{28}$$

because any *holomorphic function* $f(z)$ satisfies

$$\partial_{\bar{z}} f(z) = 0 \tag{29}$$

and therefore the derivative only acts on the Gaussian. The fact that the ground state can be any holomorphic function times a Gaussian shows again the big degeneracy of the Landau levels. At the same time for an arbitrary eigenfunction with $m = 0$ we know that $\langle (z, \bar{z}) \mid b \mid m = 0\rangle = 0$ and in an

analogous way we find that this condition is satisfied by any wave function of the form $g(z) \cdot \exp(-\frac{z\bar{z}}{4})$ with $g(z)$ some *anti-holomorphic* function. Therefore we conclude that $\langle (x,y) \mid 0,0\rangle = \frac{1}{\sqrt{2\pi}} \exp(-\frac{z\bar{z}}{4})$. This is the eigenstate with minimal energy and vanishing angular momentum, which follows directly from eqn. (22). With this knowledge we can calculate a basis of the ground state functions by acting with the $b^\dagger$ operator on the state $|0,0\rangle$

$$b^\dagger \Psi_{0,0} \propto b^\dagger \exp(-\frac{|z|^2}{4}) = \frac{1}{\sqrt{2}}(\frac{z}{2} - 2\partial_{\bar{z}}) \exp(-\frac{|z|^2}{4}) \tag{30}$$

$$\propto z \exp(-\frac{|z|^2}{4}) \tag{31}$$

Since z^m is holomorphic, $\partial_{\bar{z}}(z^m) = 0$ and we get

$$(b^\dagger)^m \Psi_{0,0}(x,y) \propto z^m \exp(-\frac{|z|^2}{4}) \tag{32}$$

and after a dimension analysis we find

$$\Psi_{0,m}(z) \propto (\frac{z}{l_B})^m \exp(-\frac{|z|^2}{4l_B^2}) \tag{33}$$

This is no surprise since any holomorphic function can be represented as a power series in z. In a similar way by acting with the $a^\dagger$ operator on a ground state eigenfunction one could obtain the excited eigenfunctions.
The Landau levels (resp. eigenfunctions) in symmetric gauge happen to be most convenient for describing the fractional quantum hall effect [1].

2.4 Degeneracy of Landau levels in symmetric and Landau gauge

In reality it is often the case that a physical system is confined to certain region. E.g. in the case of the quantum hall effect this is the expanse of the conductor sample. To learn something about the degeneracy of the system it is interesting to calculate the number of states (with the same energy) per area. In this subsection we will derive this quantity in the symmetric and the Landau gauge. We return to using *cgs* units for this purpose.

2.4.1 Wave functions in Landau gauge

In Landau gauge the vector potential $\vec{A}$ is

$$\vec{A} = xB\hat{y} \tag{34}$$

and, again confining ourselves to two dimensions, the Hamiltonian reads

$$\mathcal{H} = \frac{1}{2m}(p_x^2 + (p_y - \frac{eB}{c}x)^2) \tag{35}$$

Because of translational invariance in the y-direction the energy eigenfunctions must also be eigenfunctions of p_y, which motivates the ansatz

$$\Psi_k(x,y) = \exp(iky) f_k(x) \tag{36}$$

Acting with the Hamiltonian on Ψ_k we find the equation of a displaced harmonic oscillator in the x direction.

$$\mathcal{H}_k = \frac{1}{2m}p_x^2 + \frac{m\omega_c^2}{2}(x - kl_B)^2 \tag{37}$$

Which finally leads to the following result [1]:

$$E_n = \hbar\omega_c(n + \frac{1}{2}) \tag{38}$$

$$\Psi_{n,k}(x,y) \propto \exp(iky) H_n(x - kl_B^2) \exp(-\frac{(x - kl_B^2)^2}{2l_B^2}) \tag{39}$$

with H_n the usual Hermite polynomial wave function of the harmonic oscillator. Notice that the wave function depends on two quantum numbers $n \in \mathbb{N}_0$ and $k \in \mathbb{R}$, but the energy only depends on n which implies a huge degeneracy.
We see that the wave functions are not gauge invariant and therefore do not have any physical meaning. Also since there is a huge degeneracy we are able to cook up almost any possible wave function by taking linear combinations of the eigenfunctions.

2.4.2 Degeneracy in symmetric gauge

In the symmetric gauge we see that the basis wave functions or more specifically the modulus squared of the wave functions form concentric rings around the origin (see Figure 1). The higher the angular momentum m, the further out the ring. As the wave function (and its modulus squared) in the lowest Landau level with angular momentum m are given as

$$\Psi_{0,m}(z,\bar{z}) \propto (\frac{z}{l_B})^m \exp(-|z|^2/4l_B^2) \tag{40}$$

$$|\Psi_{0,m}(r,\phi)|^2 \propto (\frac{r}{l_B})^{2m} \exp(-r^2/2l_B^2) \tag{41}$$

they are peaked at a ring of radius $r = \sqrt{2m}l_B$. This means that in a disc shaped region of area $A = \pi R^2$ the number of states is the integer part of [1]

$$\mathcal{N} = R^2/(2l_B^2) = A/(2\pi l_B^2) = \frac{eBA}{2\pi\hbar c} \tag{42}$$

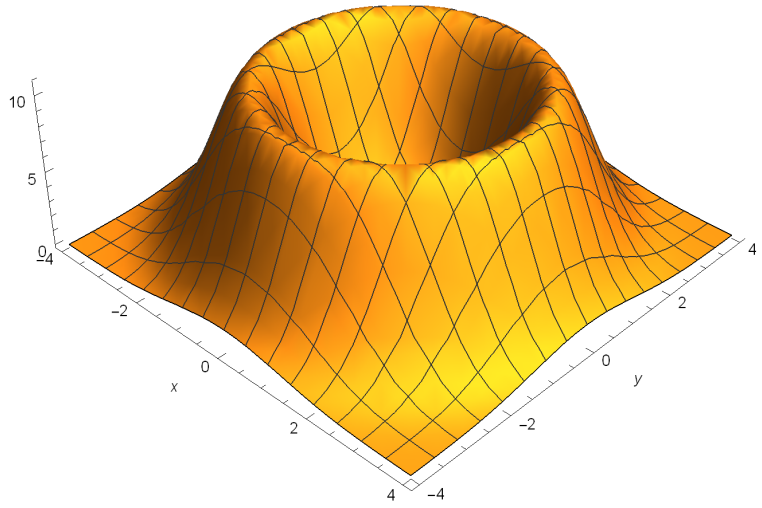

Figure 1: Plot of the non-normalized modulus squared wave function $|\Psi_{0,m}(r,\phi)|^2$ for $m = 3$ and $l_B = 1$. It is peaked at $r = \sqrt{6}$.

2.4.3 Degeneracy in Landau gauge

We would expect to get the same result in the Landau gauge and actually this is the case. In order to calculate the number of states we need to restrict ourselves to a finite region of the (x,y)-plane. We pick a rectangle with sides of lengths L_x and L_y and we want to know how many states fit inside this rectangle. Having a finite size L_y is like putting the system in a box in the y-direction. This yields a quantization of the wave vector k, namely $k = \frac{2\pi}{L_y} \cdot N$, where $N \in \mathbb{Z}$. The allowed values of N are then further restricted by the somewhat heuristic argument that the center of force of the oscillator $x_0 = kl_B^2$ must physically lie within the system $0 \leq x_0 \leq L_x$ which gives a maximum value for N which is [1]

$$\mathcal{N} = \frac{L_x L_y}{2\pi l_B^2} = \frac{eBA}{2\pi\hbar c} \tag{43}$$

Notice that in both gauges we didn't make any attempt to normalize the wave functions, since it is of no relevance for the calculation of the degeneracy and provides no further physical insight.

3 Vector potential outside of a solenoid

In this section we study the physics of a quantum particle moving around a flux tube. Motivated by the the appearance of the gauge potentials V and $\vec{A}$ in the Hamiltonian for a charged particle, we will show that to some extent the physics of a quantum particle are affected by these potentials, rather than only by the electric and the magnetic field. In the first subsection we mention two important concepts of the *proseminar on Berry phases*. We will need these concepts in the context of spectral flow and the Aharonov-Bohm phase. The definitions of these concepts are taken from [1].

3.1 Adiabatic theorem and Berry phases

1. The adiabatic theorem: The adiabatic theorem states that if we place a system in a non-degenerate energy eigenstate and vary parameters sufficiently slowly, then the system will cling to that energy eigenstate.

2. Berry phase and Berry connection: The Berry phase is given as

$$\exp(i\gamma) = \exp(-i \oint_C \mathcal{A}_i(\lambda) d\lambda^i) \tag{44}$$

where $\mathcal{A}_i$ is the Berry connection defined as

$$\mathcal{A}_i = -i \langle n \mid \frac{\partial}{\partial \lambda^i} \mid n \rangle \tag{45}$$

The Berry phase arises in the context of a general Hamiltonian, which we write as $H(x^a; \lambda^i)$. The x^a are the degrees of freedom of the system and the λ^i represent the parameters of the Hamiltonian. They are fixed and their values are determined by some external apparatus.

3.2 Spectral flow

We consider a particle moving around a solenoid. The solenoid shall be taken to expand infinitely in the z-direction and have a profile of area A, carrying a constant magnetic field $\vec{B}$ and hence magnetic flux $\Phi = BA$. Outside the solenoid the magnetic field is zero, but the vector potential is not. This follows from Stoke's theorem which tells us that the line integral outside the solenoid is given by

$$\oint \vec{A} \cdot d\vec{r} = \int \vec{B} \cdot d\vec{S} = \Phi \tag{46}$$

This is simply solved in cylindrical polar coordinates by

$$A_\phi = \frac{\Phi}{2\pi r} \tag{47}$$

Let's now consider a charged quantum particle restricted to lie in a ring of radius r outside the solenoid. The only dynamical degree of freedom is then the angular coordinate $\phi \in [0, 2\pi)$. The Hamiltonian, the energy eigenstates and the energy eigenvalues are then easy to find [1]

$$\mathcal{H} = \frac{1}{2m}(p_\phi - \frac{e}{c}A_\phi)^2 = \frac{1}{2mr^2}(-i\hbar\partial_\phi - \frac{e\Phi}{2\pi c})^2 \tag{48}$$

$$\Psi = \frac{1}{\sqrt{2\pi r}}\exp(in\phi) \tag{49}$$

$$E = \frac{1}{2mr^2}(\hbar n - \frac{e\Phi}{2\pi c})^2 = \frac{\hbar^2}{2mr^2}(n - \frac{\Phi}{\Phi_0})^2 \tag{50}$$

where $n \in \mathbb{Z}$ and $\Phi_0 = 2\pi\hbar c/e$ is the quantum of flux. We make the following observations. The energy spectrum of the particle is affected by the flux even though the particle never goes near the region with magnetic field. And if the flux is a multiple of the quantum of flux the energy spectrum is unaffected.
We can imagine the following: we turn off the solenoid and place the particle in the $n = 0$ ground state. We very slowly increase the flux. By the adiabatic theorem, the particle remains in the $n = 0$ state. But, by the time we have reached $\Phi = \Phi_0$ it is no longer in the ground state. It is now in the state that we previously labeled $n = 1$. Similarly each state n is shifted to the next state $n + 1$. This is an example of a phenomenon called spectral flow. As we increase the flux by one unit of Φ_0, the spectrum returns to itself, but individual states have morphed into each other. In this sense by a change of parameter - in this case Φ - the spectrum of the Hamiltonian "flows".

4 Aharonov-Bohm Effect

In this section we study further the physicality of the vector potential. In order to do so it is illustrative to consider a modified double slit experiment (see Figure 2). As in a usual double slit experiment, from the source region A a beam of coherent particles of charge q is emitted. The beam is split into two parts and is brought together after passing the slits. In region B a detector is placed, which measures an interference pattern. In addition an impenetrable cylinder (e.g. a thin, but long solenoid) containing constant magnetic flux is placed behind the middle wall such that each part of the beam goes on opposite sides of the solenoid. Of course *classically* we wouldn't expect to have any influence of the flux of the solenoid on the motion of a particle, since $\vec{B} = 0$ everywhere where the particle can move. However *quantum mechanically* we will find that there is an influence on the interference pattern. This physically measurable influence of the vector potential is commonly referred to as the *Aharonov-Bohm effect*.
In the first subsection we will look at a slightly different setup that will provide us more insight of the connection to Berry phases. In the second subsection we return to the setup introduced above and quantify the change of the interference pattern.

4.1 Aharonov-Bohm phase

In this subsection we consider the Aharonov-Bohm phase which turns out to be also a special case of general Berry phases. We consider the same setup as in subsection 3.2 with the only difference that now the particle is restricted to lie in a small box. Mathematically, we describe this by including a potential $U(\vec{x})$ in the Hamiltonian that is infinite outside the box. By calling the box "small" we mean that the gauge potential is approximately constant inside the box. If we place the center of the box

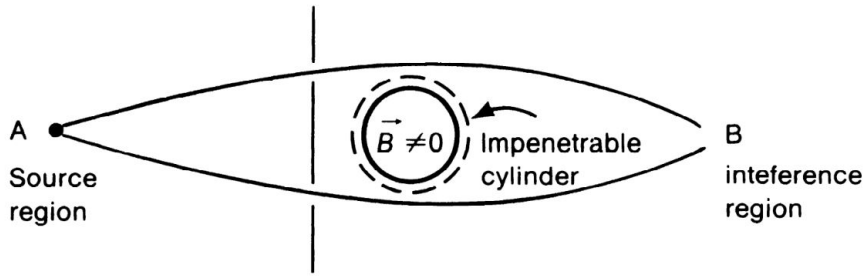

Figure 2: Sketch of the setup for the Aharonov-Bohm effect. In order for the magnetic field to approximately vanish outside the solenoid, we take the solenoid to be very large in the direction perpendicular to the image plane. Image from [3]

at position $\vec{x} = \vec{X}$, then the Hamiltonian reads

$$\mathcal{H} = \frac{1}{2m}(-i\hbar\vec{\nabla} - \frac{e}{c}\vec{A}(\vec{X}))^2 + U(\vec{x} - \vec{X}) \tag{51}$$

We now place the center of the box at a position $\vec{x} = \vec{X_0}$ where the gauge potential vanishes. This can always be done due to gauge freedom. The solution to the Hamiltonian is now just the solution of a particle in a box and we denote the ground state solution by $\Psi(\vec{x} - \vec{X_0})$. In specifying this wave function we also have made a choice of the phase. The connection to Berry phases becomes evident, when doing the following: We slowly move the box in some path in space. On this path the gauge potential experienced by the particle changes. The Schrödinger equation for the Hamiltonian (51) is then solved by the state

$$\Psi(\vec{x} - \vec{X}) = \exp\left(\frac{ie}{\hbar c}\int_{\vec{x}=\vec{X_0}}^{\vec{x}=\vec{X}} \vec{A}(\vec{x}) \cdot d\vec{x}\right) \Psi(\vec{x} - \vec{X_0}) \tag{52}$$

This is because when the gradient $\vec{\nabla}$ acts on the exponent, it pulls down a factor which cancels the $-\frac{e}{c}\vec{A}$ term. To obtain the Berry phase we take the box in a loop C and bring it back to where we started. Our original wave function thereby catches a phase $\Psi(\vec{x} - \vec{X_0}) \to \exp(i\gamma)\Psi(\vec{x} - \vec{X_0})$ with

$$\exp(i\gamma) = \exp\left(\frac{ie}{\hbar c}\oint_C \vec{A}(\vec{x}) \cdot d\vec{x}\right) \tag{53}$$

Comparing this to our general expression for the Berry phase, we see that in this particular context the Berry connection is actually identified with the electromagnetic vector potential,

$$\vec{\mathcal{A}}(\vec{X}) = -\frac{e}{\hbar c}\vec{A}(\vec{x} = \vec{X}) \tag{54}$$

and the parameter space is in this special case the position space $\mathbb{R}^3 - \{line\}$. A particle of arbitrary charge q that goes around a region containing flux Φ will pick up an Aharonov-Bohm phase

$$\exp\left(\frac{iq\Phi}{\hbar c}\right) \tag{55}$$

4.2 Path integral formulation of the Aharonov-Bohm-effect

An elegant way to derive the influence of the vector potential on the interference pattern is by using Feynman's path integral formalism. In one dimension the time evolution of a generic wave function $\Psi(x,t)$ can be expressed using the propagator $K(x, t_1; x', t_0)$:

$$\Psi(x, t_1) = \int_{-\infty}^{\infty} K(x, t_1; x', t_0)\Psi(x', t_0)dx' \tag{56}$$

In Feynman's formalism this propagator can be expressed as

$$K(x, t_1; x^{'}, t_0) = \sum_p W_p \exp\left(iS_p(t_1, t_0)/\hbar\right) \tag{57}$$

where the sum represents the sum over all possible paths, W_p is some weighting function and S is the classical action given as

$$S(t_1, t_0) = \int_{t_0}^{t_1} \mathcal{L}(\dot{x}(t), x(t))dt \tag{58}$$

Feynman showed that with this formulation the probability amplitude to find a particle at time t_1 at position x_1 that at time t_0 was at position x_0 is given as [3]

$$\langle x_1, t_1 \mid x_0, t_0 \rangle = \int_{x_0}^{x_1} \mathcal{D}\left[x(t)\right] \exp\left[i\frac{S(t_1, t_0)}{\hbar}\right] \tag{59}$$

where $\mathcal{D}$ is some kind of infinite dimensional integral operator. Eqn. (58) is known as *Feynman's path integral*. In fact the relevant part for the Aharonov-Bohm effect is

$$\langle x_1, t_1 \mid x_0, t_0 \rangle \propto \sum_p \exp\left(\frac{iS(t_1, t_0)}{\hbar}\right) \tag{60}$$

We remember the Lagrangian of a charged particle in a general EM-field.

$$\mathcal{L} = \frac{m}{2}\dot{\vec{x}}^2 + \frac{q}{c}\dot{\vec{x}} \cdot \vec{A} - qV \tag{61}$$

If we don't have any EM-field, we are back at the original two slit experiment, and we want to denote the action for this case as S_0. By turning on a magnetic field, we change the action

$$S \to S_0 + \frac{q}{c}\int \vec{A} \cdot \dot{\vec{x}}dt = S_0 + \frac{q}{c}\int \vec{A} \cdot d\vec{l} \tag{62}$$

Going back to the experiment the probability amplitude to find the particle at some point on the detector is proportional to [6]

$$\sum_{all\ paths\ through\ slit\ 1} \exp\left(iS_0/\hbar + iq/\hbar c \int \vec{A} \cdot d\vec{l}\right) + \sum_{all\ paths\ through\ slit\ 2} \exp\left(iS_0/\hbar + iq/\hbar c \int \vec{A} \cdot d\vec{l}\right) \tag{63}$$

The physically important quantity is the difference in accumulated phases between the two paths. This is given by [6]

$$\exp\left(\frac{iq}{\hbar c}\int_{slit\ 1} \vec{A} \cdot d\vec{l} - \frac{iq}{\hbar c}\int_{slit\ 2} \vec{A} \cdot d\vec{l}\right) = \exp\left(\frac{iq}{\hbar c}\oint \vec{A} \cdot d\vec{l}\right) \tag{64}$$

Using Stoke's theorem, we can write the exponent in eqn. (64) as

$$\frac{iq}{\hbar c}\oint \vec{A} \cdot d\vec{l} = \frac{iq}{\hbar c}\int_{enclosed} \vec{B} \cdot d\vec{S} = \frac{iq}{\hbar c}\Phi_{enclosed} \tag{65}$$

Notice that all line integrals above are independent of the path as long as the path does not enclose a magnetic flux. Therefore the contributions to all paths going through slit 1 (resp. slit 2) are given by a common phase factor, because in the regions where there is no magnetic flux, $\vec{A}$ can be written as the gradient of a scalar function. Furthermore it is obvious from eqn. (64) that the difference in accumulated phases is independent of the chosen gauge.

The physically observable result is a shift of the interference pattern measured on the observation screen. Since all physical influences only act locally, i.e. they are described by fields, one is lead to the conclusion that the electromagnetic vector potential has in some sense to be considered a physical quantity.
As in subsection 3.2 we find that the effect is only sensitive to non integer multiples of the quantum of flux. Because if the flux Φ is an integer multiple of the quantum of flux $\Phi_0 = 2\pi\hbar/qc$, then the phase shift is an integer multiple of 2π and is hence equivalent to no phase shift.

Including relativistic considerations, the covariance of eqn. (65) demands that there should be a similar result for the scalar potential V. I.e. from the condition that the following loop integral is Lorentz-invariant:

$$-\frac{q}{\hbar}\oint\left(V dt - \frac{\vec{A}}{c}\cdot d\vec{l}\right) \tag{66}$$

where the path of integration now goes over any closed path in space-time. A description of an experiment displaying a phase difference due to the scalar potential V can be found in [7].

5 The topological perspective: the fundamental group of $\mathbb{R}^3 - \{line\}$

In this section we look at the topological perspective of the Aharonov-Bohm effect. In particular in what way does the Aharonov-Bohm phase of the wave function of a charged particle depend on the topology of the space it moves in. In this context we come across the concept of the fundamental group of a (topological) space. In our case the space of interest is $\mathbb{R}^3 - \{line\}$ which mathematically grasps the situation we dealt with in sections 3 and 4, i.e. the real 3D-space in the absence of an infinitely expanded cylinder (the solenoid). In the first subsection we proof that topologically our space is equivalent to the space S^1 and therefore has the same fundamental group, which is the group of integer numbers $(\mathbb{Z}, +)$. In the second subsection we try to motivate this result. In the last subsection we will have a look at the implications for the Aharonov-Bohm effect.
The definitions in the subsections 5.1 & 5.2 are taken from [8].

5.1 Homotopy equivalence between S^1 and $\mathbb{R}^3 - \{line\}$

In order to proof that there exists a homotopy equivalence between S^1 and $\mathbb{R}^3 - \{line\}$, we need to introduce the topological concepts of homotopy, homotopy equivalence and deformation retraction. Two mathematical objects are said to be homotopic if one can be continuously deformed into the other. For the case of maps this is formally given by the following definition:

Definition 1. *Let $f, g : X \to Y$ be two continuous maps between two topological spaces. f and g are said to be **homotopic**, if there is a continuous map $H : X \times [0,1] \to Y$ such that $\forall x \in X : H(x,0) = f(x) \wedge \forall x \in X : H(x,1) = g(x)$ and we write $f \simeq g$. H is called a **homotopy** between f and g.*

In the case of topological spaces we need to introduce the notion of homotopy equivalence.

Definition 2. *Let X and Y be topological spaces.*
*(a) A continuous map $f : X \to Y$ is called **homotopy equivalence** if there exists a continuous map $g : Y \to X$ with $g \circ f \simeq id_X$ and $f \circ g \simeq id_Y$.*
*(b) X and Y are called **homotopy equivalent** if there exists a homotopy equivalence between them.*
*(c) X is called **contractible** if X is homotopy equivalent to the space consisting of only one point.*

Again this means two topological spaces are homotopy equivalent if one can be continuously deformed into the other. We note that homotopy equivalence of two topological spaces is indeed an equivalence relation. A frequently occurring case of homotopy equivalence is the one that one of the spaces is a subspace of the other and the homotopy equivalence is the embedding map.

Definition 3. *Let A be a subset of a topological space X. We refer to the embedding map as $i : A \to X, x \mapsto x$.*
A continuous map $f : X \to A$ with $f\mid_A = id_A$ (and therefore $f \circ i = id_A$) and $i \circ f \simeq id_X$ is called ***deformation retraction****. If such a map exists A is called a* ***deformation retract*** *of X.*

Obviously every deformation retraction is a homotopy equivalence. So a deformation retract is always homotopy equivalent to the space it is embedded in [8]. With this definitions we are now capable of giving a proof that there exists a homotopy equivalence between S^1 and $\mathbb{R}^3 - \{line\}$ or rather that S^1 is a deformation retract of $\mathbb{R}^3 - \{line\}$.

Proof. Lets denote $A = S^1$ and $X = \mathbb{R}^3 - \{line\}$ Let $i : A \to X, \vec{r} \mapsto \vec{r}$ be the embedding map. We claim that the following map is a deformation retraction between X and A.

$$f : X \to A, (x, y, z) \mapsto (\frac{x}{|x+y|}, \frac{y}{|x+y|}, 0)$$

Obviously we have $f \circ i = id_A$ but also there exists a homotopy between id_X and $i \circ f : X \to X, \vec{r} \mapsto f(\vec{r})$ which is

$$H : X \times [0,1] \to X$$
$$((x, y, z), s) \mapsto (\frac{x}{|x+y|^s}, \frac{y}{|x+y|^s}, (1-s)z)$$

and therefore $id_X \simeq i \circ f$ and f is a deformation retraction between X and A. □

5.2 The fundamental group of $\mathbb{R}^3 - \{line\}$

In this subsection we introduce the concept of the fundamental group of a topological space. This concept allows it to define a group structure on the topological space and analyze it with algebraic methods. The fundamental group can be defined in the following way:

Definition 4. *The fundamental group $\pi_1(x, X)$of a topological space X is defined as the set of all homotopy equivalence classes of paths from x to x in X.*

The group operation is the connection of paths, the identity element is the homotopy equivalence class of all null homotopic paths and for every group element the inverse element is given by the equivalence class of the inverse paths. We note that in the case of a path-connected space the fundamental group $\pi_1(x, X)$ is the same for every $x \in X$ and we simply write $\pi_1(X)$.
Now that we know what a fundamental group is we can explain why $\pi_1(\mathbb{R}^3 - \{line\}) \sim (\mathbb{Z}, +)$. With methods of algebraic topology it can be proven that $\pi_n(S^n) \sim (\mathbb{Z}, +)$. It can also be proven that homotopy equivalent (path-connected) spaces have isomorphic fundamental groups.[8] Therefore it follows that $\pi_1(\mathbb{R}^3 - \{line\}) \sim (\mathbb{Z}, +)$. Graphically this result can be understood as follows. When considering all possible closed (continuous) paths in $\mathbb{R}^3 - \{line\}$, it is clear that all loops that do not go around the line are contractible, i.e. they are in the equivalence class of the constant path corresponding to the identity element 0. In general the group element $Z \in \mathbb{Z}$ corresponds to the equivalence class of all loops with winding number Z around the line. The connection of two loops corresponds to the addition of the respective winding numbers.

5.3 Relation to the Aharonov-Bohm effect

Now the relation to the Aharanov-Bohm effect is quite obvious, when we consider the example of the Aharonov-Bohm phase where we moved a small box in space. Here the phase is proportional to the winding number of a loop in the space $\mathbb{R}^3 - \{line\}$, i.e. we get a phase $Z \cdot \frac{iQ\Phi}{\hbar c}$ where $Z \in \mathbb{Z}$ is the winding number.
Furthermore for the Aharonov-Bohm effect to occur, we need a configuration space which is not simply

connected. In the case of $\mathbb{R}^3 - \{line\}$ this shows up in the following: even though everywhere in our space we have $\vec{\nabla} \times \vec{A} = 0$ and therefore $\vec{A} = \vec{\nabla} D$ with D some scalar function, we are not allowed to conclude

$$\oint \vec{A} \cdot d\vec{l} = \oint \vec{\nabla} D \cdot d\vec{l} = 0 \tag{67}$$

if the loop is not homotopic to the constant path, i.e. if the loop encloses some flux.

References

[1] D. Tong, *The Quantum Hall Effect*, arXiv:1606.06687v2 [hep-th] (2016)

[2] J. Stack, Script to Lecture "Quantum Mechanics II", [Online] 10 February 2018 https://courses.physics.illinois.edu/phys581/sp2013/charge_mag.pdf

[3] J.J. Sakurai, *Modern Quantum Mechanics*, (Addison-Wesley, Reading MA, 1994)

[4] B.H. Bransden, C.J. Joachain, *Quantum Mechanics*, (Pearson Education Ltd., 2000), 2nd edition

[5] F. Schwabl, *Quantenmechanik (QM I)*, (Springer, 1998), 5th edition

[6] S.H. Simon, *Topological Quantum: Lecture Notes and Proto-Book*, [Online] 22 January 2018 http://www-thphys.physics.ox.ac.uk/people/SteveSimon/topological2017/LastFullDraftNov29.pdf

[7] Y. Aharonov, D. Bohm, *Significance of Electromagnetic Potentials in the Quantum Theory.* (The Physical Review. 115, No. 3, 1959), pp. 485-491.

[8] A. Gathmann, Script to Lecture "Einführung in die Topologie" TU Kaiserslautern, [Online] 20 February 2018 http://www.mathematik.uni-kl.de/~gathmann/class/topo-2017/topo-2017.pdf

[9] W. Wendelin, Script to Lecture "Topologie", ETH Zürich, [Online] 1 March 2018 https://metaphor.ethz.ch/x/2017/fs/401-2554-00L/