



POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Integer Quantum Hall Effect and Chern Simons Theories

Tesi di Laurea Magistrale in
Nuclear Engineering - Ingegneria Nucleare

Author: **Diego Bragato**

Date: **November 11, 2022**

Student ID: 625759

Advisor: Prof. Marco Beghi

Co-advisors: Lars Fritz, Mikael Fremling

Academic Year: 2022-23



Utrecht
University



do your thing

Words of Thanks

Thanks to Lars Fritz for hosting this thesis at the Institute of Theoretical Physics of the University of Utrecht and making it possible, for his encouragement, explanations of physics and patience. Thanks also to my professor Marco Beghi at the Polytechnic of Milan again for hosting this thesis. A special thanks to Mikael Fremling from the university of Utrecht for his continued support in both explaining me the parts of the physics that I didn't understand and concrete help in writing and reviewing the thesis.

Thanks to my employer ING Bank for the financial support, a final special thanks to my family again for the patience and continued support across the years.

Abstract

This thesis describes the physics of the Integer Quantum Hall effect, with brief references to its development into related research areas. This thesis was realized with the help of the Institute of Theoretical Physics of the University of Utrecht without which it would not have been possible. The objective of the thesis is to cover some of the knowledge needed to get introduced to recent research topics of condensed matter physics. In a first introductory chapter I start from a phenomenological description of the Quantum Hall Effect and the transversal conductivity. After that I develop the computations for the Landau levels in different geometries and with different gauges both with the ladder operators and with the solution of the radial problem. I cover the Aharonov Bohm effect, I analyze the gauge invariance and describe the flux quantization condition or Dirac quantization. I describe the technique of the electromagnetic Berry phase, introducing the Berry connection and the Berry curvature two concepts of differential geometry. I introduce intuitively the quantization of the conductivity with the Laughlin pump argument and describe the monopole magnetic field. I derive the TKNN formula computing the conductivity from the Kubo formula and explain the quantization of the conductivity related to the topological invariant numbers of Chern, both in the case of a simple magnetic field and in the case of particles on a lattice. It is shown how the topological Chern numbers are connected with the integral of the Berry curvature and how there is a topological Chern number for each band. I then briefly refer to topological insulators. Finally I derive the quantization of conductivity with the topological Lagrangian of Chern-Simons, I explain its relation with the Berry curvature and briefly refer to edge states and to an example of Lagrangian used to describe the fractional Quantum Hall Effect.

Keywords: Integer Quantum Hall Effect, Chern-Simons theories, TKNN formula, Berry phase, Berry Curvature, Berry connection, Aharonov Bohm effect, Landau levels

Abstract in lingua italiana

Questo elaborato non innovativo descrive la fisica dell'effetto Hall quantistico intero, con brevi accenni ai suoi sviluppi nei settori di ricerca correlati. Questa tesi è stata realizzata con l'aiuto dell'istituto di Fisica Teorica dell'università di Utrecht senza il quale essa non sarebbe stata possibile. L'obbiettivo della tesi è quello di coprire parte di quell'insieme di conoscenze necessarie per avvicinarsi a settori di ricerca recenti di materia condensata. In un primo capitolo introduttivo si parte da una descrizione della fenomenologia dell'effetto Hall quantistico e la quantizzazione della conduttività trasversale. Successivamente si sviluppano i calcoli per il problema dei livelli quantistici di Landau, nelle varie geometrie e nei differenti gauge, sia con gli operatori di scala, sia con la soluzione del problema radiale. Si affronta l'effetto Aharonov Bohm, si analizza l'invarianza di gauge e si descrive la condizione di quantizzazione del flusso o quantizzazione di Dirac. Si descrive la tecnica della fase elettromagnetica di Berry, introducendo la connessione di Berry e la curvatura di Berry due concetti di geometria differenziale. Si introduce intuitivamente la quantizzazione della conduttività con l'argomentazione della pompa di Laughlin, e si descrive il campo magnetico di monopolo. Si deriva la formula TKNN calcolando la conduttività dalla formula di Kubo e si spiega la quantizzazione della conduttività legata ai numeri invarianti topologici di Chern, sia nel caso di un semplice campo magnetico, sia nel caso di particelle su un lattice. Si fa vedere come i numeri topologici sono connessi con l'integrale della curvatura di Berry e come ci sia un numero topologico per ogni banda, e si accenna brevemente agli isolanti topologici. Infine si ri-deriva la quantizzazione della conduttività con la Lagrangiana topologica di Chern-Simons, se ne spiega la sua relazione con la curvatura di Berry e si accenna sia agli stati di bordo sia ad un esempio di Lagrangiana per descrivere l'effetto Hall quantistico frazionario.

Parole chiave: Effetto Hall quantistico intero, teorie di Chern-Simons, formula TKNN, fase di Berry, curvatura di Berry, connessione di Berry connection, effetto Aharonov Bohm, Livelli di Landau

Contents

| | |
|--|------------|
| Words of Thanks | i |
| Abstract | iii |
| Abstract in lingua italiana | v |
| Contents | vii |
| 1 Phenomenology | 1 |
| 1.1 The experimental setup of the classical Hall effect | 1 |
| 1.2 The experimental setup of the quantum Hall effect | 1 |
| 1.3 Hofstadter butterfly | 2 |
| 1.4 Integer Quantum Hall Effect | 2 |
| 1.5 Fractional Quantum Hall Effect | 3 |
| 1.6 Relevance of Quantum Hall Effect | 4 |
| 1.6.1 Relevance of the experiments | 4 |
| 1.6.2 Theoretical outline | 4 |
| 1.6.3 My specific interests | 5 |
| 2 Landau Levels | 7 |
| 2.1 Historical Introduction | 7 |
| 2.2 Introduction | 7 |
| 2.3 Problem setup | 8 |
| 2.4 Gauge invariant treatment | 8 |
| 2.5 Landau gauge | 11 |
| 2.6 Case on a cylinder | 15 |
| 2.7 Case on a rectangular finite sample | 16 |
| 2.8 Symmetric gauge | 18 |
| 2.8.1 Problem setup | 18 |
| 2.8.2 Ladder operators | 19 |
| 2.8.3 Eigen values and Eigenfunctions with the radial equation | 20 |
| 3 Aharonov bohm effect | 25 |
| 3.1 Introduction | 25 |
| 3.2 Classical motion of electrons in a magnetic field | 26 |

| | | | |
|----------|-------|--|-----------------|
| viii | 3.2.1 | Geometry | Contents |
| | 3.2.2 | Classical case with cartesian coordinates | 26 |
| | 3.2.3 | Local non inertial ref frame | 27 |
| 3.3 | | The field source - the solenoid | 28 |
| | 3.3.1 | Geometry | 28 |
| | 3.3.2 | Magnetic Field of a solenoid | 28 |
| | 3.3.3 | Magnetic Vector potential of a solenoid | 29 |
| | | General and homogenous solution | 29 |
| | | Particular solution | 31 |
| | | Conclusions | 31 |
| | | Fluid Dynamics analogy | 32 |
| | | CPT argument | 32 |
| | | A Quantum field theory perspective | 33 |
| 3.4 | | The propagating electrons | 34 |
| | 3.4.1 | Introduction | 34 |
| | 3.4.2 | Wire geometry | 34 |
| 3.5 | | Schroedinger equation invariance for local Gauge transformations | 37 |
| 3.6 | | Aharonov Bohm experiment | 38 |
| 3.7 | | Flux quantization | 41 |
| 4 | | Berry phase | 43 |
| | 4.1 | Introduction | 43 |
| | 4.2 | Adiabatic theorem | 44 |
| | 4.3 | Berry Phase | 46 |
| | | 4.3.1 Definition | 46 |
| | | 4.3.2 Magnetic analogy for the Berry Phase | 46 |
| | 4.4 | How Aharanov relates to Berry phase | 47 |
| | 4.5 | Berry connection | 49 |
| | | 4.5.1 Definition | 49 |
| | | 4.5.2 Gauge invariance | 49 |
| | 4.6 | Berry curvature | 51 |
| | | 4.6.1 Definition | 51 |
| | | 4.6.2 Magnetic analogy | 52 |
| | | 4.6.3 Quadratic formula | 52 |
| | 4.7 | Topology | 53 |
| | | 4.7.1 Euler formula | 53 |
| | | 4.7.2 Algebraic Topology | 54 |
| | | 4.7.3 The connection to differential Geometry | 54 |
| | | 4.7.4 The Gauss Bonnet theorem | 55 |

| | | |
|----------|--|-----------|
| 5 | TKNN formula | 57 |
| 5.1 | Introduction | 57 |
| 5.1.1 | Approach | 57 |
| 5.1.2 | Historical introduction | 57 |
| 5.2 | Laughlin pump argument | 57 |
| 5.3 | Berry Curvature | 60 |
| 5.3.1 | Electrons in a Brillouin zone | 60 |
| 5.3.2 | Quantization of Berry phase in a compact manifold | 61 |
| 5.4 | The Hall conductivity | 62 |
| 5.4.1 | The Kubo formula | 62 |
| 5.4.2 | Kubo formula for the current | 64 |
| 5.5 | TKNN formula | 68 |
| 5.5.1 | Electrons in a Magnetic field | 68 |
| 5.5.2 | The TKNN formula in a simple magnetic field | 69 |
| 5.5.3 | TKNN formula for particles in a lattice | 70 |
| 6 | Chern Simons theories | 73 |
| 6.1 | Introduction | 73 |
| 6.2 | Physics Introduction | 73 |
| 6.3 | Justification of the Chern-Simons terms | 74 |
| 6.3.1 | Parity and time reversal is violated | 74 |
| 6.3.2 | Long distance physics | 74 |
| 6.3.3 | Possible Lagrangian functions | 75 |
| 6.3.4 | Anyons | 76 |
| 6.4 | Dirac quantization condition | 76 |
| 6.4.1 | Introduction | 76 |
| 6.4.2 | Electro Magnetic duality | 77 |
| 6.4.3 | The monopole vector potential | 78 |
| 6.4.4 | The Dirac quantization condition | 80 |
| 6.4.5 | An aside over the existence of magnetic monopoles | 81 |
| 6.5 | Chern-Simons theories | 81 |
| 6.5.1 | The Gauge transformation | 82 |
| 6.5.2 | Case without a border | 83 |
| 6.5.3 | Quantization | 84 |
| 6.5.4 | From the Chern-Simons Lagrangian to the Hall current | 86 |
| 6.5.5 | Case with a border - edge states | 87 |
| 6.6 | Chern-Simons theories applied to Fractional Quantum Hall effects | 87 |
| 6.7 | Chern-Simons theories in Particle physics | 88 |

1 — Phenomenology

1.1. The experimental setup of the classical Hall effect

The classical Hall effect was discovered in 1879 by Edwin Hall, typically a current is propagated in a flat sample of a conducting material, with an orthogonal applied magnetic field.

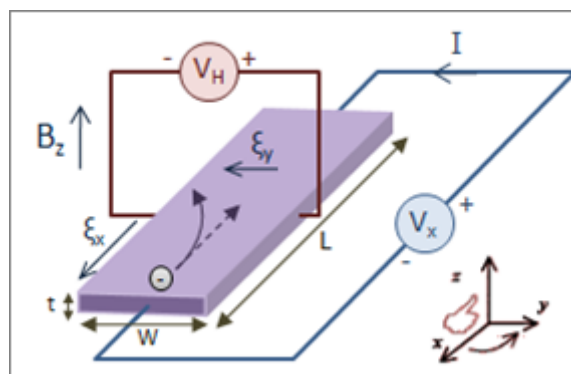


Figure 1: Classical Hall Effect

A Hall voltage transversal to the current is induced by the flow of electrons. Intuitively the electrons in a magnetic field follow circular orbits orthogonal to the field. The system is stationary with a steady current, the Hall voltage compensate for the horizontal deviation of electrons due to the magnetic field. ¹

1.2. The experimental setup of the quantum Hall effect

Historically the standard setup of the modern quantum Hall effect is to have a sample which is a planar junction between two layers of semiconductors usually implemented as a square MOSFET, within the boundary region between of the two semiconductors electrons are confined horizontally in the x,y plane. The sample is cooled down to temperatures of the order of liquid helium or less and a strong vertical magnetic field of the order of 1 Tesla is applied. Nowadays is possible to achieve measurements for temperatures of the order of milli-kelvin and fields of the order of 10 tesla or more. Different thermodynamic properties such as the Hall resistivity, i.e. the transversal resistivity, are measured for different strenghts of the magnetic field.

It was possible to achieve these new experiments given the developement of a set of technologies which were commercialized and available for scientific use in the late 1970s: first the

¹Most of this chapter is based on David Tong (2016b). *Lectures on the Quantum Hall Effect*. <http://www.damtp.cam.ac.uk/user/tong/qhe.html>. [Online; accessed 12-October-2021]

MOSFET, and high quality GaAs interfaces as their alternatives, second cryogenics technologies, third the availability of high intensity magnetic fields provided again from low temperature superconductors.

More recently the availability of graphene, which is effectively a single layer of atoms, permits to realize the 2 dimensional confinement of electrons within a monolayer of graphene itself. It is finally important to mention that there is an upper boundary to current available magnetic fields which are of the order of 100 Tesla.

1.3. Hofstadter butterfly

A precursor of the discovery of the Quantum Hall Effect is in the Ph.D. work of Douglas Hofstadter ² in 1976. Hofstadter analyzed the Hamiltonian for a set of electrons in a square lattice, in a strong orthogonal magnetic field and in a 2 dimensional confinement. The problem was already solved theoretically with the derivation of the Harper equation, Hofstadter computed numerically the spectrum diagonalizing the Hamiltonian and plotted it.

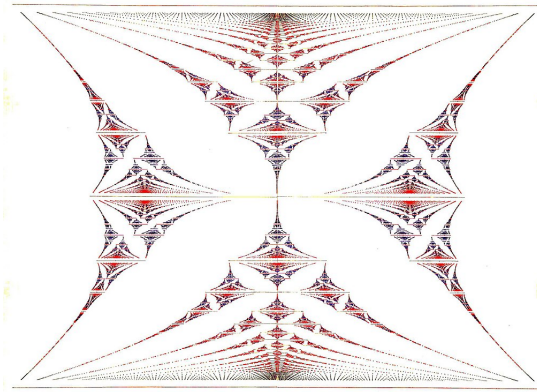


Figure 2: Hofstadter Butterfly

The result was what is now called the Hofstadter butterfly. The plot reveals that the electrons organize themselves in bands, with fractional characteristic quantities, and that the spectrum shows a peculiar fractal structure with self repeating patterns at different scales. This hints to the pervasive presence of fractions and of a hierarchical structure of states in the quantum Hall setup. Given the bizarre unexpected spectrum Hofstadter proposed in the same paper to attempt similar measurements experimentally. ³

² Douglas R. Hofstadter (1976). “Energy levels and wavefunctions of Bloch electrons in rational and irrational magnetic fields”. In: *Physical Review B*. 14.6, pp. 2239–2249

³A comprehensive introductory description of the Hofstadter butterfly is provided in Stephanie Matern (2013). “Hall-Leitfähigkeiten im Hofstadter-Schmetterling”. MA thesis. Universität zu Köln, Institut für Theoretische Physik

1.4. Integer Quantum Hall Effect

The first most famous experiment that discovered the Quantum counterpart of the Hall effect was executed by von Klitzing in 1980⁴, for which he was awarded the nobel prize in 1985.

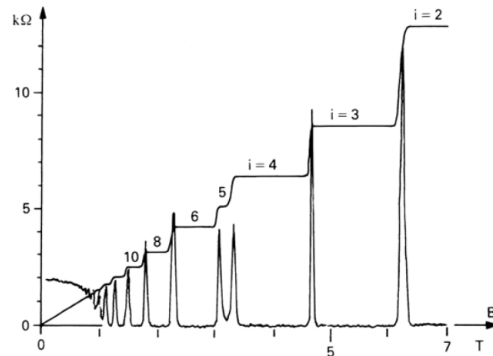


Figure 3: Integer Quantum Hall plateaus

The measurement focused on the Hall resistivity which in essence is increasing and consists of a set of plateaus where is constant, the conductivity on the reverse is almost zero when there is a plateau and it's extremely high between the plateaus.

Given the precision of the quantization of the plateaus and the fact that this is independent from the disorder in the sample, higher and higher precision measurements have been executed. The precision is recently so high that is one of the most precisely measured quantities in physics, nowadays the international standards for electrical resistivity units are based on the Integer Quantum Hall experiments.⁵

1.5. Fractional Quantum Hall Effect

As the disorder of the sample is decreased, and the magnetic field is increased the Hall integer plateaus becomes less prominent but more plateaus appear now with a fractional characteristic number.

⁴ M. Pepper K. v Klitzing G. Dorda (n.d.). “New Method for High-Accuracy Determination of the Fine Structure Constant Based on Quantized Hall Resistance”. In: *Phys. Rev. Lett.* 45 (), p. 494

⁵ B. Jeckelmann and B. Jeanneret (2001). “The quantum Hall effect as an electrical resistance standard”. In: *Rep. Prog. Phys.* 64, pp. 1603–1655

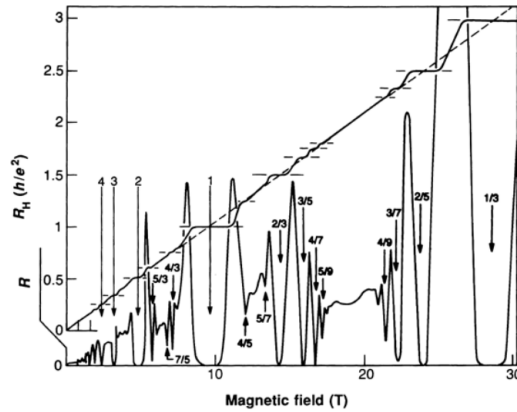


Figure 4: Fractional Quantum Hall plateaus

This was first experimentally discovered by 1982 by Tsui and Stormer⁶. The first theory was laid down by Laughlin⁷ and this led to the joint Nobel prize of Tsui, Stormer and Laughlin in 1998.

1.6. Relevance of Quantum Hall Effect

A thorough review of the current state of the art theoretical approaches and experimental results is given from Prof Hansson⁸. David J. Thouless, F. Duncan M. Haldane and J. Michael Kosterlitz were awarded the Nobel prize in 2016 for theoretical discoveries of topological phase transitions and topological phases of matter including their pioneering work on the Quantum Hall effect.

1.6.1. Relevance of the experiments

There are many reasons why these experiments are attractive from a scientific point of view: the experimental setup is essentially very simple, it is 2 dimensional and controllable, all the measured quantities are rather easy to measure with a high degree of accuracy, the problem is solvable as an exact problem in quantum mechanics, the exact solutions can be used as a benchmark for numerical simulations of more complex geometries and more complex experimental conditions. What is even more interesting is that there is quite a complex set of discrete quantities that are coming out of the experiments, which in essence triggers the use of new theoretical methods.

⁶ H. L. Stormer, D. C. Tsui and A. C. Gossard (1982). “Two-Dimensional Magnetotransport in the Extreme Quantum Limit”. In: *Phys. Rev. Lett.* 48, p. 1559

⁷ R. B. Laughlin (1983). “The Anomalous Quantum Hall Effect: An Incompressible Quantum Fluid with Fractionally Charged Excitations”. In: *Phys. Rev. Lett.* 50, p. 1395

⁸ T. H. Hansson et al. (2017). “Quantum Hall physics: Hierarchies and conformal field theory techniques”. In: *Reviews of modern physics* 89.april–june

1.6.2. Theoretical outline

From a theoretical standpoint the problem is rooted in the Landau levels problem which is a well established solved system in quantum mechanics.

In the case of the integer Quantum Hall effect the major physics properties can be in essence described as a system of free electrons, and the problem is considered essentially solved from a theoretical perspective.

This is not the case for the Fractional quantum hall which is still considered an open research problem, the major physics properties can now be described as a system of interacting electrons, and this does trigger a much more complex set of experimental results and theoretical methods.

One of the underlying and unifying concepts to interpret these measurements reflects a hierarchical structure of electron levels underneath and that is one of the leading themes of the research in this field. A second even more fundamental underlying theme is the emergence of topologically invariant numbers, which in itself is a novelty with respect to orthodox quantum mechanics.

In summary there are research hints in many directions but there is no single overarching theoretical approach that predicts all the observed hierarchies.

It is compelling the idea to be able to derive all these topological numbers from a single hierarchical structure, ideally from the hierarchical structure of the levels themselves. This has been partially achieved at least for the Integer Quantum Hall effect, with the emergence of Chern numbers which are the topological invariants of Chern classes.

1.6.3. My specific interests

Finally to conclude on what triggered my interest is the emergence of hierarchies and especially of continued fractions, in the experimental results.⁹¹⁰ These experimental results together with the intuition that both the integer and the fractional quantum Hall effects shall be rooted in a single underlying explanation from Prof. Jainendra Jain leads to be looking for some more fundamental approach.

⁹ F. D. M. Haldane (1983). “Fractional quantization of the Hall effect, A hierarchy of incompressible quantum fluid states”. In: *Phys. Rev. Lett.* 51, pp. 605–608

¹⁰ B. I. Halperin (1984). “Statistics of quasiparticles and the hierarchy of fractional quantized Hall states”. In: *Phys. Rev. Lett.* 52, pp. 1583–1586

References

- D. C. Tsui, H. L. Stormer and A. C. Gossard (1982). “Two-Dimensional Magnetotransport in the Extreme Quantum Limit”. In: *Phys. Rev. Lett.* 48, p. 1559.
- Haldane, F. D. M. (1983). “Fractional quantization of the Hall effect, A hierarchy of incompressible quantum fluid states”. In: *Phys. Rev. Lett.* 51, pp. 605–608.
- Halperin, B. I. (1984). “Statistics of quasiparticles and the hierarchy of fractional quantized Hall states”. In: *Phys. Rev. Lett.* 52, pp. 1583–1586.
- Hansson, T. H. et al. (2017). “Quantum Hall physics: Hierarchies and conformal field theory techniques”. In: *Reviews of modern physics* 89.april–june.
- Hofstadter, Douglas R. (1976). “Energy levels and wavefunctions of Bloch electrons in rational and irrational magnetic fields”. In: *Physical Review B*. 14.6, pp. 2239–2249.
- Jeckelmann, B. and B. Jeanneret (2001). “The quantum Hall effect as an electrical resistance standard”. In: *Rep. Prog. Phys.* 64, pp. 1603–1655.
- K. v Klitzing G. Dorda, M. Pepper (n.d.). “New Method for High-Accuracy Determination of the Fine Structure Constant Based on Quantized Hall Resistance”. In: *Phys. Rev. Lett.* 45 (), p. 494.
- Laughlin, R. B. (1983). “The Anomalous Quantum Hall Effect: An Incompressible Quantum Fluid with Fractionally Charged Excitations”. In: *Phys. Rev. Lett.* 50, p. 1395.
- Matern, Stephanie (2013). “Hall-Leitfähigkeiten im Hofstadter-Schmetterling”. MA thesis. Universität zu Köln, Institut für Theoretische Physik.

Readings and online resources

Tong, David (2016b). *Lectures on the Quantum Hall Effect*. <http://www.damtp.cam.ac.uk/user/tong/qhe.html>. [Online; accessed 12-October-2021].

2 — Landau Levels

2.1. Historical Introduction

All the forms of magnetism in condensed matter are essentially quantum phenomena, from a pure classical standpoint, the Bohr–Van Leeuwen theorem states that all the induced currents in atoms will have a thermal average which is zero. What’s more the Drude model on it’s own is not able to do predictions in regards to magnetism in materials, only the Sommerfeld free electron model is able to do some partially correct predictions with the introduction of the Fermi statistics.

Alkaline metals have electrons which are not tightly bound and behave as free electrons, these have intrinsic spin which aligns to the external magnetic field and provides what is now called Pauli paramagnetism¹. Free electrons have also orbital motions, and the induced magnetic field tends to compensate the first effect, this is what is now called Landau diamagnetism.

In order to improve on the Pauli model for Alkaline metals Landau² solved the Schroedinger equation for the case of electrons in a vertical magnetic field, and discovered the degeneracy of the electron Levels. This is what is now called the Landau problem and the levels are called Landau Levels. In his work Landau mentioned that the case for low temperatures and high magnetic field would deserve a different treatment, in particular there will be a non linear dependency between the magnetization and the external field, and specifically that it will have a strong oscillating behaviour and an effective averaged field. This leads to the Haas - van Alphen effect, to a divergent periodic dependency in the inverse of the magnetic field, still discovered by Landau³, and to the experimental setup of the Quantum Hall Effect.

2.2. Introduction

Once we try to solve the 2 dimensional quantum problem for Landau levels we are left to a problem that reduces to a one dimensional quantum oscillator, one degree of freedom is not present in the quantum numbers and the levels are degenerate. Given the degeneracy of Landau levels and the gauge freedom we can yes build the full set of solutions as linear combinations and as an Hilbert state, but we are left with the different basis of the Hilbert space, the parametrization of the gauge and the mathematical structures of the wave functions within those basis. Although they are all linearly equivalent, each of these basis actually contains some piece of extra information. This is a bit like looking at an image behind a diffraction

¹ W. Pauli (1927). “Über Gasentartung und Paramagnetismus”. In: *Z. Phys.* 41, p. 81

² L. Landau (1930). “Diamagnetismus der metalle”. In: *Z. Phys.* 64, p. 629

³ D. Ter Haar (1965). *Men of Physics L.D. Landau, Low temperature and solid state physics* p. 36

grating, or one may argue at the bottom of a well, where according to each angle we can see part of the truth, but not the full truth. The problem with this is that there is no one way, there are too many ways, and we hope to analyze it in this section, from the different angles to get a better insight and grip of the physics.

2.3. Problem setup

We start our journey with a constant vertical magnetic field in the z direction and therefore orthogonal to the x, y plane. The magnetic field will be treated as positive when it is in the same direction of the vertical axis.

$$\mathbf{B} = \begin{pmatrix} 0 \\ 0 \\ B_z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ B \end{pmatrix} \quad (2.1)$$

To analyze this in a larger perspective it is important to notice that this field can be considered in a first semi-classical approximation an externally imposed magnetic field that does not create an induced field in the sample and the electrons, but from a more precise quantum field theory point of view it is the actual result of the combination of the external field with the induced field. This combination is also averaged with respect to multiple length scales, for example very small (e.g. much smaller than an atom) and very large length scales (e.g. macroscopic) cannot be always significant, in other words there must also be some upper and lower cutoff scales.

2.4. Gauge invariant treatment

The most simple possible treatment of the problem is without a gauge choice that is given here.^{4 5} Let's start from the classical Lagrangian:

$$L = \frac{1}{2}m\dot{\mathbf{x}}^2 + q\dot{\mathbf{x}} \cdot \mathbf{A} \quad (2.2)$$

If we write the Euler-Lagrange equations for it we obtain the Newton equation with the Lorentz force

$$\begin{aligned} m\ddot{\mathbf{x}} &= q\dot{\mathbf{x}} \wedge \mathbf{B} \\ \mathbf{B} &= \nabla \wedge \mathbf{A} \end{aligned} \quad (2.3)$$

Under a generic gauge transformation the Lagrangian changes only by a total derivative

$$\begin{aligned} \mathbf{A} &\rightarrow \mathbf{A} + \nabla\alpha \\ L &\rightarrow L + q\dot{\alpha} \end{aligned} \quad (2.4)$$

⁴Most of this section is taken from The lecture notes David Tong (2016b). *Lectures on the Quantum Hall Effect*. <http://www.damtp.cam.ac.uk/user/tong/qhe.html>. [Online; accessed 12-October-2021] (pp. 14-18)

⁵There are other resources on this treatment such as Dr. Pekka Pietiläinen Professor Tapash Chakraborty (1995). *The Quantum Hall Effects - Integral and Fractional* ch. 1.2, pag 4

Therefore the equation of motion will be invariant with respect to the gauge transformation. We can introduce the classical canonical momentum

$$\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{x}}} = m\dot{\mathbf{x}} + q\mathbf{A} = \boldsymbol{\pi} + q\mathbf{A} \quad (2.5)$$

where the $\boldsymbol{\pi} = m\dot{\mathbf{x}}$ is the mechanical momentum.

And shift to the Classical Hamiltonian formulation:

$$H = \dot{\mathbf{x}} \cdot \mathbf{p} - L = \frac{1}{2m}(\mathbf{p} - q\mathbf{A})^2 = \frac{1}{2m}\boldsymbol{\pi}^2 \quad (2.6)$$

This has the same structural form of the Hamiltonian in the case without magnetic field

$$H = \frac{1}{2m}\boldsymbol{\pi}^2 \quad (2.7)$$

This structural invariance is also a statement that means that the magnetic field does not work, or also that the magnetic field always acts orthogonally to the direction of motion. From a coordinate perspective \mathbf{x}, \mathbf{p} are now canonical variables where instead the standard coordinates $\mathbf{x}, \boldsymbol{\pi}$ are not, therefore if we introduce the Poisson brackets, these have canonical commutation relations:

$$\{x_i, p_j\} = \delta_{ij}, \{x_i, x_j\} = 0, \{p_i, p_j\} = 0 \quad (2.8)$$

Where instead the mechanical momenta are not canonical variables and do have other commutation relations:

$$\{\pi_i, \pi_j\} = \{p_i - qA_i, p_j - qA_j\} = q\left(\frac{\partial A_j}{\partial x^i} - \frac{\partial A_i}{\partial x^j}\right) = q\varepsilon_{ijk}B_k \quad (2.9)$$

We now want to quantize the problem reusing directly the Poisson brackets of the classical problem, We replace the canonical momentum with the momentum operator and we have the Quantum Hamiltonian

$$\hat{H} = \frac{1}{2m}(\hat{\mathbf{p}} - q\mathbf{A})^2 \quad (2.10)$$

It is possible to solve this problem without any gauge fixing if we introduce the canonical momentum, the benefit of this approach is that is purely algebraic and we can get to the eigenvalues, where instead we will not have be able to deduce the wave functions in terms of position and momentum which will also contain a wealth of relevant information.

$$\hat{\boldsymbol{\pi}} = \hat{\mathbf{p}} - q\mathbf{A} \quad (2.11)$$

Our Hamiltonian is now reduced to:

$$\hat{H} = \frac{1}{2m} \hat{\boldsymbol{\pi}}^2 \quad (2.12)$$

In order to now quantize we need to quantize the canonical relationships of the mechanical momentum

$$[\hat{\pi}_x, \hat{\pi}_y] = iq\hbar B_z \quad (2.13)$$

We have chosen a special coordinate system where the field is in the same direction of the z axis, by comparison it's also possible to generalize this for a generic direction of the field and we see the direct formal correspondence with the Poisson brackets that we have seen earlier 2.9:

$$[\hat{\pi}_i, \hat{\pi}_j] = iq\hbar \varepsilon_{ijk} B_k \quad (2.14)$$

Now we can notice that the Hamiltonian is a quadratic form, this is enough to justify to try linear transformations that can simplify the problem and reduce it to an harmonic oscillator type of problem. We therefore introduce raising and lowering operators as new operators:

$$\begin{aligned} \hat{a} &= \frac{1}{\sqrt{2q\hbar B}} (\hat{\pi}_x + i\hat{\pi}_y) \\ \hat{a}^\dagger &= \frac{1}{\sqrt{2q\hbar B}} (\hat{\pi}_x - i\hat{\pi}_y) \end{aligned} \quad (2.15)$$

and the commutation relations are now becoming:

$$[a, a^\dagger] = 1 \quad (2.16)$$

Where we have chosen the initial linear combination in such a way to obtain the commutation relations for the harmonic oscillator.

If we rewrite the Hamiltonian now

$$\hat{H} = \frac{1}{2m} \hat{\boldsymbol{\pi}}^2 = \hbar\omega_B \left(\hat{a}\hat{a}^\dagger + \frac{1}{2} \right) \quad (2.17)$$

where we introduced the cyclotron frequency $\omega_B = \frac{qB}{m}$. Seeing the Hamiltonian is now formally equivalent to an harmonic oscillator, we can therefore state that we transformed the initial 2 dimensional problem into a one dimensional harmonic oscillator problem.

We can now construct an Hilbert space as is usual from the ladder operators we introduce a ground state $|0\rangle$ which is defined from

$$\hat{a}|0\rangle = 0 \quad (2.18)$$

and we build the rest of the Hilbert space from the relations

$$\begin{aligned}\hat{a}^\dagger|n\rangle &= \sqrt{n+1}|n+1\rangle = 0 \\ \hat{a}|n\rangle &= \sqrt{n-1}|n-1\rangle = 0\end{aligned}\tag{2.19}$$

The state has now eigenvalues

$$E_n = \hbar\omega_B\left(n + \frac{1}{2}\right); n \in \mathbb{N}\tag{2.20}$$

These levels are called Landau levels⁶.

Now we moved from a two dimensional problem and two degrees of freedom to an harmonic oscillator problem in one quantum number, and the energy becomes equally spaced, which looks much like an harmonic oscillator in one dimension. Where the extra degree of freedom ended up ?

As usual there should be one quantum number per degree of freedom and when the energy does not depends on the full set of quantum it means that there is a degeneracy of the levels, and in essence in this chapter we will describe multiple ways to describe this degeneracy.

We can notice that we achieved all of this with no information about the gauge, and ultimately the spectrum does not depend on the gauge, cause the gauge should not include "physical" information, and is factored out, as is typically the case. We in fact computed the eigenvalues, but we only formally computed the eigenfunctions, to achieve a full explicit derivation of the eigenfunctions we need to fix the Gauge. We can then also re-analyze the same problem in different gauges so that we will have more insights on the wave functions and their structure.

The core objective is to use this study of wave functions as foundational in regards to the further study of other global and local invariants built from the wave functions, which play a preminent role in the overall analysis of the quantum hall effect, given the degeneracy of the spectrum.

2.5. Landau gauge

First we consider the problem in the full plane, later we will restrict it to subsets of the plane, as usual there is an orthogonal and constant magnetic field across the full plane. We now also start from a specific gauge for the vector potential, this will be linear in the coordinates so that the first derivatives of the potential are constant. In fact we want a constant magnetic field which is also dependent only on the first derivatives. This will help in simplifying the equations and is called the Landau gauge:

$$A_L^{(x)} = yB \begin{vmatrix} 1 \\ 0 \\ 0 \end{vmatrix}\tag{2.21}$$

⁶ L. Landau (1930). "Diamagnetismus der metalle". In: *Z. Phys.* 64, p. 629

We write down the Schroedinger for the two dimensional problem

$$\frac{1}{2m}[(\hat{p}_x - qyB)^2 + \hat{p}_y^2]\psi(\vec{x}, t) = i\hbar \frac{\partial}{\partial t}\psi(\vec{x}, t) \quad (2.22)$$

And the stationary case

$$\frac{1}{2m}[(\hat{p}_x - qyB)^2 + \hat{p}_y^2]\psi(x, y) = E_n\psi(x, y) \quad (2.23)$$

There is a plane wave solution in the x direction , i.e. there is a k_x continuous quantum number, in fact the Hamiltonian commute with the momentum operator:

$$[\hat{H}, \hat{p}_x] = 0 \quad (2.24)$$

Given they commute, there shall be a common basis of the operators, more precisely the spectrum in x is a set of plane waves and there can be an integral superpositions of these. i.e.

$$\psi(\vec{x}) = \int_{\mathbb{R}^2} a(k_x)\psi(y)e^{ik_x x} dk_x \quad (2.25)$$

And

$$\hat{p}_x\psi(\vec{x}) = \hbar k_x\psi(\vec{x}) \quad (2.26)$$

And the schroedinger simplify to:

$$\frac{1}{2m}[(\hbar k_x - qyB)^2 + \hat{p}_y^2]\psi = E_n\psi \quad (2.27)$$

And using the first term as a potential energy we can give it an harmonic oscillator form

$$\frac{1}{2m}p_y^2\psi(y) + \frac{m}{2}\left(\frac{qB}{m}\right)^2\left(y - \frac{\hbar k_x}{qB}\right)^2\psi(y) = E_n\psi(y) \quad (2.28)$$

i.e we introduce

$$\begin{aligned} \omega_B &= \frac{qB}{m} \\ y_0 &= \frac{\hbar k_x}{qB} \end{aligned} \quad (2.29)$$

And we obtain

$$\frac{1}{2m}p_y^2\psi(y) + \frac{1}{2}m\omega_B^2(y - y_0)^2\psi(y) = E_n\psi(y) \quad (2.30)$$

This is an oscillator on y that is centered on y_0 , the peculiar thing here is that the coordinate x has no role in the Hamiltonian and therefore no role in the energy spectrum that we see in Figure 2.1. There is in fact no kinetic energy operator $\frac{\hat{p}_x^2}{2m}$ in x, neither a plane wave energy

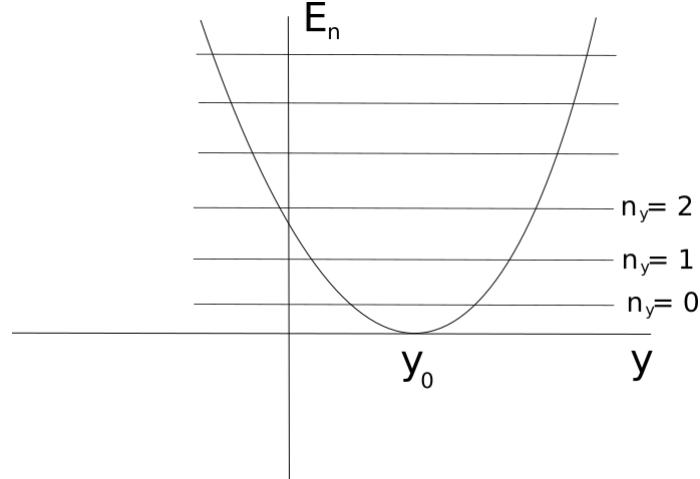


Figure 2.1: Energy spectrum

component $\frac{\hbar^2 k_x^2}{2m}$ in the spectrum. This essentially means that for each level in n_y there is an infinite degeneracy of levels in k_x . Given the sample is infinite the eigenvalue k_x is continuous, i.e. not quantized, and can take any value.

The magnetic length of the oscillator is defined as:

$$l_B = \sqrt{\frac{\hbar}{m\omega}} = \sqrt{\frac{\hbar}{qB}} \quad (2.31)$$

$$y_0 = k_x l_B^2$$

The magnetic length gives the order of magnitude of the width of the level, around the central coordinate y_0 .

This has eigenvalues in the discrete n_y quantum number and for each n_y there is a continuous set of states in k_x which is degenerate (i.e the spectrum does not depend on k_x):

$$E_{n_y, k_x} = \hbar \left(\frac{qB}{m} \right) \left(n_y + \frac{1}{2} \right) \quad (2.32)$$

i.e. we expect a set of equally spaced states in energy like in the case of the harmonic oscillator, that we also see in Figure 2.1.

We can solve the equation 2.30 re-using the solutions for the Harmonic oscillator which are built from Hermite polynomials:

$$\hat{H} = \frac{\hat{p}_x^2}{2m} + \frac{1}{2} m \omega^2 x^2 \quad (2.33)$$

$$\psi_n(x) \sim H_n\left(\frac{x}{l}\right) e^{-\frac{x^2}{2l^2}}; l = \sqrt{\frac{\hbar}{m\omega}}$$

If we now write down the eigenfunctions for our problem in full we then have

$$\psi_{n,k}(x, y) \sim e^{ik_x x} H_n(y - k_x l_B^2) e^{-\frac{(y - k_x l_B^2)^2}{2l_B^2}}; n \in \mathbb{N}; k_x \in \mathbb{R} \quad (2.34)$$

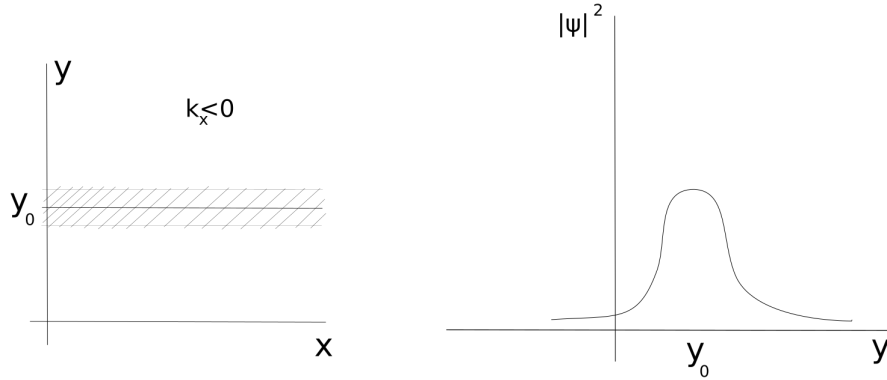


Figure 2.2: Wave functions: infinite sample

In terms of the wave functions (see Figure 2.2) there is still a planar wave component in the x direction and therefore it is possible to have for example wave packets in the x direction as a superposition of plane wave components in x . Graphically we can see this as an harmonic oscillator with states centered in y_0 .

And the full state itself we can write it as a linear combination of the eigenfunctions:

$$|\psi\rangle = \sum_n \int_{\mathbb{R}} a_n(k_x) e^{ik_x x} |n_y\rangle dk_x \quad (2.35)$$

Where the $|n_y\rangle$ are the harmonic oscillator eigenfunctions in y .

There is a second way to express the Landau gauge which is over the y coordinate:

$$A_L^{(y)} = xB \begin{vmatrix} 0 \\ -1 \\ 0 \end{vmatrix} \quad (2.36)$$

It's important to stop for a moment and notice the extra minus sign in the gauge choice, this is introduced given the the equation is not invariant by just swapping the coordinates given there is the magnetic field. The magnetic field itself in fact is not parity invariant and reverses the sign with a parity transformation, in fact if we look in detail to the swap of coordinates it can be decomposed into a proper rotation which does not change the magnetic field and a parity transformation which does.

It is now possible to redo the same computation with this new gauge and and we obtain exactly the same type of eigenvalues and eigenfunctions but this time over the opposite coordinate, in fact we find an oscillator centered over a coordinate x_0 instead of y_0 . Also the eigenfunctions remains of gaussian type but this time in the x direction and are centered around the coordinate x_0 , as before there are planar waves but this time along the y direction.

In this case the wave functions maintain the same form, in this case gaussian, despite the swap of coordinates, but in general if we change to other gauges the new eigenfunctions we find

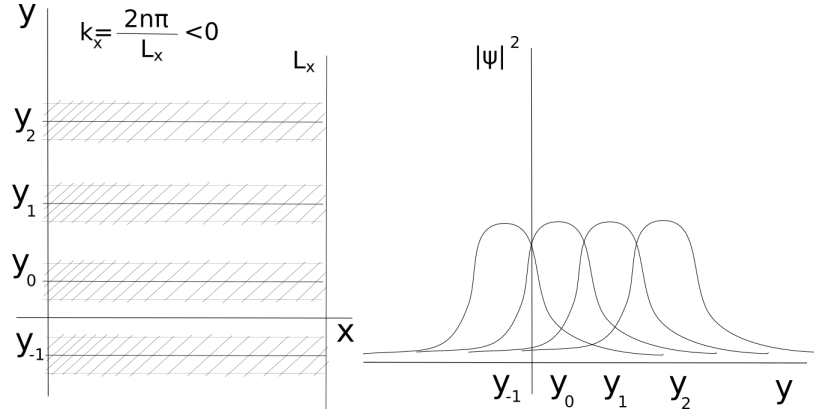


Figure 2.3: Wave functions: Case on a Cylinder

may also have a different form. In any case the wave functions are a complete set in the Hilbert space and therefore by infinite linear combinations can generate any other function in the same Hilbert space and therefore any other complete set of eigenfunctions.

2.6. Case on a cylinder

We now analyze the same problem but this time on an infinite strip with periodic boundary conditions:

$$\psi(0) = \psi(L_x) \tag{2.37}$$

this case and geometry is also called problem on a cylinder to be precise where the term infinite strip is for boundary conditions like the particle in a box case:

$$\begin{aligned} \psi(0) &= 0 \\ \psi(L_x) &= 0 \end{aligned} \tag{2.38}$$

Another way to look at this difference geometrically is to imagine the infinite strip in the y direction replicated across the full plane in the x direction, in the particle in a box case the wave functions are exactly zero on the junction lines for any value of y , where in the case of periodic boundary conditions case they just need to be the same on both side of the strip, but they may have different values for each y .

Constraining the system along one direction means also that the previously continuous eigenvalues become discrete and quantized. On a finite strip along the y direction, the k_x is quantized as we see in Figure 2.3. In fact If we impose 2.37 we obtain:

$$k_x = \frac{2\pi n}{L_x} \tag{2.39}$$

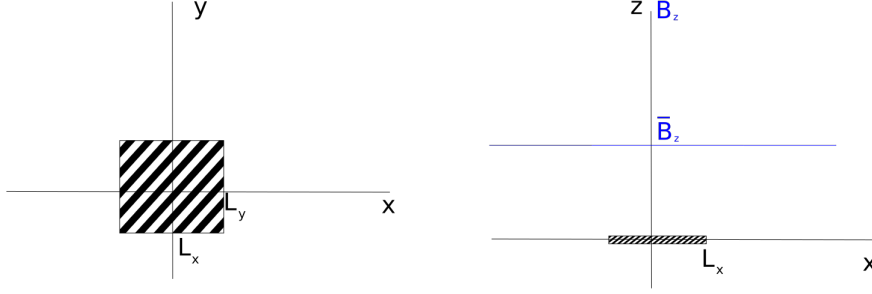


Figure 2.4: Geometry: finite rectangular sample

and therefore we have an infinite numerable set of levels located at coordinates

$$y_n = -\frac{\hbar}{qB} \frac{2\pi n_x}{L_x} = \frac{2\pi n_x}{L_x} l_B^2 \quad (2.40)$$

where it is important to notice that the spacing of the levels in the y coordinate, is comparable if not smaller than the magnetic length and this gives the intuition that the orbitals are almost overlapping in space. In fact $l_B \ll L_x$ and therefore:

$$\Delta y = \frac{2\pi l_B^2}{L_x} \ll l_B \sim \text{orbital width} \quad (2.41)$$

Being the levels also degenerate in n_x , i.e. exactly with the same energy, even with the smallest amount of external energy (i.e. a gapless excitation) is possible to have one electron shifting from one level to another.

2.7. Case on a rectangular finite sample

Given the same constant field along z direction, we will now analyze yet another geometry, the case of a rectangular sample of size L_x, L_y confining the electrons (See figure 2.4) to do this we will use the periodic boundary conditions in both x and y

$$\begin{aligned} \psi(x + L_x, y) &= \psi(x, y) \\ \psi(x, y + L_y) &= \psi(x, y) \end{aligned} \quad (2.42)$$

We want to achieve now that the wave function $e^{ik_x x}$ is periodic in $x = x + L_x$, the condition on the x coordinate is the same condition of the previous scenario and developed gives as before:

$$k_x L_x = 2\pi n_x; n_x \in \mathbb{N} \quad (2.43)$$

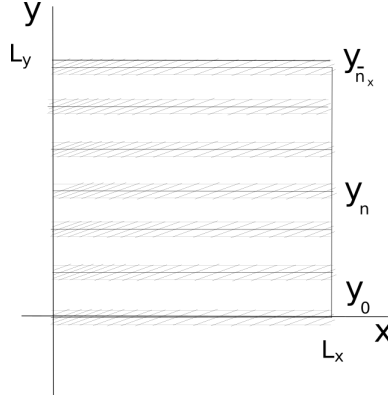


Figure 2.5: Wavefunctions: finite rectangular sample

We know all levels are parallel to the x axis, that the center of each level is $y_n = -k_x l_B^2$ and that center must be within the sample, i.e $0 \leq y_n \leq L_y$ (See figure 2.5). For y to be positive we must have $n_x \leq 0$, for $y = 0$ we have $n_x = 0$, and we have also a maximum value \bar{n}_x for $y = L_y$ such that:

$$-\bar{n}_x \leq n_x \leq 0; n_x \in \mathbb{N} \quad (2.44)$$

n_x is still a periodic index that represent the degeneracy, but if before the degeneracy could assume any natural number, now the degeneracy is bounded and we end up with a large finite number of discrete degenerate levels. We can enforce $y_{\bar{n}_x} = L_y$ on the upper y boundary and we replace using 2.43

$$\begin{aligned} y_{\bar{n}_x} &= -k_x l_B^2 = L_y \\ L_y &= -\frac{2\pi\bar{n}_x}{L_x} l_B^2 \end{aligned} \quad (2.45)$$

and we get the Landau Degeneracy \bar{n}_x i.e. the maximum number of Landau Levels that can fit in the sample

$$Degeneracy = \bar{n}_x = \frac{L_y L_x}{2\pi l_B^2} = \frac{A}{\hbar c} q B \frac{1}{2\pi} = \frac{AB}{\frac{2\pi\hbar c}{q}} \quad (2.46)$$

Or summarizing:

$$Degeneracy = \frac{\Phi}{\Phi_0} \quad (2.47)$$

where $\Phi = BA$ is the total magnetic flux across the sample and $\Phi_0 = \frac{2\pi\hbar c}{q}$ is the flux quantum. The Landau degeneracy can therefore be also described as the total number of flux quanta in the external magnetic field.

To get the idea of the number of levels for a field of 1 Gauss (i.e. 10^{-4} Tesla):

$$\begin{aligned} \Phi_0 &= 2 * 10^{-7} \text{Gauss} \cdot \text{cm}^2 \\ \text{Deg} &= \frac{\Phi}{\Phi_0} = \frac{1 \text{Gauss} \cdot \text{cm}^2}{2 * 10^{-7} \text{Gauss} \cdot \text{cm}^2} = 5 * 10^6 \end{aligned} \quad (2.48)$$

Where the typical order of magnitude of fields for Quantum hall experiment is of the order of 1 Tesla.⁷

Finally to conclude we can also define the filling factor that is the number of filled levels with respect to the total: ⁸

$$\nu = \frac{n_x}{n_x}. \quad (2.52)$$

2.8. Symmetric gauge

2.8.1. Problem setup

Symmetric gauge

$$A_L^{(xy)} = \frac{1}{2} A_L^{(x)} + \frac{1}{2} A_L^{(y)} = \frac{B}{2} \begin{vmatrix} y \\ -x \\ 0 \end{vmatrix} \quad (2.53)$$

The Schroedinger

$$\frac{1}{2m} [(\hat{p}_x - \frac{qB}{2}y)^2 + (\hat{p}_y + \frac{qB}{2}x)^2 + \hat{p}_z^2] \psi(\vec{x}, t) = E_n \psi(\vec{x}, t) \quad (2.54)$$

Where now we only have the null commutator:

$$[\hat{H}, \hat{p}_z] = 0 \quad (2.55)$$

⁷ Barton Zwiebach (2018a). *L14.4 Landau levels (continued). Finite sample. MIT 8.06 Quantum Physics III.* <https://www.youtube.com/watch?v=Uux0VkJaoxY&t=318s>. [Online; accessed 12-October-2021]

⁸ Dr. Pekka Pietiläinen Professor Tapash Chakraborty (1995). *The Quantum Hall Effects - Integral and Fractional* pag 6.

$$N_s = \frac{L_x L_y}{2\pi l_0^2} \quad (2.49)$$

$$N_s = \frac{e}{2\pi\hbar c} \Phi = \frac{\Phi}{\Phi_0} \quad (2.50)$$

Landau degeneracy is the total number of flux quanta in the external magnetic field

$$\nu = 2\pi l_0^2 n_0 \quad (2.51)$$

dimensionless density of electrons expressed as the filling factor of the Landau Level

And we can solve the problem in just x and y . We can reduce it to the adimensional equation

$$\left[(-i\frac{\partial}{\partial x} - \frac{1}{2}y)^2 + (-i\frac{\partial}{\partial y} + \frac{1}{2}x)^2\right]\psi(\vec{x}, t) = \varepsilon_n\psi(\vec{x}, t) \quad (2.56)$$

We introduce two complex variables

$$\begin{aligned} z &= x + iy \\ \bar{z} &= x - iy \end{aligned} \quad (2.57)$$

And obtain

$$H = \left[-4\frac{\partial^2}{\partial z\partial\bar{z}} + z\frac{\partial}{\partial z} - \bar{z}\frac{\partial}{\partial\bar{z}} + \frac{1}{4}z\bar{z}\right] \quad (2.58)$$

2.8.2. Ladder operators

In the case of the one dimensional harmonic oscillator, we usually introduce the ladders operators to compute the eigenvalues directly. In this case there are four ladders operators, given there are four base operators $\hat{p}_x, \hat{p}_y, x, y$, four degrees of freedom, and 2 constants of motion \hat{H}, \hat{L} .

We introduce:

$$\begin{aligned} b &= \frac{1}{\sqrt{2}}\left(\frac{1}{2}\bar{z} + 2\frac{\partial}{\partial z}\right); b^\dagger = \frac{1}{\sqrt{2}}\left(\frac{1}{2}z - 2\frac{\partial}{\partial\bar{z}}\right) \\ a &= \frac{1}{\sqrt{2}}\left(\frac{1}{2}z + 2\frac{\partial}{\partial\bar{z}}\right); a^\dagger = \frac{1}{\sqrt{2}}\left(\frac{1}{2}\bar{z} - 2\frac{\partial}{\partial z}\right) \end{aligned} \quad (2.59)$$

where

$$\begin{aligned} [a, a^\dagger] &= [b, b^\dagger] = 1 \\ [a, b] &= [a, b^\dagger] = [a^\dagger, b] = [a^\dagger, b^\dagger] = 0 \end{aligned} \quad (2.60)$$

The hamiltonian gets the form

$$H = a^\dagger a + \frac{1}{2} \quad (2.61)$$

and the eigenstates

$$\begin{aligned} |n, m\rangle &= \frac{b^{\dagger m+n} a^{\dagger n}}{\sqrt{(m+n)!} \sqrt{n!}} |0, 0\rangle \\ H|n, m\rangle &= \left(n + \frac{1}{2}\right) |n, m\rangle \end{aligned} \quad (2.62)$$

Given

$$\begin{aligned} a|0, 0\rangle &= 0 \\ b|0, 0\rangle &= 0 \end{aligned} \quad (2.63)$$

We get

$$|0, 0\rangle = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{4}z\bar{z}} \quad (2.64)$$

and for the first Landau level we have

$$|0, m\rangle \approx b^\dagger{}^m e^{-\frac{1}{4}z\bar{z}} \approx \left(\frac{z}{2} - 2\frac{\partial}{\partial\bar{z}}\right)^m e^{-\frac{1}{4}z\bar{z}} \approx z^m e^{-\frac{1}{4}z\bar{z}} = \eta_{0m} \quad (2.65)$$

This is named as a hole cause J.J inverted z and \bar{z} More generally we can build linear combinations of these and we arrive to the general case

$$\sum a_m |0, m\rangle \approx f(z) e^{-\frac{1}{4}z\bar{z}} = \eta_{0m} \quad (2.66)$$

where $f(z)$ is analytic

m is the angular momentum because

$$\begin{aligned} L = L_z &= -i\hbar \frac{\partial}{\partial\theta} = -\hbar \left(z \frac{\partial}{\partial z} - \bar{z} \frac{\partial}{\partial\bar{z}} \right) = -\hbar (b^\dagger b - a a^\dagger) \\ -\hbar (b^\dagger b - a a^\dagger) |n, m\rangle &= -\hbar m |0, 0\rangle \\ m &= -n, -n + 1, \dots \end{aligned} \quad (2.67)$$

The b operators are degenerate and these two show the degeneracy of the landau levels, there is an extra quantum number which is degenerate and is actually the angular momentum quantum number. The angular momentum here is conserved and the total angular momentum L is the same as the L_z component given the geometry.

2.8.3. Eigen values and Eigenfunctions with the radial equation

Let's now compute wavefunctions and eigenvalues directly⁹

$$\hat{H} = \frac{1}{2m^*} (\hat{\mathbf{p}} - q\mathbf{A})^2 \quad (2.68)$$

We remain with

$$\hat{H} = \frac{1}{2m^*} \hat{\mathbf{p}}^2 + \frac{1}{2} \omega_C \hat{L}_z + \frac{q^2 B^2}{8m^*} (x^2 + y^2) \quad (2.69)$$

⁹ Dr. Pekka Pietiläinen Professor Tapash Chakraborty (1995). *The Quantum Hall Effects - Integral and Fractional*, Appendix A

¹⁰ where we introduced the cyclotron frequency

$$\omega_C = \frac{qB}{m^*} \quad (2.70)$$

¹¹ in Cylindrical coordinates

$$-\frac{\hbar^2}{2m^*} \left[\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial \psi}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 \psi}{\partial \phi^2} + \frac{\partial^2 \psi}{\partial z^2} \right] - i\hbar\omega_C \frac{\partial \psi}{\partial \phi} = \left(E - \frac{1}{8} m^* \omega_C^2 r^2 \right) \psi \quad (2.71)$$

Given angular momentum, which is same as L_z , and z-momentum are conserved we substitute:

$$\psi = \frac{1}{\sqrt{2\pi}} f(r) e^{-im\phi} e^{ik_z z} \quad (2.72)$$

Then we have

$$\frac{\hbar^2}{2m^*} \left[\frac{d^2 f(r)}{dr^2} + \frac{1}{r} \frac{df(r)}{dr} - m^2 \frac{f(r)}{r^2} \right] + \left(E - \frac{1}{8} m^* \omega_C^2 r^2 - \frac{\hbar^2}{2m^*} k_z^2 + \frac{1}{2} m\hbar\omega_C \right) f(r) = 0 \quad (2.73)$$

We introduce an adimensional length and energy

$$x = \frac{m^* \omega_C}{2\hbar} r^2 = \frac{r^2}{2l_B^2} = [M][T]^{-1}[M]^{-1}[L]^{-2}[T][L]^2 \quad (2.74)$$

$$\beta_m = \frac{1}{\hbar\omega_C} \left(E - \frac{\hbar^2}{2m^*} k_z^2 \right) + \frac{1}{2} m$$

We get

$$x \frac{d^2 f}{dx^2} + \frac{df}{dx} + \left[\beta_m - \frac{m^2}{4x} - \frac{1}{4} x \right] f(x) = 0 \quad (2.75)$$

We simplify with:

$$f = x^{-\frac{1}{2}} R(x) \quad (2.76)$$

And we remain with the Whittaker equation:

$$\frac{d^2 R}{dx^2} + \left[-\frac{1}{4} + \frac{\beta_m}{x} + \frac{1-m^2}{4x^2} \right] R = 0 \quad (2.77)$$

This equation has two linearly independent fundamental solutions, the Whittaker functions:

$$W_{\beta, \frac{1}{2}m}(x), W_{-\beta, \frac{1}{2}m}(-x) \quad (2.78)$$

There are explicit formulas for these.

From a mathematical standpoint the Whittaker equation is also known as a form for the Confluent hypergeometric equation, this equation has various equivalent forms, one of the most

¹⁰in C.G.S. the minimal coupling is $\mathbf{p} - \frac{q}{c} \mathbf{A}$

¹¹in C.G.S this is $\omega_C = \frac{qB}{mc}$

well known forms is the Kummer equation.

$$zw''(z) + (b - z)w'(z) - aw(z) = 0 \quad (2.79)$$

that we are going to derive here.

Substituting:

$$R(x) = e^{-kx}x^n w(x) \quad (2.80)$$

We obtain the differential equation:

$$xw'' + (|m| + 1 - x)w' + (\beta - \frac{|m| + 1}{2})w = 0 \quad (2.81)$$

Which is the confluent hypergeometric equation the solution of which is the confluent hypergeometric function

$$w = F(-(\beta - \frac{|m| + 1}{2}), |m| + 1, x) \quad (2.82)$$

If the wavefunction is everywhere finite, $\beta - \frac{|m| + 1}{2}$ is a positive integer n .

The energy levels are then:

$$E = \hbar\omega_C(n + \frac{1}{2}|m| - \frac{1}{2}m + \frac{1}{2}) + \frac{\hbar^2 k_z^2}{2m^*} \quad (2.83)$$

The corresponding solutions are:

$$R_{n,m}(r) = \frac{1}{l_B^{|m|+1}|m|!} \left[\frac{(|m| + n)!}{2^{|m|}n!} \right]^{\frac{1}{2}} e^{-\frac{r^2}{4l_B^2}} r^{|m|} F(-n, |m| + 1, \frac{r^2}{2l_B^2}) \quad (2.84)$$

with

$$F(-n, |m| + 1, \frac{r^2}{2l_B^2}) = \frac{\Gamma(n + 1)\Gamma(|m| + 1)}{\Gamma(n + |m| + 1)} L_n^{|m|}(\frac{r^2}{2l_B^2}) = \frac{n!m!}{(n + |m|)!} L_n^{|m|}(\frac{r^2}{2l_B^2}) \quad (2.85)$$

where the $L_n(x)^\alpha$ are the Laguerre polynomials.

The wave function:

$$\psi_{n,m}(r) = \left[\frac{n!}{2\pi l_B^2 2^m (n + |m|)!} \right]^{\frac{1}{2}} e^{-im\phi - \frac{r^2}{4l_B^2}} \left(\frac{r}{l_B} \right)^{|m|} L_n^{|m|} \left(\frac{r^2}{4l_B^2} \right) \quad (2.86)$$

And the lowest eigenvector i.e. the Lowest Landau Level is:

$$\psi_m(z) = \left[\frac{1}{2\pi l_B^2 2^m m!} \right]^{\frac{1}{2}} \left(\frac{z}{l_B} \right)^m e^{-\frac{|z|^2}{4l_B^2}} \quad (2.87)$$

Given the equation is linear, it is also important to notice that we can build linear combinations of these solutions

$$\psi_m(z) = f(z) e^{-\frac{|z|^2}{4l_B^2}} \quad (2.88)$$

where $f(z)$ is a polynomial of order m . This will form the basis of the Laughlin Ansatz.

We refer to the literature for further properties of the different types of hypergeometric equations.

It's possible also to analyze more complex versions of this problem with an harmonic or elliptic confinement potential ¹²

¹²See for example Marieke Van Beest (2015). *Fock-Darwin States for an Elliptical Spin-Orbit Coupled Quantum Well*. <https://nbi.ku.dk/english/theses/bachelor-theses/mariek-van-beest/BachelorThesisMariekeVanBeest.pdf>. [Online; accessed 12-October-2021]

References

- Haar, D. Ter (1965). *Men of Physics L.D. Landau, Low temperature and solid state physics*.
- Landau, L. (1930). “Diamagnetismus der metalle”. In: *Z. Phys.* 64, p. 629.
- Pauli, W. (1927). “Uber Gasentartungund Paramagnetismus”. In: *Z. Phys.* 41, p. 81.
- Professor Tapash Chakraborty, Dr. Pekka Pietiläinen (1995). *The Quantum Hall Effects - Integral and Fractional*.

Readings and online resources

- Beest, Marieke Van (2015). *Fock-Darwin States for an Elliptical Spin-Orbit Coupled Quantum Well*. <https://nbi.ku.dk/english/theses/bachelor-theses/mariek-van-beest/BachelorThesisMariekeVanBeest.pdf>. [Online; accessed 12-October-2021].
- Zwiebach, Barton (2018a). *L14.4 Landau levels (continued). Finite sample. MIT 8.06 Quantum Physics III*. <https://www.youtube.com/watch?v=Uux0VkJaoxY&t=318s>. [Online; accessed 12-October-2021].

3 — Aharonov bohm effect

3.1. Introduction

The standard experimental setup to achieve macroscopic constant magnetic field is to use a very long vertical solenoid, more precisely within the solenoid the field is constant and outside is zero, and the longer the solenoid the more this approximation is correct. Outside the solenoid the magnetic vector potential instead is not zero as we will see and this induces phase shifts on the wave functions.

The standard assumption in classical electrodynamics is that the vector potential is purely a mathematical artifact, without mathematical consequences, and without a physical interpretation. Only the field itself can be measured directly, and can be considered a physical quantity. This is only partially true in the quantum case, in essence the vector potential can induce phase shifts in the wave function of an electron propagating through it. This is in parallel with the fact that a gauge transformation induces a phase shift in the wave function. Global phase shifts per se have no also no physical consequences to the extent of quantum mechanics but once we consider a finite loop we end up with a discrete phase shift and this creates a measurable quantity.

When this was discovered by Aharonov, Bohm and an earlier paper from Eherenberg Siday ^{1 2 3} it created a big sensation given there was for the first time a measurable effect caused by the presence of the potentials and in absence of a magnetic field, where until then, the potentials were considered purely a mathematical artifact. This shifted the attention from the fields to the potentials as being more fundamental. In the treatment with the field only, given the field is not zero only within the solenoid and radially "far away" outside the solenoid the field is zero the field would induce a non local effect on the wave function.

The treatment with vector potentials instead permits to safeguard the principle of locality and that the wave function has a local effect, i.e. a local phase shift, due to a local vector potential.

Finally another important logical consequence of this effect is for action principles, given they are also dependent directly from the potentials, as is the case for the Schroedinger equation, they are also as fundamental as the vector potential itself. ⁴

¹ D Aharonov Y; Bohm (1959). "Significance of electromagnetic potentials in quantum theory". In: *Physical Review* 115 (3). doi:10.1103/PhysRev.115.485, pp. 485–491

² D Aharonov Y; Bohm (1961). "Further Considerations on Electromagnetic Potentials in the Quantum Theory". In: *Physical Review* 123 (4). doi:10.1103/PhysRev.123.1511, pp. 1511–1524

³ RE Ehrenberg W; Siday (1949). "The Refractive Index in Electron Optics and the Principles of Dynamics". In: *Proceedings of the Physical Society B*. 62 (1). doi:10.1088/0370-1301/62/1/303, pp. 8–21

⁴A lot of this chapter is derived from Griffiths D.J. (2005). *Introduction to quantum mechanics* Ch. 10.2.3

3.2. Classical motion of electrons in a magnetic field

The underlying motion for electrons moving in a magnetic field is circular, without acceleration and orthogonal to the magnetic field, this basic law of motion is underlying all the classical and quantum physics of electron moving in magnetic fields including Landau levels and the aharonov Bohm effect.

3.2.1. Geometry

Let's consider one electron moving in a vertical constant magnetic field that is present all across the space

$$\mathbf{B} = \begin{vmatrix} 0 \\ 0 \\ \overline{B_z} \end{vmatrix} \quad (3.1)$$

Simplifying notation

$$B = B_z = \overline{B_z} \quad (3.2)$$

3.2.2. Classical case with cartesian coordinates

If we write the Lorentz force in the S.I. system:

$$\vec{\mathbf{F}} = q\vec{v} \wedge \vec{\mathbf{B}} = m \frac{d\vec{v}}{dt}; \quad (3.3)$$

5

The equation in z is trivial and describes an inertial motion, the equations of motion in x, y are:

$$\frac{d}{dt} \begin{vmatrix} v_x \\ v_y \end{vmatrix} = \begin{vmatrix} 0 & -\omega \\ \omega & 0 \end{vmatrix} \begin{vmatrix} v_x \\ v_y \end{vmatrix} \quad (3.4)$$

This can be simplified to two independent harmonic oscillators

$$\begin{aligned} \frac{d^2}{dt^2} v_x + \omega^2 v_x &= 0 \\ \frac{d^2}{dt^2} v_y + \omega^2 v_y &= 0 \end{aligned} \quad (3.5)$$

If we introduce simple initial conditions for the v_x and v_y , i.e constant initial velocity and no

⁵in C.G.S. system q would be substituted with $\frac{q}{c}$

acceleration:

$$\begin{aligned} v_x \Big|_{t=0} &= 0; \dot{v}_x \Big|_{t=0} = 0 \\ v_y \Big|_{t=0} &= v^0; \dot{v}_y \Big|_{t=0} = 0 \end{aligned} \quad (3.6)$$

We obtain:

$$\begin{aligned} v_x &= v^0 \sin \omega t \\ v_y &= v^0 \cos \omega t \end{aligned} \quad (3.7)$$

Therefore the motion is planar and the initial angular velocity v_0 is conserved, if there is an initial velocity along the magnetic field the motion will be elicoidal in the same z direction.

3.2.3. Local non inertial ref frame

We can write the same problem in the rotating frame B together with the electron, where Ω is the rotational speed, if we decompose the acceleration:

$$\mathbf{a}_A = \mathbf{a}_B + 2\boldsymbol{\Omega} \times \mathbf{v}_B + \frac{d\boldsymbol{\Omega}}{dt} \times \mathbf{x}_B + \boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \mathbf{x}_B) . \quad (3.8)$$

The second term is the coriolis force and is zero if we assume an initial radial velocity zero, the third term is the Euler force that is also zero given is a steady rotation and Ω is constant, the last term which is the centrifugal force is instead significant, if we write the equation of motion:

$$\mathbf{F} = q\vec{v} \wedge \vec{\mathbf{B}} = m \frac{d\vec{v}}{dt} + m \left[\frac{v_\varphi^2}{\rho} \right] \hat{\rho}; \quad (3.9)$$

The treatment here becomes simple, we can assume stationary conditions (i.e. long term behaviour with no acceleration):

$$\vec{\mathbf{F}} = -qv_\rho B_z \hat{\varphi} + qv_\varphi B_z \hat{\rho} \quad (3.10)$$

$$\begin{aligned} qv_\varphi B_z &= m \frac{v_\varphi^2}{\rho} \\ -qv_\rho B_z &= 0 \end{aligned} \quad (3.11)$$

We introduce the cyclotron frequency

$$\omega = \frac{qB_z}{m} \quad (3.12)$$

and we remain again with a constant angular velocity:

$$v_\varphi = \frac{qB_z}{m} \rho = \omega \rho \quad (3.13)$$

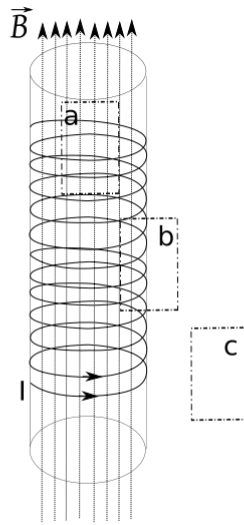


Figure 3.1: Solenoid

3.3. The field source - the solenoid

We will treat the solenoid as a field generating component classically, and the empty space surrounding it also classically, a complete treatment does need to quantize both the magnetic field sources, i.e the moving electrons, and the empty space surrounding the solenoid as a set of quantum harmonic oscillators.

3.3.1. Geometry

We start from an infinitely long vertical solenoid as drawn in figure 3.1

3.3.2. Magnetic Field of a solenoid

Applying the Ampere law to the solenoid

$$\oint B \cdot d\mathbf{l} = \mu_0 i \quad (3.14)$$

we can compute that it creates a constant internal magnetic field as is depicted in Figure 3.1, we will compute in sequence the circuitation to the loops a, b, c .

First of all following the right hand rule we can state that if the current is rotating anti clock wise looking from above the field inside the solenoid is vertical and going upwards, outside the solenoid is instead vertical and going downward.

If we apply the ampere law to a circuit a inside the solenoid there is no current through it

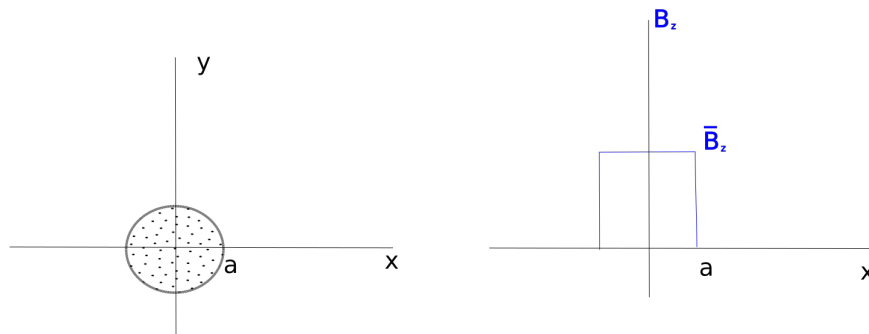


Figure 3.2: Magnetic Field

and therefore the circulation must be zero. If we compute the circulation we will have for the upper and lower side a contribution of zero given the field is orthogonal to the circuit, and two opposite cancelling contributions for the left and right side for a total of zero.

Given the loop can be as big or as small as we wish as far as it does not intersect the solenoid, the field is therefore constant for any point inside.

Considering an outer circuit c the same considerations apply and the field must be constant, if we add the hypothesis that the field is zero at large distances we end up concluding that the field must be zero for any point outside the solenoid.

Considering a circuit b that intersect the boundary there is a net flowing current through the loop given by Ni where N is the number of spirals and i is the current in the solenoid. For the circulation we have again an upper and lower part of the loop that do not contribute, a right side (i.e. external part) where the field is zero and does not contribute and an internal part which gives the only contribution. We can then apply the Ampere law:

$$Bl = \mu_0 Ni \quad (3.15)$$

Where l is the vertical length of the loop, this is the general law for the solenoids.

Summarizing we can argue that inside the solenoid the field is constant and going upwards and outside is zero as we see in figure 3.2

3.3.3. Magnetic Vector potential of a solenoid

General and homogenous solution

We want now to compute the magnetic vector potential given the magnetic field of the solenoid. We start from the definition of the vector potential which outside the solenoid is a generic equation of an irrotational field, and inside the solenoid there is constant microscopic

circulation:

$$\begin{aligned}\nabla \wedge \mathbf{A} &= 0; \rho > a \\ \nabla \wedge \mathbf{A} &= B_z; \rho \leq a\end{aligned}\tag{3.16}$$

We choose the gauge condition

$$\nabla \cdot \mathbf{A} = 0\tag{3.17}$$

We can distinguish two spatial domains of this equation one inside the solenoid and one outside, on the boundary between the two we expect the potential to be at least continuous but not differentiable given the discontinuity in the magnetic field.

We notice that the equations outside of the solenoid $\rho \geq a$ is homogenous and reduces to the equations of an ideal zero viscosity incompressible fluid where the vector potential takes the role of the velocity of the fluid. This is what is called a fluid dynamics to magnetostatic analogy, it's actually possible to build multiple type of analogies mapping the magnetic quantities to fluid quantities in a different way, typically magnetostatic analogies are based directly on the magnetic field and not on the vector potential as is done here. ⁶

$$\begin{aligned}\nabla \wedge \mathbf{A} &= 0 \\ \nabla \cdot \mathbf{A} &= 0\end{aligned}\tag{3.18}$$

Given the irrotational condition we can introduce a scalar potential $\phi_{\mathbf{A}}$ and the equation can be reduced to the Laplace equation and to classical potential theory

$$\begin{aligned}\mathbf{A} &= \nabla \phi_{\mathbf{A}} \\ \nabla^2 \phi_{\mathbf{A}} &= 0\end{aligned}\tag{3.19}$$

The problem has cylindrical symmetry, we can therefore introduce cylindrical coordinates and this equation will have cylindrical harmonics as solutions:

$$\phi_{\mathbf{A}}(\rho, \varphi, z) = \sum_n \int d|k| A_n(k) P_n(k, \rho) \Phi_n(\varphi) Z(k, z)\tag{3.20}$$

More precisely this is a linear combination in the coefficients $A_n(k)$ with a discrete quantum number n and a continuous one in k and where $P_n(k, \rho)$ cylindrical Bessel and modified Bessel functions radially. ⁷

Inside the solenoid $\rho \leq a$ the equations are linear and non homogenous, therefore the general solution will be built from a general solution of the homogenous equation plus a particular solution of the non homogenous equation. For the homogenous equation we again have a linear combination cylindrical harmonics as solutions.

⁶see for example T.E. Faber (1995). *Fluid Dynamics for Physicists* Ch. 4.5

⁷ William R. Smythe (1989). *Static and Dynamic Electricity* Ch 5.29

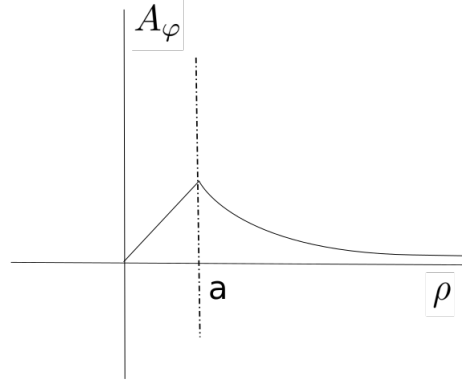


Figure 3.3: Vector potential for a Solenoid

Particular solution

For the particular solution we are left with:

$$\begin{aligned}
 A_z &= 0 \\
 \rho \leq a : A_\varphi &= \frac{\Phi}{2\pi a^2} \rho; \rho > a : A_\varphi = \frac{\Phi}{2\pi \rho} \\
 A_\rho &= 0
 \end{aligned} \tag{3.21}$$

Which is depicted in figure 3.3

Conclusions

In general the Laplace equation has a set of uniqueness theorems for dirichlet and von neumann boundary conditions which are applicable here, this domain is not simply connected, and namely punctured, therefore the uniqueness is not trivial but it is considered to be taken care by potential theory.

Outside the solenoid the magnetic field here is zero but the potential is not, consequence of this is that the wave function outside the solenoid will depend locally on the potential and non locally on the magnetic field.

Given one solution that we have found now we can generalize this with gauge transformations and find other solutions of the irrotational equation, with other gauge conditions.

Given the linearity of the equations, the role of the cylindrical harmonics sits as a background field, which is linearly superposed with the particular solution, more precisely the average of the homogenous solution will be zero, and therefore this leads to an averaged field which is the same as the particular solution.

This is the same as the infinite spindle solution (with radius a) in 2D fluid dynamics where the core of the spindle is linear, i.e. the density of the vortex is constant (e.g tornado with

simple core)^{8 9 10}

Fluid Dynamics analogy

We just solved the equations 3.18 in 2 dimensions thanks to the cylindrical symmetry, this is the same as the case of an ideal fluid in 2 dimensional potential theory. We can also expand on the magnetostatic to fluid dynamics analogy and we can intuitively compare the magnetic vector potential to the averaged velocity of a fluid, where the generated fluid velocity field is equivalent to the solutions of potential theory for a vertical line source, i.e. a spindle, spinning with a finite core. This is also what is called a Rankine model of a vortex. To make a vivid fluid dynamics image of these solutions they are the same as the ones for the case for example of an ideal tornado (or an ideal vortex) with a finite core where the edge of the tornado drags along the surrounding air.

Given the direction of the current is parallel and has the same direction of the vector potential we can also imagine intuitively that the current in the solenoid is essentially dragging the vector potential around the solenoid as it would be for a fluid dynamics case where the field source (i.e. the spindle or tornado) will drag along the fluid (i.e. the surrounding air). Historically this interpretation can be traced back to the complete drag hypothesis of Michael Faraday.

CPT argument

To show that this analogy is persistent to change of charge we introduce a CPT argument: First we reverse the charge of the carriers from positive carriers to negative electrons, then reverse the time direction therefore reverse the direction of motions for the charge carriers, the electrons will be moving downwards spiralling clock wise looking from above, and finally we need to add a parity transform.

If we add the parity to the system in the mirror image the spiral is with the opposite chirality the electrons are now moving downwards but now anticlockwise this means that if we apply the right hand rule we can see that the magnetic field remains pointing upwards. In the mirror image we need to substitute the right hand rule with the left hand rule the vector potential with a left hand rule pointing upwards is then directed clock wise looking from above, i.e. still in the same direction of the moving electrons.

If we add the parity to the coordinate reference frame one of the coordinate swap direction, at that point we are in a left handed reference frame where the wedge product is defined with a minus sign but then $\nabla \wedge B = 0 \mapsto -\nabla \wedge B = 0$ the magnetic field therefore does not reverse direction. Again the wedge product for the definition of the vector potential has an extra minus sign and the vector potential is instead with opposite direction $B = \nabla \wedge A \mapsto B = -\nabla \wedge A$ and therefore clockwise looking from above, i.e. again consistent with the electrons.

⁸ Richard Fitzpatrick (2018). *Theoretical Fluid Mechanics*

⁹ Richard Fitzpatrick (n.d.). *Two-Dimensional Vortex Filaments*. utexas.edu. [Online; accessed 2-February-2022]

¹⁰Wikipedia *Elementary flow* (n.d.). Wikipedia. [Online; accessed 2-February-2022]

A Quantum field theory perspective

Starting again from the fluid dynamics analogy, from a quantum field theory perspective we can then also interpret the vector potential as a set of gapless excitations, namely as an ideal fluid of virtual photons in the empty space surrounding the solenoid. In this context we can make also a parallel with the Casimir effect where these same virtual photons lead to measurable effects.

In this model of fluid we see that there is an averaged field which corresponds to the rankine model, on top there are also the cylindrical harmonics that we discussed previously which take the substitute role of the quantum harmonic oscillators and these solutions have also a similar status of being a fluid of virtual photons, and of being an additional background field.

We also introduced a concept of velocity, but we introduced it from an electrodynamics standpoint, therefore we cannot give a global role to this velocity as it would be for a galilean velocity, or worse a velocity with respect to an external absolute reference frame as in the case of a newtonian velocity, we can only use local reference frames and local velocities in the sense of special relativity, the best reference frame we actually have is the one solidal with the movement of the charge carriers and from there we can only define local velocities and local drag nearby to the charge carriers. This is then again consistent with the complete drag hypothesis.

Still considering velocity, there is no movement in this case given the field function do not depend on time, but this is the same as the case for the quantum harmonic oscillator wave functions, which in the standard coordinate representation don't really oscillate. By comparison we can reintroduce a semi classical concept of oscillation with coherent states, both in the case of harmonic oscillators and here for photonic states from the work of Glauber. Another way to look at this from a modern perspective is to compare it with a boson or fermi liquid where also the definition of velocity is essentially local.

It is important to distinguish drag from viscosity, a standard temperature macroscopic experimental fluid has viscosity, namely tangential viscosity, and drag is directly associated with viscosity. This is instead an ideal fluid which has a viscosity zero, and therefore we cannot associate a microscopic model of tangential forces as is typically done for macroscopic solids and liquids, a probably better intuition can be given with liquid Helium 4 which is bosonic and has viscosity zero.

Another microscopic model of this fluid that we can immediately relate is brownian motion, which is directly associated to the stochastic interpretation of the Laplace equation, but also to the Feynmann Kac formula and the stochastic interpretation of path integrals, in this model there are interactions and bumps between particles but again there is no viscosity. ^{11 12 13 14}

¹¹Faraday was convinced that Aether had complete drag (find ref.), this would be the reference frame of the moving object

¹²Despite it was the basis for Maxwell to build the equations. All mechanical models of aether failed, and were stripped off from literature

¹³Lorentz leaves to Aether his last mechanical property "immobility" Albert Einstein (1920). *Sidelights on relativity*. gutenberg.org. [Online; accessed 2-February-2022] and gives the role to Aether to conduct EM waves independently from mass motion

¹⁴Einstein comments that Special relativity removes also this "immobility" Albert Einstein (1920). *Sidelights on relativity*. gutenberg.org. [Online; accessed 2-February-2022] he gives the role of aether to the geometry of

3.4. The propagating electrons

3.4.1. Introduction

We will now treat the propagating electrons with standard quantum mechanics, a complete treatment would require also considerations about the surrounding empty space which should be modeled with quantum mechanical oscillators representing the field, and the interactions of these oscillators with the electrons.

$$\frac{1}{2m} [-i\hbar\nabla - q\mathbf{A}]^2 \psi(\vec{x}, t) = i\hbar \frac{\partial \psi}{\partial t}(\vec{x}, t) \quad (3.22)$$

$$\frac{1}{2m} \left\{ -\hbar^2 \nabla^2 + 2i\hbar q \mathbf{A} \cdot \nabla + iq\hbar [\nabla \cdot \mathbf{A}] + q^2 \mathbf{A}^2 \right\} \psi(\vec{x}, t) = i\hbar \frac{\partial \psi}{\partial t}(\vec{x}, t) \quad (3.23)$$

using gauge $\nabla \cdot \mathbf{A} = 0$

$$\frac{1}{2m} \left\{ -\hbar^2 \nabla^2 + 2i\hbar q \mathbf{A} \cdot \nabla + q^2 \mathbf{A}^2 \right\} \psi(\vec{x}, t) = i\hbar \frac{\partial \psi}{\partial t}(\vec{x}, t) \quad (3.24)$$

3.4.2. Wire geometry

Instead of solving the complete scattering problem of a beam of electron incidents towards the solenoid, in order to make the computation more simple we make a simplification and consider the electron confined to a wire and having a fixed radius as from Figure 3.4). The complete problem would be solved by separation of variables, the radial part would be a Bessel equation, but the angular part of the equations will remain essentially the same as the wire geometry and this permit us to analyze the significant part of the behaviour of the electrons. Namely the solution is composed by three parts a beam on the left and one on the right of the solenoid plus a set of bound states around the solenoid, which are also solutions of the wire geometry. These three sets are essentially coupled together and we can imagine electrons that may sit for a while circling around the solenoid and interfering with the two beams.

the electron is considered confined on the wire, and the wave function also, by simmetry the wave function now only depends on φ and t

Cylindrical:

$$\nabla f = \frac{\partial f}{\partial \rho} \hat{\rho} + \frac{1}{\rho} \frac{\partial f}{\partial \varphi} \hat{\varphi} + \frac{\partial f}{\partial z} \hat{z} \quad (3.25)$$

space time: "the hypothesis of aether in itself is not in conflict with the special theory of relativity. Only we must be on our guard against ascribing a state of motion to the Aether"

¹⁵Dirac gives the role of Aether to the ensemble of virtual particles

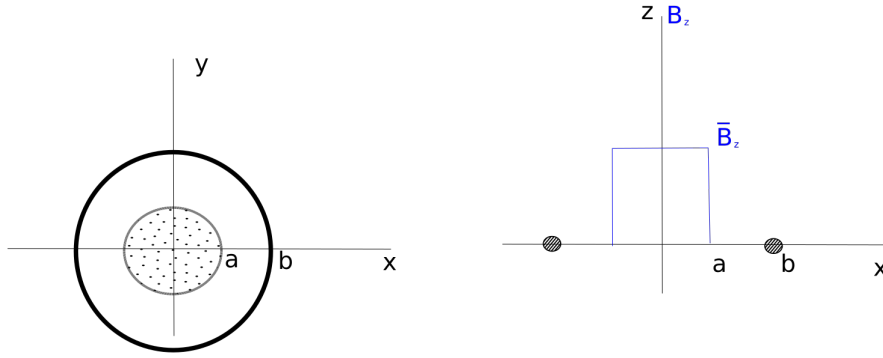


Figure 3.4: Electron on a wire

$$\nabla^2 f = \frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\rho \frac{\partial f}{\partial \rho} \right) + \frac{1}{\rho^2} \frac{\partial^2 f}{\partial \varphi^2} + \frac{\partial^2 f}{\partial z^2} \quad (3.26)$$

Here from the solutions of A before:

$$\mathbf{A} = A_\varphi \hat{\varphi} = \frac{\Phi}{2\pi b} \hat{\varphi} \quad (3.27)$$

Therefore:

$$\frac{1}{2m} \left\{ -\frac{\hbar^2}{b^2} \frac{\partial^2 \psi(\varphi, t)}{\partial \varphi^2} + 2i\hbar q A \frac{1}{b} \frac{\partial \psi(\varphi, t)}{\partial \varphi} + q^2 A^2 \psi(\varphi, t) \right\} = i\hbar \frac{\partial \psi}{\partial t}(\varphi, t) \quad (3.28)$$

And the eigenvalue equation

$$\frac{1}{2m} \left\{ -\frac{\hbar^2}{b^2} \frac{\partial^2 \psi_n(\varphi)}{\partial \varphi^2} + 2i\hbar q A \frac{1}{b} \frac{\partial \psi_n(\varphi)}{\partial \varphi} + q^2 A^2 \psi_n(\varphi) \right\} = E_n \psi_n(\varphi) \quad (3.29)$$

Given $A = \frac{\Phi}{2\pi b}$; $b > a$:

$$\frac{1}{2m} \left\{ -\frac{\hbar^2}{b^2} \frac{\partial^2 \psi_n(\varphi)}{\partial \varphi^2} + 2i\hbar \left(\frac{q\Phi}{2\pi b} \right) \frac{1}{b} \frac{\partial \psi_n(\varphi)}{\partial \varphi} + \left(\frac{q\Phi}{2\pi b} \right)^2 \psi_n(\varphi) \right\} = E_n \psi_n(\varphi) \quad (3.30)$$

$$\frac{\partial^2 \psi_n(\varphi)}{\partial \varphi^2} - 2i \left(\frac{q\Phi}{2\pi \hbar} \right) \frac{\partial \psi_n(\varphi)}{\partial \varphi} - \left(\frac{q\Phi}{2\pi \hbar} \right)^2 \psi_n(\varphi) = -\frac{2mb^2}{\hbar^2} E_n \psi_n(\varphi) \quad (3.31)$$

We introduce a first adimensional characteristic quantity which is the charge times the quantum of flux :

$$\beta = \left(\frac{q\Phi}{2\pi \hbar} \right) \quad (3.32)$$

And a second adimensional reduced energy:

$$\epsilon_n = \frac{2mb^2}{\hbar^2} E_n - \beta^2 \quad (3.33)$$

The equation reduces to

$$\frac{d^2\psi_n(\varphi)}{d\varphi^2} - 2i\beta\frac{d\psi_n(\varphi)}{d\varphi} + \epsilon_n\psi_n(\varphi) = 0 \quad (3.34)$$

This is a fixed coefficients, 2nd order ordinary differential equation, which is solved with an ansatz

$$\psi_n(\varphi) = \sum a_{nm} e^{C_m \varphi} \quad (3.35)$$

Where generically $a_{nm}, C_m \in \mathbb{C}$ and imposing normalization

$$\psi_n(\varphi) = \sum a_{nm} e^{i\lambda_m \varphi}; \lambda_m \in \mathbb{R}; \sum a_{nm} a_{mk}^\dagger = \delta_{mk} \quad (3.36)$$

Where a_{nm} is unitary.

The equation becomes second order polynomial equation in λ_n

$$-\lambda_n^2 + 2\beta\lambda_n + \epsilon_n = 0 \quad (3.37)$$

The determinant is

$$\Delta = 4\beta^2 + 4\epsilon_n \quad (3.38)$$

And the solutions:

$$\lambda_n = \beta \pm \sqrt{\beta^2 + \epsilon_n} \quad (3.39)$$

i.e.

$$\lambda_n = \left(\frac{q\Phi}{2\pi\hbar} \right) \pm \frac{b}{\hbar} \sqrt{2mE_n} \quad (3.40)$$

We impose continuity for ψ in $\phi = 0$ and $\phi = 2\pi$ and obtain

$$E_n = \frac{\hbar^2}{2mb^2} \left(n - \frac{q\Phi}{2\pi\hbar} \right)^2 \quad (3.41)$$

The positive n corresponds to particles travelling in the same direction as the current in the solenoid and they have lower energy by a factor $\frac{q\Phi}{\pi\hbar}$ than the ones travelling in the opposite direction of the current of the solenoid

3.5. Schroedinger equation invariance for local Gauge transformations

Let's assume again that our particle is outside of the solenoid where there is a non zero vector potential but a zero magnetic field, we have again

$$\begin{aligned}\nabla \wedge \mathbf{A} &= 0 \\ \mathbf{A} &= \nabla \phi_{\mathbf{A}}\end{aligned}\tag{3.42}$$

We also assume for simplicity that the magnetic potential is static. We can notice that the equation above is written in a gauge independent manner because we have a very restrictive external condition of the magnetic field to be zero, where by comparison 3.19 was not gauge independent but in a more general setting of non zero magnetic fields, but ultimately the scalar $\phi_{\mathbf{A}}$ potential generating the vector potential is the same in both.

Given the field is irrotational We can also rewrite the potential as

$$g(\mathbf{r}) = \frac{q}{\hbar} \phi_{\mathbf{A}}(\mathbf{r}) = \frac{q}{\hbar} \int_{\mathcal{O}}^{\mathbf{r}} \mathbf{A}(\mathbf{r}') \cdot d\mathbf{r}'\tag{3.43}$$

Where \mathcal{O} is starting point from which to compute the integral, this definition is only possible on irrotational fields, in the general case of non irrotational fields the potential will depend on the path along which the integral was computed from \mathcal{O} to the point \mathbf{r}

$$g(\mathbf{r}, Path(\mathbf{r})) = \frac{q}{\hbar} \int_{\mathcal{O}, Path(\mathbf{r})}^{\mathbf{r}} \mathbf{A}(\mathbf{r}') \cdot d\mathbf{r}'\tag{3.44}$$

Let's assume now that a particle is moving in such a region of space and write down the time dependent Schroedinger equation.

$$\left[\frac{1}{2m} \left(-i\hbar\nabla - q\mathbf{A} \right)^2 + V \right] \Psi = i\hbar \frac{\partial \Psi}{\partial t}\tag{3.45}$$

A special case of this is when the potential is actually an Electric field for example $V = q\varphi$ but these considerations are generic for any potential.

We can simplify the Schroedinger with the substitution:

$$\begin{aligned}\Psi &= e^{ig(\mathbf{r})} \Psi' \\ g(\mathbf{r}) &= \frac{q}{\hbar} \int_{\mathcal{O}}^{\mathbf{r}} \mathbf{A}(\mathbf{r}') \cdot d\mathbf{r}'\end{aligned}\tag{3.46}$$

And we obtain back again the Schroedinger but this time without magnetic field

$$\left[\frac{1}{2m} (-i\hbar\nabla)^2 + V \right] \Psi' = i\hbar \frac{\partial \Psi'}{\partial t}\tag{3.47}$$

Therefore if we can solve this last equation we can solve the initial one with an extra irrotational

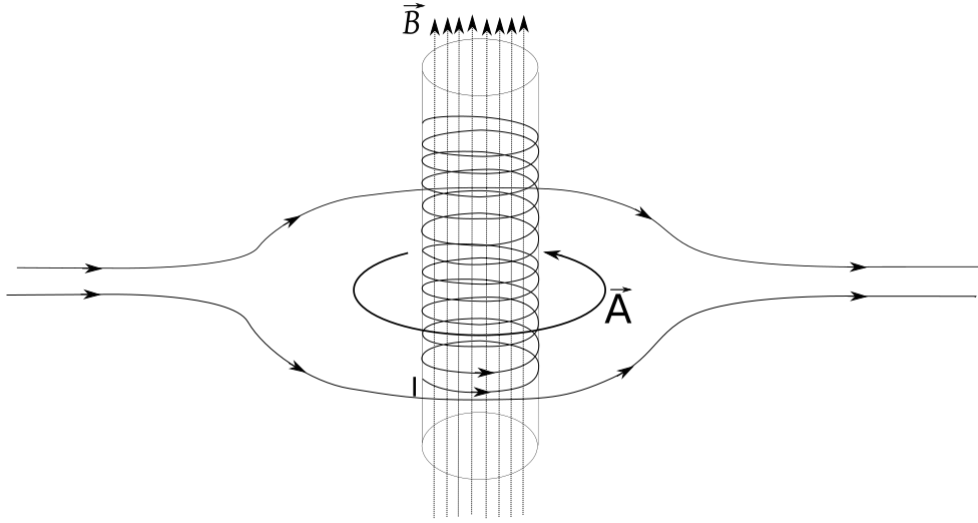


Figure 3.5: Aharonov bohm ideal experiment

vector potential, and this will come back useful later.

A second way to look at it is that we can interpret $g(\mathbf{r})$ as a local gauge transformation generating function $\zeta(\mathbf{x})$ and we invoke the fact that a gauge transformation on the field induces a gauge transform on the wave function:

$$\begin{aligned}\psi(\mathbf{x}) &\rightarrow \psi(\mathbf{x})e^{i\zeta(\mathbf{x})} \\ \mathbf{A}(\mathbf{x}) &\rightarrow \mathbf{A}(\mathbf{x}) + \partial_{\mathbf{x}}(\zeta(\mathbf{x}))\end{aligned}\tag{3.48}$$

we can essentially get rid of the vector potential field with a gauge transformation,

We can apply it here for the gauge transformation for this zero magnetic field and irrotational vector potential, but this is valid in general for non zero magnetic fields.

In order to be complete we shall mention here the parallel with Yang mills theories where a non local Gauge transformations can generate the field, and a parallel with General relativity where in the same manner the coordinate transformation by the equivalence principle are completely equivalent to the gravitational field.

3.6. Aharonov Bohm experiment

We finally consider the combined problem of sources, fields and propagating electrons and assemble the Aharonov Bohm ideal experiment. We imagine a beam of electrons propagating left to right and arranged in such a way that half of the beam goes left and half goes right of the solenoid as from figure 3.5

We now want to compute the g for two beams one going around on the left and one on the right of the solenoid, For the vector potential of the solenoid that we computed previously in 3.21. The path on the right has the electron moving in the same direction of the vector potential field and therefore the total integral is positive:

$$g_{Right} = \frac{q}{\hbar} \int \mathbf{A} \cdot d\mathbf{r} = \frac{q}{\hbar} \int_{Right} \frac{\Phi}{2\pi\rho} \hat{\phi} \cdot r \hat{\phi} d\varphi = +\frac{q\Phi}{2\hbar} \quad (3.49)$$

The path on the left has the electron moving in the opposite direction than the vector potential and the integral is negative:

$$g_{Left} = \frac{q}{\hbar} \int \mathbf{A} \cdot d\mathbf{r} = \frac{q}{\hbar} \int_{Left} \frac{\Phi}{2\pi\rho} \hat{\phi} \cdot r \hat{\phi} d\varphi = -\frac{q\Phi}{2\hbar} \quad (3.50)$$

and the difference between the two is proportional to the magnetic field included between the paths

$$g_{Right} - g_{Left} = \frac{q\Phi}{\hbar} \quad (3.51)$$

We can call this a phase shift given the same gauge transformation also induces a phase shift in the wave function

$$\psi(\mathbf{x}) \mapsto e^{iq\frac{\Phi}{\hbar}} \psi(\mathbf{x}) \quad (3.52)$$

the phenomenon is measurable and was measured first by Chambers and others later providing better and better evidence of the complete shielding from the magnetic field.¹⁶

To visualize a concrete experiment we depicted it in figure 3.6 we shall imagine a double slit experiment where in the middle of the two slits just behind the screen we positioned a vertical solenoid. Given outside of the solenoid the field is zero but the magnetic potential is not we do have the Aharonov Bohm effect. More precisely we can imagine the wave function is combined by two parts one passing through the right slit and one through the left slit and

$$|\psi\rangle = |Left\rangle + |Right\rangle \quad (3.53)$$

Once the solenoid current is switched on an additional phase shift $\frac{e\Phi}{\hbar}$ will be introduced between the two waves and the interference pattern will change.

For further online resources to learn about the Aharanov Bohm effect one can see¹⁷¹⁸

¹⁶ R.G Chambers (1960). “Shift of an Electron Interference Pattern by Enclosed Magnetic Flux”. In: *Physical Review Letters* 5 (1). doi:10.1103/PhysRevLett.5.3, pp. 3–5

¹⁷ Yakir Aharonov (2011). *Nonlocal Phenomena in Quantum Mechanics*. YouTube. [Online; accessed 2-February-2022]

¹⁸ Murray Peshkin (n.d.). *Things I Do and Do Not Understand About the Aharonov-Bohm Effect*. YouTube. [Online; accessed 2-February-2022]

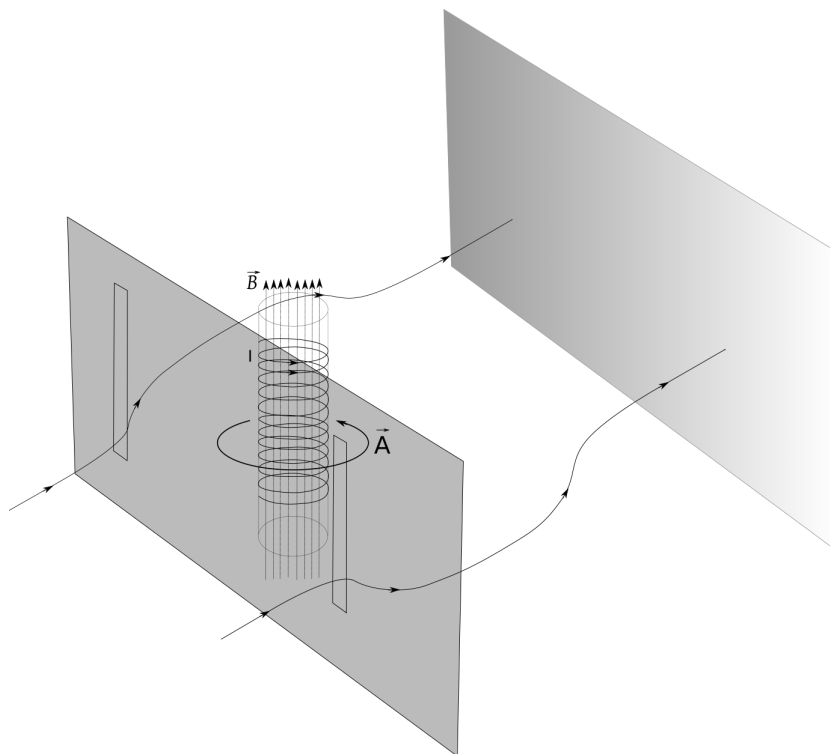


Figure 3.6: Aharonov Bohm experimental setup

3.7. Flux quantization

Let's now analyze more in detail the integral:

$$\frac{q}{\hbar} \int_{Path} \mathbf{A} \cdot d\mathbf{r} \quad (3.54)$$

This quantity is a type of electromagnetic phase, i.e. phase difference of the vector potential along a path, same as the Berry phase that we will see later. This integral needs to be analyzed along different paths.

Let's consider as our integration contour a closed path of the particle making a circle around the solenoid, the region is not simply connected, we already computed the two halves of the integral earlier in 3.51, therefore:

$$\frac{q}{\hbar} \oint \mathbf{A} \cdot d\mathbf{r} = \frac{q\Phi}{\hbar} \quad (3.55)$$

And the wave function after one loop around the solenoid changes by a constant phase as we have seen in 3.52

As far as we consider this integral on a closed contour this quantity is Gauge invariant, but in somewhat of a peculiar manner. If we now introduce a Gauge transformation:

$$\begin{aligned} \psi(\mathbf{x}) &\rightarrow \psi(\mathbf{x})e^{i\zeta(\mathbf{x})} \\ A(\mathbf{x}) &\rightarrow A(\mathbf{x}) + \partial_{\mathbf{x}}(\zeta(\mathbf{x})) \end{aligned} \quad (3.56)$$

we obtain:

$$\frac{q}{\hbar} \oint \mathbf{A}' \cdot d\mathbf{r} = \frac{q}{\hbar} \int \mathbf{A}(\mathbf{x}) \cdot d\mathbf{r} + \frac{q}{\hbar} \oint \partial_{\mathbf{x}}(\zeta(\mathbf{x})) \cdot d\mathbf{r} = \frac{q}{\hbar} \oint \mathbf{A}(\mathbf{x}) \cdot d\mathbf{r} - [\zeta(\mathbf{x}_i) - \zeta(\mathbf{x}_f)] \quad (3.57)$$

The quantity $[\zeta(\mathbf{x}_i) - \zeta(\mathbf{x}_f)]$ is called boundary term, given it is the difference between two values on the boundary of the integration interval and it generalizes in multiple dimensions to a generic integral on the boundary of the integration domain. Now there are two cases for the boundary term in fact ζ can be single valued and multivalued. In the case ζ is single valued the quantity is zero and therefore the phase is gauge invariant.

In the second case we start from the hypothesis that the wave function is single valued:

$$\psi'(\mathbf{x}_i) = \psi'(\mathbf{x}_f) \quad (3.58)$$

If we introduce a Gauge transform:

$$\psi(\mathbf{x}_i)e^{i\zeta(\mathbf{x}_i)} = \psi(\mathbf{x}_f)e^{i\zeta(\mathbf{x}_f)} \quad (3.59)$$

And if we use again the hypothesis of being single valued 3.58 on $\psi(\mathbf{x})$ we obtain:

$$\zeta(\mathbf{x}_i) = \zeta(\mathbf{x}_f) + 2\pi n; n \in \mathbb{Z} \quad (3.60)$$

Therefore we can summarize the two cases in the statement

$$\frac{q}{\hbar} \oint \mathbf{A}' \cdot d\mathbf{r} = \frac{q}{\hbar} \oint \mathbf{A} \cdot d\mathbf{r} + 2\pi n; n \in \mathbb{Z} \quad (3.61)$$

In general the Aharonov Bohm phase is not quantized, and the above is the maximum we can say once there is a gauge transformation.

Which is always true if the phase is quantized:

$$\frac{q}{\hbar} \oint \mathbf{A} \cdot d\mathbf{r} = 2\pi n; n \in \mathbb{Z} \quad (3.62)$$

Rewording the computation process we can state that the single value condition of the wave function together with the Gauge invariance induces a quantization condition on the phase. Another even quicker way to rephrase it is also to say that the boundary term drops off (or better becomes degenerate) when the contour is closed, and the gauge invariance is still preserved.

This is also called flux quantization because for a generic region of space we can apply stokes theorem and transform the circuitation integral into a flux integral on the magnetic field \mathbf{B}

$$\frac{q}{\hbar} \oint \mathbf{A} \cdot d\mathbf{r} = \frac{q}{\hbar} \int (\nabla \wedge \mathbf{A}) \cdot \mathbf{n} dS = \frac{q}{\hbar} \int \mathbf{B} \cdot \mathbf{n} dS = \frac{q}{\hbar} \Phi \quad (3.63)$$

And we can rewrite it as a quantization condition on the magnetic flux ¹⁹

$$\frac{q}{\hbar} \Phi = 2\pi n; n \in \mathbb{Z} \quad (3.64)$$

We will see this same mathematics of the boundary term is happening in multiple places. Specifically in the Berry phase and in the TKNN formula, different physics but same mathematics with the boundary term which becomes quantized and drops out for closed loops.

¹⁹ *Flux Quantization and the Aharonov-Bohm Effect* (n.d.). fitzpatrick. [Online; accessed 2-February-2022]

References

- Aharonov Y; Bohm, D (1959). “Significance of electromagnetic potentials in quantum theory”.
In: *Physical Review* 115 (3). doi:10.1103/PhysRev.115.485, pp. 485–491.
- (1961). “Further Considerations on Electromagnetic Potentials in the Quantum Theory”.
In: *Physical Review* 123 (4). doi:10.1103/PhysRev.123.1511, pp. 1511–1524.
- Chambers, R.G (1960). “Shift of an Electron Interference Pattern by Enclosed Magnetic Flux”.
In: *Physical Review Letters* 5 (1). doi:10.1103/PhysRevLett.5.3, pp. 3–5.
- D.J., Griffiths (2005). *Introduction to quantum mechanics*.
- Ehrenberg W; Siday, RE (1949). “The Refractive Index in Electron Optics and the Principles of Dynamics”. In: *Proceedings of the Physical Society B*. 62 (1). doi:10.1088/0370-1301/62/1/303, pp. 8–21.
- Faber, T.E. (1995). *Fluid Dynamics for Physicists*.
- Fitzpatrick, Richard (2018). *Theoretical Fluid Mechanics*.
- Smythe, William R. (1989). *Static and Dynamic Electricity*.

Readings and online resources

Aharonov, Yakir (2011). *Nonlocal Phenomena in Quantum Mechanics*. YouTube. [Online; accessed 2-February-2022].

Einstein, Albert (1920). *Sidelights on relativity*. gutenberg.org. [Online; accessed 2-February-2022].

Elementary flow (n.d.). Wikipedia. [Online; accessed 2-February-2022].

Fitzpatrick, Richard (n.d.). *Two-Dimensional Vortex Filaments*. utexas.edu. [Online; accessed 2-February-2022].

Flux Quantization and the Aharonov-Bohm Effect (n.d.). fitzpatrick. [Online; accessed 2-February-2022].

Peshkin, Murray (n.d.). *Things I Do and Do Not Understand About the Aharonov-Bohm Effect*. YouTube. [Online; accessed 2-February-2022].

4 — Berry phase

4.1. Introduction

As we have seen for the Aharonov Bohm effect, we can associate the electromagnetic vector potential to physical consequences, as far as all physical consequences are bound to be gauge independent. The electromagnetic vector potential is therefore a vector field and a set of transformations for which the field changes. The physical consequences of the electromagnetic potential do not change, as far as the electromagnetic field is the same.

We are familiar to this situation for example in special relativity, where all laws of nature are the same for all inertial reference frames and for all coordinate transformations between them. In the same manner, we can state that the laws of electromagnetism and all physical consequences are the same for any gauge transformations.

In this sense, we then consider the electromagnetic potential as a legitimate basic fundamental field. We can also state that as far the underlying group is a local $U(1)$ transformation, this can be parametrized with a local angle or phase, and this can be considered a local degree of freedom.

The defining property of the vector potential actually is gauge invariance, and as we have seen also in the Aharonov-Bohm chapter, gauge invariance directly implies a local $U(1)$ transformation on the wave function and again a local degree of freedom which is now the phase of the wave function. Therefore the wave function is deeply interlinked with the electromagnetic potential through the gauge transformation.

It is now possible from the wavefunctions and the eigenstates to build a path dependent integral, which is called the Berry phase¹. If the path is closed, the Berry phase will be gauge invariant and will be an observable. The Berry phase is a type of electromagnetic phases, and there are others similar electromagnetic phases which are path dependent integrals of the electromagnetic field. Those have become influential in condensed matter physics and beyond².

It is also possible to build two geometric quantities from the wave function. Namely a first differential 1-form which will be the Berry connection, and a second differential 2-form which will be the Berry curvature, from these it is then possible to build the Berry phase³. The relation between connection and curvature is in close analogy with what is possible to do with the electromagnetic potential and the electromagnetic field. We will see how these quantities are

¹ Michael Victor Berry (Mar. 1984). “Quantal phase factors accompanying adiabatic changes”. In: 392 (1802)

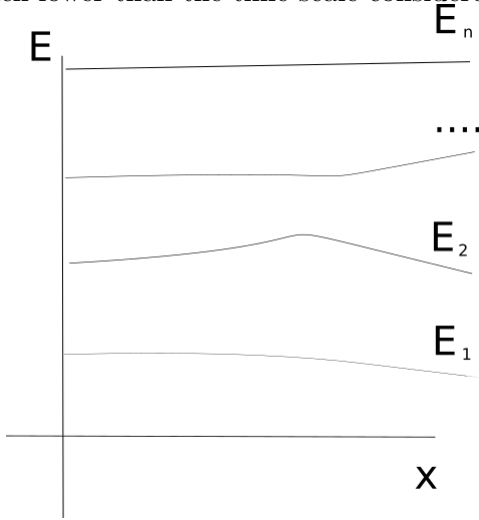
² Bouchard F. et al. Cohen E. Larocque H. (2019). “Geometric phase from Aharonov–Bohm to Pancharatnam–Berry and beyond”. In: *Nat Rev Phys* (1), pp. 437–449

³ Barton Zwiebach (2018b). *L17.2 Berry’s phase and Berry’s connection. MIT 8.06 Quantum Physics III*. <https://www.youtube.com/watch?v=iGG9EG3SNz0>. [Online; accessed 12-October-2021]

gauge invariant, that their integrals can be measured and that they have a complex interplay between gauge invariance and quantization.

4.2. Adiabatic theorem

Let's start from a Hamiltonian $H(\mathbf{R}(t))$ that depends on the parameters \mathbf{R} which are slowly varying in time. The definition of slow here is defined by the distance in energy by the eigenstates. Through the uncertainty principle, we can define a timescale that shall be always much lower than the time scale considered.



This way we clearly also identify that while slowly varying the parameters, the energy of the eigenstates also changes, but the eigenstates remain clearly separated in energy. We will then generalize this to the case of bands and these bands shall remain clearly separated, or in other words, gapped. Given these bands do not intersect, the states are ordered, and together with a robust degeneracy of the bands such as the one of Landau levels, this will lead to the concept of topological order and to topological insulators.

We start from the instantaneous Schroedinger equation:

$$H(\mathbf{R}(t))|\psi_m(t)\rangle = E_m(t)|\psi_m(t)\rangle, \quad (4.1)$$

and its instantaneous eigenstates:

$$\langle\psi_m(t)|\psi_n(t)\rangle = \delta_{mn}. \quad (4.2)$$

A generic state can be expanded as:

$$|\Psi(t)\rangle = \sum a_n(t)|\psi_n(t)\rangle. \quad (4.3)$$

When we plug $|\Psi(t)\rangle$ in the full Schroedinger we obtain

$$i\hbar\partial_t|\Psi(t)\rangle = \hat{H}(\mathbf{R}(t))|\Psi(t)\rangle. \quad (4.4)$$

Finally multiplying by a generic eigenvector:

$$\langle \psi_m(t) | i\hbar \partial_t | \Psi(t) \rangle = \langle \psi_m(t) | H(\mathbf{R}(t)) | \Psi(t) \rangle. \quad (4.5)$$

Expanding $|\Psi(t)\rangle$ into a linear combination of eigenstates:

$$\dot{a}_m + \sum_n \langle \psi_m(t) | \partial_t | \psi_n(t) \rangle a_n = -\frac{i}{\hbar} E_m(t) a_m. \quad (4.6)$$

If we introduce the adiabatic approximation

$$|\langle \psi_m(t) | \partial_t | \psi_n(t) \rangle a_n| \ll |a_m|, \quad \forall m \neq n, \quad (4.7)$$

we have

$$\dot{a}_m \simeq -\langle \psi_m(t) | \partial_t | \psi_m(t) \rangle a_m - \frac{i}{\hbar} E_m(t) a_m \quad (4.8)$$

and

$$a_m(t) = e^{-\frac{i}{\hbar} \int_{t_0}^t E_m(t') dt'} e^{i\gamma_m(t)} a_m(t_0). \quad (4.9)$$

From (4.9) we identify the dynamic phase as:

$$\theta_m(t) = -\frac{1}{\hbar} \int_{t_0}^t E_m(t') dt'. \quad (4.10)$$

The second phase can be expanded into:

$$\gamma_m(t) = i \int_{t_0}^t \langle \psi_m(t') | \partial_{t'} | \psi_m(t') \rangle dt' = i \int_{t_0}^t \langle \psi_m(\mathbf{R}) | \partial_{\mathbf{R}} | \psi_m(\mathbf{R}) \rangle \dot{\mathbf{R}} dt' = i \int_C \langle \psi_m(\mathbf{R}) | \partial_{\mathbf{R}} | \psi_m(\mathbf{R}) \rangle d\mathbf{R}, \quad (4.11)$$

where we removed the time integral with a change of variables in the last step. Given we removed the dependency with time, it may take a microsecond or a million years to move along the path but still the integral will be the same, and the integral just depends on the path. We thus write $\gamma_m[C]$ instead of $\gamma_m(t)$ to emphasize it.

This phase is called Berry phase and is defined as:

$$\gamma_m[C] = i \int_C \langle \psi_m(\mathbf{R}) | \partial_{\mathbf{R}} | \psi_m(\mathbf{R}) \rangle d\mathbf{R}, \quad (4.12)$$

where C is the path (either open or closed) in the parameter space taken by \mathbf{R} . It is important to note that being $\gamma_m[C]$ the argument of an exponential, the Berry phase is real.

To summarize: if a wave function starts in a eigenstate $\psi_m(t_0)$ it will remain in the same

state acquiring a set of phase factors:

$$\psi_m(t) = e^{i\theta_m(t)} e^{i\gamma_m(t)} \psi_m(t_0) \quad (4.13)$$

This is the statement of the adiabatic theorem and we can phrase it in terms of the full wave function as in (4.13) or in terms of the coefficients of the total wave function and its initial state as in (4.9).

4.3. Berry Phase

4.3.1. Definition

We now want to analyze some properties of the Berry Phase, and we recall its general definition:

$$\gamma_m[C] = i \int_{\mathbf{R}_i, C}^{\mathbf{R}_f} \langle \psi_m(\mathbf{R}) | \partial_{\mathbf{R}} \psi_m(\mathbf{R}) \rangle d\mathbf{R}, \quad (4.14)$$

where we considered a path C and an initial point \mathbf{R}_i and a final one \mathbf{R}_f .

The Berry phase (4.14) in the general case is path dependent, and therefore it acts as a geometric quantity. More precisely, the integrand, which is called the Berry connection, is a local geometric quantity of the parameter space.

Typically, the Berry phase is analyzed for holonomic systems, where the initial state is the same as the final state, therefore the path in the parameter space is closed:

$$\gamma_m[C] = i \oint_C \langle \psi_m(\mathbf{R}) | \partial_{\mathbf{R}} \psi_m(\mathbf{R}) \rangle d\mathbf{R}, \quad (4.15)$$

however the phase still does depend on the path.

As we will see later, the Berry connection not only is gauge invariant, it is also a physical quantity, which can be in principle be measured along a specific path of the parameter space. For example, in the case of optics the parameter space can be the physical space of coordinates, and the path can be a the real physical path of the ray tracing approximation.⁴

4.3.2. Magnetic analogy for the Berry Phase

In the case of a space parameter of 3 dimensions, we can make an analogy between the Berry Phase and the magnetic field⁵. We start from the definition of the Berry connection:

$$\mathbf{\Gamma}_n(\mathbf{R}) = i \langle \psi_n(\mathbf{R}) | \partial_{\mathbf{R}} \psi_n(\mathbf{R}) \rangle. \quad (4.16)$$

⁴ Bouchard F. et al. Cohen E. Larocque H. (2019). “Geometric phase from Aharonov–Bohm to Pancharatnam–Berry and beyond”. In: *Nat Rev Phys* (1), pp. 437–449

⁵ Griffiths D.J. (2005). *Introduction to quantum mechanics* pp 381

If we associate the Berry connection to the magnetic vector potential, we will see here that we can retrieve consistently the properties of the magnetic field, and later we will also see that this is rooted in the gauge invariance. The formal identification is the following

$$\Gamma_n(\mathbf{R}) \rightarrow \mathbf{A} \quad (4.17)$$

. We can then write an analogy with the magnetic field

$$\nabla_{\mathbf{R}} \wedge \Gamma_n(\mathbf{R}) \rightarrow \mathbf{B} \quad (4.18)$$

and a second analogy with the magnetic flux :

$$\gamma_n[C] = \oint_C \Gamma_n(\mathbf{R}) d\mathbf{R} \rightarrow \Phi. \quad (4.19)$$

We can see that the analogy is consistent with the definition of magnetic flux, in fact:

$$\Phi = \int \mathbf{B} d\mathbf{S} = \int \nabla \wedge \mathbf{A} d\mathbf{S} = \oint \mathbf{A} d\mathbf{R}. \quad (4.20)$$

and the analogous quantities:

$$\gamma_n[C] = \int \nabla_{\mathbf{R}} \wedge \Gamma_n(\mathbf{R}) dS = \oint_C \Gamma_n(\mathbf{R}) d\mathbf{R}. \quad (4.21)$$

4.4. How Aharonov relates to Berry phase

As Berry pointed out, the Aharonov Bohm effect can also be understood as an example of a geometric phase. We can recall the setup in the Aharonov Bohm experiment and start from a particle confined in a box around a physical point \mathbf{R} outside the solenoid. We write down the Schroedinger equation:

$$\frac{1}{2m} \left[\left(-i\hbar\nabla - q\mathbf{A} \right)^2 + V(\mathbf{r} - \mathbf{R}) \right] \psi_n = E_n \psi_n. \quad (4.22)$$

where \mathbf{R} is a parameter that spans the coordinate space.

We can get rid of the vector potential with a local gauge transform $g(\mathbf{r}, \mathbf{R})$

$$\begin{aligned} \psi_n(\mathbf{r}, \mathbf{R}) &= e^{ig} \psi'_n(\mathbf{r} - \mathbf{R}) \\ g(\mathbf{r}, \mathbf{R}) &= \frac{q}{\hbar} \int_{\mathbf{R}}^{\mathbf{r}} \mathbf{A}(\mathbf{r}') d\mathbf{r}' \end{aligned} \quad (4.23)$$

where we use as the reference point to compute the integral the center of the box \mathbf{R} . We get

$$\frac{1}{2m} \left[\left(-i\hbar\nabla \right)^2 + V(\mathbf{r} - \mathbf{R}) \right] \psi'_n = E_n \psi'_n. \quad (4.24)$$

Let's now carry the Box around the solenoid by letting the parameter \mathbf{R} make a circle around the solenoid, and compute the Berry phase.

First we compute:

$$\begin{aligned} \nabla_{\mathbf{R}} \psi_n &= \nabla_{\mathbf{R}} \left[e^{ig} \psi'_n(\mathbf{r} - \mathbf{R}) \right] = i[\nabla_{\mathbf{R}} g] e^{ig} \psi'_n(\mathbf{r} - \mathbf{R}) + e^{ig} \nabla_{\mathbf{R}} \psi'_n(\mathbf{r} - \mathbf{R}) = \\ &= -i\frac{q}{\hbar} \mathbf{A}(\mathbf{R}) e^{ig} \psi'_n(\mathbf{r} - \mathbf{R}) + e^{ig} \nabla_{\mathbf{R}} \psi'_n(\mathbf{r} - \mathbf{R}), \end{aligned} \quad (4.25)$$

then

$$\begin{aligned} \langle \psi_n | \nabla_{\mathbf{R}} \psi_n \rangle &= \int d^3\mathbf{r} e^{-ig} [\psi'_n(\mathbf{r} - \mathbf{R})]^* e^{ig} \left[-i\frac{q}{\hbar} \mathbf{A}(\mathbf{R}) \psi'_n(\mathbf{r} - \mathbf{R}) + \nabla_{\mathbf{R}} \psi'_n(\mathbf{r} - \mathbf{R}) \right] = \\ &= -i\frac{q}{\hbar} \mathbf{A}(\mathbf{R}) + \int d^3\mathbf{r} [\psi'_n(\mathbf{r} - \mathbf{R})]^* [\nabla_{\mathbf{R}} \psi'_n(\mathbf{r} - \mathbf{R})] \end{aligned} \quad (4.26)$$

The last integral is the average expectation value of the momentum for an eigenstate of the Hamiltonian of (4.22). Given $|\psi_n\rangle$ is spatially localized,

$$- \int d^3\mathbf{r} [\psi'_n(\mathbf{r} - \mathbf{R})]^* [\nabla \psi'_n(\mathbf{r} - \mathbf{R})] = \frac{i}{\hbar} \langle \psi_n | \hat{\mathbf{p}} | \psi_n \rangle = 0, \quad (4.27)$$

this is actually zero. Collecting the pieces we now have:

$$\langle \psi_n | \nabla_{\mathbf{R}} \psi_n \rangle = -i\frac{q}{\hbar} \mathbf{A}(\mathbf{R}). \quad (4.28)$$

Finally we use the Berry phase definition:

$$\gamma_n[C] = i \oint_C \langle \psi_n | \nabla_{\mathbf{R}} \psi_n \rangle \cdot d\mathbf{R}, \quad (4.29)$$

to compute the integral as:

$$\gamma_n[C] = \frac{q}{\hbar} \oint_C \mathbf{A}(\mathbf{R}) \cdot d\mathbf{R}. \quad (4.30)$$

We apply Stokes theorem

$$\gamma_n[C] = \frac{q}{\hbar} \int_{\Sigma} \nabla \wedge \mathbf{A}(\mathbf{R}) dS = \frac{q}{\hbar} \int_{\Sigma} \mathbf{B} dS, \quad (4.31)$$

where C is the border of the surface Σ . In the last step we used the definition of the magnetic field. Now we can use the definition of flux across the boundary C which encloses the solenoid $\Phi = \int_{\Sigma} \mathbf{B} dS$, and where the non null contribution is only the one inside the solenoid. This leads

to:

$$\gamma_n = \frac{q}{\hbar} \Phi \quad (4.32)$$

which is the same as the result of the Aharonov Bohm calculation.

4.5. Berry connection

4.5.1. Definition

We introduced the Berry connection earlier, it is now time to analyze it better. We now define the Berry connection or Berry vector potential:

$$\mathbf{A}_n(\mathbf{R}) = i \langle \psi_n(\mathbf{R}) | \partial_{\mathbf{R}} | \psi_n(\mathbf{R}) \rangle \quad (4.33)$$

where \mathbf{R} belongs to a m -dimensional generic parameter manifold with eventually a non-trivial topology. We can also introduce the index notation for the Berry connection:

$$A_\nu^n(\mathbf{R}) = i \langle \psi_n(\mathbf{R}) | \frac{\partial}{\partial R^\nu} | \psi_n(\mathbf{R}) \rangle \quad (4.34)$$

where ν is a vector index, and n is just a generic index that does not transform as a vector. As a first property, it is important to stress that since the Berry phase is real also the Berry connection is real.

The fact that is called a connection is a terminology from differential geometry, which means that it transforms according to a set of rules after an underlying transformation in the parameter space. In the current case of the Berry phase, the underlying transformation is actually a gauge transformation.

It is not an invariant with the set of transformations as is the case for tensors, but it still transforms consistently. A definition of a connection is usually algebraic, but is dual to a geometric notion of parallel transport and it is also dual to a geometric notion of tangent bundles. Sometimes for connections we use the symbol Γ to underline the relationship with the Christoffel symbols $\Gamma_{\nu\rho}^\mu$ of General Relativity. These are also defining algebraically a connection, and in that case the underlying transformations are space-time coordinate transformations.

4.5.2. Gauge invariance

We informally associated the Berry connection to the vector potential in (4.17) and showed that we can derive an analogy with the Magnetic field. This was not by chance, in fact, the Berry connection has a gauge like freedom, and this is the sole defining property of a vector potential. To be specific by applying a gauge transformation $e^{i\zeta(\mathbf{R})}$ to the wave function

$$\psi_n(\mathbf{R}) \rightarrow \psi_n(\mathbf{R}) e^{i\zeta(\mathbf{R})}, \quad (4.35)$$

the vector potential changes as:

$$A_n(\mathbf{R}) \rightarrow A_n(\mathbf{R}) - \partial_{\mathbf{R}}(\zeta(\mathbf{R})). \quad (4.36)$$

In fact,

$$\begin{aligned} A_n(\mathbf{R})' &= i\langle\psi_n(\mathbf{R})|e^{i\zeta(\mathbf{R})}\partial_{\mathbf{R}}|\psi_n(\mathbf{R})e^{i\zeta(\mathbf{R})}\rangle = ie^{-i\zeta(\mathbf{R})}\langle\psi_n(\mathbf{R})|\partial_{\mathbf{R}}|\psi_n(\mathbf{R})\rangle e^{i\zeta(\mathbf{R})} + \\ &ie^{-i\zeta(\mathbf{R})}\langle\psi_n(\mathbf{R})|\psi_n(\mathbf{R})\rangle e^{i\zeta(\mathbf{R})}\partial_{\mathbf{R}}(i\zeta(\mathbf{R})) = A_n(\mathbf{R}) - \partial_{\mathbf{R}}(\zeta(\mathbf{R})) \end{aligned} \quad (4.37)$$

As a notation we will use $\mathbf{R}_i, \mathbf{R}_f$ for the initial and final states, where we usually consider a long time period T such that we return to the original state. Therefore the initial and final states will be equivalent to:

$$\begin{aligned} \mathbf{R}_i &= \mathbf{R}(0) \\ \mathbf{R}_f &= \mathbf{R}(T) \end{aligned} \quad (4.38)$$

We will now prove that on a closed path the Berry phase (i.e the integral of the connection) is gauge invariant. Let's consider how the Berry phase changes when a gauge transformation is applied:

$$\gamma_n' = \oint_C \mathbf{A}_n' \cdot d\mathbf{R} = \oint_C \mathbf{A}_n \cdot d\mathbf{R} - \oint_C \partial_{\mathbf{R}}(\zeta(\mathbf{R}))d\mathbf{R} = \oint_C \mathbf{A}_n \cdot d\mathbf{R} - [\zeta(\mathbf{R}_i) - \zeta(\mathbf{R}_f)] \quad (4.39)$$

Now on a closed path $\mathbf{R}_i = \mathbf{R}_f$ and there are two cases for $[\zeta(\mathbf{R}_i) - \zeta(\mathbf{R}_f)]$. These are when ζ is single-valued or multivalued. In the most simple case $\zeta(\mathbf{R})$ is smooth and single valued, therefore $[\zeta(\mathbf{R}_i) - \zeta(\mathbf{R}_f)] = 0$ and the phase is gauge invariant.

In the second case ζ can be multivalued, given ζ is not a physical quantity this is perfectly possible. When $\zeta(\mathbf{R})$ is multivalued, there can be a case where we can still maintain gauge invariance of the eigenstates, and of the Berry phase itself. In these cases we will obtain that the Berry phase is only defined up to 2π .

In fact if we consider the eigenstates single valued, before and after the parameter variation we shall return to the same value:

$$|\mathbf{n}(\mathbf{R}(T))\rangle = |\mathbf{n}(\mathbf{R}(0))\rangle. \quad (4.40)$$

We gauge transform on both sides

$$|\mathbf{n}(\mathbf{R}(T))e^{i\zeta(\mathbf{R}(T))}\rangle = |\mathbf{n}(\mathbf{R}(0))e^{i\zeta(\mathbf{R}(0))}\rangle \quad (4.41)$$

where we can extract the exponents cause these are complex constants

$$e^{i\zeta(\mathbf{R}(T))}|\mathbf{n}(\mathbf{R}(T))\rangle = e^{i\zeta(\mathbf{R}(0))}|\mathbf{n}(\mathbf{R}(0))\rangle \quad (4.42)$$

and we get

$$\zeta(\mathbf{R}(T)) - \zeta(\mathbf{R}(0)) = 2\pi m \quad (4.43)$$

for $m \in \mathbb{Z}$, and this confirms that ζ is multivalued.

Therefore to still have gauge invariance in these cases we need to have an extra condition on the Berry phase. This condition is:

$$\oint \mathbf{A}'_n d\mathbf{R} = \oint \mathbf{A}_n d\mathbf{R} - 2\pi m, \quad (4.44)$$

which tells us that the Berry phase is only well defined up to modulo 2π , or that you can always do a gauge transformation to shift the Berry phase by a constant value multiple of 2π . It is important to distinguish this condition from a quantization condition given the phase itself can assume any value. This is sometimes also written as:

$$\gamma_n[C](\psi) = \gamma_n[C](\psi') \pmod{2\pi}. \quad (4.45)$$

4.6. Berry curvature

4.6.1. Definition

We consider now a closed loop in parameter space. From this loop we can evaluate the holonomy of the connection, i.e. how much parallel transport along the loop in the parameter space does not return to the original state. Therefore in a very long time T , we will slowly change the parameters and we will come back to the original state. Given the loop is closed we can use the Stokes theorem:

$$\oint_C \mathbf{v} \cdot d\mathbf{s} = \int_A \nabla \wedge \mathbf{v} \cdot d\mathbf{A} = \int_A \epsilon_{ijk} \nabla_j v_k dS_i. \quad (4.46)$$

when we apply it to the Berry connection, we get:

$$\gamma_m = i \oint_C \langle \psi_m(\mathbf{R}) | \partial_{\mathbf{R}} | \psi_m(\mathbf{R}) \rangle \cdot d\mathbf{R} = i \int_A \nabla_{\mathbf{R}} \wedge \langle \psi_m(\mathbf{R}) | \partial_{\mathbf{R}} | \psi_m(\mathbf{R}) \rangle \cdot d\mathbf{S} \quad (4.47)$$

$$\begin{aligned} \gamma_m &= i \oint_A dS_i \epsilon_{ijk} \frac{\partial}{\partial R^j} \langle \psi_n(\mathbf{R}) | \frac{\partial}{\partial R^k} | \psi_n(\mathbf{R}) \rangle \\ &= i \oint_A dS_i \epsilon_{ijk} [\langle \frac{\partial}{\partial R^j} \psi_n(\mathbf{R}) | \frac{\partial}{\partial R^k} | \psi_n(\mathbf{R}) \rangle + \langle \psi_n(\mathbf{R}) | \frac{\partial^2}{\partial R^j \partial R^k} | \psi_n(\mathbf{R}) \rangle]. \end{aligned} \quad (4.48)$$

Now the second term is symmetric in i and j and disappears as far as the metric of the parameter space is flat i.e.:

$$[\frac{\partial}{\partial R^j}, \frac{\partial}{\partial R^k}] = 0 \quad (4.49)$$

Summarizing we obtain:

$$\gamma_n = i \oint_A dS_i \epsilon_{ijk} \langle \frac{\partial}{\partial R^j} \psi_n(\mathbf{R}) | \frac{\partial}{\partial R^k} \psi_n(\mathbf{R}) \rangle. \quad (4.50)$$

We can now introduce the Berry curvature as the integrand:

$$\Omega_{\mu\nu}^n(\mathbf{R}) = \frac{\partial}{\partial R^\mu} A_\nu^n(\mathbf{R}) - \frac{\partial}{\partial R^\nu} A_\mu^n(\mathbf{R}) \quad (4.51)$$

in such a way that:

$$\gamma_n = \int \boldsymbol{\Omega} \cdot d\mathbf{S} \quad (4.52)$$

4.6.2. Magnetic analogy

To see explicitly the relationship with magnetism, we can rewrite the integral (4.50) into:

$$\gamma_n = i \oint_A \langle \nabla_{\mathbf{R}} \psi_n(\mathbf{R}) | \wedge | \nabla_{\mathbf{R}} \psi_n(\mathbf{R}) \rangle d\mathbf{S} \quad (4.53)$$

where we introduced a shorthand vectorial notation for the Berry curvature:

$$\begin{aligned} \boldsymbol{\Omega}_n(\mathbf{R}) &= \langle \nabla_{\mathbf{R}} \psi_n(\mathbf{R}) | \wedge | \nabla_{\mathbf{R}} \psi_n(\mathbf{R}) \rangle \\ \Omega_{ij}^n &= \langle \frac{\partial}{\partial R_i} \psi_n(\mathbf{R}) | \frac{\partial}{\partial R_j} \psi_n(\mathbf{R}) \rangle - \langle \frac{\partial}{\partial R_j} \psi_n(\mathbf{R}) | \frac{\partial}{\partial R_i} \psi_n(\mathbf{R}) \rangle \end{aligned} \quad (4.54)$$

This notation is introduced to stress the fact that we can consider the berry curvature as the curl of the Berry potential, or as the infinitesimal circuitation of the Berry Potential.

Another way to justify the definition (4.51) is by direct analogy with the definition of electromagnetic field as the antisymmetric tensor built from the first order derivatives of the potential:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \quad (4.55)$$

where also in the electromagnetic case the vector potential has the role of a connection, and the electromagnetic field has the role of curvature.

By our previous magnetic analogy in 3 dimensions the Berry curvature is equivalent to the Magnetic field generated by the Berry potential in parameter space, in the same way that the Magnetic field is the curl of the magnetic vector potential.

4.6.3. Quadratic formula

There is a second way to write the expression of the Berry curvature in fact we can transform the definition (4.54) into:

$$\Omega^n(\mathbf{R}) = \langle \nabla_{\mathbf{R}} \psi_n(\mathbf{R}) | \wedge | \nabla_{\mathbf{R}} \psi_n(\mathbf{R}) \rangle = \sum_{m \neq n} \langle \nabla_{\mathbf{R}} n(\mathbf{R}) | m(\mathbf{R}) \rangle \wedge \langle m(\mathbf{R}) | \nabla_{\mathbf{R}} n(\mathbf{R}) \rangle \quad (4.56)$$

where we simplified the notation using $|n(\mathbf{R})\rangle$ instead of $|\psi_n(\mathbf{R})\rangle$. Starting from the Schroedinger equation:

$$\hat{H}(\mathbf{R})|n(\mathbf{R})\rangle = E_n|n(\mathbf{R})\rangle \quad (4.57)$$

we can write the off diagonal elements:

$$\langle m(\mathbf{R}) | \nabla_{\mathbf{R}} | n(\mathbf{R}) \rangle = \frac{\langle m(\mathbf{R}) | \nabla_{\mathbf{R}} \hat{H}(\mathbf{R}) | n(\mathbf{R}) \rangle}{E_n - E_m}; n \neq m \quad (4.58)$$

and finally

$$\Omega^n(\mathbf{R}) = i \sum_{m \neq n} \frac{\langle n(\mathbf{R}) | \nabla_{\mathbf{R}} \hat{H}(\mathbf{R}) | m(\mathbf{R}) \rangle \wedge \langle m(\mathbf{R}) | \nabla_{\mathbf{R}} \hat{H}(\mathbf{R}) | n(\mathbf{R}) \rangle}{(E_m(\mathbf{R}) - E_n(\mathbf{R}))^2} \quad (4.59)$$

or

$$\Omega_{\nu\mu}^n(\mathbf{R}) = i \sum_{n \neq n'} \frac{\langle n | \frac{\partial H}{\partial R_\mu} | n' \rangle \langle n' | \frac{\partial H}{\partial R_\nu} | n \rangle - \langle n | \frac{\partial H}{\partial R_\nu} | n' \rangle \langle n' | \frac{\partial H}{\partial R_\mu} | n \rangle}{(\varepsilon_n - \varepsilon_{n'})^2} \quad (4.60)$$

We will use this formula in the derivation of the TKNN formula in the upcoming chapter.

4.7. Topology

This section will give a short historical introduction to topology and mentions some important results

4.7.1. Euler formula

Euler was the first to recognize the importance of topology, with the Königsberg bridge problem. While solving the problem, he showed that there are properties of sets and spaces which are not local but depends on what is connected to what. This ultimately resulted in the Euler characteristic formula which can be written as:

$$V - E + F = \chi \quad (4.61)$$

where V is the number of vertexes, E the number of edges, and F the number of faces, χ is the Euler characteristic. Euler studied the simple case where the geometry of the base manifold

was simply connected such as a flat plane, a platonic solid or a sphere. In all these cases $\chi = 2$.

It's now possible to generalize the Euler formula to generic smooth surfaces such as the torus and the Euler characteristic become a special number summarizing the properties of the full manifold. The Euler characteristic is the sum of the Betti numbers.

$$\chi = \sum_n (-1)^n b_n \quad (4.62)$$

In the case of a 2D-surface embedded in a 3D space the zero Betti number is the number of disconnected parts of a the manifold and the first Betti number is the number of two dimensional holes in the manifold.

4.7.2. Algebraic Topology

Poincaré invented the modern field of topology in the Analysis Situs. He introduced homotopy, homology, cohomology and proved the duality between simplicial homology and it's dual cohomology. The work of Poincaré was invented based on symplcial complexes (i.e. generalized n-dimensional triangles) and was heavily anchored on geometric intuition. Again intuitively homology can be described as a way to count n-dimensional holes where the primary example for Poincaré were the Betti numbers.⁶

Emmy Noether was the first one to realize that the simplicial complexes of Poincaré, and the algebraic geometry problems of Hilbert had something in common. They had a group theory counterpart which are now called homology groups, and the Betti numbers were the ranks (i.e. the number of independent generators) of these groups. The hierarchies of topological characteristic numbers are representatives of the underlying topological group structure, in the same way quantum numbers represent symmetries of linear operators. This evolved into the work of Whitney and the definition of singular and cellular homology and their dual cohomologies which was given a solid axiomatic footing by Eilenberg and Steenrod. Singular homology builds upon the notion of continuity defined in point set topology and then generalizes from the case of smooth manifolds of Poincaré into generic topological spaces.⁷

4.7.3. The connection to differential Geometry

De Rham studied differential forms, and he proved an isomorphism between cohomology of differential forms and singular cohomology. Another important piece of the puzzle was the duality between the exterior derivative and integration, which culminates in the generalized stokes theorem of Cartan, and it's topological counterpart which is the generalized stokes theorem on chains.⁸ Cohomology classes have additional product structures which are called cup and cap products.

Chern realized that all different types of cohomology and all characteristic classes (such as

⁶ Allen Hatcher (2001). *Algebraic Topology* Ch 2.1

⁷ Allen Hatcher (2001). *Algebraic Topology* Ch 2.3

⁸ Loring W. Tu Raoul Bott (1982). *Differential forms in algebraic topology*

Euler, pontjagrin and stiefel-whitney classes) have a similar structure. Characteristic classes are a technique grounded in cohomology and that can generalize to different areas of mathematics⁹, showing similar structures across areas such as differential geometry (differential cohomology)¹⁰, and more recently algebraic geometry (motivic cohomology).

Chern classes are cohomology classes in the sense of de Rham cohomology, Chern and Weil discovered an homomorphism between the C-*algebras of principal G bundles and the De Rham cohomology, effectively describing a duality between differential geometry and Algebraic topology. Chern numbers are a way to catalogue Chern classes a bit in a similar way as we have seen Betti numbers earlier.

4.7.4. The Gauss Bonnet theorem

The Gauss Bonnet theorem is a relationship in two dimensions between integrals and the Euler characteristic:

$$\int_M K dA + \int_{\partial M} k_g ds = 2\pi\chi(M), \quad (4.63)$$

Here K is the Gaussian curvature, i.e. the product of the minimum and maximum curvature $K = k_1k_2$, and k_g is the geodesic curvature, i.e. the curvature along the tangent in the sense of the Frenet-Serret formulas. In the case the manifold is compact and without boundaries the second integral is zero. The equation then reduces to:

$$\int_M K dA = 2\pi\chi(M), \quad (4.64)$$

As an application of Chern Classes, Chern proved a generalized form of the Gauss bonnet theorem, this was a bridge between algebraic topology and differential geometry.

In general, for any closed C^∞ orientable n-dimensional M, the Chern theorem is:

$$\chi(M) = (e(TM), [M]) \quad (4.65)$$

where $\chi(M)$ is the Euler characteristic and the above pairing $(,)$ denotes the cap product with $e(TM)$, the Euler class of the tangent bundle of M .

In the case where M is a 2n-dimensional manifold the Euler class becomes:

$$e(\Omega) = \frac{1}{(2\pi)^n} \text{Pf}(\Omega), \quad (4.66)$$

where $\text{Pf}(\Omega)$ is the Pfaffian i.e. an invariant polynomial in the curvature Ω . The Chern theorem can then be written as:

$$\chi(M) = \frac{1}{(2\pi)^n} \int_M \text{Pf}(\Omega). \quad (4.67)$$

⁹ James D. Stasheff John W. Milnor (1974). *Characteristic classes*

¹⁰ Shiing-Shen Chern (1995). *Complex Manifolds Without Potential Theory*

To conclude, with a topic for a more in-depth study, an important generalization of the Chern-Gauss-Bonnet theorem is the Atiyah–Singer index theorem which has various applications in physics.

We already cited the mathematical literature. To learn more about Chern numbers one may look at ¹¹. For a visual example of fibrations with special points and with physical applications one may look at Skyrmions ¹² and their magnetic version¹³.

¹¹ Taylor L. Hughes B. Andrei Bernevig (2013). *Topological Insulators and Topological Superconductors*

¹² Wikipedia (2022b). *Skyrmions*. <https://en.wikipedia.org/wiki/Skyrmion>. [Online; accessed 12-October-2022]

¹³ Wikipedia (2022a). *Magnetic Skyrmions*. https://en.wikipedia.org/wiki/Magnetic_skyrmion. [Online; accessed 12-October-2022]

References

- B. Andrei Bernevig, Taylor L. Hughes (2013). *Topological Insulators and Topological Superconductors*.
- Berry, Michael Victor (Mar. 1984). “Quantal phase factors accompanying adiabatic changes”. In: 392 (1802).
- Chern, Shiing-Shen (1995). *Complex Manifolds Without Potential Theory*.
- Cohen E. Larocque H., Bouchard F. et al. (2019). “Geometric phase from Aharonov–Bohm to Pancharatnam–Berry and beyond”. In: *Nat Rev Phys* (1), pp. 437–449.
- Hatcher, Allen (2001). *Algebraic Topology*.
- John W. Milnor, James D. Stasheff (1974). *Characteristic classes*.
- Raoul Bott, Loring W. Tu (1982). *Differential forms in algebraic topology*.

Readings and online resources

- Wikipedia (2022a). *Magnetic Skyrmions*. https://en.wikipedia.org/wiki/Magnetic_skyrmion. [Online; accessed 12-October-2022].
- (2022b). *Skyrmions*. <https://en.wikipedia.org/wiki/Skyrmion>. [Online; accessed 12-October-2022].
- Zwiebach, Barton (2018b). *L17.2 Berry's phase and Berry's connection*. *MIT 8.06 Quantum Physics III*. <https://www.youtube.com/watch?v=iGG9EG3SNz0>. [Online; accessed 12-October-2021].

5 — TKNN formula

5.1. Introduction

5.1.1. Approach

In this chapter we will derive the TKNN formula. The name is due to the names of D. J. Thouless, M. Kohmoto, M. P. Nightingale and M. den Nijs. It's one of the leading results in the motivation for the Nobel prize of 2016 in physics and is a foundational result in regards to topological insulators. The TKNN formula gives an explanation of the quantization of the Integer Quantum Hall Effect. To achieve that we will first give a brief intuitive explanation of what is the Integer quantum Hall with the Laughlin pump argument. Then, we will compute the left side of the TKNN formula with the Berry curvature and the right side, i.e. the Hall conductivity, through the Kubo formula. Finally we will assemble the formula, discuss it and see it in different variations.

5.1.2. Historical introduction

As we have discussed the paper of Laughlin¹ argued that there was a quantization in the scope of the Integer Quantum Hall Effect, the paper of Thouless was the first one to fully compute the presence of quantized numbers. The objective of the paper was to describe both the integer fractions of the Integer Quantum Hall Effect and the spectrum of the Hofstadter butterfly through the Harper equation². Historically is important to note that the Berry phase was published slightly later in 1984³ and the link between the two through differential geometry and the theory of Fiber Bundles was published in the same period by mathematical physicist Barry Simon⁴.

5.2. Laughlin pump argument

The basic intuition of the quantization in the integer quantum hall effect can be explained with the Laughlin pump argument⁵. A simple way to explain this is through the corbino

¹ Laughlin (1981). “Quantized Hall conductivity in two dimensions”. In: *Physical Review Letters B* 23, p. 5632

² D. J. Thouless et al. (1982). “Quantized Conductance in a Two-Dimensional Periodic Potential”. In: *Physical Review Letters* 49, p. 405

³ Michael Victor Berry (Mar. 1984). “Quantal phase factors accompanying adiabatic changes”. In: 392 (1802)

⁴ Barry Simon (Dec. 1983). “Holonomy, the Quantum Adiabatic Theorem, and Berry’s Phase”. In: *Phys. Rev. Lett.* (51), p. 2167

⁵ Laughlin (1981). “Quantized Hall conductivity in two dimensions”. In: *Physical Review Letters B* 23, p. 5632

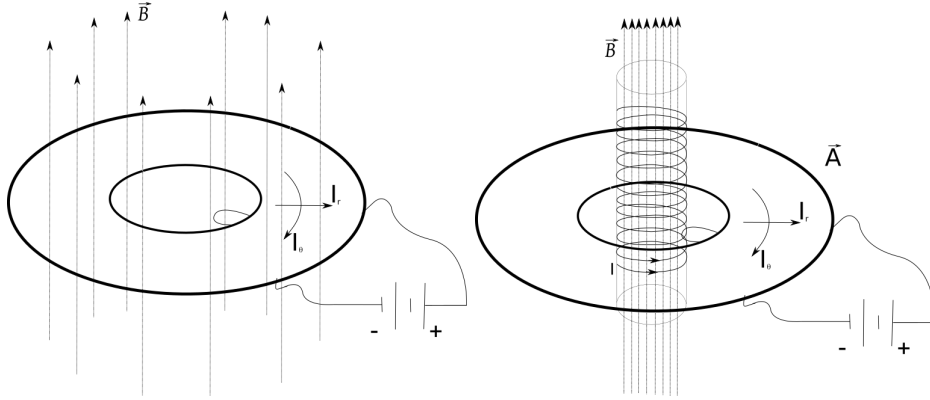


Figure 5.1: Corbino disk

geometry.

Let's start with the simple Corbino disk, which is depicted in figure (5.1) on the left. In this case we have electrical leads attached at the two edges of the disk and through them we push a current. When the magnetic field is switched on the Lorentz force comes into play and introduces an effect called magnetoresistance, which is an effective resistance on the initial current.

We analyze now the case on the right where the magnetic field is created by a vertical magnetic flux line at the center of the disk. As in the case of the Aharonov-Bohm effect, the magnetic field intersecting the disk is null but the potential is not null. We switch the field on adiabatically and we would like to see what happens when we vary it slowly from $\Phi = 0$ first to a generic flux $\Phi(t)$ and then to one Quantum of flux Φ_0 . First of all, by Faraday induction we can write that there is an induced transversal electric field:

$$\frac{\partial \Phi}{\partial t} = - \int \mathbf{E} \cdot d\mathbf{l} = -E_\varphi 2\pi r \quad (5.1)$$

This electric field generates a radial current density

$$j_r = \sigma_{r\varphi} E_\varphi, \quad (5.2)$$

where $\sigma_{r\varphi}$ is the transversal conductivity.

If we compute the total displaced charge:

$$\begin{aligned} \Delta Q &= \int I_r dt = \int 2\pi r j_r dt = - \int 2\pi r \sigma_{r\varphi} E_\varphi dt \\ &= - \int \frac{\partial \Phi}{\partial t} \sigma_{r\varphi} dt = -\sigma_{r\varphi} [\Phi(t) - \Phi(0)] = -\sigma_{r\varphi} \Phi(t) \end{aligned} \quad (5.3)$$

Therefore in the case of one quantum of flux we have a directly proportional value:

$$\Delta Q = -\sigma_{r\varphi} \Phi_0 \quad (5.4)$$

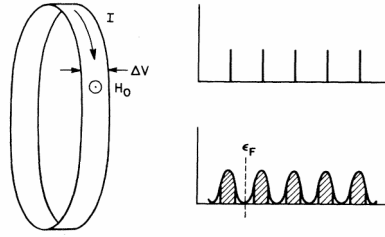


FIG. 1. Left: Diagram of metallic loop. Right: Density of states without (top) and with (bottom) disorder. Regions of delocalized states are shaded. The dashed line indicates the Fermi level.

Figure 5.2: Laughlin pump argument - image retrieved from Laughlin 1981

Now, we introduce the spectrum which is the one of Landau levels, in the case of a pure sample this is a set of Dirac functions, in the case of a limited amount of disorder we can imagine that these levels, gets smeared out like in figure (5.2) from the original Laughlin publication. We can imagine a fixed number of electrons per level which are all filled up to the Fermi energy, and the Fermi energy lying between two levels. We may now ask when we add an extra quantum of flux what happens to the occupation numbers?

We are moving electrons from the inner side of the the disk of figure (5.1) to the outer side given that the levels are ordered by radius. Or if one prefers from the left side of the cylinder in figure (5.2) to the right side.

First of all, after one quantum of flux is added we can state that by gauge transformation the Hamiltonian is the same and therefore the available levels are the same. Second we can say that we can only move a discrete number of electrons between levels, therefore also the displaced charge must be quantized:

$$\Delta Q = -\sigma_{r\varphi} \Phi_0 = -ne; n \in \mathbb{Z}. \quad (5.5)$$

Finally by consequence we can write that the conductivity is actually quantized:

$$\sigma_{r\varphi} = n \frac{e^2}{2\pi\hbar}; n \in \mathbb{Z}. \quad (5.6)$$

While we are adding the adiabatic perturbation from zero up to one quantum of flux, the process is continuous and the states move continuously as in figure (5.3) situation (a). Once we added one quantum of flux all levels are shifted as in figure (5.3) situation (b). The levels are the same, but the situation is equivalent to having moved one electron from the lowest state to the upper one.

This means in essence that the adding of the flux is a periodic signal that goes back to the beginning after every quantum of flux. The parameter space in this case is then periodic where the varying parameter is the flux.

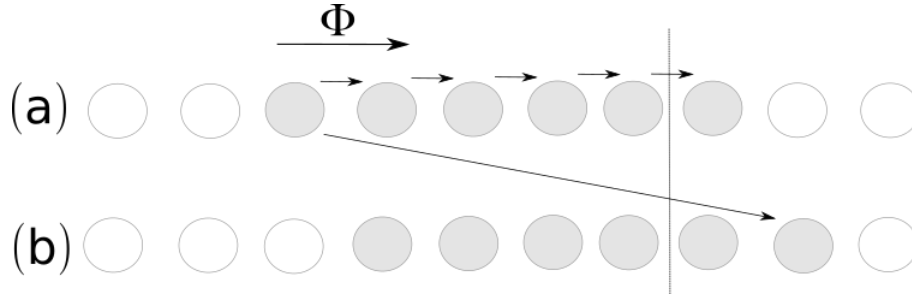


Figure 5.3: pump argument levels

5.3. Berry Curvature

5.3.1. Electrons in a Brillouin zone

We now want to use the Berry curvature in the context of the Band theory. Let's imagine to have electrons on a lattice and therefore also a notion of a Brillouin zone in the reciprocal space.

To make it concrete, we take a rectangular lattice, where the lattice is periodic in \mathbf{x} and the reciprocal space is periodical in \mathbf{k} , more precisely if we start from a rectangular lattice, we end up with a rectangular Brillouin zone.

To describe the lattice we have periodic boundary conditions both in \mathbf{x} and \mathbf{k} space, and therefore we can quickly summarize this by saying that both the \mathbf{x} and the reciprocal \mathbf{k} space are a tori. Let's now introduce 3 conditions:

- The spectrum decomposes into bands parametrized by the momentum \mathbf{k} and therefore that we can apply the Bloch theorem. There is also a slight variation to this in the case of the presence of a magnetic field where the Brillouin zone is modified into a magnetic Brillouin zone but the results are still valid.
- The electrons are not interacting and the multiple particle spectrum is obtained just by filling the single particle spectrum taking into account the Pauli exclusion principle.
- There is an energy gap between bands and the bands don't touch, the Fermi energy value is just within the band gap, all bands below the Fermi Energy are filled and all bands above are empty, i.e. in a canonical band theory this will be an insulator.

As we we will see with these conditions we are able to assign a topological invariant integer to each band which is a Chern number.

It's important to understand that this approach can then be generically extended not only to bands but to generic gapped systems. Most of the physics in a gapped system happens in a precisely defined finite energy scale and it therefore happens in a precise time scale and a finite range of values for \mathbf{k} .

The momentum \mathbf{k} is not a physical parameter that we can change, as we used to do with the adiabatic theorem, \mathbf{k} is a parameter in the sense that for every \mathbf{k} there is a different Hilbert

space, therefore we can do without the adiabatic theorem, and reuse the Berry phase machinery in any case.

Finally, we can also state that a major property of the parameter space \mathbf{k} is that it is actually closed and bounded, or more precisely compact, which means in essence there is both an infrared and ultraviolet limit (i.e. lower and upper meaningful energies), and that this is also true for a non rectangular Brillouin zone.

Given we can apply the Berry phase machinery, we can compute it on finite loops into our parameter space, and in combination, given the periodicity, we can also use the Bloch theorem.

From the Bloch theorem we can write:

$$\psi_{n,\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{n,\mathbf{k}}(\mathbf{r}), \quad (5.7)$$

where $u_{n,\mathbf{k}}$ are the Bloch wave functions.

Therefore we can also write the Berry connection as:

$$A_i^n(\mathbf{k}) = i \langle \psi_n(\mathbf{k}) | \frac{\partial}{\partial k_i} | \psi_n(\mathbf{k}) \rangle = i \langle u_n(\mathbf{k}) | \frac{\partial}{\partial k_i} | u_n(\mathbf{k}) \rangle, \quad (5.8)$$

where the exponent cancels out in the first term.

We can write the Berry curvature with the quadratic formula (4.60):

$$\Omega_{ij}^n(\mathbf{k}) = i \sum_{n \neq n'} \frac{\langle n | \frac{\partial \hat{H}}{\partial k_i} | n' \rangle \langle n' | \frac{\partial \hat{H}}{\partial k_j} | n \rangle - \langle n | \frac{\partial \hat{H}}{\partial k_j} | n' \rangle \langle n' | \frac{\partial \hat{H}}{\partial k_i} | n \rangle}{(\varepsilon_n - \varepsilon_{n'})^2} \quad (5.9)$$

where our parameter space R_μ has become the \mathbf{k} space.

Finally we can also write the Berry curvature as:

$$\begin{aligned} \Omega_{ij}^n(\mathbf{k}) &= \int d^d r \frac{\partial \psi_n^*(\mathbf{k})}{\partial k_i} \frac{\partial \psi_n(\mathbf{k})}{\partial k_j} - \frac{\partial \psi_n^*(\mathbf{k})}{\partial k_j} \frac{\partial \psi_n(\mathbf{k})}{\partial k_i} \\ &= \int d^d r \frac{\partial u_n^*(\mathbf{k})}{\partial k_i} \frac{\partial u_n(\mathbf{k})}{\partial k_j} - \frac{\partial u_n^*(\mathbf{k})}{\partial k_j} \frac{\partial u_n(\mathbf{k})}{\partial k_i}. \end{aligned} \quad (5.10)$$

Here d is the number of dimensions we are integrating upon. In the case of integer quantum Hall Effect these will be 2, and the integral is extended to the full space.

Finally applying Stokes theorem we can write for the 2 dimensional case:

$$\int_A dk^2 \Omega_{ij}^n(\mathbf{k}) = \oint_{\partial A} dk_j \int d^2 r \left(u_n^* \frac{\partial u_n}{\partial k_j} - \frac{\partial u_n^*}{\partial k_j} u_n \right), \quad (5.11)$$

where the integral is done over a closed loop in the reciprocal space.

5.3.2. Quantization of Berry phase in a compact manifold

We would need now to prove that the integral of the Berry curvature is quantized, and this is thanks to the fact that the integral is over a compact manifold and there is no boundary

term. The proof of it's similar in spirit to the proof of the Gauss Bonnet theorem.

$$\int K dA = 2\pi\chi \quad (5.12)$$

Where K is the Gaussian curvature given by $\frac{1}{R_1 R_2}$ where R_1 and R_2 are the curvature radii.

The same happens for the Berry curvature, that when integrated over a closed manifold without border is quantized. When the Berry curvature is integrated over the full space of momenta it has a special name and is called Chern number C_n

$$\int d^2\mathbf{k} \Omega_n(\mathbf{k}) = 2\pi C_n. \quad (5.13)$$

5.4. The Hall conductivity

5.4.1. The Kubo formula

We will now derive the Hall conductivity from the Kubo formula. We start from a generic perturbation problem:

$$\begin{aligned} \hat{H}_0 &= \frac{1}{2m}(\hat{\mathbf{p}} - e\mathbf{A})^2 \\ \delta\hat{H}(t) &= V(\vec{r}, t) \end{aligned} \quad (5.14)$$

The core additional assumption is that in the past(for $t < t_0$) the system was at equilibrium, we are describing what happens after switching on this perturbation that moves the system slightly out of equilibrium ⁶.

Therefore before t_0 the system was in it's set of unperturbed eigenstates:

$$\hat{H}_0|\mathbf{n}(t_0)\rangle = E_{\mathbf{n}}^0|\mathbf{n}(t_0)\rangle. \quad (5.15)$$

The expectation value for a generic observable \hat{O} at equilibrium is:

$$\begin{aligned} \langle\hat{O}\rangle &= \frac{1}{Z_0} \sum_n \langle\mathbf{n}(t_0)|\hat{O}|\mathbf{n}(t_0)\rangle e^{-\beta E_{\mathbf{n}}^0} \\ Z_0 &= \sum_n e^{-\beta E_{\mathbf{n}}^0} \end{aligned} \quad (5.16)$$

We now switch on the perturbation at t_0 with a heavyside step function $\theta(t - t_0)$, later we will also assume that this t_0 is far away in the past with respect to t and we want to compute

⁶we can be actually quite far away of equilibrium, but there is a concept of smooth transitions, where there derivatives are still meaningful, there are no jump processes and we are essentially not in a turbulent regime

the long term state after the perturbation is applied:

$$\begin{aligned}\hat{H}(t) &= \hat{H}_0 + \delta\hat{H}(t) \\ \delta\hat{H}(t) &= \theta(t - t_0)\delta\hat{H}'(t).\end{aligned}\tag{5.17}$$

We can imagine from that moment t_0 onwards, some time-based evolution of both the Hamiltonian and the eigenstates, somewhat in a similar manner to what happens for adiabatic processes.

$$\hat{H}(t)|\mathbf{n}(t)\rangle = i\hbar\partial_t|\mathbf{n}(t)\rangle\tag{5.18}$$

To linear order in the perturbation $\delta\hat{H}(t)$ ⁷

$$\hat{U}(t, t_0) \approx 1 - \frac{i}{\hbar} \int_{t_0}^t \delta\hat{H}(t') dt'\tag{5.19}$$

Given the perturbation will be small we will use the interaction picture, we write down here the generic Kubo formula⁸ for the observable \hat{O} in order to use it later:

$$\begin{aligned}\langle \mathbf{n}(t) | \hat{O}_I(t) | \mathbf{n}(t) \rangle &\approx \langle \mathbf{n}(t_0) | \hat{O}_{0,I}(t) | \mathbf{n}(t_0) \rangle \\ &- \frac{i}{\hbar} \int_{t_0}^t dt' \frac{1}{Z_0} \sum_n e^{-\beta E_n} \langle \mathbf{n}(t_0) | \hat{O}_I(t) \delta\hat{H}_I(t') - \delta\hat{H}_I(t') \hat{O}_I(t) | \mathbf{n}(t_0) \rangle.\end{aligned}\tag{5.20}$$

Introducing the notation $[\dots, \dots]_-$ for the commutator:

$$\langle \mathbf{n}(t) | \hat{O}_I(t) | \mathbf{n}(t) \rangle \approx \langle \mathbf{n}(t_0) | \hat{O}_{0,I}(t) | \mathbf{n}(t_0) \rangle - \frac{i}{\hbar} \int_{t_0}^t dt' \langle \mathbf{n}(t_0) | [\hat{O}_I(t), \delta\hat{H}_I(t')]_- | \mathbf{n}(t_0) \rangle_0.\tag{5.21}$$

A second notational aspect is that we imploded the statistical mechanics average and the quantum mechanics one with the conventional notation:

$$\langle \dots \rangle_0 = \frac{1}{Z_0} \sum_n e^{-\beta E_n} \langle \dots \rangle\tag{5.22}$$

A key aspect of this approximation is that the observable is computed at time t where the Hamiltonian perturbation is computed at time t' . The idea is that we can then expand the perturbation with respect to its initial value at t_0 . In general there are also time ordering considerations, with respect to the different time points taken into account. These are essential in order to maintain causality, but here we will not analyze them in detail.

In the general scenario of using the Kubo formula, but also for other type of quantum statistical mechanics problems, it is safer to take this statistical average (or in other words thermal average) into account at last. This is typically done after the rest of the computations, in order to take along thermodynamic considerations in a parametric manner, as an extra "thermodynamic spread" of the variables, and finalize this parametric freedom at the end.

⁷This is an approximated definition of the time propagator

⁸Ryogo Kubo (June 1957). "Statistical-Mechanical Theory of irreversible processes I.". In: *Journal of the physical society of japan* 12.6, p. 570

There is finally a more general way to write the Kubo formula as a generic linear response kernel, given a set of source terms $\phi_j(\vec{x}, t)$, which induces an Hamiltonian coupling:

$$H_{source}(t) = \int d^{d-1}\vec{x} \phi_i(\vec{x}, t) O_i(\vec{x}, t). \quad (5.23)$$

The sources induce a linear response:

$$\delta\langle O_i(\vec{x}, t) \rangle = \int d^d\vec{x}' dt' \chi_{ij}(\vec{x}, t, \vec{x}', t') \phi_j(\vec{x}', t') \quad (5.24)$$

which is now a convolution integral in the sources. The kernel of which is given by:

$$\chi_{ij}(\vec{x}, \vec{x}', t - t') = -i\theta(t - t') \langle [O_i(\vec{x}, t), O_j(\vec{x}', t')] \rangle \quad (5.25)$$

Where the $\theta(t - t')$ reflect that this is the retarded causal response after the effect at time t'

5.4.2. Kubo formula for the current

A big part of this section is derived from Tong Quantum Hall lectures⁹ and the Kinetic theory lectures¹⁰.

To compute the conductivity we introduce a small extra variable electric field and we will compute the linear part of it's relationship with the current, another way is that we can imagine this process also as probe electric field that generates a current, or equivalently a probe current where we measure the resistivity (which is the inverse matrix of the conductivity), and we want to compute it's linear response, i.e first order perturbation.

We use the temporal gauge where the electric potential is zero $A_0 = \phi = 0$ and therefore:

$$\mathbf{E} = -\partial_t \mathbf{A} \quad (5.26)$$

To compute the conductivity we introduce an electric AC field:

$$\mathbf{E}(t) = \mathbf{E} e^{-i\omega t} \quad (5.27)$$

This is the generic case to analyze conductivity, we are ultimately interested to know what happens in the case that we have a DC current and therefore in the limit $\omega \rightarrow 0$. We will develop the formulas carrying over ω as a parameter and taking the limit at the end before the thermal average. In essence there will be infinities for certain frequencies and those shall be carefully evaluated.

⁹ David Tong (2016b). *Lectures on the Quantum Hall Effect*. <http://www.damtp.cam.ac.uk/user/tong/qhe.html>. [Online; accessed 12-October-2021] 2.2.3

¹⁰ David Tong (2016a). *Lectures on the Kinetic theory*. <http://www.damtp.cam.ac.uk/user/tong/kinetic.html>. [Online; accessed 1-March-2021] Ch. 4.4

Given the choosen gauge, the potential is:

$$\mathbf{A} = \frac{1}{i\omega} \mathbf{E} e^{-i\omega t} \quad (5.28)$$

Ultimately, we are looking for the conductivity tensor $\sigma(\omega)$ that will give us the first order response based on the frequency of the probe signal.

$$j^i = \sigma_k^i(\omega) E^k \quad (5.29)$$

As a first approximation what we actually do is that we start from a classical field and electrons are treated quantum mechanically. This is reflected in the Langrangian density:

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - A_\mu j^\mu \quad (5.30)$$

where j^μ is the current density. We consider the perturbation originated by the fact that the probing current is small. Therefore we have an Hamiltonian density \mathcal{H} for the perturbation that we can write as:

$$\delta\mathcal{H} = -\mathbf{j} \cdot \mathbf{A}. \quad (5.31)$$

If we now write the kubo formula for the current density \mathbf{j} we have:

$$\langle \mathbf{n}(t) | \hat{\mathbf{j}}_I(t) | \mathbf{n}(t) \rangle \approx \langle \mathbf{n}(t_0) | \hat{\mathbf{j}}_0 | \mathbf{n}(t_0) \rangle - \frac{i}{\hbar} \int_{t_0}^t dt' \langle \mathbf{n}(t_0) | [\hat{\mathbf{j}}_I(t), \delta\hat{H}_I(t')]_- | \mathbf{n}(t_0) \rangle_0. \quad (5.32)$$

The first term is the current without electric field and we assume that this terms vanish because this is a probe current generated by a probe electric field and there is no pre-existing current due to other effects. In general this also means that we can do this probing in a generic manner given an already existing background field or background current, but we will assume that these are zero. We can now expand the interaction picture operators into

$$\begin{aligned} \hat{j}_{iI}(t) &= e^{i\hat{H}_0 t/\hbar} \hat{j}_i(t) e^{-i\hat{H}_0 t/\hbar} \\ \delta\hat{H}_I(t) &= e^{i\hat{H}_0 t/\hbar} \delta\hat{H}(t) e^{-i\hat{H}_0 t/\hbar}. \end{aligned} \quad (5.33)$$

Therefore we can remove the interaction picture operators and remain with

$$\langle \mathbf{n}(t) | \hat{j}_i(t) | \mathbf{n}(t) \rangle \approx -\frac{i}{\hbar} \int_{t_0}^t dt' \langle \mathbf{n}(t_0) | [\hat{j}_i(t), \delta\hat{H}(t')]_- | \mathbf{n}(t_0) \rangle_0. \quad (5.34)$$

If we substitute the Hamiltonian with

$$\delta H(t') = -\frac{1}{i\omega} j_k(t') E^k e^{-i\omega t'} \quad (5.35)$$

We have:

$$\langle \mathbf{n}(t) | \hat{j}_i(t) | \mathbf{n}(t) \rangle \approx \frac{1}{\hbar\omega} \int_{t_0}^t \theta(t' - t_0) dt' \langle \mathbf{n}(t_0) | [\hat{j}_i(t), \hat{j}_k(t')]_- | \mathbf{n}(t_0) \rangle_0 E^k e^{-i\omega t'}. \quad (5.36)$$

We change variables in the integral with the substitution

$$\begin{aligned} t'' &= t - t'; dt'' = -dt' \\ t' &= t_0; t'' = t - t_0 \\ t' &= t; t'' = 0 \end{aligned} \quad (5.37)$$

We can separate out of the integral the time dependence and obtain:

$$\langle \mathbf{n}(t) | \hat{j}_i(t) | \mathbf{n}(t) \rangle \approx \frac{1}{\hbar\omega} \left(\int_0^{t-t_0} dt'' e^{i\omega t''} \langle \mathbf{n}(t_0) | [\hat{j}_i(t), \hat{j}_k(t-t'')]_- | \mathbf{n}(t_0) \rangle_0 \right) E^k e^{-i\omega t}. \quad (5.38)$$

We can then consider t_0 a time that was much before the time t as if there was a large amount of time to go at regime and as if the perturbation was almost always on, from this condition we can impose the limit $t - t_0 \rightarrow +\infty$. Therefore the conductivity in first order is then:

$$\sigma_{ik}(\omega) = \frac{1}{\hbar\omega} \int_0^\infty dt e^{i\omega t} \langle \mathbf{n}(t_0) | [\hat{j}_i(t_0), \hat{j}_k(t)]_- | \mathbf{n}(t_0) \rangle_0 \quad (5.39)$$

and the linear response equation can be written as:

$$\langle \mathbf{n}(t) | \hat{j}_i(t) | \mathbf{n}(t) \rangle \approx \sigma_{ik}(\omega) E^k e^{-i\omega t}. \quad (5.40)$$

We can here see in the formula on the right there is the electric field, and an input frequency on the probing electric field implies the same frequency in the output computed (or measured) current. This preservation of frequency between input and output is the essence of linear response, and is general to linear system theory. Any non-linearity would crop in when there are effects across multiple frequencies.

For a two dimensional system, the Hall conductivity is the off-diagonal:

$$\sigma_{xy}(\omega) = \frac{1}{\hbar\omega} \int_0^\infty dt e^{i\omega t} \langle \mathbf{n}(t_0) | [\hat{j}_x(t_0), \hat{j}_y(t)]_- | \mathbf{n}(t_0) \rangle_0. \quad (5.41)$$

We can use the fact that the operators $\hat{j}_i(t)$ evolves in time with

$$\begin{aligned} \hat{j}_x(t) &= e^{i\hat{H}_0 \frac{t}{\hbar}} \hat{j}_x(t_0) e^{-i\hat{H}_0 \frac{t}{\hbar}} \\ \hat{j}_y(t) &= e^{i\hat{H}_0 \frac{t}{\hbar}} \hat{j}_y(t_0) e^{-i\hat{H}_0 \frac{t}{\hbar}} \end{aligned} \quad (5.42)$$

We insert a complete basis $|\mathbf{n}'(t_0)\rangle \langle \mathbf{n}'(t_0)|$ of H_0 in the integral:

$$\hat{H}_0 |\mathbf{n}'(t_0)\rangle = E_{\mathbf{n}'} |\mathbf{n}'(t_0)\rangle \quad (5.43)$$

and we can drop all t_0 given that all operators $j_x(t_0), j_y(t_0)$ and states $|\mathbf{n}(t_0)\rangle, |\mathbf{n}'(t_0)\rangle$ are evaluated in t_0 . This gives us:

$$\sigma_{xy}(\omega) = \frac{1}{\hbar\omega} \int_0^\infty dt e^{i\omega t} \sum_{\mathbf{n}'} [\langle \mathbf{n} | \hat{j}_x | \mathbf{n}' \rangle_0 \langle \mathbf{n}' | \hat{j}_y | \mathbf{n} \rangle_0 e^{i(E_{\mathbf{n}'} - E_{\mathbf{n}}) \frac{t}{\hbar}} - \langle \mathbf{n} | \hat{j}_y | \mathbf{n}' \rangle_0 \langle \mathbf{n}' | \hat{j}_x | \mathbf{n} \rangle_0 e^{i(E_{\mathbf{n}} - E_{\mathbf{n}'}) \frac{t}{\hbar}}]. \quad (5.44)$$

We extend the integral domain by introducing a heaviside step function:

$$\theta(t) = \frac{1}{2\pi i} \lim_{\varepsilon \rightarrow 0} \int_{-\infty}^\infty \frac{e^{itx}}{x - i\varepsilon} dx \quad (5.45)$$

$$\sigma_{xy}(\omega) = \frac{1}{\hbar\omega} \int_{-\infty}^\infty dt \theta(t) e^{i\omega t} \sum_{\mathbf{n}'} [\langle \mathbf{n} | \hat{j}_x | \mathbf{n}' \rangle_0 \langle \mathbf{n}' | \hat{j}_y | \mathbf{n} \rangle_0 e^{i(E_{\mathbf{n}'} - E_{\mathbf{n}}) \frac{t}{\hbar}} - \langle \mathbf{n} | \hat{j}_y | \mathbf{n}' \rangle_0 \langle \mathbf{n}' | \hat{j}_x | \mathbf{n} \rangle_0 e^{i(E_{\mathbf{n}} - E_{\mathbf{n}'}) \frac{t}{\hbar}}] \quad (5.46)$$

We integrate first over time, then over the x variable of the $\theta(t)$, take the limit over ε and the equation then reads:

$$\sigma_{xy}(\omega) = \frac{1}{i\omega} \sum_{\mathbf{n}' \neq \mathbf{n}} \left[\frac{\langle \mathbf{n} | \hat{j}_y | \mathbf{n}' \rangle_0 \langle \mathbf{n}' | \hat{j}_x | \mathbf{n} \rangle_0}{\hbar\omega + E_{\mathbf{n}'} - E_{\mathbf{n}}} - \frac{\langle \mathbf{n} | \hat{j}_x | \mathbf{n}' \rangle_0 \langle \mathbf{n}' | \hat{j}_y | \mathbf{n} \rangle_0}{\hbar\omega + E_{\mathbf{n}} - E_{\mathbf{n}'}} \right]. \quad (5.47)$$

The states where $\mathbf{n}' = \mathbf{n}$ don't contribute to the sum, given first and second term become the same. We are now finally looking for the continuum current DC limit $\omega \rightarrow 0$ and we can expand for small ω :

$$\frac{1}{\hbar\omega + E_{\mathbf{n}'} - E_{\mathbf{n}}} \approx \frac{1}{E_{\mathbf{n}'} - E_{\mathbf{n}}} - \frac{\hbar\omega}{(E_{\mathbf{n}'} - E_{\mathbf{n}})^2} + \mathcal{O}(\omega^2) \quad (5.48)$$

and equivalently for the other term.

For a general computation, some of these terms may diverge, and the higher order terms $\mathcal{O}(\omega^2)$ may also contain divergences. If we write down the first term in the case for transversal conductivity (i.e. the Hall conductivity) it cancels out.

We can then write the second term, for the Hall conductivity and remain with:

$$\sigma_{xy} \approx i\hbar \sum_{\mathbf{n}' \neq \mathbf{n}} \frac{\langle \mathbf{n} | \hat{j}_y | \mathbf{n}' \rangle_0 \langle \mathbf{n}' | \hat{j}_x | \mathbf{n} \rangle_0 - \langle \mathbf{n} | \hat{j}_x | \mathbf{n}' \rangle_0 \langle \mathbf{n}' | \hat{j}_y | \mathbf{n} \rangle_0}{(E_{\mathbf{n}} - E_{\mathbf{n}'})^2}. \quad (5.49)$$

Finally after the limit in the frequency we can do the statistical mechanics average in the temperature $\langle \dots \rangle_0$ taking the limit in which the temperature goes to zero $T \rightarrow 0$ and therefore having one single state from classical statistical mechanics standpoint:

$$\sigma_H = \sigma_{xy} = i\hbar \sum_{\mathbf{n}' \neq \mathbf{n}} \frac{\langle \mathbf{n} | \hat{j}_y | \mathbf{n}' \rangle \langle \mathbf{n}' | \hat{j}_x | \mathbf{n} \rangle - \langle \mathbf{n} | \hat{j}_x | \mathbf{n}' \rangle \langle \mathbf{n}' | \hat{j}_y | \mathbf{n} \rangle}{(E_{\mathbf{n}} - E_{\mathbf{n}'})^2}, \quad (5.50)$$

which is the final version of the Kubo formula for the Hall conductivity.

5.5. TKNN formula

This section is derived from Haldane Nobel lecture¹¹ which contains some further insights about topological insulators and Tong¹²

5.5.1. Electrons in a Magnetic field

Instead of considering a probe electric field there is a second way to introduce the perturbation, i.e. is by introducing a magnetic field as a perturbing field. This will also induce a current and therefore we can extend our computation of conductivity given a source term which is now a magnetic field. More precisely, we can start from the geometry of a conductor in the shape of a torus and we can then imagine two sources of magnetic field. A first one as a solenoid with infinitesimal radius along the z direction, and a second source which is a similar solenoid along the angular direction of the torus, where the magnetic flux is completely embedded within the solenoid and therefore within the torus. Again there is no magnetic field outside the solenoid, only vector potential.

Such an ideal solenoid source is also called a flux line (or tube) and if the characteristic lengths of the torus are L_x and L_y we can use the magnetic flux as the basic unit instead of the magnetic field, and we can then use the quantum of flux as the unit of measurement.

We start from a Landau gauge, an unperturbed state $|\mathbf{n}_0\rangle$ and we add a perturbation with these two extra flux tubes. If we write down these two as sources we can write:

$$\begin{aligned} A_x &= \frac{\Phi_x}{L_x} \\ A_y &= \frac{\Phi_y}{L_y} + Bx \end{aligned} \quad (5.51)$$

This means that our perturbation Hamiltonian is now

$$\Delta H = -\mathbf{J} \cdot \mathbf{A} = -\sum \frac{\Phi_i}{L_i} J_i \quad (5.52)$$

where we considered only first order perturbations in the variation of the potential \mathbf{A} . To first order in perturbation we can write:

$$|\mathbf{n}_0\rangle' = |\mathbf{n}_0\rangle + \sum_{\mathbf{n} \neq \mathbf{n}_0} \frac{\langle \mathbf{n} | \Delta H | \mathbf{n}_0 \rangle}{E_n - E_0} |\mathbf{n}\rangle \quad (5.53)$$

but we can write the same as a variation in the Φ_i :

$$|\mathbf{n}_0\rangle' = |\mathbf{n}_0\rangle + \sum_i \Phi_i \left| \frac{\partial \mathbf{n}_0}{\partial \Phi_i} \right\rangle \quad (5.54)$$

¹¹<https://www.nobelprize.org/uploads/2018/06/haldane-lecture-slides.pdf>

¹²<http://www.damtp.cam.ac.uk/user/tong/qhe/qhe.pdf>

and therefore we can write

$$\left| \frac{\partial \mathbf{n}_0}{\partial \Phi_i} \right\rangle = -\frac{1}{L_i} \sum_{\mathbf{n} \neq \mathbf{n}_0} \frac{\langle \mathbf{n} | J_i | \mathbf{n}_0 \rangle}{E_n - E_0} |\mathbf{n}\rangle \quad (5.55)$$

Then we can use this expression as building block for the previous Hall conductivity (5.50) which can be rewritten as:

$$\begin{aligned} \sigma_{xy} &= i\hbar L_x L_y \sum_{\mathbf{n} \neq \mathbf{n}_0} \frac{\langle \mathbf{n}_0 | J_y | \mathbf{n} \rangle \langle \mathbf{n} | J_x | \mathbf{n}_0 \rangle - \langle \mathbf{n}_0 | J_x | \mathbf{n} \rangle \langle \mathbf{n} | J_y | \mathbf{n}_0 \rangle}{(E_n - E_0)^2} \\ &= i\hbar \left[\left\langle \frac{\partial \mathbf{n}_0}{\partial \Phi_y} \middle| \frac{\partial \mathbf{n}_0}{\partial \Phi_x} \right\rangle - \left\langle \frac{\partial \mathbf{n}_0}{\partial \Phi_x} \middle| \frac{\partial \mathbf{n}_0}{\partial \Phi_y} \right\rangle \right] \\ &= i\hbar \left[\frac{\partial}{\partial \Phi_y} \langle \mathbf{n}_0 | \frac{\partial \mathbf{n}_0}{\partial \Phi_x} \rangle - \frac{\partial}{\partial \Phi_x} \langle \mathbf{n}_0 | \frac{\partial \mathbf{n}_0}{\partial \Phi_y} \rangle \right] \end{aligned} \quad (5.56)$$

5.5.2. The TKNN formula in a simple magnetic field

We start from a set of electron in a magnetic field on a toroidal geometry, we can now state that the varying parameters of the fluxes are also here periodic as from the reasoning of the Laughlin pump argument, therefore we can consider them periodic and we can introduce the angles to describe the parameter space:

$$\theta_i = \frac{2\pi\Phi_i}{\Phi_0}; \theta_i \in [0, 2\pi) \quad (5.57)$$

and $\Phi_0 = \frac{2\pi\hbar c}{e}$ is the flux quantum. At this point the parameter space is a 2 dimensional torus \mathbf{T}_{Φ}^2 , which is compact closed and without border. We can write the Berry curvature for the base state \mathbf{n}_0 :

$$\Omega_{xy} = i \left[\frac{\partial}{\partial \theta_y} \langle \mathbf{n}_0 | \frac{\partial \mathbf{n}_0}{\partial \theta_x} \rangle - \frac{\partial}{\partial \theta_x} \langle \mathbf{n}_0 | \frac{\partial \mathbf{n}_0}{\partial \theta_y} \rangle \right] \quad (5.58)$$

If we average over all fluxes we can then write:

$$\sigma_{xy} = -\frac{e^2}{\hbar} \int_{\mathbf{T}_{\Phi}^2} \frac{d^2\theta}{(2\pi)^2} \Omega_{xy} \quad (5.59)$$

and given the integral of the Berry curvature is quantized as from (5.13) the Hall conductivity is quantized:

$$\sigma_{xy} = -\frac{e^2}{\hbar} \int_{\mathbf{T}_{\Phi}^2} \frac{d^2\theta}{(2\pi)^2} \Omega_{xy} = -\frac{e^2}{2\pi\hbar} C; C \in \mathbb{Z} \quad (5.60)$$

Where C is the first Chern number, this is a first simplified version of the TKNN formula, valid for the base state \mathbf{n}_0 and this is of course the statement of the quantization for the integer Quantum Hall Effect.

5.5.3. TKNN formula for particles in a lattice

We want now to assemble a more general version of the TKNN formula, applying it for particles in a lattice, and this same approach is key to understand topological insulators. For this first part we will follow the original TKNN paper¹³.

We start from a electrons in a periodic potential $U(x, y)$ which has periods a and b in the directions x and y , and a uniform vertical magnetic field. $U(x, y)$ is a periodic perturbation potential which can be then tuned to a Landau Levels model for small periodic potentials $|U| \ll \hbar\omega_C$, and to the tight binding limit for strong periodic potentials.

Given is periodic this potential leads to a band structure and the number of bands will depend on the filling factor $\phi = ab\frac{eB}{hc}$, the filling factor can be interpreted also as the number of flux quanta per unit cell. We take ϕ to be a rational number $\frac{p}{q}$ and we use the Landau gauge with components $(0, eBx)$. We can then use the generalized Bloch condition:

$$\psi_{k_1, k_2}(x + qa, y)e^{-2\pi i\frac{py}{b} - ik_1qa} = \psi_{k_1, k_2}(x, y + b)e^{-ik_2b} = \psi_{k_1, k_2}(x, y) \quad (5.61)$$

where $k_1 \bmod \frac{2\pi}{aq}$ and $k_2 \bmod \frac{2\pi}{b}$ are good quantum numbers. We can now define the functions $u_{k_1, k_2} = \psi_{k_1, k_2}e^{-ik_1x - ik_2y}$ which satisfy the generalized periodic boundary conditions:

$$u_{k_1, k_2}(x + qa, y)e^{-2\pi i\frac{py}{b}} = u_{k_1, k_2}(x, y + b) = u_{k_1, k_2}(x, y) \quad (5.62)$$

If we substitute in the Hamiltonian of the Landau levels with a small perturbation U we obtain:

$$\hat{H}(k_1, k_2) = \frac{1}{2m}(-i\hbar\frac{\partial}{\partial x} + \hbar k_1)^2 + \frac{1}{2m}(-i\hbar\frac{\partial}{\partial x} + \hbar k_2 - eBx)^2 + U(x, y) \quad (5.63)$$

Ultimately we can see that it is a bit like we have introduced an effective crystal momentum $\hbar k_1, \hbar k_2$, which is the effect of what is called a magnetic translation operator¹⁴.

The lattice here is given by the periodic perturbation potential, and therefore it can be realized in different ways such as a crystal lattice or a periodic electromagnetic field. The fact that the potential is periodic grants translation invariance, and the definition of \mathbf{k} only makes sense thanks to this periodicity and the fact that \mathbf{k} is conserved. In the case of a generic magnetic field, translation invariance will be broken, and therefore momentum is not conserved, again in that case we would need a periodicity to achieve the same translation invariance.

If the particles are in a lattice their momentum lies again on a torus \mathbf{T}^2 which is the Brillouin zone. One can again define a Berry connection on the Brillouin zone and the integral for each band will be again a Chern number which determines the Hall conductivity.

¹³ D. J. Thouless et al. (1982). “Quantized Conductance in a Two-Dimensional Periodic Potential”. In: *Physical Review Letters* 49, p. 405

¹⁴ J. Zak (1964a). “Magnetic translation Group”. In: *Phys. Rev.* 134 (6A), A1602 J. Zak (1964b). “Magnetic translation Group II. Irreducible Representations”. In: *Phys. Rev.* 134 (6A), A1607

We can then sum across all filled bands and obtain:

$$\sigma_{xy} = \frac{e^2}{2\pi\hbar} \sum_n C_n, \quad (5.64)$$

which is again the TKNN formula for the Hall conductivity.

To prove it we start from the Hall conductivity (5.50). We are in the case where the parameter space is the reciprocal \mathbf{k} space and the sum (or integration) is extended over the Brillouin zone.

Is important to remember that \mathbf{n} corresponds to two quantum numbers n, \mathbf{k} therefore we can split the sum across the two quantum numbers

$$\sigma_{xy} = i\hbar \sum_{n' \neq n} \sum_{\mathbf{k}', \mathbf{k}} \frac{\langle n, \mathbf{k} | \hat{j}_y | n', \mathbf{k}' \rangle \langle n', \mathbf{k}' | \hat{j}_x | n, \mathbf{k} \rangle - \langle n, \mathbf{k} | \hat{j}_x | n', \mathbf{k}' \rangle \langle n', \mathbf{k}' | \hat{j}_y | n, \mathbf{k} \rangle}{(E_{n, \mathbf{k}} - E_{n', \mathbf{k}'})^2}, \quad (5.65)$$

notice that \mathbf{k}, \mathbf{k}' are independent because they are summed in the scope of their respective band n, n' . We know that:

$$\langle n, \mathbf{k} | \hat{j}_y | n', \mathbf{k}' \rangle = \langle n, \mathbf{k} | \hat{j}_y | n', \mathbf{k} \rangle \delta_{\mathbf{k}\mathbf{k}'} \quad (5.66)$$

It remains

$$\sigma_{xy} = i\hbar \sum_{n' \neq n} \sum_{\mathbf{k}} \frac{\langle n, \mathbf{k} | \hat{j}_y | n', \mathbf{k} \rangle \langle n', \mathbf{k} | \hat{j}_x | n, \mathbf{k} \rangle - \langle n, \mathbf{k} | \hat{j}_x | n', \mathbf{k} \rangle \langle n', \mathbf{k} | \hat{j}_y | n, \mathbf{k} \rangle}{(E_{n, \mathbf{k}} - E_{n', \mathbf{k}'})^2}, \quad (5.67)$$

Due to translation invariance \mathbf{k} is conserved and therefore:

$$\hat{\mathbf{j}} = \frac{e}{\hbar} \frac{\partial \hat{H}}{\partial \mathbf{k}}, \quad (5.68)$$

We can write the Hall conductivity:

$$\sigma_{xy} = \frac{ie^2}{\hbar} \sum_{n' \neq n} \sum_{\mathbf{k}} \frac{\langle n, \mathbf{k} | \frac{\partial \hat{H}}{\partial k_y} | n', \mathbf{k} \rangle \langle n', \mathbf{k} | \frac{\partial \hat{H}}{\partial k_x} | n, \mathbf{k} \rangle - \langle n, \mathbf{k} | \frac{\partial \hat{H}}{\partial k_x} | n', \mathbf{k} \rangle \langle n', \mathbf{k} | \frac{\partial \hat{H}}{\partial k_y} | n, \mathbf{k} \rangle}{(E_{n, \mathbf{k}} - E_{n', \mathbf{k}'})^2} \quad (5.69)$$

Where we can already recognize the sum on the left to be similar to one of the forms of writing the Berry curvature as from (5.9).

Given for large systems \mathbf{k} will be continuous we can write the sum as an integral:

$$\sigma_{xy} = \frac{e^2}{\hbar} \sum_{n' \neq n} \int_{BZ} \Omega_{xy}^n(\mathbf{k}) \frac{d^2 \mathbf{k}}{(2\pi)^2}, \quad (5.70)$$

If we substitute the quantization condition (5.13) we obtain:

$$\sigma_{xy} = \frac{e^2}{2\pi\hbar} \sum_{n' \neq n} C_n, \quad (5.71)$$

which is the TKNN formula as expected.

To give a bit more physical intuition We can also rewrite this formula making explicit the sum as from the original Thouless paper:

$$\sigma_H = \frac{ie^2}{\hbar} \sum_{E_n < E_F} \sum_{E_{n'} > E_F} \frac{(\frac{\partial \hat{H}}{\partial k_y})_{nn'} (\frac{\partial \hat{H}}{\partial k_x})_{n'n} - (\frac{\partial \hat{H}}{\partial k_x})_{nn'} (\frac{\partial \hat{H}}{\partial k_y})_{n'n}}{(E_n - E_{n'})^2} \quad (5.72)$$

The Fermi Level E_F is in the middle of two bands which are gapped, the first sum is achieved on the set of fully occupied levels in the bands lower than E_F and the second sum is on the levels occupied in the upper bands. This is also one of the definitions of the term topological insulators given the Fermi level is in the middle of the gapped bands.

To conclude analyzing the TKNN formula (5.71) we can again state that for each band there is a Chern number C_n which is a topological characteristic number of the band, that each band has a conductivity proportional to it's chern number and that the total Hall conductivity is quantized.

References

- Kubo, Ryogo (June 1957). “Statistical-Mechanical Theory of irreversible processes I.” In: *Journal of the physical society of japan* 12.6, p. 570.
- Laughlin (1981). “Quantized Hall conductivity in two dimensions”. In: *Physical Review Letters* B 23, p. 5632.
- Simon, Barry (Dec. 1983). “Holonomy, the Quantum Adiabatic Theorem, and Berry’s Phase”. In: *Phys. Rev. Lett.* (51), p. 2167.
- Thouless, D. J. et al. (1982). “Quantized Conductance in a Two-Dimensional Periodic Potential”. In: *Physical Review Letters* 49, p. 405.
- Zak, J. (1964a). “Magnetic translation Group”. In: *Phys. Rev.* 134 (6A), A1602.
- (1964b). “Magnetic translation Group II. Irreducible Representations”. In: *Phys. Rev.* 134 (6A), A1607.

Readings and online resources

Tong, David (2016a). *Lectures on the Kinetic theory*. <http://www.damtp.cam.ac.uk/user/tong/kinetic.html>. [Online; accessed 1-March-2021].

6 — Chern Simons theories

6.1. Introduction

This chapter gives a brief introduction to Chern-Simons theory, a form of Gauge theory. From a field theoretic perspective this shall be studied in the context of Yang Mills theory again a Gauge theory. Both of the two theories can be considered as extensions to Quantum Electrodynamics. We will see how the Chern-Simons theory is interconnected with the Integer Quantum Hall Effect and with topology.

Given the breath of the subjects and techniques involved we refer to other publications and resources for further study. In general it's important to mention the availability of other resources online such as the lecture notes of Dan Freed¹, Gregory Moore² and Dunne³.

6.2. Physics Introduction

All across this chapter we will adopt a field theoretic Lagrangian approach, where the equation of motions are just a consequence of the choosen Lagrangian. Typical Lagrangian terms encountered in field theory couple with the metric. As an example we provide here the free field lagrangian of Electromagnetism:

$$\mathcal{L}_{EM} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu}. \quad (6.1)$$

We can quickly generalize this Lagrangian to a curved manifold by introducing the metric $g^{\mu\nu}$:

$$\mathcal{L}_{EM} = -\frac{1}{4g^2}\sqrt{-g}g^{\mu\rho}g^{\nu\sigma}F_{\mu\nu}F_{\rho\sigma}. \quad (6.2)$$

The novelty of Chern-Simons terms is that is possible to build Lagrangians which do not couple to the metric and it that sense these Chern-Simons terms are called topological. A key part of the behaviour of topological Lagrangian terms is driven by boundary conditions of the derived equations and boundary integrals while expanding the Lagrangian with integration by

¹ Dan Freed (2022). *Lectures*. <https://ncatlab.org/nlab/show/Chern-Simons+theory>. [Online; accessed 1-April-2022]

² Gregory W. Moore (2022). *TASI Lectures*. <https://www.physics.rutgers.edu/~gmoore/TASI-ChernSimons-StudentNotes.pdf>. [Online; accessed 1-April-2022]

³ Dunne (1998). *Lectures at the 1998 Les Houches Summer School*. <https://arxiv.org/pdf/hep-th/9902115.pdf>. [Online; accessed 1-April-2022]

parts. The most simple of such terms is the Chern-Simons term:

$$\mathcal{L} = \frac{K}{4\pi} \varepsilon^{\mu\nu\rho} A_\mu \partial_\nu A_\rho. \quad (6.3)$$

This is also called the "Abelian" Chern-Simons term, this is the most simple topological term available in 3 dimensions. It is important to see how these terms generalize. In fact still in 3 dimension there is a second "Non-Abelian" Chern-Simons term:

$$\mathcal{L}_{CS} = \frac{K}{4\pi} \varepsilon^{\mu\nu\rho} \text{Tr}[A_\mu \partial_\nu A_\rho + \frac{2}{3} A_\mu A_\nu A_\rho]. \quad (6.4)$$

One key aspect of these terms is that they are antisymmetric multi-linear forms in the potentials and the derivatives, which is a reason why they do not couple to the metric. Given it is possible to build Quantum Field Theories from such Lagrangians, this has led to applications in particle physics first and then to applications in condensed matter.

The "Non-Abelian" term link us to the mathematical motivations of these terms which are the Chern-Simons Forms, but we will not cover it here. ^{4 5 6}

6.3. Justification of the Chern-Simons terms

It is always possible to invent a Lagrangian term as an Ansatz, try to validate it against a set of criterias such as symmetries, and then derive some experimental results from the field equations. In this section we will see that Chern-Simons terms are meaningful and significant with respect to the physics of the Quantum Hall Effect that we want to describe.

6.3.1. Parity and time reversal is violated

If we want to describe the Quantum Hall Effect, there is always a magnetic field, and whenever there is a magnetic field both parity and time reversal are violated. The Chern-Simons terms have this same property because they respect rotational invariance but break both parity and time reversal. There are also other 2 dimensional systems, e.g based on graphene, where parity is violated. In such systems time reversal may be also broken due to an external magnetic field, a circularly polarized light, extra terms such as the Haldane next nearest neighbour hopping term ⁷, or terms based on the spin. Also here Chern-Simons theory can be applied.

⁴ "Characteristic Forms and Geometric Invariants" (Jan. 1974). In: *Annals of Mathematics, Second Series* 99.1, pp. 48–69

⁵ Edward Witten (1988). "Topological Quantum Field Theory". In: *Commun. Math. Phys.* 117, pp. 353–386

⁶ Albert Schwarz (2000). *Topological Quantum Field Theories*. <https://arxiv.org/abs/hep-th/0011260>

⁷ F. D. M. Haldane (2015–2018). "Model for a Quantum Hall Effect without Landau Levels: Condensed-Matter Realization of the "Parity Anomaly."". In: *Physical Review Letters*, 61(18)

6.3.2. Long distance physics

For a generic Electromagnetic field we can write the Lagrangian density as:

$$\mathcal{L} = \frac{1}{2}\epsilon E^2 - \frac{1}{2\mu}B^2 + \mathcal{O}(E, B)^4 \quad (6.5)$$

where $\mathcal{O}(E, B)^4$ are higher order terms. A Lagrangian such as this one which is built from the fields is automatically gauge invariant, because the fields are gauge invariants. The terms in the fields of the Lagrangian are second order in the derivatives and second order in the potentials. If we now compare the electromagnetic field term with a Chern-Simons term this is instead second order in the potential and first order in the derivatives. This means that at large distances the Chern-Simons term can be considered the lowest order, of total order 3, and therefore the most important one. In this sense the Chern-Simons theory can be considered an effective field theory.

6.3.3. Possible Lagrangian functions

Another way too see that the Chern-Simons Lagrangians are meaningful is to consider all possible Lagrangian(s) terms available and their combinations, restrict this set with some reasonable constraints and see if we obtain something similar to Chern-Simons terms.

In order to have a Lagrangian that does not couple to the metric we will need to have an anti-symmetric term in the potentials and derivatives. Therefore we will need to rely on linear combinations based on the Levi Civita symbol. If we then want to build a Lagrangian of third order the last term must be a potential and we can only have combinations like:

$$\varepsilon^{\mu\nu\rho}\partial_\mu A_\nu A_\rho, \quad \varepsilon^{\mu\nu\rho}\partial_\mu\partial_\nu A_\rho, \quad \varepsilon^{\mu\nu\rho}A_\mu\partial_\nu A_\rho, \quad \varepsilon^{\mu\nu\rho}A_\mu A_\nu A_\rho. \quad (6.6)$$

We are here using the Einstein summation convention. It's clear that the first two are zero given they are symmetric in two terms which are summed together by the Levi civita symbol. The third one, which is the Abelian Chern-Simons one, is one actual possible term. Finally the last one is again zero.

If we would want to do this for a order four term, adding either a ∂_ξ or a A_ξ we will always have at least two symmetric terms, and this would actually lead again the cancellation with the levi civita symbol $\varepsilon^{\mu\nu\rho\xi}$ of order four, the only term that may not cancel out is $\varepsilon^{\mu\nu\rho\xi}\partial_\mu A_\nu\partial_\rho A_\xi$ but it is possible to prove that it still does.

In the case of order five we need to add one of the set ∂_ξ, A_ξ and one of ∂_χ, A_χ , in total we will obtain terms like $\varepsilon^{\mu\nu\rho\xi\chi}A_\mu\partial_\nu A_\rho\partial_\xi A_\chi$, out of which some of them will be not zero.

This same reasoning can then be applied to any order to conclude the following:

- Only the odd orders have non trivial terms, and this is the case also for the Chern-Simons terms.
- A lot of these Lagrangian terms actually ends up to have the same order in the derivatives

and the potentials of Chern-Simons terms of higher order.

- Lagrangians are built out of linear combinations of terms, if we carefully look at higher order Chern-Simons terms, they contain coefficients that this approach is not capable to predict.

6.3.4. Anyons

A peculiarity of the 2 dimensional physics at play is summarized, in the special particle statistics, which is a mix between fermionic and bosonic behaviours, these type of fields are what we call Anyons.

Let now a_μ be an emergent gauge field, not just the external electromagnetic field A_μ , the Chern-Simons Lagrangian density, which exists only in 2+1 dimensions:

$$\mathcal{L} = \frac{K}{4\pi} \varepsilon^{\mu\nu\rho} a_\mu \partial_\nu a_\rho. \quad (6.7)$$

The standard quantum field theory quantization procedure leads to field operators \hat{a}_μ , and to a basis of creation and annihilation operators that have peculiar particle statistics.

It's possible to arrive to the statistics with an heuristic argument about particle exchange⁸ but also with a more robust argument on the topological structure of the configuration space⁹ More in general Chern-Simons theories are generalizations of Yang Mills gauge theories in odd dimensions¹⁰.

6.4. Dirac quantization condition

In order to derive the quantization of the Quantum Hall current, we need to prove the fact that the parameter k of the Chern-Simons Lagrangian is quantized. In order to do that we will now introduce the Dirac quantization condition, which is a constraint between the quantization of charge and the quantization of magnetic flux. This is a novelty with respect to classical electrodynamics due to the way the quantum mechanics momentum operator couples directly with the electromagnetic potential, as in the minimal coupling substitution $\hat{\mathbf{p}} \rightarrow \hat{\mathbf{p}} - q\mathbf{A}$.

6.4.1. Introduction

In classical electrodynamics the electromagnetic potential can be always removed from the equations, and the field equations can always be expressed in terms of the derivatives of the potential. In quantum mechanics this is no longer the case, because the electromagnetic potential is explicitly in the Schroedinger equation.

⁸ David Tong (2016b). *Lectures on the Quantum Hall Effect*. <http://www.damtp.cam.ac.uk/user/tong/qhe.html>. [Online; accessed 12-October-2021], 90-91

⁹ Jon Magne; Myrheim Leinaas (Jan. 1977). "On the theory of identical particles". In:

¹⁰ Ademola Adeifeoba (n.d.). *Introduction to Chern-Simons Theory*. <https://osf.io/preprints/africarxiv/am7pt/download>

If there are points, surfaces or other regions where the potential or its derivative does not exist, or is multivalued, these may lead to trouble for the existence and uniqueness of the solutions. Such points may lead to constraints and quantization conditions of the magnetic flux and the electromagnetic charges.

In the previous chapters we have seen such cases for the Berry phase and Aharonov Bohm effect. Considering a generic manifold covered by an atlas, and the atlas made of separate charts for different regions of the manifold. The potential was defined differently on different charts and made continuous with a gauge transformation on the border of the regions where multiple charts are available.

This was analyzed by Dirac in 1931¹¹ and is now known as Dirac monopole argument. It can be considered the precursor of what we have seen in the case of the Aharonov Bohm and the Berry phase. Ultimately Dirac was asking why the electron charge is quantized, and if that quantization can be reduced to something similar to observables of standard quantum mechanics where the charge is some eigenvalue of some equation. In more modern terms we may ask if the quantization of charge is attached to something more fundamental about the electromagnetic field and potentials or it comes out for example from some geometrical or topological argument.

Dirac understood that the phase, which is what we now call the Berry phase, depends on the path and is directly related with the introduction of the electromagnetic field and the gauge transformations. He then achieved to describe the mechanics of gauge transformations on the border of the charts. From there he deduced that if there is a duality between electric and magnetic charges, then it follows that the quantization condition of one leads to the quantization condition of the other.

6.4.2. Electro Magnetic duality

The first important consideration is that classical electrodynamics without sources is completely symmetric in the electric and magnetic fields. This is possible to see directly from the free field Maxwell equations:

$$\begin{aligned} \nabla \cdot \mathbf{E} &= 0 & \nabla \cdot \mathbf{B} &= 0 \\ \nabla \wedge \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} & \nabla \wedge \mathbf{B} &= \frac{\partial \mathbf{E}}{\partial t}. \end{aligned} \tag{6.8}$$

Therefore a part from Lorentz invariance we have what is called Electromagnetic duality. The free field Maxwell equations are invariant under the transformation:

$$(\mathbf{E}, \mathbf{B}) \rightarrow (\mathbf{B}, -\mathbf{E}) \tag{6.9}$$

A natural way to generalize Maxwell equations is to introduce Magnetic sources μ , i.e. the

¹¹P. A. M. Dirac, Quantised singularities in the electromagnetic field, Proc. R. Soc. A133 (1931), 60–72.

magnetic monopoles and the monopoles currents \mathbf{K} :

$$\begin{aligned} \nabla \cdot \mathbf{E} &= \rho & \nabla \cdot \mathbf{B} &= \mu \\ \nabla \wedge \mathbf{E} &= \mathbf{K} - \frac{\partial \mathbf{B}}{\partial t} & \nabla \wedge \mathbf{B} &= \mathbf{J} + \frac{\partial \mathbf{E}}{\partial t}. \end{aligned} \quad (6.10)$$

We have for simplicity embedded the permittivity inside the currents, meaning that the currents are actually effective currents including the effect of the propagating medium.

In tensor notation we can introduce the Hodge dual tensor:

$$*F^{\alpha\beta} = \frac{1}{2}\varepsilon^{\alpha\beta\gamma\delta}F_{\gamma\delta} \quad (6.11)$$

Which is actually the same as the substitution (6.9). Still using tensor notation, the generalization of Maxwell equations becomes:

$$\begin{cases} \partial_\alpha F^{\alpha\beta} = j^\beta \\ \partial_\alpha *F^{\alpha\beta} = 0 \end{cases} \rightarrow \begin{cases} \partial_\alpha F^{\alpha\beta} = j^\beta \\ \partial_\alpha *F^{\alpha\beta} = k^\beta \end{cases}. \quad (6.12)$$

We have the original Maxwell equations on the left and the generalized ones on the right. The k^β is the quad-current of the newly introduced sources i.e. the magnetic monopoles. Of course the generalized version reduces to the standard one as far as there are no monopoles i.e for $k^\beta = 0$ ¹²

Introducing magnetic monopoles is actually a big logical step. It is a fully separate hypothesis and our arguments about quantization, should ideally work also without them. Another way to look at this is that generalizing the electromagnetic field this way we are modifying the homogenous Maxwell equations. From a tensor calculus standpoint these are the Bianchi identities, which should come out automatically from the geometry and therefore should not be generalized. It is to be seen then what the Bianchi identities are for the generalized equations, how they can be interpreted, and if the generalization did not remove an essential constraint in the Maxwell equations.

If we go for a minimalistic approach without introducing magnetic monopoles, Maxwell equations maybe be correct as they are. We accept the asymmetry between electric and magnetic charges or more precisely the non-existence of monopoles, since in that case the magnetic field is just a dynamic counterpart of the electric field. This is consistent with special relativity, and electric charges are dual then just to magnetic fluxes, which was the original spirit of Gauss.

Using the duality between electric charges and magnetic fluxes, the Dirac quantization condition will then become that if the electric charge is quantized by consequence the flux is quantized. This is exactly what we have seen as the flux quantization condition in the aharanov-bohm Chapter, and whenever we use the term magnetic monopole in what follows we can also call it a quantum of flux.

¹²Electromagnetic Duality for Children JM Figueroa-O’Farrill <https://www.maths.ed.ac.uk/~jmf/Teaching/Lectures/EDC.pdf>

6.4.3. The monopole vector potential

Let's start to ask ourselves what is logically the most simple field for a magnetic monopole. For an isolated point like electric charge q the field goes with a radial square law. This is a fundamental solution of the Gauss law for the electric field. It's then logical to expect that the most simple possibility for a magnetic charge is again to be point alike. If we imagine a point like static magnetic charge of intensity g we will write the generalized Gauss law of the magnetic field as

$$\nabla \cdot \mathbf{B}(r) = g. \quad (6.13)$$

The solution to this equation is again a radial field whose intensity goes with a square law. Let's try to write it down as:

$$\mathbf{B}(r) = \frac{g}{4\pi} \frac{1}{r^2} \mathbf{e}_r. \quad (6.14)$$

In this conventions the magnetic charge g is the exact same as the magnetic flux

$$g = \int_{\Sigma} \mathbf{B} \cdot \mathbf{n} \, dS. \quad (6.15)$$

We want now to find a magnetic potential $\mathbf{B} = \nabla \wedge \mathbf{A}$ for this field. We can start from an ansatz:

$$\begin{aligned} \theta \neq \pi : \mathbf{A}_+ &= \frac{g}{4\pi r} \frac{1 - \cos \theta}{\sin \theta} \hat{\varphi} \\ \theta \neq 0 : \mathbf{A}_- &= -\frac{g}{4\pi r} \frac{1 + \cos \theta}{\sin \theta} \hat{\varphi} \end{aligned} \quad (6.16)$$

This solution is not defined as a single function on \mathbb{R}^3 but is defined piece-wise on two charts: \mathbf{A}_+ is defined on the chart that is the full \mathbb{R}^3 without the negative z-axis (i.e. $\theta \neq \pi$), and \mathbf{A}_- is defined on the full \mathbb{R}^3 without the positive z-axis (i.e. $\theta \neq 0$). The combination of the two charts covers the full \mathbb{R}^3 and the combination of \mathbf{A}_+ and \mathbf{A}_- is our actual function definition. It is possible to prove that there is no single function that can be defined across the full \mathbb{R}^3 that respects the equations, and the only type of solutions that can be built are defined piece-wise on multiple charts as (6.16).

This half of an axis of singularity is what is now known as Dirac string. Some people interpret this as an infinitely thin solenoid that starts at the center of the coordinate system and continue to infinity along the negative z axis. A magnetic monopole in general can be interpreted as the extremity of such a solenoid that lies at the center of the coordinate system. The Dirac string can also be interpreted as a type of branch cut, i.e a path in the domain of definition of a complex function along which the function is not properly defined, either being multi-valued or divergent. Typically branch cuts may end on to singular points as is the case here.

On the contrary of what one might expect, this solution is perfectly legitimate. In fact in the regions where both \mathbf{A}_+ and \mathbf{A}_- are defined there is always a Gauge transformation from one to the other. With this technique we can always split the space in such a way that we can define piece-wise a vector potential that is single valued and continuous. In fact in this case their common region of definition is the complement of the z axis with respect to \mathbb{R}^3 . It is important to notice that this region is not simply connected. On this region the curl of the difference is zero and there can always be a gauge transformation between the two:

$$\begin{aligned}\nabla \wedge (\mathbf{A}_+ - \mathbf{A}_-) &= 0 \\ \mathbf{A}_+ - \mathbf{A}_- &= \nabla \chi\end{aligned}\tag{6.17}$$

for some scalar function χ . The difference of the two potentials is then irrotational, but given the domain is not simply connected, the same difference is not a conservative field, and therefore there is no unique function whose gradient can lead to the difference. In other words the gauge transformation function χ cannot be unique across the full domain.

For our monopole field we can see that:

$$\mathbf{A}_+ - \mathbf{A}_- = \frac{g}{2\pi r} \hat{\varphi} = \nabla \chi = \frac{1}{r \sin \theta} \frac{\partial \chi}{\partial \varphi} \hat{\varphi}.\tag{6.18}$$

As an example we can choose for the border of the two charts the condition $\theta = \frac{\pi}{2}$ which is the x and y plane. This reduces to:

$$\mathbf{A}_+ - \mathbf{A}_- = \frac{g}{2\pi r} \hat{\varphi} = \frac{1}{r} \frac{\partial \chi}{\partial \varphi} \hat{\varphi}\tag{6.19}$$

On this border we can then do a gauge transformation to match the two definitions of \mathbf{A}_+ and \mathbf{A}_- and this will be generated by

$$\chi = \frac{g}{2\pi} \varphi.\tag{6.20}$$

given φ is an angle this is a multivalued function.

6.4.4. The Dirac quantization condition

We finally now compute an integral over a sphere Σ surrounding the monopole, where Σ_{\pm} are the upper and lower hemisphere and E is the circle in between the two. We can write:

$$\begin{aligned}g &= \int \mathbf{B} \cdot \mathbf{n} \, dS = \int_{\Sigma_+} (\nabla \wedge \mathbf{A}_+) \cdot \mathbf{n} \, dS + \int_{\Sigma_-} (\nabla \wedge \mathbf{A}_-) \cdot \mathbf{n} \, dS = \\ &= \int_E \mathbf{A}_+ \cdot d\mathbf{l} - \int_E \mathbf{A}_- \cdot d\mathbf{l} = \int_E \nabla \chi \cdot d\mathbf{l} = \chi(2\pi) - \chi(0).\end{aligned}\tag{6.21}$$

Now we introduce the schroedinger equation:

$$\frac{1}{2m}(-i\hbar\nabla + q\mathbf{A})^2\psi = i\hbar\frac{\partial\psi}{\partial t}. \quad (6.22)$$

We are going to solve it for the vector potential of the monopole (6.16).

The equation shall be invariant by any gauge transform:

$$\begin{aligned} \mathbf{A} &\rightarrow \mathbf{A} + \nabla\chi' \\ \psi &\rightarrow e^{-i\frac{q}{\hbar}\chi'}\psi \end{aligned} \quad (6.23)$$

To maintain the wave function single valued we need then to enforce that the two phases are the same on a common border of the two charts. The first phase is acquired by gauge transformation as from (6.23). The second phase is due to the jump between the two different charts as from (6.20). For the border of the chart we use again $\theta = \frac{\pi}{2}$. We thus require that:

$$e^{-i\frac{q}{\hbar}\chi'} = e^{-i\frac{q}{\hbar}\frac{g\varphi}{2\pi}}, \quad (6.24)$$

and therefore:

$$\frac{q}{\hbar}\chi' = \frac{q}{\hbar}\frac{g\varphi}{2\pi} + 2\pi n; \quad n \in \mathbb{Z}. \quad (6.25)$$

Given we already computed $g = \chi(2\pi) - \chi(0)$ with our integral, we can reduce it to the Dirac quantization condition:

$$\frac{q}{\hbar}g = 2\pi n; \quad n \in \mathbb{Z}. \quad (6.26)$$

If one prefers the version without magnetic monopoles:

$$\frac{q}{\hbar} \int_{S^2} \mathbf{B} \cdot \mathbf{n} \, dS = 2\pi n; \quad n \in \mathbb{Z}. \quad (6.27)$$

This does not mean that Gauss theorem for the magnetic field is not correct, and the divergence of the field is not zero. Instead it means that there is a complex interplay between two effects. One is the gauge transform, the other is the non uniqueness of the potential across the full space, and the two effects compensate each other exactly to maintain the divergence zero.

6.4.5. An aside over the existence of magnetic monopoles

Introducing magnetic monopoles is quite well accepted in the research community, because it can be said that until we find them all the equations are still standing, and is just enough to have one single monopole existing at the other end of the universe to justify this quantization argument. This is a weak logical argument because what you typically do here is an integral over a sphere including the monopole at the center and the surface of the sphere is where you locally

measure the field, that would imply some shape of assumption of the large scale structure of the universe such as being open, simply connected and essentially flat on the large scale, which is far from trivial.

6.5. Chern-Simons theories

6.5.1. The Gauge transformation

Let's now restrict our self to 2+1 dimensions. In this case is possible to introduce a Chern-Simons Lagrangian density term which is:

$$\mathcal{L}_{CS}[A] = \frac{K}{4\pi} \epsilon^{\mu\nu\rho} A_\mu \partial_\nu A_\rho; \quad \mu, \nu, \rho = 0, 1, 2 \quad (6.28)$$

As we have seen in (6.6), this term is the only significant one in 3 dimensions. In more generality as we have seen in section 6.3.4 Chern-Simons characteristic classes are available only in odd dimensions, and the case of 3 dimensions is the first one that leads to non trivial physics.

Yet another way is to say the same is that these type of Lagrangians are a type of topological quantum field theory. These were first proposed by Schwartz and further expanded by Witten. These are theories where the topology of the manifold is important. The Lagrangian is not dependent on the metric, and the integrals at the border are relevant for the physics.

Finally, is important to distinguish two notations available in the literature: often the A_μ is reserved for external magnetic fields where instead a_μ is reserved for quantized fields, or dynamical degrees of freedom. The same Chern-Simons Lagrangian can be then constructed by a combination of both of the two.

This type of term violates parity, and this is always the case when there is a magnetic field.

Let's prove how this term behave with respect to Gauge invariance. Given it is explicitly dependent on the vector potential A_μ , it's not at all evident that the action is gauge invariant.

Under a Gauge transformation the field and the Lagrangian transforms as

$$\begin{aligned} A_\mu(\mathbf{r}, t) &\rightarrow A_\mu(\mathbf{r}, t) + \partial_\mu \zeta(\mathbf{r}, t) \\ L'_{CS}[A] &= L_{CS}[A] + \frac{K}{4\pi} \int d^2x \epsilon^{\mu\nu\rho} \partial_\mu (\zeta \partial_\nu A_\rho). \end{aligned} \quad (6.29)$$

Note that there is a global derivative that plays a role here. More precisely, to evaluate that term we need to specify on which manifold we are integrating. Given the integrand is a total derivative we can simplify the integral through integration by parts, to an integral on the border.

In two spatial dimensions, there are now two cases: there are manifolds with a border such as a disk, and without border such as a sphere or torus. In regards to the time coordinate also here there are two cases, one when the time coordinate is a circle and therefore the system is periodic in time and a second case where the time coordinate is an open interval. Being periodic

in time means that the system after a certain amount of time returns to the initial state. In this case the system is holonomic, as we saw in the case of the adiabatic transformations in the Berry phase chapter.

6.5.2. Case without a border

We will now analyze the first case where the manifold does not have a border and therefore is compact and closed both in space and time. Examples of this can be a circle, a sphere or a torus. Having no borders in the time direction means being holonomic and typically this is reflected in an closed loop integral over the time coordinate.

The action functional that we build from the Lagrangian density is:

$$S_{CS}[A] = \frac{K}{4\pi} \int dt \int d^2x \epsilon^{\mu\nu\rho} A_\mu(\mathbf{x}) \partial_\nu A_\rho(\mathbf{x}) \quad (6.30)$$

At this point, if we write down the quantum field theory partition function from this action we have:

$$\mathcal{Z}[\mathbf{A}] = \int e^{i\frac{S_{CS}[\mathbf{A},\varphi]}{\hbar}} \mathcal{D}[\varphi]. \quad (6.31)$$

Here the vector potential \mathbf{A} is considered an external source, the fields φ are the dynamical degrees of freedom, and the integral $\mathcal{D}[\varphi]$ is a variation over all possible fields configurations on all space time.

A generic expectation value for a measurable quantity F is given by:

$$\langle F \rangle = \frac{1}{\mathcal{Z}[\mathbf{A}]} \int e^{i\frac{S_{CS}[\mathbf{A},\varphi]}{\hbar}} F[\mathbf{A}, \varphi] \mathcal{D}[\varphi], \quad (6.32)$$

where $F[\mathbf{A}, \varphi]$ is a generic functional in the fields φ , having the sources \mathbf{A} as parameters.

The vector potential \mathbf{A} and a generic measurable quantity F must be single valued. We can say that even if the action S_{CS} or the Lagrangian \mathcal{L}_{CS} is not single valued, the quantity $e^{i\frac{S_{CS}[A]}{\hbar}}$ must be single valued.

If we consider the gauge transformation from (6.29) we would still require that:

$$S'_{CS}[A] = S_{CS}[A] + 2\pi n\hbar; n \in \mathbb{Z}. \quad (6.33)$$

We then get a consistency condition:

$$2\pi n\hbar = \frac{K}{4\pi} \int dt \int d^2x \epsilon^{\mu\nu\rho} \partial_\mu (\zeta \partial_\nu A_\rho); n \in \mathbb{Z}. \quad (6.34)$$

With a bit of differential topology it is possible to say that the second integral is quantized, and therefore K is quantized. This is in essence due to the Chern-Gauss-Bonnet theorem and the Atiyah-Singer index theorem.

This goes back again to the idea that the ζ is multivalued and it's in essence the same mathematics of the border integral of the berry phase and the flux quantization. The case without border leads to a quantization condition of K . The case with border will lead to the physics of edge states, i.e the quantization of the border states, that we will not cover here.

6.5.3. Quantization

It is possible to prove that from equation (6.34) we can derive the quantization of K in general. We will instead follow David Tong and show that this is true in a very special case with a set of limiting hypothesis.

Let's start to expand the Action into the potential and the fields:

$$S_{CS} = \frac{K}{4\pi} \int d^3x \epsilon^{\mu\nu\rho} A_\mu \partial_\nu A_\rho = \frac{K}{4\pi} \int d^3x [A_0 F_{12} + A_1 F_{20} + A_2 F_{01}]. \quad (6.35)$$

We now introduce a simplified assumption of a constant magnetic field. This gives the constraints

$$\begin{aligned} \partial_t F_{12} &= 0 \\ \partial_t (\partial_1 A_2 - \partial_2 A_1) &= 0 \end{aligned} \quad (6.36)$$

Which is always true if also the potential does not depend on time. We can therefore set

$$\begin{aligned} \partial_0 A_2 &= 0 \\ \partial_0 A_1 &= 0 \end{aligned} \quad (6.37)$$

which leads to the following simplifications

$$\begin{aligned} F_{12} &= \partial_1 A_2 - \partial_2 A_1 \\ F_{20} &= \partial_2 A_0 - \partial_0 A_2 = \partial_2 A_0. \\ F_{01} &= \partial_0 A_1 - \partial_1 A_0 = -\partial_1 A_0 \end{aligned} \quad (6.38)$$

We now substitute (6.38) into (6.35) and then integrate by parts in the spatial derivatives

$$\begin{aligned} S_{CS} &= \frac{K}{4\pi} \int d^3x [A_0 F_{12} + A_1 \partial_2 A_0 - A_2 \partial_1 A_0] = \\ &= \frac{K}{4\pi} \int d^3x [A_0 F_{12} - \partial_2 (A_1) A_0 + \partial_1 (A_2) A_0] \\ &= \frac{K}{4\pi} \int d^3x [A_0 F_{12} + A_0 (\partial_1 (A_2) - \partial_2 (A_1))] \\ &= \frac{K}{2\pi} \int d^3x A_0 F_{12} \end{aligned} \quad (6.39)$$

We assume a manifold which is closed without a border (such as a sphere or a torus). Therefore the border terms of the integration by parts are zero.

Still to simplify our proof we will specify the gauge. We choose a gauge where the electric

potential is constant, $A_0 = a_0$ and this gives

$$S_{CS} = a_0 \frac{K}{2\pi} \int d^3x F_{12} = a_0 \frac{K}{2\pi} \int dt \int d^2x F_{12}. \quad (6.40)$$

As earlier we have done, we introduce the holonomy condition over time, i.e the fact that the manifold is closed in the time coordinate, therefore the integral over time is given by the periodicity over time:

$$\int dt = \beta \quad (6.41)$$

And we introduce the quantization condition (6.27) and obtain:

$$S_{CS} = a_0 \beta \frac{K}{2\pi} \frac{2\pi \hbar l}{e} = a_0 \beta \frac{K \hbar l}{e}; l \in \mathbb{Z} \quad (6.42)$$

Now we introduce a large Gauge transformation:

$$\begin{aligned} A_\mu &\rightarrow A_\mu + \partial_\mu \zeta \\ \zeta &= \tau \frac{2\pi \hbar m}{e\beta}; m \in \mathbb{Z} \\ A_0 &\rightarrow A_0 + \frac{2\pi \hbar m}{e\beta}; m \in \mathbb{Z} \end{aligned} \quad (6.43)$$

We call this large because in essence we use a multivalued gauge transformation function ζ , which is multivalued due to the integer m , and in patches the gauge transformation compensates the multivaluedness and permits to achieve a single valued vector potential field, in this sense the transformation can be larger than one patch.

Together with a Gauge transformation the wave function undergoes a local U(1) transformation, i.e. it is invariant by $\psi' \rightarrow \psi e^{i\zeta}$, in the same manner a quantized field $\hat{\mathbf{A}} \rightarrow \hat{\mathbf{A}} e^{i\zeta}$

The fact that the Gauge transformation is large is generic, but we used here a gauge transformation function ζ that is linear in the time τ , and that is due to our simplifying assumptions.

This induces a change in the action as:

$$\begin{aligned} S_{CS} &= a_0 \beta \frac{K \hbar l}{e}; l \in \mathbb{Z} \rightarrow \\ S'_{CS} &= \frac{\beta \hbar K l}{e} \left(a_0 + \frac{2\pi \hbar m}{e\beta} \right) = \\ &= S_{CS} + \frac{\beta \hbar K l}{e} \frac{2\pi \hbar m}{e\beta} = S_{CS} + \frac{2\pi \hbar^2 K l m}{e^2}; l, m \in \mathbb{Z} \end{aligned} \quad (6.44)$$

If we now compare this result with (6.33)

$$\begin{aligned} S'_{CS} &= S_{CS} + \frac{2\pi \hbar^2 K l m}{e^2}; l, m \in \mathbb{Z} \\ S'_{CS}[A] &= S_{CS}[A] + 2\pi n \hbar; n \in \mathbb{Z} \end{aligned} \quad (6.45)$$

we obtain a condition on n , given l and m arbitrary integers:

$$\frac{2\pi\hbar^2 K l m}{e^2} = 2\pi n \hbar; l, m \in \mathbb{Z} \quad (6.46)$$

Solving for n , gives us the equation:

$$n = \frac{\hbar K l m}{e^2} \quad (6.47)$$

and the only way to guarantee that n is an integer for arbitrary l and m is if K is a multiple of $\frac{e^2}{\hbar}$. Note that K carries the same units as σ_{xy} , so this is finally again the quantization condition of the integer quantum hall effect as we will see in the following section.

The fact that the Lagrangian density is not gauge invariant is not much of a problem given it is not a physical quantity. What we are really interested in are physical quantities, and how they behave in regards to gauge transforms.

As before, in the case of the Berry phase, the Berry connection is non physical and not gauge invariant, but the Berry phase is gauge invariant instead. What we have seen in the case of the Berry phase is that the gauge invariance of the Berry phase actually induces the quantization of the Berry phase, given that the gauge generating function is multivalued.

What we saw here is that the Lagrangian density is not Gauge invariant but the Action shall be Gauge invariant. That induces quantization of the parameters in the Lagrangian densities. Again there is a total derivative at play within the integral as in (6.29), and the gauge generating function is multivalued and defined differently on different charts.

6.5.4. From the Chern-Simons Lagrangian to the Hall current

From the Chern-Simons action we can compute the current

$$J_\mu = \frac{\partial S_{CS}[a]}{\partial A_\mu} = \frac{K}{2\pi} \varepsilon^{\mu\nu\rho} \partial_\nu A_\rho \quad (6.48)$$

If we identify the physical field we get:

$$J_i = -\frac{K}{2\pi} \epsilon_{ij} E_j \quad (6.49)$$

This means that we are describing a Hall conductivity

$$\sigma_{xy} = \frac{K}{2\pi} \quad (6.50)$$

If we identify the Chern factor as

$$K = \frac{e^2 \nu}{\hbar} \quad (6.51)$$

, this is a Hall conductivity of ν filled Landau levels. Given ν is quantized, as we have seen

from (6.47), we can identify this with the Landau levels.

6.5.5. Case with a border - edge states

In the case of a disk instead, or if we want a square sample, then the integral will have a border and that contribution will lead to edge states, i.e quantized states that live on the border of the manifold.

From a general perspective this can also be said that the border information is very relevant for the physics as it carries the topological information. Most Lagrangians considered in field theory are local, and metric dependent, and typically all the border terms on the integrations by parts are null. In this case, since the Lagrangian is not dependent on the metric and it carries some information on the border, the integration by parts then leads to a border contribution that is not zero.

6.6. Chern-Simons theories applied to Fractional Quantum Hall effects

As a brief entrée to applications and extensions of Chern-Simons theories is important to describe how Lagrangian terms of Chern-Simons type can be put into action in real theories.

Let's start from a Chern-Simons Lagrangian density (6.28). We can decompose the total field into two pieces: an external field A^μ and an internal field a^μ . Intuitively, the external field is the big field that we have for example in the case of the Quantum hall effect, but this field is taken to be constant. The internal field is instead is the oscillating field or coupling field. We then get the following equations:

$$\begin{aligned} A_T^\mu &= A^\mu + a^\mu \\ \partial_\nu A^\mu &= 0 \end{aligned} \quad (6.52)$$

We now want that the external field is not to be quantized and the internal field will be quantized. This is also reflected in the coupling term where the currents don't couple with the external fixed field:

$$\mathcal{L}_{coupling} = a_\mu j^\mu + A_\mu j^\mu \approx a_\mu j^\mu \quad (6.53)$$

If we now expand the Lagrangian density we have:

$$\mathcal{L} = -\frac{K}{4\pi} \epsilon^{\mu\nu\rho} a_\mu \partial_\nu a_\rho - \frac{K}{2\pi} \epsilon^{\mu\nu\rho} A_\mu \partial_\nu a_\rho + a_\mu j^\mu \quad (6.54)$$

Which is the Lagrangian density for the integer quantum hall effect. We can easily generalize the above Lagrangian by making the two insertions of K different introducing an integer value t :

$$\mathcal{L}_{wz} = -\frac{1}{4\pi} K \varepsilon^{\mu\nu\sigma} a_\mu \partial_\nu a_\sigma - \frac{e}{2\pi} t A_\mu \varepsilon^{\mu\nu\sigma} \partial_\nu a_\sigma + a_\mu j^\mu \quad (6.55)$$

This Lagrangian permits to describe one or more fluids with filling fractions $\frac{t^2}{K}$. For $t = 1$ this describes the long range physics for the Laughlin states. The fact that $t \neq K$ is important to allow a description of fractionalized charges.

We can now generalize again this Lagrangian (6.55) introducing n Chern-Simons fields a_μ^α the coupling constant K becomes now a matrix $K_{\alpha\beta}$ where $\alpha, \beta = 1, \dots, n$, and n is also the rank of the matrix, this to summarize the linear combination in the n fields. We introduce also a second set of linear coupling constants t^α to generalize the second term.

$$\mathcal{L}_{wz} = -\frac{1}{4\pi} K_{\alpha\beta} \varepsilon^{\mu\nu\sigma} a_\mu^\alpha \partial_\nu a_\sigma^\beta - \frac{e}{2\pi} t_\alpha A_\mu \varepsilon^{\mu\nu\sigma} \partial_\nu a_\sigma^\alpha + a_\mu^\alpha j_\alpha^\mu \quad (6.56)$$

Intuitively we can imagine these extra coupling fields as n quasiparticles, but we did not make any precise statement if these fields are localized or not. In the same spirit the coupling currents are generalized from a single current j^μ to a current for each extra field j_q^μ .

It's important to mention that $K^{\alpha\beta}$ and t^α is not just a set of continuous coefficients, but they encode only topological information (i.e. they don't couple to the local metric). More precisely K is a symmetric matrix with only integer coefficients (odd on the diagonal and even elsewhere). These coefficients also encode the symmetry of the modular group $SL(2, \mathbf{Z})$ and this leads therefore to the emergence of continued fractions, in the experimental results¹³.

We finally introduce two extra generalizations to arrive to the Wen Zee Lagrangian¹⁴:

$$\mathcal{L}_{wz} = -\frac{1}{4\pi} K_{\alpha\beta} \varepsilon^{\mu\nu\sigma} a_\mu^\alpha \partial_\nu a_\sigma^\beta - \frac{e}{2\pi} t_\alpha A_\mu \varepsilon^{\mu\nu\sigma} \partial_\nu a_\sigma^\alpha - \frac{s_\alpha}{2\pi} \omega_i \varepsilon^{i\nu\sigma} \partial_\nu a_\sigma^\alpha + a_\mu^\alpha l_\alpha^q j_q^\mu \quad (6.57)$$

l_α^q is a generalized coupling with the currents where the case $l_\alpha^q = \delta_\alpha^q$ is the case of minimal coupling, an example of the more general case here is a rotation of the coupling effect with respect to the direction of the currents.

We include n spinorial fields (one for each Chern-Simons field introduced previously) s^α and ω^i as a spin connection (i.e. a linear combination or more precisely linear form) where $i = 1, 2, 3$. These fields come into play and are not zero, when we solve the schrodinger equation on a surface that is not flat such as a sphere or a torus.

This Lagrangian can now model all abelian Quantum Hall states as an effective field theory. These Lagrangians are meaningful typically in the study of the Fractional Quantum Hall Effect.¹⁵

¹³ David Tong (2016b). *Lectures on the Quantum Hall Effect*. <http://www.damtp.cam.ac.uk/user/tong/qhe.html>. [Online; accessed 12-October-2021] pp 166-167

¹⁴ Zee Wen (July 1992). "Classification of Abelian quantum hall states and matrix formulation of topological fluids". In: *Phys Rev B* 46.4 (2)

¹⁵ T. H. Hansson et al. (2017). "Quantum Hall physics: Hierarchies and conformal field theory techniques". In: *Reviews of modern physics* 89.april-june

6.7. Chern-Simons theories in Particle physics

As yet another application it's important to mention that there is an equivalent of the Chern-Simons theory in the scope of particle physics: these are called Chiral anomalies. Again there is a symmetry at play and a current that is conserved such as the massless chiral currents of the Dirac term of the QED Lagrangian. In this case we do an adiabatic variation of the electromagnetic potential to obtain at the end of the variation the same initial set of states.

After the adiabatic variation, we actually introduced one particle less on the left chiral states and one particle more on the right chiral states (or also one quasi-hole state and one quasi-particle state in condensed matter terms). If we consider this chiral currents symmetry together with the Gauge symmetry we also again have a special boundary terms that is quantized and that compensate the Gauge symmetry on the border so that the total Gauge symmetry is still valid.¹⁶¹⁷

¹⁶ Hitoshi Murayama (n.d.[b]). *232B online lecture Perturbation Theory Anomalies on April 21, 2020*. <https://www.youtube.com/watch?v=IZwPr5nneEE&t=2100s>

¹⁷ Hitoshi Murayama (n.d.[a]). *232B online discussion section Chern-Simons term and Quantum Hall Effect on April 23, 2020*. <https://www.youtube.com/watch?v=ku0eLRDUqyE&t=1421s>

References

- “Characteristic Forms and Geometric Invariants” (Jan. 1974). In: *Annals of Mathematics, Second Series* 99.1, pp. 48–69.
- Haldane, F. D. M. (2015–2018). “Model for a Quantum Hall Effect without Landau Levels: Condensed-Matter Realization of the “Parity Anomaly.””. In: *Physical Review Letters*, 61(18).
- Leinaas, Jon Magne; Myrheim (Jan. 1977). “On the theory of identical particles”. In:
- Wen, Zee (July 1992). “Classification of Abelian quantum hall states and matrix formulation of topological fluids”. In: *Phys Rev B* 46.4 (2).
- Witten, Edward (1988). “Topological Quantum Field Theory”. In: *Commun. Math. Phys.* 117, pp. 353–386.

Readings and online resources

- Adeifeoba, Ademola (n.d.). *Introduction to Chern-Simons Theory*. <https://osf.io/preprints/africarxiv/am7pt/download>.
- Dunne (1998). *Lectures at the 1998 Les Houches Summer School*. <https://arxiv.org/pdf/hep-th/9902115.pdf>. [Online; accessed 1-April-2022].
- Freed, Dan (2022). *Lectures*. <https://ncatlab.org/nlab/show/Chern-Simons+theory>. [Online; accessed 1-April-2022].
- Moore, Gregory W. (2022). *TASI Lectures*. <https://www.physics.rutgers.edu/~gmoore/TASI-ChernSimons-StudentNotes.pdf>. [Online; accessed 1-April-2022].
- Murayama, Hitoshi (n.d.[a]). *232B online discussion section Chern-Simons term and Quantum Hall Effect on April 23, 2020*. <https://www.youtube.com/watch?v=ku0eLRDUqyE&t=1421s>.
- (n.d.[b]). *232B online lecture Perturbation Theory Anomalies on April 21, 2020*. <https://www.youtube.com/watch?v=IZwPr5nneEE&t=2100s>.
- Schwarz, Albert (2000). *Topological Quantum Field Theories*. <https://arxiv.org/abs/hep-th/0011260>.