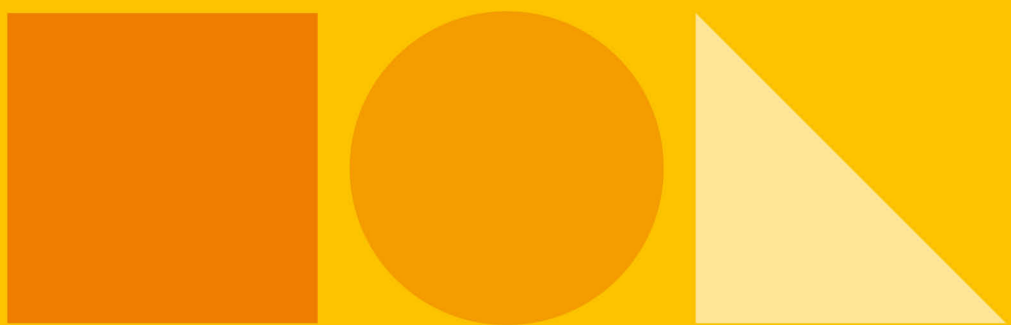
IOP Concise Physics | A Morgan & Claypool Publication

A Brief Introduction to Topology and Differential Geometry in Condensed Matter Physics

Antonio Sergio Teixeira Pires



IOP ebooks

A Brief Introduction to Topology and Differential Geometry in Condensed Matter Physics

A Brief Introduction to Topology and Differential Geometry in Condensed Matter Physics

Antonio Sergio Teixeira Pires
Universidade Federal de Minas Gerais, Belo Horizonte, Brazil

Morgan & Claypool Publishers

ISBN 978-1-64327-374-7 (ebook)
ISBN 978-1-64327-371-6 (print)
ISBN 978-1-64327-372-3 (mobi)

DOI 10.1088/2053-2571/aaec8f

Version: 20190301

IOP Concise Physics
ISSN 2053-2571 (online)
ISSN 2054-7307 (print)

A Morgan & Claypool publication as part of IOP Concise Physics
Published by Morgan & Claypool Publishers, 1210 Fifth Avenue, Suite 250, San Rafael, CA, 94901, USA

IOP Publishing, Temple Circus, Temple Way, Bristol BS1 6HG, UK

To Rosangela, Henrique and Guilherme

Contents

Appendices

Preface

In recent years there have been great advances in the applications of topology and differential geometry to problems in condensed matter physics. Concepts drawn from topology and geometry have become essential to the understanding of several phenomena in the area. Physicists have been creative in producing models for actual physical phenomena which realize mathematically exotic concepts, and new phases have been discovered in condensed matter in which topology plays a leading role. An important classification paradigm is the concept of topological order, where the state characterizing a system does not break any symmetry, but it defines a topological phase in the sense that certain fundamental properties change only when the system passes through a quantum phase transition.

The main purpose of this book is to provide a brief, self-contained introduction to some mathematical ideas and methods from differential geometry and topology, and to show a few applications in condensed matter. It conveys to physicists the bases for many mathematical concepts, avoiding the detailed formality of most textbooks. The reader can supplement the description given here by consulting standard mathematical references such as those listed in the references.

There are many good books written about the subject, but they present a lot of material and demand time to gain a full understanding of the text. Here, I present a summary of the main topics, which will provide readers with an introduction to the subject and will allow them to read the specialized literature.

Very little in this text is my original contribution since the goal of the book is pedagogy rather than originality. It was mainly collected from the literature. Some time ago, I used to teach differential geometry in a graduate course about classical mechanics and wrote a book (in Portuguese) on the topic. Now, I have adapted that material and included ideas that appeared in the last years, to write the present book.

Chapter 1 is an introduction to path integrals and it can be skipped if the reader is familiar with the subject. Chapters 2–4 are the core of the book, where the main ideas of topology and differential geometry are presented. In chapter 5, I discuss the Dirac equation and gauge theory, mainly applied to electrodynamics. In chapters 6–8, I show how the topics presented earlier can be applied to the quantum Hall effect and topological insulators. I will be mainly interested in the technical details because there are already excellent books and review articles dealing with the physical aspects. In chapter 9, I treat the application of topology to one- and two-dimensional antiferromagnets and the XY model. The framework presented here can also be used to study other systems, such as topological superconductors and quasi-metals. The appendices, although important for the application of differential geometry to some problems in condensed matter, are more specific.

Acknowledgements

It is a pleasure to acknowledge Joel Claypool, Melanie Carlson, Susanne Filler, Karen Donnison, Chris Benson, and others in the editorial and production teams of Morgan & Claypool and IOP Publishing for all their support and help. I also would like to acknowledge the Conselho Nacional de Pesquisa (CNPQ) for partial finantial support. I am grateful to all who gave permission to reproduce images included in the figures.

Author biography

Antonio Sergio Teixeira Pires

Antonio Sergio Teixeira Pires (born 18 November 1948) is a Professor of Physics in the Physics Department at the Universidade Federal de Minas Gerais, Belo Horizonte, Brazil. He received his PhD in Physics from University of California in Santa Barbara in 1976. He works in quantum field theory applied to condensed matter. He is a member of the Brazilian Academy of Science, was an Editor of the *Brazilian Journal of Physics* and a member of the Advisory Board of the *Journal of Physics*: *Condensed Matter*.

IOP Concise Physics

A Brief Introduction to Topology and Differential Geometry in Condensed Matter Physics

Antonio Sergio Teixeira Pires

Chapter 1

Path integral approach

1.1 Path integral

A convenient tool to treat topological quantum effects in quantum field theory is the path integral technique, and in this chapter, I am going to present the basic ideas (following mainly Ashok 1993). For more details I refer the reader to the references (Altland and Simons 2010, Fradkin 2013, Kogut 1979, Schwartz 2014, Tsvelik 1996, Wen 2004). Readers familiar with the subject can skip this chapter. I will start by establishing the path integral approach for the single particle in quantum mechanics in one dimension. The formalism can then be easily generalized to arbitrary spatial dimensions.

In path integral formalism the aim is to calculate the probability amplitude that a particle that starts at the position x_i at a time t_i ends up at a position x_f at a time t_f, with $t_f > t_i$. From quantum mechanics we know that this is given by the time-evolution operator $U(t_f, x_f; t_i, x_i)$ which in the Heisenberg picture is written as

$$U(t_f, x_f; t_i, x_i) = \left\langle x_f, t_f \,\middle|\, x_i, t_i \right\rangle, \tag{1.1}$$

where $|x, t\rangle$ is a coordinate basis for every time t. We divide the time interval between the initial and final time into N infinitesimals steps of length

$$\Delta t = \frac{t_f - t_i}{N}. \tag{1.2}$$

Any intermediate time can be written as $t_n = t_i + n\Delta t$, with $n = 1, 2, \ldots, (N-1)$. Considering time ordering from left to right, we can write equation (1.1) as (see figure 1.1)

$$\begin{aligned} U(t_f, x_f; t_i, x_i) = \lim_{\Delta t \to 0, N \to \infty} \int & dx_1 \ldots dx_{N-1} \langle x_f, t_f \mid x_{N-1}, t_{N-1} \rangle \\ & \langle x_{N-1}, t_{N-1} \mid x_{N-2}, t_{N-2} \rangle \ldots \langle x_1, t_1 \mid x_i, t_i \rangle. \end{aligned} \tag{1.3}$$

doi:10.1088/2053-2571/aaec8fch1

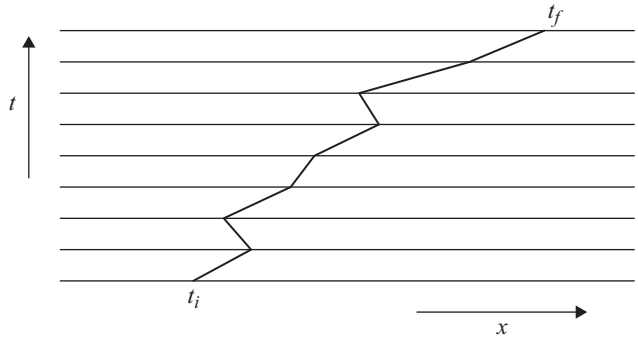


Figure 1.1. A discrete time axis and a path in quantum mechanics.

We know that

$$| x, t\rangle = e^{iHt} | x\rangle, \tag{1.4}$$

where I have set $\hbar = 1$, and we should remember to put it back if we are going to perform calculations. Therefore, we can write

$$\begin{aligned} \langle x_n, t_n | x_{n-1}, t_{n-1}\rangle &= \langle x_n | e^{-it_nH}e^{it_{n-1}H} | x_{n-1}\rangle = \langle x_n | e^{-i(t_n-t_{n-1})H} | x_{n-1}\rangle \\ &= \langle x_n | e^{-i\Delta tH} | x_{n-1}\rangle. \end{aligned} \tag{1.5}$$

Using the result

$$\langle x_2 | H | x_1\rangle = \int \frac{dp}{2\pi} e^{-ip(x_1-x_2)} H(x, p), \tag{1.6}$$

we find

$$\langle x_n, t_n | x_{n-1}, t_{n-1}\rangle = \int \frac{dp_n}{2\pi} e^{ip_n(x_n-x_{n-1})-i\Delta tH\left(\frac{x_n+x_{n-1}}{2}, p_n\right)}, \tag{1.7}$$

where to get a Weyl ordered Hamiltonian I wrote H using the mid-point prescription. Taking equation (1.7) into (1.3), and identifying $x_0 = x_i$, $x_n = x_f$ we can write

$$\begin{aligned} U(t_f, x_f; t_i, x_i) &= \lim_{\Delta t\to 0, N\to\infty} \int dx_1 \ldots dx_{N-1} \frac{dp_1}{2\pi} \ldots \frac{dp_N}{2\pi} \\ &\exp\left\{ i \sum_{n=1}^{N} \left[p_n(x_n - x_{n-1}) - \Delta tH\left(\frac{x_n + x_{n-1}}{2}, p_n\right)\right]\right\}. \end{aligned} \tag{1.8}$$

Let us now consider a Hamiltonian of the type

$$H(x, p) = \frac{p^2}{2m} + V(x). \tag{1.9}$$

This Hamiltonian covers a wide class of problems; however, some important applications, as will be shown in the next section, do not fit into this framework. Using equation (1.9) in (1.8) leads to

$$\begin{aligned} U(t_f, x_f; t_i, x_i) &= \lim_{\Delta t\to 0, N\to\infty} \int dx_1 \ldots dx_{N-1} \frac{dp_1}{2\pi} \ldots \frac{dp_N}{2\pi} \\ &\exp\left\{ i\Delta t \sum_{n=1}^{N} \left[p_n \left(\frac{x_n - x_{n-1}}{\Delta t} \right) - \frac{p_n^2}{2m} - V\left(\frac{x_n + x_{n-1}}{2} \right) \right] \right\}. \end{aligned} \tag{1.10}$$

Performing the momentum integrals using the result for Gaussian integration

$$\int_{-\infty}^{\infty} dp e^{-\frac{ap^2}{2} + bp} = \sqrt{\frac{2\pi}{a}} e^{\frac{b^2}{2a}}, \tag{1.11}$$

we obtain

$$\begin{aligned} U(t_f, x_f; t_i, x_i) &= \lim_{\Delta t\to 0, N\to\infty} \left(\frac{m}{2\pi i \Delta t} \right)^{N/2} \int dx_1 \ldots dx_{N-1} \\ &\exp\left\{ i\Delta t \sum_{n=1}^{N} \left[\frac{m}{2} \left(\frac{x_n - x_{n-1}}{\Delta t} \right)^2 - V\left(\frac{x_n + x_{n-1}}{2} \right) \right] \right\}. \end{aligned} \tag{1.12}$$

Taking $N \to \infty$, while keeping $(t_f - t_i) = N\Delta t$ fixed, we can substitute the sum by an integral

$$\Delta t \sum_{n=1}^{N} \to \int_{t_i}^{t_f} dt, \tag{1.13}$$

and write equation (1.12) as

$$U(t_f, x_f; t_i, x_i) = \int Dx \exp\left\{ i \int_{t_i}^{t_f} dt \left[\frac{1}{2} m \left(\frac{dx}{dt} \right)^2 - V(x) \right] \right\} = \int Dx e^{iS[x]}, \tag{1.14}$$

where

$$S[x] = \int_{t_i}^{t_f} dt L\left(x, \frac{dx}{dt} \right), \tag{1.15}$$

L is the classical Lagrangian, $S[x]$ is the action, and we have introduced the integration measure

$$\int D[x(t)] = \lim_{N\to\infty} \prod_{n=1}^{N-1} \int \frac{dx_n}{\xi}, \tag{1.16}$$

with $\xi = (i2\pi\Delta t/m)^{1/2}$. In some cases, more care must be applied in taking the continuum limit, but here I am considering only the essential details. Equation (1.14)

is the path integral for the probability amplitude of a particle in quantum mechanics. Feynman's idea of introducing the technique was that a particle going from A to B takes every possible trajectory, with each trajectory contributing with a complex factor e^{iS}.

Each path is weighted by its classical action, there are no quantum mechanical operators in the path integral. The quantum effects are present by the fact that the integration extends over all paths and is not just the subset of solutions of the classical equations of motion.

Following the same procedure, we can show that in quantum field theory with a Lagrangian density $L\ (\phi,\ \partial_\mu\phi)$ (where $\mu = t,\ x,\ y,\ z$) the amplitude transition from the state $\phi_i(r)$ to $\phi_f(r)$ is given by

$$\int D\phi(r,\ t)e^{iS[\phi(t,r)]}, \tag{1.17}$$

where the action is now given by

$$S[\phi] = \int d^4x\ L\ (\phi,\ \partial_\mu\phi) \tag{1.18}$$

In the path integral expression, the integration is performed over all possible paths in which ϕ, which at an initial time took the configuration $\phi_i(r)$, evolves at the final time t_f into the configuration $\phi_f(r)$. The field ϕ in condensed matter is in general an order parameter for a system, such as a superconductor or a ferromagnet.

1.2 Spin

One important application of the path integral approach in condensed matter is in magnetic systems. However, in the integrand of the path integral formalism one has an exponential of the classical action. But the spin is a fundamentally quantum object and the mechanics of a classical spin cannot be expressed within the standard formulation of Hamiltonian mechanics. We must resort to the coherent state formalism. I will illustrate this for the spin 1/2 case. For a spin 1/2 particle, we have only two states $|s_z\rangle$, $s_z = \pm 1$, with zero energy, and $s_z(t)$ is not a continuous function. To use the path integral approach, we use the coherent states $|\vec{n}\rangle$ where $\vec{n}$ is a unit vector and $|\vec{n}\rangle$ describes different states. $|\vec{n}\rangle$ is an eigenstate of the spin operator in the $\vec{n}$ direction: $\vec{n}.\ \vec{S}|\vec{n}\rangle = S|\vec{n}\rangle$.

We write

$$|\ \vec{n}\rangle = |\ z\rangle = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}, \tag{1.19}$$

with $|z_1|^2 + |z_2|^2 = 1$. The total phase of z is not determined, so that we can write

$$z = \begin{pmatrix} e^{-i\phi}\cos(\theta/2) \\ \sin(\theta/2) \end{pmatrix}, \tag{1.20}$$

where (θ, ϕ) are the polar coordinates of $\vec{n}$. The coherent states $|\vec{n}\rangle$ are complete, so that we can write

$$\int \frac{d^2\vec{n}}{2\pi} \mid \vec{n}\rangle\langle\vec{n} \mid = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{1.21}$$

Now we can calculate the amplitude $\langle\vec{n}_2|U(t, 0)|\vec{n}_1\rangle$ that a state $|\vec{n}_1\rangle$ at a time $t = 0$ evolves to the state $|\vec{n}_2\rangle$ at time t. Since $H = 0$, we have $U(t, 0) = 1$. Inserting

$$\int \frac{d^2\vec{n}}{2\pi} \mid \vec{n}\rangle\langle\vec{n} \mid, \tag{1.22}$$

into $\langle\vec{n}_2|\vec{n}_1\rangle$ we obtain the path integral

$$\langle\vec{n}_2|\vec{n}_1\rangle = \lim_{N\to\infty} \int \prod_{i=1}^{N} \frac{d^2\vec{n}(t_i)}{2\pi} \langle\vec{n}(t)|\vec{n}(t_N)\rangle \ldots \langle\vec{n}(t_2)|\vec{n}(t_1)\rangle\langle\vec{n}(t_1)|\vec{n}(0)\rangle. \tag{1.23}$$

Now

$$\langle\vec{n}(\delta t)|\vec{n}(0)\rangle = z^+(\delta t)z(0), \tag{1.24}$$

but, $z^+(\delta t)z(\delta t) = 1$, so we can write

$$\begin{aligned} \langle\vec{n}(\delta t)|\vec{n}(0)\rangle &= 1 - z^+(\delta t)[z(\delta t) - z(0)] = 1 - z^+(\delta t)\left[\frac{z(\delta t) - z(0)}{\delta t}\right]\delta t \\ &= 1 - z^+(\delta t)\frac{\partial z(\delta t)}{\partial t}\delta t \approx \exp\left(-z^+\frac{\partial z}{\partial t}\delta t\right), \end{aligned} \tag{1.25}$$

which leads to

$$\langle\vec{n}_2(t)|\vec{n}_1(t)\rangle = \int D^2\left(\frac{\vec{n}(t)}{2\pi}\right) e^{iS[\vec{n}(t)]}, \tag{1.26}$$

(where D is the measure) with the action

$$S[\vec{n}(t)] = i\int_0^t dt z^+\frac{\partial z}{\partial t}. \tag{1.27}$$

This is an interesting result, despite $H = 0$, we have obtained a non-zero action. The term e^{iS} is here purely a quantum effect and is called the Berry phase. Berry phases will be treated in more detail in chapter 6. We can also write equation (1.27) as

$$S(\theta, \phi) = \frac{1}{2}\int dt(1 - \cos\theta)\frac{\partial\phi}{\partial t}. \tag{1.28}$$

If we have a spin $\vec{S}$ in a constant magnetic field $\vec{B} = -B\vec{n}$, and the ground state energy is denoted by E_0, the action in a time interval T is given by $-E_0T$. Let us consider what happens when the orientation of $\vec{B}$ changes slowly in time, writing

$\vec{B} = -B\vec{n}(t)$. The ground state now evolves as $|\vec{n}(t)\rangle$, and the amplitude probability is given by

$$\langle \vec{n} \mid \exp\left[-i\int_0^T dt\vec{B}(t).\ \vec{S}\right] \mid \vec{n}\rangle = e^{iS}. \tag{1.29}$$

Inserting many equation (1.22) terms into the time interval $[0,\ T]$ we find

$$\langle \vec{n} \mid \exp\left[-i\int_0^T dt\vec{B}(t).\ \vec{S}\right] |\vec{n}\rangle = e^{-iE_0T} \exp\left[i\int_0^T dti\langle\vec{n}(t)|\ \frac{d}{dt}\ |\vec{n}(t)\rangle\right], \tag{1.30}$$

and the action can be written as

$$S = -E_0T + i\int_0^T dtz^+\frac{dz}{dt}. \tag{1.31}$$

We can see there is an extra term given by the Berry phase. As we will see later, this is a topological term, and I will denote it by S_{top} to distinguish it from the spin S.

For general spin S, equation (1.28) can be written as

$$S_{\text{top}}[\theta,\ \phi] = iS\int d\text{t}(1\ -\ \cos\theta)\frac{\partial\phi}{\partial\text{t}}. \tag{1.32}$$

If the motion of $\vec{n}(t)$ is such that its orientation coincides at the beginning and the end of the time interval, and considering that in the spherical coordinate system $(\hat{e}_r,\ \hat{e}_\theta,\ \hat{e}_\phi)$ we have

$$\frac{d\vec{n}}{dt} = \frac{d\theta}{dt}\hat{e}_\theta + \frac{d\phi}{dt}\sin\theta\hat{e}_\phi, \tag{1.33}$$

we can write equation (1.32) as

$$S_{\text{top}}[\theta,\ \phi] = iS\oint_\gamma d\text{t}\frac{d\vec{n}}{d\text{t}}.\ \vec{A} = iS\oint_\gamma d\vec{n}.\ \vec{A}, \tag{1.34}$$

where we have defined

$$\vec{A} = \frac{1\ -\ \cos\theta}{\sin\theta}\hat{e}_\phi. \tag{1.35}$$

Using Stokes's theorem, we have

$$S_{\text{top}}[\vec{n}] = iS\oint_\gamma d\vec{n}.\ \vec{A} = iS\oint_{A_\gamma} d\vec{\sigma}.\ (\vec{\nabla}\times\vec{A}), \tag{1.36}$$

but $\vec{\nabla}\times\vec{A} = \hat{e}_r$, which leads to

$$S_{\text{top}}[\vec{n}] = iS\oint_{A_\gamma} d\vec{\sigma}.\ \vec{e}_r = iSA_\gamma, \tag{1.37}$$

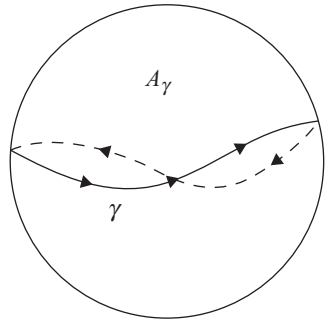


Figure 1.2. Region of integration in equation (1.37).

where A_γ is the region in the sphere S^2 which has the curve γ as its boundary and contains the north pole (see figure 1.2). The action S_{top} is thus a measure of the area bounded by the curve γ: $t \to \vec{n}(t)$.

Using $\vec{B} \equiv \vec{\nabla} \times \vec{A}$, equation (1.37) can be interpreted as the action for a particle moving in a radial magnetic field of a magnetic monopole of strength 4π located at the origin of the sphere.

If we had taken $\vec{A} = -\frac{1-\cos\theta}{\sin\theta}\hat{e}_\phi$, the newly defined vector potential would be non-singular in the southern hemisphere, and we would have got

$$S_{\text{top}}[\vec{n}] = -iSA'_\gamma \tag{1.38}$$

where A'_γ is the area of a surface bounded by γ but covering the south pole of the sphere. The minus sign is due to the outward orientation of the surface A'_γ. We can see that the difference between the northern and the southern parts is given by $4\pi iS$, having in mind that the intersection between the two surfaces is the sphere. We will come back to this subject in chapter 9, when we will discuss magnetic models.

1.3 Path integral and statistical mechanics

In statistical mechanics, the equilibrium properties of a system can be obtained from the partition function $Z = \text{tr}\exp(-\beta H)$, where 'tr' denotes a summation over all possible configurations of the system. For a single particle we have

$$Z = \text{tr}[e^{-\beta H}] = \int dx \langle x \mid e^{-\beta H} \mid x \rangle. \tag{1.39}$$

The partition function can be interpreted as a trace over the transition amplitude $\langle x|e^{-iHt}|x\rangle$ evaluated at an imaginary time $t = -i\beta$. The transformation $t = -i\tau$ is called a Wick rotation. Although mathematically this can be a highly nontrivial procedure, the formal prescription is simple. First, we make the substitution $t = -i\tau$, and then we define the imaginary time action S_E using the real time action S_M through the correspondence

$$e^{iS_M}\big|_{t=-i\tau} \equiv e^{-S_E}, \tag{1.40}$$

where the subscripts E and M stand for Euclidean and Minkowskian space–time. For a field $\phi(t, r)$ in quantum field theory we have

$$Z = \int_{\phi_i=\phi_f} D\phi(\tau, r)e^{-S[\phi(\tau,r)]}. \tag{1.41}$$

Here we are summing over a path in which the field $\phi(\tau, r)$ obeys periodic boundary conditions in the imaginary-time direction. In equation (1.41) we integrate over all trajectories with the sole requirement $\phi_i = \phi_f$, with no constraint on what the starting point is. All we must impose is that the field comes back to where it started after Euclidean time τ. We can think of τ as parameterizing a circle.

While all bosonic fields are periodic in the time direction, fermionic fields should be made anti-periodic: they pick up a minus sign as we go around the circle.

Following Tanaka and Takayoshi (2015) we define a topological term S_{top} as the portion of the action which arises in addition to the kinetic action coming directly from the Hamiltonian H. When using the imaginary time, the term S_{top} is purely imaginary and hence contributes a phase factor to the Boltzmann weight e^{-S} (this leads to nontrivial quantum interference effects). The total action is generally of the form: $S = S_{\text{kin}} + S_{\text{top}}$.

Another way to introduce topological terms is the following. The symmetric stress–energy tensor $T_{\mu\nu}$ can be defined as a variation of the action with respect to the metric tensor $g^{\mu\nu}$. More precisely, an infinitesimal variation of the action can be written as

$$\delta S = \int dx \sqrt{g}\, T_{\mu\nu}\delta g^{\mu\nu}, \tag{1.42}$$

where $\sqrt{g}\,dx$ is an invariant volume of space (see chapter 4). We define topological terms as the metric-independent terms in the action. It follows that topological terms do not contribute to the stress–energy tensor. We will study topological terms in more detail later in the text.

1.4 Fermion path integral

A path integral over fermions is basically the same as for bosons, but we must consider that fermions anti-commute. However, we cannot directly write a Lagrangian for fermions, since they have no classical analogue. To implement the path integral, we need the notion of anti-commuting classical variables that are called *Grassmann* variables (Ashok 1993, Altland and Simons 2010).

A Grassmann algebra is a set of objects θ_i with the following properties:

(a) They anti-commute $\theta_i\theta_j + \theta_j\theta_i = 0$. This implies $\theta_i^2 = 0$ for any i.
(b) $\theta_i + \theta_j = \theta_j + \theta_i$.
(c) They can be multiplied by complex numbers $a \in C$.
(d) There is an element 0 such that $\theta_i + 0 = \theta_i$.

For any θ, the most general element of the algebra is

$$g = a + b\theta, \quad \text{with } a, b \in c. \tag{1.43}$$

For two θ the most general element is

$$g = a + b\theta_1 + c\theta_2 + d\theta_1\theta_2, \tag{1.44}$$

and so on. In defining a derivative, the direction in which the derivative operates must be specified. For a right derivative we have

$$\frac{\partial}{\partial\theta_i}(\theta_j\theta_k) = \theta_j\left(\frac{\partial\theta_k}{\partial\theta_i}\right) - \left(\frac{\partial\theta_j}{\partial\theta_i}\right)\theta_k = \delta_{ik}\theta_j - \delta_{ij}\theta_k. \tag{1.45}$$

For a left derivative the result is

$$\frac{\partial}{\partial\theta_i}(\theta_j\theta_k) = \left(\frac{\partial\theta_j}{\partial\theta_i}\right)\theta_k - \theta_j\left(\frac{\partial\theta_k}{\partial\theta_i}\right) = \delta_{ij}\theta_k - \delta_{ik}\theta_j. \tag{1.46}$$

Here I will use left derivatives. Note that we have

$$\frac{\partial}{\partial\theta_i}\frac{\partial}{\partial\theta_j} + \frac{\partial}{\partial\theta_j}\frac{\partial}{\partial\theta_i} = 0. \tag{1.47}$$

For a fixed i we have

$$\left(\frac{\partial}{\partial\theta_i}\right)^2 = 0. \tag{1.48}$$

If D represents the operation of differentiation with respect to one Grassmann variable and I represents the operation of integration, we must have

$$ID = DI = 0. \tag{1.49}$$

So, using equation (1.48) we see that the integration can be identified with differentiation: $I = D$.

For a function we have

$$\int d\theta f(\theta) = \frac{\partial f(\theta)}{\partial\theta}, \tag{1.50}$$

which gives

$$\int d\theta = \theta, \quad \int \theta d\theta = 1. \tag{1.51}$$

If we write $\theta' = a\theta$ with $a \neq 0$, we find

$$\int d\theta f(\theta) = \frac{\partial f(\theta)}{\partial\theta} = a\frac{\partial f(\theta'/a)}{\partial\theta'} = a\int d\theta' f(\theta'/a). \tag{1.52}$$

For many Grassmann variables, if $\theta'_i = a_{ij}\theta_j$ (where we sum over repeated indices) with $\det a_{ij} \neq 0$, we get

$$\int \prod_{i=1}^{n} d\theta_i f(\theta_i) = (\det a_{ij}) \int \prod_{i=1}^{n} d\theta'_i f(a_{ij}^{-1}\theta'_j). \tag{1.53}$$

We define a delta function as

$$\delta(\theta) = \theta. \tag{1.54}$$

We can verify that it satisfies

$$\int d\theta \delta(\theta) = \int d\theta\theta = 1. \tag{1.55}$$

For a function $f(\theta) = a + b\theta$, we have

$$\int d\theta \delta(\theta) f(\theta) = \int d\theta\theta f(\theta) = \int d\theta\theta(a + b\theta) = \int d\theta\theta a = \frac{\partial(\theta a)}{\partial\theta} = a = f(0). \tag{1.56}$$

For path integral calculations, we need Gaussian integrals. For two θ_i we have

$$\int d\theta_1 d\theta_2 e^{-\theta_1 A_{12}\theta_2} = \int d\theta_1 d\theta_2 (1 - A_{12}\theta_1\theta_2) = A_{12}, \tag{1.57}$$

where we have expanded the exponential in a Taylor series. The variable θ does not need to be small; rather the exponential is defined by its Taylor expansion.

Let us now consider two sets of independent Grassmann variables $(\theta_1, \ldots, \theta_n)$ and $(\overline{\theta}_1, \ldots, \overline{\theta}_n)$. We want to calculate the integral

$$I = \int \prod_{i,j} d\overline{\theta}_i d\theta_j e^{-\overline{\theta}_i A_{ij}\theta_j}. \tag{1.58}$$

We have

$$I = \int \prod_{i,j} d\overline{\theta}_i d\theta_j \left[1 - \overline{\theta}_i A_{ij}\theta_j + \frac{1}{2}(\overline{\theta}_i A_{ij}\theta_j)(\overline{\theta}_k A_{kl}\theta_l) + \ldots\right]. \tag{1.59}$$

The only non-zero term in this expansion is the one with all $n\theta_i$ and all $n\overline{\theta}_i$. This will give

$$I = \frac{1}{n!} \sum_{\text{permutations}\{i_n\}} \pm A_{i_1 i_2} \ldots A_{i_{n-1} i_n}. \tag{1.60}$$

If A_{ij} is a matrix, equation (1.59) is a sum over all elements $\{i, j\}$ where we choose each row and column once, with the sign from the ordering. But this is just the determinant. So the result is:

$$I = \det(A) \tag{1.61}$$

It is easy now to show that

$$\int \prod_{i,j} d\bar{\theta}_i d\theta_j e^{-\bar{\theta}_i A_{ij} \theta_j + c_i^* \theta_i + \bar{\theta}_i c_i} = \det A \exp(c_i^* A_{ij}^{-1} c_j). \tag{1.62}$$

That is all we need for the fermion path integral.

References and further reading

Ashok D 1993 *Field Theory: A Path Integral Approach* (Singapore: World Scientific)

Altland A and Simons B 2010 *Condensed Matter Field Theory* (Cambridge: Cambridge University Press)

Fradkin E 2013 *Field Theories of Condensed Matter Physics* (Cambridge: Cambridge University Press)

Kogut J 1979 An introduction to lattice gauge theory and spin systems *Rev. Mod. Phys.* **51** 659

Schwartz M D 2014 *Quantum Field Theory and the Standard Model* (Cambridge: Cambridge University Press)

Tanaka A and Takayoshi S 2015 A short guide to topological terms in the effective theories of condensed matter *Sci. Technol. Adv. Mater.* **16** 014404

Tsvelik A M 1996 *Quantum Field Theory in Condensed Matter Physics* (Cambridge: Cambridge University Press)

Wen X G 2004 *Quantum Field Theory of Many-Body Systems* (Oxford: Oxford University Press)

IOP Concise Physics

A Brief Introduction to Topology and Differential Geometry in Condensed Matter Physics

Antonio Sergio Teixeira Pires

Chapter 2

Topology and vector spaces

2.1 Topological spaces

In topology two objects are considered equivalent if they can be continuously deformed into one another through bending, twisting, stretching, and shrinking while avoiding tearing apart or gluing parts together. In topology, we are interested in the properties of objects that remain unchanged by such continuous deformations. Topology differs from geometry in that geometrically equivalent objects often share numerically measured quantities, such as lengths or angles, while topologically equivalent objects resemble each other in a more qualitative sense. The size of an object does not matter in topology, since we do not measure distances.

For instance, a cup can be continuously transformed into a torus, and therefore they are topologically equivalent (figure 2.1), but we cannot deform a cup into a double torus as shown in figure 2.2.

In topology, the idea of closeness, or limits, is described in terms of relationships between sets rather than in terms of distance. Other types of spaces like metric spaces and manifolds are generalizations of topological spaces with some extra constraints or structures. A collection of objects is called a set. We denote by $\mathfrak{R}$ the set of all real numbers and by $\mathfrak{R}^n$ the set of all n-tuples.

If A is a subset of a set X, then every point in X is one of just two types in relation to $A \subset X$ either (i) x belongs to A; or (ii) it does not, in which case it belongs to the complement A^C of A defined as: $A^C = \{x \in X | x \notin A\}$.

Here I will provide a quick introduction to some key ideas in topology. For more information the reader is referred to the references (Flanders 1963, Hatcher 2002, Isham 1999, Kelly 1970).

Definiton 1. Let X be a set. A topology on X is a collection T of subsets of X satisfying the following conditions:

doi:10.1088/2053-2571/aaec8fch2

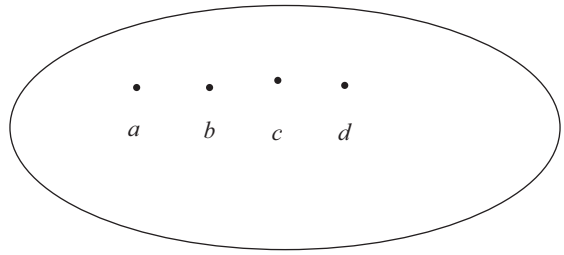
Figure 2.1. A cup can be continuously transformed in a torus.

Figure 2.2. A double torus.

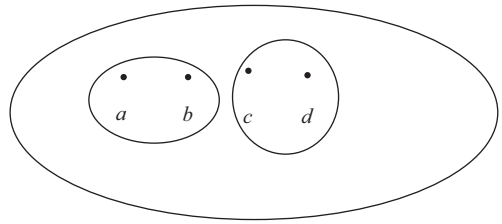
(1) T contains $\varnothing$ and X (where $\varnothing$ is the empty set).
(2) T is closed under arbitrary unions. That is, the union of any elements of subcollections of T is in T.
(3) T is closed under finite intersections. That is, if $U_1, U_2 \in T$ then $U_1 \cap U_2 \in T$.

***Example* 1.** Let X be a set of four elements $X = \{a, b, c, d\}$. There are several possible topologies in the set X:

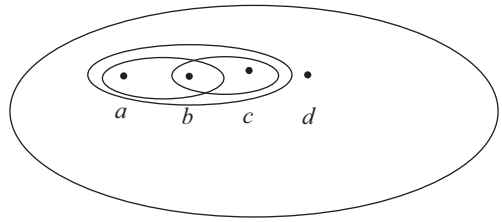
(a) $T = \{\varnothing, X\}$



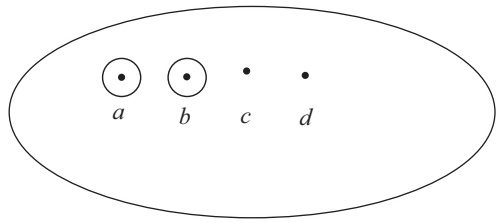
(b) $T = \{\varnothing, \{a, b\}, \{c, d\}, X\}$



(c) $T = \{\varnothing, \{a, b\}, \{b, c\}, \{b\}, \{a, b, c\}\}$



The configuration is not a topology since $\{a\} \cup \{b\} = \{a, b\} \notin T$.



***Example* 2.** Now I am going to show that the collection

$$T = \{\varnothing, X, \{a\}, \{c, d\}, \{a, c, d\}, \{b, c, d, e\}\}$$

defines a topology on the set $X = \{a, b, c, d, e\}$.

(i) The calculation of the unions of members of T gives:

$$\{a\} \bigcup \{c, d\} = \{a, c, d\} \in T, \{a\} \bigcup \{a, c, d\} \in T,$$

$$\{a\} \cup \{b, c, d, e\} = \{a, b, c, d, e\} = X \in T, \{c, d\} \cup \{a, c, d\} = \{a, c, d\} \in T,$$

$$\{c, d\} \bigcup \{b, c, d, e\} = \{b, c, d, e\} \in T,$$

$$\{a, c, d\} \bigcup \{b, c, d, e\} = \{a, b, c, d, e\} = X \in T.$$

(ii) The calculation of the intersections of the members of T gives

$$\{a\} \bigcap \{c, d\} = \varnothing \in T, \{a\} \bigcap \{a, c, d\} = \{a\} \in T,$$

$$\{a\} \bigcap \{b, c, d, e\} = \varnothing \in T, \{c, d\} \bigcap \{a, c, d\} = \{c, d\} \in T$$

$$\{c, d\} \bigcap \{b, c, d, e\} = \{c, d\} \in T, \{a, c, d\} \bigcap \{b, c, d, e\} = \{c, d\} \in T.$$

Thus all conditions for T to be a topology are satisfied.

Definition 2. A set X together with a topology T on it, is called a topological space $\{X, T\}$. The elements of T are called open subsets of X. A subset $F \subseteq X$ is called closed if its complement F^C is open. A subset N containing an element $x \in X$ is called a neighborhood of x if there is an open subset $U \subseteq N$ with $x \in U$. Thus, an open neighborhood of x is simply an open subset containing x.

***Example* 3.** In example 2, the open sets are

$$\varnothing, X, \{a\}, \{c, d\}, \{a, c, d\}, \{b, c, d, e\},$$

and hence the closed sets are

$$X, \varnothing, \{b, c, d, e\}, \{a, b, e\}, \{b, e\}, \{a\}.$$

The subset $\{a,b\}$ is neither open nor closed. The subset $\{a\}$ is both open and closed.

***Example* 4.** A set V in the plane is a neighborhood of p if we can draw a circle around p inside V (figure 2.3).

A rectangle is not a neighborhood of any point in its border (figure 2.4).

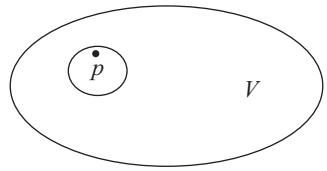


Figure 2.3.

Figure 2.4.

Definition 3. A point p is called a limit point of the set X if every open set containing p also contains some point (s) of X different from p (p does not need to lie in X). So, a set is closed if it contains all its limit points.

Definition 4. A topological space X is said to be *Hausdorff* if for any two distinct points $x, y \in X$ there exist two disjoint open subsets U, V ($U \cap V = \varnothing$) such that $x \in U$ and $y \in V$.

Let A be a subset of the topological space X. An open cover of A is a collection C of open sets whose union contains A. A subcover derived from the open cover C is a subcollection C$''^{*}$ of C whose union contains A. A topological space X is said to be *compact* if every open cover of X has a finite subcover. This says that however we write X as a union of open sets, there is always a finite subcollection of those sets whose union is X. Any space consisting of a finite number of points is compact.

The open interval (0, 1) is not compact. An open cover of (0, 1) is given by

$$\left\{\left(\frac{1}{n}, 1\right) \mid n = 2, \ldots, \infty\right\}.$$

However, no finite subcollection of these sets will cover (0, 1). On the other hand, the proof that [0, 1] is compact is quite elaborate. For our purposes we can say that compact sets are the sets which are closed and bounded. Compact means intuitively that the region R does not 'go to infinity' and does not have 'holes cut out of it' nor have 'bits of it's boundary removed'. The surface of a sphere, the torus, and the set of points lying within or on the unit circle are compact. The infinite Euclidian plane, the open unit disc and the closed disc with the center removed are not compact.

A mapping ϕ: $X \to Y$ between two topological spaces is called continuous if, for any open set $U \subset Y$, the set $\phi^{-1}(U) \subset X$ is open in X.

A map is called a *homeomorphism* (an isomorphism in the context of general topology) if ϕ is a bijection and ϕ and ϕ^{-1} are continuous.

(Note: a bijection is a mapping between the elements of two sets, where each element of one set is paired with exactly one element of the other set, and each element of the other set is paired with exactly one element of the first set.)

One of the main problems of topology is to understand when two topological spaces X and Y are similar or dissimilar and to classify the different families of spaces that are not equivalent under a continuous deformation.

2.2 Group theory

A group G is a set of elements a, b, c, … such that a form of group 'multiplication' (i.e. a rule for combining any two elements) may be defined which associates with a pair of elements of the set a third element in the set (Lang 1968, Tinkham 1964, Tung 1985). This multiplication must satisfy the following requirements:

(a) The product of any two elements of the set is in the set (i.e. the set is closed under group multiplication).
(b) The multiplication is associative; for example, $a(bc) = (ab)c$.
(c) There is a unit element e such that $ea = ae = a$.
(d) There is an inverse a^{-1} to each element a such that $aa^{-1} = a^{-1}a = e$.

If the multiplication is commutative, so that $ab = ba$, for all a and b, the group is said to be Abelian. The number of elements in the group is said to be the order of the group.

***Example* 1.** The set of integers ….−3, −2, −1, 0, 1, 2, 3… together with the addition is a group called Z.

***Example* 2.** The cyclic group of order 2, with two elements e and x such that $ex = xe = x$ and $e^2 = x^2 = e$. An example is the multiplicative group comprising 1 and −1.

***Example* 3.** The *circle group* T, is the multiplicative group of all complex numbers with absolute value 1. (It is also the group U(1) of 1×1 complex-valued unitary matrices.) It can be parameterized by an angle θ: $z = e^{i\theta} = \cos\theta + i \sin\theta$.

***Example* 4.** The six matrices below, if ordinary matrix multiplication is used as the group-multiplication operation, form a group

$$E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad B = \frac{1}{2}\begin{pmatrix} -1 & \sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix},$$

$$C = \frac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ -\sqrt{3} & 1 \end{pmatrix}, \quad D = \frac{1}{2}\begin{pmatrix} -1 & \sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix}, \quad F = \frac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}.$$

A subset of a group G, which is itself a group is called a subgroup of G.

Let $S = e, s_2, s_3, \ldots, s_g$ be a subgroup of order g of a larger group G of order h. We call the set of g elements ex, s_2x, s_3x, $\ldots$, s_gx a right coset Sx if x is not in S. Similarly, we define the set xS as being a left coset. These cosets cannot be subgroups, since they cannot include the identity element. In fact, a coset Sx contains no elements in common with the subgroup S.

***Example* 5.** Consider the subgroup of integers divided by 3. This forms a subgroup of the additive group of integers with elements (…−9, −6, −3, 0, 3, 6, 9…). By adding 1 to each member of the subgroup we get the coset (…−8, −5, −2, 1, 4, 7, 10…).

Let G be a group and H a subgroup having the property that $xH = Hx$ for all $x \in G$. If aH and bH are cosets of H, then the product $(aH)(bH)$ is also a coset, and the collections of cosets is a group, the product being defined as above. The group of the above cosets is called the factor (or quotient) group of G by H, and denoted G/H.

A homomorphism from a group G to another group G' is a mapping which preserves $g'_1g'_2 = g'_3$, if $g_1g_2 = g_3$. If there exists a one-to-one correspondence between the elements of G and G' in the above mapping we have an isomorphism.

A *topological group* G is a topological space which is also a group such that the group's operations and group inverse function are continuous functions with respect to the topology. A topological group is called locally compact if the underlying topological space is locally compact and Hausdorff.

If G is a locally compact Abelian group, a character of G is a continuous group homomorphism from G with value in the circle group T. The set of all characters on G can be made into a locally compact abelian group, called the dual group of G and denoted $\hat{G}$.

2.3 Cocycle

Let G be a group and $g_i \in G$. Suppose that g transforms a variable q into q_g. We associate with g an operator $U(g)$ defined to act on a function $f(q)$ according to

$$U(g)f(q) = f(q_g), \tag{2.1}$$

and to satisfy the composition law

$$U(g_1)U(g_2) = U(g_{12}), \tag{2.2}$$

if $g_1g_2 = g_{12}$. Now we can generalize equations (2.1) and (2.2) to

$$U(g)f(q) = e^{-i2\pi\theta(q,g)}f(q_g), \tag{2.3}$$

where θ is a phase factor called a 1-*cocycle* satisfying (consistent with equation (2.2))

$$\theta(q, g_1) + \theta(q, g_2) - \theta(q, g_{12}) = 0 \text{ (mod integer)}. \tag{2.4}$$

As we will see in chapter 7, for electrons in a two-dimensional lattice subject to a perpendicular magnetic field, the magnetic translations in each lattice direction x

and y with $x, y \in Z^2$ do not commute. Instead, $T_x T_y = \sigma(x, y) T_{x+y}$, where the cocycle $\sigma(x, y)$ is proportional to the magnetic field strength.

2.4 Vector spaces

Near the end of the 19th Century, Gibbs developed the vector calculus to treat objects such as force and velocity. Afterwards, the concept was generalized by mathematicians as we show below, and the general theory found applications in physics as well.

A *vector space* V is a set of objects that can be summed together (and we must define how they are summed) and multiplied by scalars such that the sum of two elements of V is an element of V, the product of an element of V by a scalar is an element of V, and the following properties are satisfied:

1. If u, v, w are elements of V, we have: $u + v = v + u$, $(u + v) + w = u + (v + w)$.
2. There is an element of V, called 0, such that $0 + u = u + 0 = u$.
3. Given an element u of V, the element $(-1)\, u$ is such that $u + (-1)\, u = 0$.
4. For all elements $u \in V$ we have: $1.u = u$.
5. If a, b and c are scalars, we have: $c\,(u + v) = cu + cv$, $(a + b)\, v = av + bv$, $(ab)\, v = a(bv)$.

A *basis* of V is a sequence of elements $(v_1, v_2, \ldots, v_n)$ which generate V and are linearly independent. If an element of V is written as a linear combination

$$v = x_1 v_1 + x_2 v_2 + \ldots + x_n v_n, \tag{2.5}$$

of the elements of the basis, the elements of V can be represented by the n numbers $(x_1, \ldots, x_n)$, called the coordinates of v with respect to that basis. The number of elements of the basis is the dimension of V. We say that the n-tuple $X = (x_1, x_2, \ldots, x_n)$ is the representative of v in relation to the above basis.

We can see that the standard vectors in physics obey the above rules and of course $\Re^n$ has a natural vector space structure. As another example we consider all matrices 2×2, with the standard rules for matrices addition. One basis is given by the matrices:

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \tag{2.6}$$

If V is a vector space, and U and W are subspaces of V, we define the sum of U and W to be the subset of V consisting of all sums $u + w$ with $u \in U$ and $w \in W$. We denote this sum by $U + W$. It is a subspace of V. If $U + W = V$ and if $U \cap W = \{0\}$ then V is the direct sum of U and V and we write $V = U \oplus W$.

An N-graded vector space is a vector space V which decomposes into a direct sum of the form $V = \oplus_{n \in N} V_n$, where V_n is a vector space, and N the set of non-negative integers.

For a given n, the elements of V_n are called homogeneous elements of degree n. A graded linear map between two graded vector spaces $f: V \to W$ is a map that preserves the grading of homogeneous elements.

Let V be a vector space and W a subspace of V. For each $v \in V$, we denote by $v + W$ the following subset of V: $v + W = \{v + w | w \in W\}$. So $v + W$ is the set of all vectors in V we get by adding v to elements of W. Note that v itself is in $v + W$ since $v + 0 = v$ and $0 \in W$. A coset of W in V is a subset of the form $v + w$. The set V / W is the set defined by $V / W = \{v + W | v \in V\}$. That is, V / W is the collection of cosets of W in V.

The n-dimensional projective space, denoted $\Re P^n$, is the space of one-dimensional subspaces (lines) in $\Re^{n+1}$. The Grassmannian, denoted $Gr(k, n)$, is the space of all k-dimensional subspaces of an n-dimensional vector space. Note that this is a generalization of projective space, since $Gr(1, n + 1) \cong \Re P^n$.

Here I will be leading mainly with vector spaces over the real numbers, but the above definitions apply also to vector spaces defined over the complex numbers.

2.5 Linear maps

A mapping F (or map) from a set A to a set B is a rule that each element of A associates with an element of B. We write $F: A \to B$. If x is an element of A, we write $F(x)$ or Fx for the element of B associated with x by F, $F(x)$ is the image of x over F. The set of all elements $F(x)$ when x ranges over all elements of A is called the image of F.

Let V and W be two vector spaces. A linear mapping $F: V \to W$ is a mapping that satisfies the following properties:

1. For all elements u and v in V we have: $F(u + v) = F(u) + F(v)$.
2. For $v \in V$ and c a scalar, we have: $F(cv) = cF(v)$.

Let F be a mapping of a set A into a set B. We say that F is injective if for $x \in A$, $y \in B$ and $x \neq y$ we have $F(x) \neq F(y)$. We say that F is surjective if for each $y \in B$ there is at least one element $x \in A$ such that $f(x) = y$. If F is injective and surjective we say that F is bijective. If f is injective and bijective it has an inverse, and in such a case, it is called an isomorphism.

Let $E_1, \ldots, E_p$ be sets. We denote by $E_1 \times E_2 \times \ldots \times E_p$ the set of all p-tuple $(x_1, x_2, \ldots, x_p)$ with x_j in E_j, for $j = 1, \ldots, p$. This set is called the Cartesian product of E_j. If F is a mapping of this set into a set A, we write $F(x_1, x_2, \ldots, x_p)$ for the image of the element $(x_1, x_2, \ldots, x_p)$. If $V_1, .., V_p, W$, are vector spaces, a mapping $F: V_1 \times \ldots \times V_p \to W$ is called linear at index j if the mapping $x_j \to F(x_1, \ldots, x_j, \ldots, x_p)$ is a linear mapping of V_j into W, for any choice of the remaining $p - 1$ variables $x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_p$.

2.6 Dual space

Let U be a vector space. We denote by U^* the set of all linear mappings of U in the set of scalars K. We know that U^* is a vector space, since we can add linear mappings and multiply them by scalars. The space U^* is called the dual space of U. Elements of this space are called functional, covectors, linear forms or 1-forms.

Let $B = \{u_n\}$ be a basis for U. Then an arbitrary vector x in U can be written uniquely as:

$$x = \xi^1 u_1 + \ldots + \xi^n u_n, \tag{2.7}$$

in terms of the basis B. The numbers ξ^i are the components of x relative to B. We will denote by $u^i(x)$ the ith component of x relative to B, that is $u^i(x) = \xi^i$. We see that:

$$u^i(x + y) = u^i(x) + u^i(y),\ u^i(ax) = au^i(x), \tag{2.8}$$

where a is a scalar. So u^i is a linear mapping of U into K and therefore an element of U^*. We have:

$$u^i(u_j) = \delta^i_j, \tag{2.9}$$

where $\delta^i_j = 0$ if $i \neq j$ and $\delta^i_j = 1$ if $i = j$. Since a linear mapping is determined completely by its values in the vectors of the basis it follows that these equations completely determine u^i. Denoting by u^i a linear form in U determined by the condition (2.9) we see that u^i carries an arbitrary vector x in U in its ith component $u^i(x)$ relative to the basis B. It can be shown that $B^* = \{u^n\}$ is a basis for the dual space, called the dual basis of B, and therefore U and U^* have the same dimension.

2.7 Scalar product

If f is a linear form in U (this is, $f \in U^*$) and if u is a vector in U we designate the value $f(u)$ by the symbol $\langle f|u\rangle$. That is, $f(u) \equiv \langle f|u\rangle$. This symbol, linear on both sides, is called the *scalar product* between u and f.

Some authors introduce the scalar product as follows: A scalar product in a vector space U is a rule that to the pair of elements v, w belonging to U associates a real number indicated by (v, w) satisfying the conditions:

1. $(v,\ w) = (w\ ,\ v)$,
2. $(u,\ v\ +\ w) = (u,\ v) + (u,\ w)$,
3. $(au,\ v) = a\ (u,\ v)$, $(u,\ av) = a(u,\ v)$,

where $u, v, w \in U$ and a is a real number. It can be shown that the two definitions are equivalent. (In the case of vector spaces over the complex numbers, we define a Hermitian product, as shown in appendix B.)

2.8 Metric space

A set R is called a *metric space* if a positive number $d(x, y)$ exists associated with any pair of elements in R such that

1. $d(x, y) = 0$ only if $x = y$,
2. $d(x, y) = d(y, x)$,
3. $d(x, y) + d(y, z) \geqslant d(x, z)$.

The number $d(x, y)$ is called the distance between x and y. If V is a vector space with a scalar product and if $v, u \in V$ we can define a distance by $d(u, v) = \sqrt{(u - v, u - v)}$.

Another example is the 'trivial distance' in a discrete set defined by

$$d_{ij} = 1 - \delta_i^{\ j} = \begin{cases} 0 & i = j \\ 1 & i \neq j \end{cases}.$$

2.9 Tensors

Let $U_1, \ldots, U_p, W$, be vector spaces and f: $U_1 \times \ldots \times U_p \rightarrow W$ a multilinear mapping. We call f a tensor in $U_1, \ldots, U_p$ with values in W. That is, a tensor is a multilinear function of vectors. It is easy to verify, using the definition of a vector space presented previously, that the tensors f: $U_1 \times \ldots \times U_p \rightarrow W$, form a vector space.

Let $U_1, \ldots, U_p, V_1, \ldots, V_q$ be vector spaces and f: $U_1 \times \ldots \times U_p \rightarrow \Re$, h: $V_1 \times \ldots \times V_q \rightarrow \Re$, where f and h are linear in each variable. We call the *tensor product* of f and h the function $f \otimes h$: $U_1 \times \ldots \times U_p \times V_1 \times \ldots \times V_q \rightarrow \Re$ defined by:

$$(f \otimes h)(x_1, \ldots, x_p, y_1, \ldots, y_q) = f(x_1, \ldots, x_p)h(y_1, \ldots, y_q), \tag{2.10}$$

for x_i in U_i and y_j in V_j. Since $f \otimes h$ is linear in each variable, this function is a tensor in $U_1, \ldots, U_p, V_1, \ldots, V_q$ with values in $\Re$.

Let V and W be vector spaces, with $v \in V$, $w \in W$, $\phi \in V^*$, $\psi \in W^*$. Let T_2 be the space of bilinear transformations:

$$T_2\text{: } V \times R \rightarrow \Re. \tag{2.11}$$

We define the tensor product $\phi \otimes \psi \in T_2$ by:

$$\phi \otimes \psi(v, w) = \phi(v)\psi(w) = \langle \phi | v \rangle \langle \psi | w \rangle. \tag{2.12}$$

Similarly, we can consider v as a linear transformation $V^* \rightarrow \Re$ of the dual space, and the same for w. We can then define the tensor product $v \otimes w$. It is the bilinear transformation $V^* \times W^* \rightarrow \Re$ that transforms a pair of forms $\phi \in V^*$, $\psi \in W^*$ into the scalar $\langle \phi | v \rangle \langle \psi | w \rangle$, this is:

$$v \otimes w(\phi, \psi) = \langle \phi | v \rangle \langle \psi | w \rangle = \phi(v)\psi(w). \tag{2.13}$$

We can extend the definition to multiple vectors and forms.

The transformation $V \times W \rightarrow V \otimes W$ given by $(v, w) \rightarrow v \otimes w$ is bilinear. We can show that if $(v_1, \ldots, v_r)$ is a basis for V and $(w_1, \ldots, w_s)$ a basis for W, then the products $v_i \otimes w_j$ is a basis for $V \otimes W$. Then if f: $V \otimes W \rightarrow \Re$, we can write

$$f = f^{ij} v_i \otimes w_j, \tag{2.14}$$

where

$$f^{ij} = f(v^i, w^j), \tag{2.15}$$

where $\{v^i\}$ and $\{w^j\}$ are the dual bases of the bases in V and W.

Let U and V be vector spaces of finite dimensions. We define a tensor of the type $\binom{p}{q}$ in U with values in V, by the mapping:

$$f\colon U^*\times\ldots\times U^*\times U\times\ldots\times U\to V \tag{2.16}$$

with U^* taken p-times and U-q times. This transformation is linear in each of the $p+q$ terms. We indicate the space of all tensors of this type with V taken as $\Re$ by U_q^p. If f is a tensor U_q^p we have

$$f=f^{i_1\ldots i_p}_{j_1\ldots j_q}u_{i_1}\otimes\ldots\otimes u_{i_p}\otimes u^{j_1}\otimes\ldots\otimes u^{j_q}, \tag{2.17}$$

with

$$f^{i_1\ldots i_p}_{j_1\ldots j_q}=f(u^{i_1},\ldots,u^{i_p},u_{j_1}\ldots,u_{j_q}). \tag{2.18}$$

If $f\in U_q^p$ and the value of $f(x_1,\ldots,x_p,y,\ldots,y^q)$, for x in U and y in U^*, does not change, when the indices in x_n or in y^m are exchanged, we say that f is a symmetric tensor. If one change of sign occurs, we say that the tensor is antisymmetric in this argument.

2.10 *p*-vectors and *p*-forms

A transposition τ is a permutation that changes the position of only two numbers in a set. Every permutation σ can be written as the product of transpositions. We say that σ is even if it can be expressed as the product of an even number of transpositions or if it is the identity. We say that σ is odd if it can be expressed as an odd number of transpositions. We can show that for any permutation σ, it is possible to attribute a sign + 1 or −1, denoted by sgn(σ), such that: (i) sgn($\sigma\sigma'$) = sgn(σ)sgn(σ'), (ii) if τ is a transposition, then sgn(τ) = −1. We can see that sgn(σ) = +1 if the permutation is even and sgn(σ) = −1 if the permutation is odd.

Definition 5. A p-vector of the vector space U ($p=0,1,2,\ldots$) is an anti-symmetric element of U_0^p. The space of all such p-vectors will be denoted by Λ^pU, with $\Lambda^0U=\Re$ by definition. A p-form in U is an anti-symmetric element of U_p^0. The space of all such p-forms will be denoted by Λ^pU^*, with $\Lambda^0U^*=\Re$ by definition. Thus a p-form w (also called an external form of degree p) is a function of p-vectors, which is p-linear and anti-symmetric, that is

$$w(v_{i_1},\ldots,v_{i_p})=(-1)^n(v_1,\ldots,v_p), \tag{2.19}$$

where $n=0$ if the permutation of $i_1,\ldots,i_p$ is even, and $n=1$ if it is odd. Thus p-forms in U are the same as p-vectors in U^* and vice versa.

Let $f\in U_0^p$. For any permutation σ of (1,..,p) we define a tensor σf in U by

$$\sigma f(\xi^1, \ldots, \xi^p) = f(\xi^{\sigma(1)}, \ldots, \xi^{\sigma(p)}), \tag{2.20}$$

where $\xi^1, \ldots, \xi^p \in U^*$. The term σf is obtained from f by permuting $\xi^1, \ldots, \xi^p$ according to σ.

We can write from a tensor f, an anti-symmetric tensor, using the operator of anti-symmetrization given by

$$Af = \frac{1}{p!}\sum_{\sigma} \mathrm{sgn}(\sigma) f, \tag{2.21}$$

where the sum is over all the $p!$ permutations σ of $(1, \ldots, p)$. For instance, given a tensor f with components f_{ijk}, we can write an anti-symmetric tensor with components $f_{[ijk]}$ as

$$f_{[ijk]} = \frac{1}{3!}(f_{ijk} + f_{jki} + f_{kij} - f_{jik} - j_{ikj} - f_{kji}). \tag{2.22}$$

2.11 Edge product

By means of the operator A we can define a new product, called the *edge product* (or exterior product), for anti-symmetric tensors. Let $f \in \Lambda^p U$ and $g \in \Lambda^q U$. We define $f \wedge g \in \Lambda^{p+q} U$ by

$$f \wedge g = \frac{(p+q)!}{p!q!} A(f \otimes g). \tag{2.23}$$

In the same way for $\phi \in \Lambda^p U^*$ and $\psi \in \Lambda^q U^* \psi \in \Lambda^q U^*$ we define $\phi \wedge \psi \in \Lambda^{p+q} U^*$ by

$$\phi \wedge \psi = \frac{(p+q)!}{p!q!} A(\phi \otimes \psi). \tag{2.24}$$

For instance, if u and v lie in U, we have

$$u \wedge v = u \otimes v - v \otimes u. \tag{2.25}$$

It can be shown that if $\{e_n\}$ is a basis for U, then $e_{i_1} \wedge \ldots \wedge e_{i_p}$, with $(i_1 < \ldots < i_p)$, is a basis for $\Lambda^p U$ $(p = 1, \ldots, n)$.

***Example* 1.** Let $\{e_n\}$ with $n = 1, 2, 3$ be a basis for a vector space with dimension 3. If we have the vectors $v = x_1e_1 + x_2e_2 + x_3e_3$, $u = y_1e_1 + y_2e_2 + y_3e_3$, then

$$v \otimes u = x_1y_1e_1 \otimes e_2 + x_1y_2e_1 \otimes e_2 + \ldots + x_3y_3e_3 \otimes e_3. \tag{2.26}$$

The space $W = U \otimes U$ has dimension 9. The space $\Lambda^2 U$ of anti-symmetric vectors has dimension 3. We have

$$v \wedge u = v \otimes u - u \otimes v = (x_1y_2 - x_2y_1)e_1 \wedge e_2 + (x_1y_3 - x_3y_1)e_1 \wedge e_3 + (x_2y_3 - x_3y_2)e_2 \wedge e_3$$

As we can see $e_1 \wedge e_2$, $e_1 \wedge e_3$, $e_2 \wedge e_3$ is a basis for $\Lambda^2 U$.

Let $\{e_1, \ldots, e_n\}$ be a basis of a vector space V of dimension n, ordered according to the sequence above. Let α be a mapping taking this basis in another ordered basis $\{v_1, \ldots, v_n\}$. If we change $\{e_n\}$ continuously into $\{v_n\}$ we can write $w_i(t) = \alpha_i^j(t)e_j$, with $0 \leqslant t \leqslant 1$, $\alpha_i^j(0) =$ identity, $w_i(0) = e_i$ and $w_i(1) = \alpha_i^j(1)e_j = v_i$.

If $\{w_i(t)\}$ remains a basis for all t in the above interval, det $\alpha(t)$ will not change its sign during the process because if there is a sign change there will be a t' where det $\alpha(t') = 0$ and for this value of t' the set of vectors $(w_i(t'), \ldots, w_n(t'))$ will be linearly dependent and hence will no longer be a basis. As det $\alpha(0) > 0$ we say that two ordered bases of a vector space V define the same orientation (or are similarly oriented) if the determinant of the matrix that carries one basis on the other is positive. A vector-oriented space is a vector space along with a choice of orientation. Note that each vector space allows exactly two orientations.

It can be shown that two bases $\{v_n\}$ and $\{u_n\}$ define the same orientation of a vector space V if and only if $u^1 \wedge \ldots \wedge u^n = \lambda v^1 \wedge \ldots \wedge v^n$ where λ is a positive number.

The set of forms of all degrees in U together with the edge product is called the Grassmann algebra of the vector space U.

2.12 Pfaffian

If A is a $n \times n$ anti-symmetric matrix, the determinant of A vanishes when n is odd, but if n is even the determinant can be written as the square of an object called Pfaffian

$$Pf(A)^2 = \det(A). \tag{2.27}$$

The Pfaffian is a polynomial of degree $n/2$ in the elements of the matrix, with integer coefficients.

Examples:

$$A = \begin{bmatrix} 0 & a \\ -a & 0 \end{bmatrix}, \rightarrow Pf(A) = a.$$

$$A = \begin{bmatrix} 0 & a & b & c \\ -a & 0 & d & e \\ -b & -d & 0 & f \\ -c & -e & -f & 0 \end{bmatrix}, \rightarrow Pf(A) = af - be + dc.$$

We can associate with any $2n \times 2n$ anti-symmetric matrix $A = \{a_{ij}\}$ the 2-vector

$$w = \sum_{i<j} a_{ij} e_i \wedge e_j, \tag{2.28}$$

where $\{e_i\}$ is the standard basis of $\Re^{2n}$. We have the following result

$$\frac{1}{n!} w^n = Pf(A) e_1 \wedge e_2 \wedge \ldots \wedge e_{2n}, \tag{2.29}$$

where w^n denotes the wedge products of n copies of w. The Pfaffian has the following proprieties:

(a) For a $2n \times 2n$ anti-symmetric matrix A

$$Pf(A^T) = (-1)^n PfA,\ Pf(\lambda A) = \lambda^n Pf(A). \tag{2.30}$$

(b) For an $2n \times 2n$ arbitrary matrix B

$$Pf(BAB^T) = \det(B) Pf(A). \tag{2.31}$$

Pfaffians appear in the expression of certain multiparticle wave equations in fractional quantum Hall effect (Moore and Read 1991).

References and further reading

Flanders H 1963 *Differential Forms with Applications for Physical Sciences* (New York: Academic)

Hatcher A 2002 *Algebraic Topology* (Cambridge: Cambridge University Press)

Isham C J 1999 *Modern Differential Geometry for Physicists* (Singapore: World Scientific)

Kelly J L 1970 *General Topology* (London: Van Nostrand)

Lang S 1968 *Linear Algebra* (Reading: Addison-Wesley)

Moore G and Read N 1991 Nonabelions in the fractional quantum Hall effect *Nucl. Phys.* B **360** 362

Tinkham M 1964 *Group Theory and Quantum Mechanics* (New York: McGraw-Hill)

Tung W K 1985 *Group Theory in Physics* (Singapore: World Scientific)

IOP Concise Physics

A Brief Introduction to Topology and Differential Geometry in Condensed Matter Physics

Antonio Sergio Teixeira Pires

Chapter 3

Manifolds and fiber bundle

3.1 Manifolds

A real n-dimensional *manifold* X is a Hausdorff topological space which looks like $\Re^n$ around each point. More precisely, a manifold is defined by introducing a set of neighborhoods U_i covering X, where each U_i is a subspace of $\Re^n$. Thus, a manifold is constructed by pasting together many pieces of $\Re^n$. The topology of a manifold is in general different from that of a vector space, and hence—in particular—it cannot be covered by a single coordinate system (Carrol 2004, Chouquet-Bruhat *et al* 1982, Curtis and Miller 1985, Eguchi *et al* 1980, Isham 1999, Lee 2003, Warner 1983).

A map (U, φ) of a manifold X is an open set U in X, called the map domain, along with a homeomorphism φ: $U \to V$ of U in the open set V in $\Re^n$. The coordinates $(x^1, \ldots, x^n)$ of the image $\varphi(x)$ of the point $x \in X$ are called coordinates of x on the map (U, φ) (local coordinates of x). A map is also called a local coordinate system and coordinates maps are called charts of X. We say that the manifold X has dimension n (see figure 3.1). Given two maps of X, φ_i: $U_i \to V_i$, and φ_j: $U_j \to V_j$, let us consider the sets (see figure 3.2):

$$V_{ij} = \varphi_i(U_i \bigcap U_j), \qquad V_{ji} = \varphi_j(U_j \bigcap U_i), \tag{3.1}$$

and the mappings φ_{ij}: $V_{ij} \to V_{ji}$

$$\phi_{ij}(y) = \phi_j(\phi_i^{-1}(y)), \qquad y \in V_{ij}. \tag{3.2}$$

The maps φ_i and φ_j are called C^r compatible (a function is said to be C^r if all its partial derivatives up to and including order r exist) if $V_i \cap V_j$ is the empty set or if $V_i \cap V_j$ is not empty but the mappings φ_{ij} and φ_{ji} are diffeomorphisms (that is, invertible and differentiable) of class C^r $(1 \leqslant r \leqslant \infty)$.

We call the *atlas* of X any set of maps of X (compatible two by two) and whose domains of definition constitute a covering of X (i.e. $\bigcup_i U_i$ covers X). We say that two

doi:10.1088/2053-2571/aaec8fch3

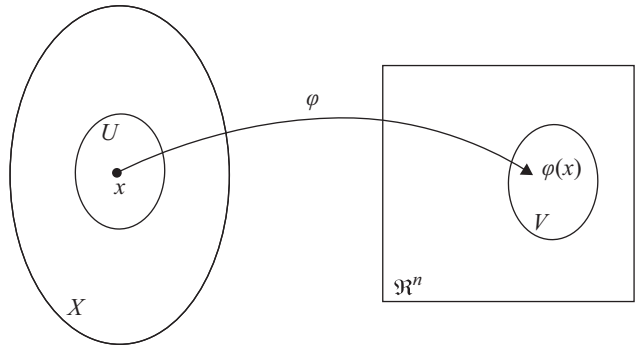


Figure 3.1.

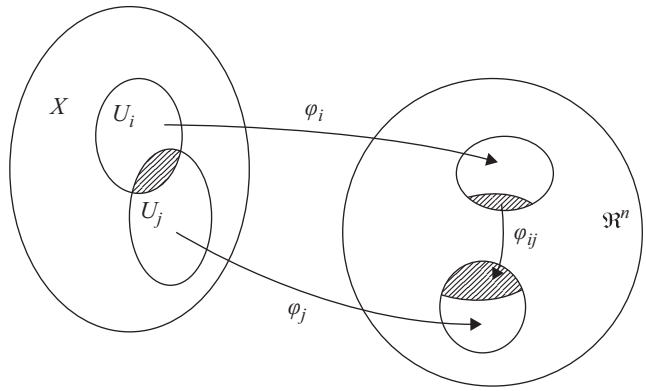


Figure 3.2.

atlases of X are compatible if the maps of these atlases constitute together an atlas of X. Thus an atlas (of dimension n) in a manifold X is therefore a collection of (n-dimensional) coordinate systems such that:

(a) Each point of X is contained in the domain of one of the coordinate systems.
(b) Two coordinate systems in the atlas overlap smoothly.

The existence of a proper atlas is, by definition, equivalent to the statement that X is a differential manifold. Two atlases are equivalent if and only if their union leads back to an atlas. For example, a geography atlas gives a set of maps of various portions of the earth and this provides a very good description of what the earth is, without actually imagining the earth embedded in three-space. It is obvious that $\Re^n$, or more generally any open set in $\Re^n$ is a manifold.

***Example* 1.** A simple example is the circle S^1

$$\{(x, y) \in \Re^n \mid x^2 + y^2 = 1\}.$$

One possible coordinate is a pair of overlapping angular coordinates. Another one is given by

$$\begin{aligned} U_1 &= \{(x, y)|x > 0\}, \quad \phi_1(x, y) = y \\ U_2 &= \{(x, y)|x < 0\}, \quad \phi_2(x, y) = y \\ U_3 &= \{(x, y)|y > 0\}, \quad \phi_3(x, y) = x \\ U_4 &= \{(x, y)|y < 0\}, \quad \phi_4(x, y) = x \end{aligned}$$

with the constraint $x^2 + y^2 = 1$.

***Example* 2.** The torus T^2 (figure 3.3) can be parameterized locally by specifying the values of two angles and it can be covered by an atlas of four mappings by

$$0 < \alpha < 2\pi, -\pi < \alpha < \pi, 0 < \phi < 2\pi, \ -\pi < \phi < \pi.$$

***Example* 3.** The sphere S^2 (figure 3.4) can be given a differential structure by means of two stereographic projections from the north and south poles using maps ϕ_1 and ϕ_2. Let P and Q be the north and south poles, respectively. We can consider the mapping (U, ϕ_1) with $U = S^2 - \{Q\}$ and $\phi_1(p) = (x^1, x^2)$, where p is a point in the surface S^2 and (x^1, x^2) is a Cartesian coordinate system in the tangent plane to the sphere at Q. The stereographic projections are given by

$$x^1 = \frac{2\xi^1}{1 - \xi^3}, \ x^2 = \frac{2\xi^2}{1 - \xi^3},$$

where (ξ^1, ξ^2, ξ^3) are the coordinates of p in $\Re^3$. Taking the tangent plane at P we get the mapping ϕ_2.

More abstractly, a set of continuous transformations such as rotations in $\Re^n$ forms a manifold. Two cones stuck together at their vertices is not a manifold, since the point at the vertices does not look locally like a Euclidean space.

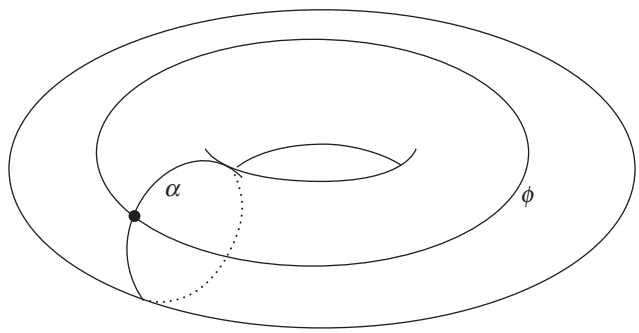


Figure 3.3.

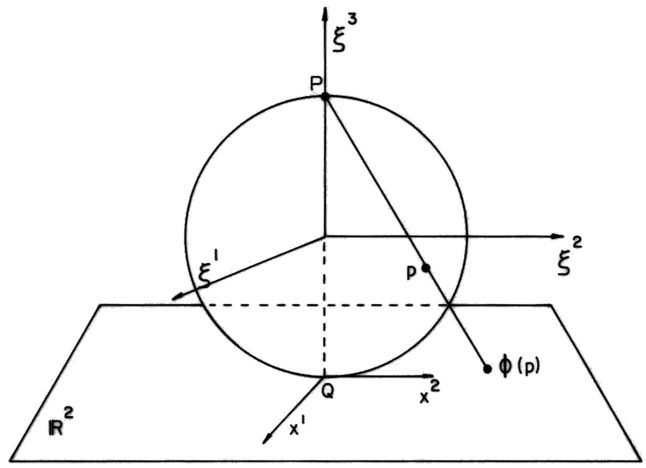


Figure 3.4.

Let X be a manifold of dimension n and $Y \subset X$. We say that Y is a submanifold of dimension $m(m < n)$ if for each $y \in Y$ there is a map (U, ϕ) in X such that if $y \in U$, the element $\varphi(x) \in \mathfrak{R}^n$ and

$$\phi(U \cap Y) = \phi(U)\mathfrak{R}^n \times 0 = \{z \in \phi(U) | \; z^{m+1} = z^{m+2} = \ldots = z^n = 0\}.$$

***Example* 4.** A cone with the origin excluded is a two-dimensional submanifold of $\mathfrak{R}^3$.

3.2 Lie algebra and Lie group

A Lie algebra is a real vector space E with a bilinear map $E \times E \to E$ denoted $(A, B) \to [A, B]$, and called commutators, which satisfies the following conditions:

(a) $[A, A] = 0$ for all $A \in E$,
(b) $[A, B] = -[B, A]$, for all $A, B \in E$,
(c) $[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0$, for all $A, B, C \in E$.

Property (c) is referred to as the Jacobi identity. A subalgebra of a Lie algebra E is a subspace of E which is closed under the bracket operation. Two Lie algebras E and F are isomorphic if there exist a linear isomorphism between them which preserves the bracket. An example of a Lie algebra is the set of all $n \times n$ real matrices, with $[A, B] = AB - BA$.

If $t^1, t^2, \ldots, t^n$ is a basis of E, and since the elements of the algebra are closed over commutation, the commutation of any two vectors of the basis can be written as a linear combination

$$[t^a, t^b] = f^{abc} t^c, \tag{3.3}$$

where f^{abc} are called the structure constants. The Lie algebra is completely determined by its structure constants.

The definition of a group was given in chapter 2. A Lie group is a group of symmetries where the symmetries are continuous and therefore we have an infinite number of elements (Mathews and Walker 1964). The group elements may be labeled by real parameters, which vary continuously; a typical group element can be written as $g(x_1, x_2, \ldots, x_n)$. For instance, a circle has a continuous group of symmetries; it can be rotated by any amount and it looks the same. To be more precise: a Lie group is a smooth manifold obeying the group properties and that satisfies the additional condition that the group operations are satisfied. If G and H are Lie groups, a Lie group homomorphism ϕ: $G \to H$ is a smooth mapping which is also a homomorphism of the abstract groups.

Definition 1. Let G be a Lie group, and $s \in G$. The left translation by s is the map L_s: $G \to G$ given by $L_s(t) = st$ for every $t \in G$. The right translation is defined analogously.

Example. The group of all $N \times N$ unitary matrices with determinant 1. This group is called $SU(N)$.

A Lie group can be parameterized by a set of continuous parameters α_i, with $i = 1, \ldots, n$, where n is the number of parameters on which the group depends. We denote the group elements by $g(\alpha_i)$. We will take

$$g(\alpha_i)|_{\alpha_i=0} = e, \tag{3.4}$$

the identity element. If D_n are $n \times n$ matrices that constitute a representation for the group, we have

$$D_n(g(\alpha_i))|_{\alpha_i=0} = I, \tag{3.5}$$

where I is the identity matrix. Expanding D_n in the neighborhood of the origin we have

$$D_n(g(\delta\alpha_i)) = I + \delta\alpha_i \frac{\partial D_n(g(\alpha_i))}{\partial \alpha_i}\bigg|_{\alpha_i=0} + \ldots \tag{3.6}$$

where $\delta\alpha_i \ll 1$. Now we define

$$X_i \equiv -i\frac{\partial D_n}{\partial \alpha_i}\bigg|_{\alpha_i=0}, \tag{3.7}$$

where I have written $-i$ in equation (3.7), such that X_i is Hermitian. Thus

$$D_n(\delta\alpha_i) = I + i\delta\alpha_i X_i + \ldots$$

Writing $\alpha_i = N\delta\alpha_i$, with $N \to \infty$, we have for finite values of α_i

$$\lim_{N\to\infty} (1 + i\delta\alpha_i X_i)^N = \lim_{N\to\infty} \left(1 + i\frac{\alpha_i}{N} X_i\right)^N. \tag{3.8}$$

Now we can use the result

$$\lim_{N\to\infty}\left(1 + i\frac{\alpha_i}{N}X_i\right)^N = e^{i\alpha_i X_i}, \tag{3.9}$$

to write

$$D_n(\alpha_i) = e^{i\alpha_i X_i}. \tag{3.10}$$

The elements X_i are called generators of the group. There is one generator for each parameter. For instance, for the group of rotations SO(3) we need three angles θ, φ, ψ to specify an element of the group. We have then three generators X_θ, X_φ and X_ψ.

For a linear displacement of a function $f(x)$ in one dimension of a distance a we have

$$(T_a f)(x) = f(x + a) = \left(1 + a\frac{d}{dx} + \frac{a^2}{2!}\frac{d^2}{dx^2} + \ldots\right) f(x) = e^{a\frac{d}{dx}} f(x). \tag{3.11}$$

Thus

$$T_a = \exp\left(a\frac{d}{dx}\right), \tag{3.12}$$

is the generator.

The generators of a Lie group X_i form a Lie algebra defined through the commutation relations

$$[X_a, X_b] = if^{abc}X_c. \tag{3.13}$$

If $f^{abc} = 0$, the Lie group is abelian, otherwise it is non-abelian.

Suppose that the Hamiltonian of a condensed matter system has some continuous symmetry given by a Lie group G. Then it is possible that at some values of the parameters of the Hamiltonian the ground state of the system breaks the symmetry up to some subgroup H of G. The ground state is then invariant under H, but not under the remaining elements of G, which are denoted as a coset and written as G/H. The coset is not a subgroup of G (for example, it does not contain the identity element).

3.3 Homotopy

A property of a topological space that is invariant under homeomorphisms (remember, a map is a homeomorphism if it is both continuous and has an inverse which is also continuous) is called a topological invariant (for instance, the dimension of a manifold and the orientability of a connected manifold are topological invariants) (Manton and Sutcliffe 2004, Mombelli 2018). If some topological invariant is different for two topological spaces X and Y the two spaces are homeomorphic. If we introduce a third space Z, then we can verify that if X is homeomorphic to Y and Y is homeomorphic to Z, then by composing the two homeomorphisms, X is homeomorphic to Z. This means that we are able to divide

all topological spaces up into equivalence classes. A pair of spaces X and Y belong to the same equivalence class if they are homeomorphic. A more detailed testing of equivalence of X and Y can be performed when we have more invariants. The homotopy theory constructs infinitely many topological invariants to characterize a given topological space and show how to compare topological spaces.

Let p_1 and p_2 be points in a manifold X. If there is a curve C in X that goes from p_1 to p_2, we say that X is connected by paths. If, in addition, X is such that given any two curves C_1 and C_2 going from any point p_1 to any point p_2, C_1 can be continuously deformed into C_2 in X; that is, if there exists a continuous function $p(t, s)$ such that for each s in the range [0, 1], $p(t, s)$ describes a curve from p_1 to p_2, when t varies, and this curve coincides with C_1 for $s = 0$ and with C_2 for $s = 1$, then X is called simply connected (or only connected). Two curves C_0 and C_1 in a manifold X, both having the same starting point p_0 and the same final point p_1 are called homotopic if one of them can be deformed in the other by continuous deformations in X. Thus, a manifold is connected if any two curves in it, having the same initial and final points, are homotopic. Given the points p_0 and p_1, the class of all curves that are homotopic at a given curve from p_0 to p_1 form an equivalence class, or homotopy class, at X.

Let X and Y be two manifolds. A map ψ_0: $X \to Y$, with $\psi_0(x_0) = y_0$, where $x_0 \in X$ and $y_0 \in Y$ is said to be homotopic to another such map ψ_1, if ψ_0 can be continuously deformed into ψ_1 (x_0 and y_0 are fixed points called base points). We can also say that ψ_0 is homotopic to ψ_1 if there is a continuous map

$$\tilde{\psi}\colon X \times [0, 1] \to Y, \tag{3.14}$$

with $t \in [0, 1]$ such that $\tilde{\psi}(x_0, t) = y_0$ for all t and $\tilde{\psi}|_{t=0} = \psi_0$, $\tilde{\psi}|_{t=1} = \psi_1$. The maps ψ are symmetric, transitive and reflexive, thus they can be classified into homotopy classes.

Let us consider the case where X is an n-sphere S^n (that is, the set of points in $\Re^{n+1}$ at unit distance from the origin). The set of homotopy classes of maps ψ: $S^n \to Y$ is denoted by $\pi_n(Y)$ (figure 3.5). For $n \geqslant 1$, the set $\pi_n(Y)$ forms a group, called the nth

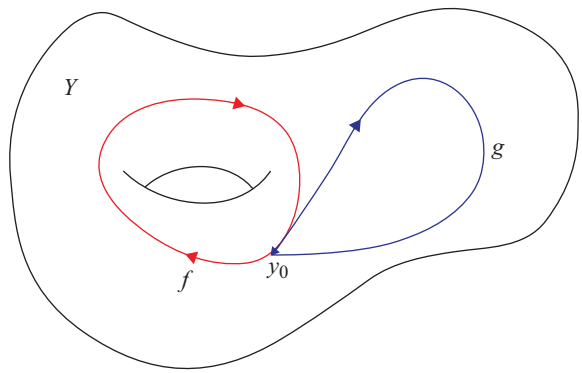


Figure 3.5. We can view the mapping $\pi_1(Y)$, f: $S^1 \to Y$, as a mapping from the interval [0, 1], such that $f(0) = f(1) = y_0$. These maps are associated with path in Y, beginning and ending at the point y_0.

homotopy group of Y. A map $S^1 \to Y$ is called a loop. The class of the constant map $S^1 \to y_0$ is the identity element of the group $\pi_1(Y)$, called the fundamental group of Y. This group is generally non-abelian. If Y is connected and $\pi_1(Y) = I$, where I denotes the trivial group with just the identity element, the space Y is said to be simply connected. In this case, every loop is contractible (can be continuously deformed to a point), i.e. homotopic to the trivial loop. If $\Re^d$ has the origin as base point, any loop ψ_0: $S^1 \to \Re^d$ is contractible (parameterize S^1 by $\theta \in [0, 2\pi]$, and define $\tilde{\psi}(\theta, t) = (1 - t)\psi_0(\theta)$, therefore $\pi_1(\Re^d) = I$). If there is a hole in the space $\Re^n$, the loops can be divided into classes, each one characterized by the number of times the loop winds around the hole.

A simple case is the mapping of circles into circles, i.e. $S^1 \to S^1$. We can parameterize the circle using an angle θ defined modulo 2π. A mapping can be defined by a continuous function $\Lambda(\theta)$ modulo 2π. As an example, let us consider two such mappings as

$$\Lambda_0(\theta) = 0 \text{ for all } \theta \tag{3.15}$$

and

$$\tilde{\Lambda}_0(\theta) = \begin{cases} t\theta & \text{for } 0 \leqslant \theta < \pi \\ t(2\pi - \theta) & \text{for } \pi \leqslant \theta < 2\pi \end{cases}, \tag{3.16}$$

where $t \in [0, 1]$. By varying t continuously down to zero the second mapping can be continuously deformed into the first. These two mappings therefore belong to the same homotopy class.

Let us now consider the mapping

$$\Lambda_1(\theta) = \theta \text{ for all } \theta. \tag{3.17}$$

As θ completes a full circle so does Λ_1. But now it cannot be continuously deformed into equation (3.15) or (3.16), since in equation (3.17) the second circle is wound once around the first circle, whereas in equations (3.15) and (3.16) it is effectively wound zero times. Thus equation (3.17) belongs to a different homotopy class from equation (3.15) or (3.16). The integer distinguishing the two classes is the winding number defined by

$$W = \frac{1}{2\pi} \int_0^{2\pi} \frac{d\Lambda}{d\theta} d\theta. \tag{3.18}$$

Therefore, $\Lambda_n(\theta) = n\theta$ is the prototype mapping belonging to the $W = n$ class (negative values of W are obtained by doing the winding in the opposite sense). The winding number W is the net number of times that the image $\Lambda(\theta)$ winds around the target as θ goes once around the domain. The product $\gamma_2\gamma_1$ of path γ_1 and γ_2 characterized by winding numbers W_1 and W_2, respectively, has winding number $W = W_1 + W_2$.

Let us now consider the non-singular mappings of a sphere S^2 into another sphere S^2. As was said above, these mappings can be classified into homotopy sectors. A mapping in one sector can be continuously deformed into another, whereas

mappings from two different sectors cannot. There is a denumerable infinity of such homotopy sectors or classes, which can be characterized by integer numbers (positive, negative and zero). That is, these homotopy classes form a group which is isomorphic to the group of integers. We can write in a compact form

$$\pi_2(\mathrm{S}^2) = Z, \tag{3.19}$$

where $\pi_n(S^m)$ means the homotopy group associated with the mappings $S^n \to S^m$ and Z is the group of integers. The integer characterizing the homotopy classes of $S^2 \to S^2$ is the number of times one of the spheres has been wrapped around the other. We will find several examples of this case later in the text. We also have:

$$\pi_n(S^n) = Z,\ \pi_n(S^m) = 0 \text{ for } n < m,\ \pi_n(S^1) = 0 \text{ for } n > 1. \tag{3.20}$$

As an example, we have $\pi_1(S^2) = 0$ since any loop in a sphere can be deformed to a point. The calculation of homotopy groups $\pi_n(S^m)$ for $n > m$ is a very difficult problem. One highly non-trivial result is $\pi_3(S^2) = Z$. The integer number labeling homotopy classes in this case is called the Hopf invariant (see section B.3). The so-called Hopf insulator, is a three-dimensional topological insulator possessing a non-zero Hopf number. The mapping of a circle into the two-dimensional torus is labeled by two integers—two winding numbers of circle around torus cycles. We have the general result

$$\pi_1(T^d) = Z \times \underset{d}{\ldots} \times Z. \tag{3.21}$$

As was said above, we are generally interested in comparing two manifolds X and Y, but instead of comparing these manifolds directly, one uses a 'test manifold' M and compares mappings of X and Y into M. Studying the homotopy classes of those mappings one can compare X with Y. It is convenient to take as the 'test manifolds' M spheres S^n, since in this case one can endow the homotopy classes of those mappings with group structures.

In section 4.16, after introducing some more mathematical concepts, I will present the degree of a map in homotopy theory.

3.4 Particle in a ring

To give a simpler example where the winding number is used in condensed matter, let us consider a Lagrangian of the form (Altland and Simons 2010):

$$L\left(\varphi, \frac{\partial \varphi}{\partial t}\right) = \frac{1}{2}\left(\frac{\partial \varphi}{\partial t}\right)^2 - iA\frac{\partial \varphi}{\partial t}. \tag{3.22}$$

This Lagrangian describes a quantum particle on a ring threaded by a magnetic flux. Here $\varphi \in [0, 2\pi]$ with periodic boundary conditions. Using the formalism of chapter 1, we write the imaginary time path integral as

$$Z = \int D\phi e^{-S[\phi]}, \tag{3.23}$$

where

$$S[\varphi] = \int_0^\beta d\tau L\left(\varphi, \frac{\partial\varphi}{\partial\tau}\right), \tag{3.24}$$

with the boundary condition $\phi(\beta) - \phi(0) \in 2\pi Z$. Using the Euler–Lagrange equation we obtain

$$\frac{\partial^2\phi}{\partial t^2} = 0. \tag{3.25}$$

There exists a whole family of solutions $\phi_w(\tau) = 2\pi W\tau T$ and the action $S[\phi_w]_{A=0} = \frac{1}{2}T(2\pi W)^2$ varies discontinuously with W. From the mathematical point of view the field ϕ is a mapping ϕ: $S^1 \to S^1$, $\tau \to \phi(\tau)$, from the unit circle (imaginary time with periodic boundary condition) into another circle. The integration over all functions $\phi(\tau)$ can be written as integration over functions $\phi(\tau)$ of different winding numbers (that is, different topological sectors):

$$Z = \sum_W \int D\phi^{-S[\phi]} = \sum_W e^{2\pi i W A} \int D\phi \exp\left(-\frac{1}{2}\int d\tau\left(\frac{\partial\phi}{\partial\tau}\right)^2\right). \tag{3.26}$$

The A-dependent term in the action

$$S_{\text{top}}[\phi] \equiv iA\int_0^\beta d\tau\frac{\partial\phi}{\partial t} = A(\phi(\beta) - \phi(0)) = i2\pi WA, \tag{3.27}$$

involves only the index of the topological sector of ϕ. We can see that S_{top} cannot affect the equation of motion, however it does affect the result of the integration: it plays the role of a W-dependent 'phase' weighting the contribution of different sectors to the path integral.

3.5 Functions on manifolds

A function f on a manifold X takes a point $x \in X$ at the point $f(x) \in \Re$. The function f: $X \to \Re$ is differentiable at the point x of a differentiable manifold X if on a map φ at x, $f\varphi^{-1}$ is differentiable at $\varphi(x)$ (see figure 3.6). This definition does not depend on the chosen map. It is important to remember the distinction between a function and its set of image points.

If $(u^1, \ldots, u^n)$ are the coordinates of the point $u \in \Re^n$, the functions β^i defined by $\beta^i(u) = u^i$ are called the coordinate functions in $\Re^n$. If $x \in X$ and (U, φ) is a map, we saw that

$$\varphi(x) = (x^1, \ldots, x^n), \phi(x) \in \Re^n. \tag{3.28}$$

Then we have

$$\beta^i.\ \varphi(x) = \beta^i(x^1, \ldots, x^n) = x^n. \tag{3.29}$$

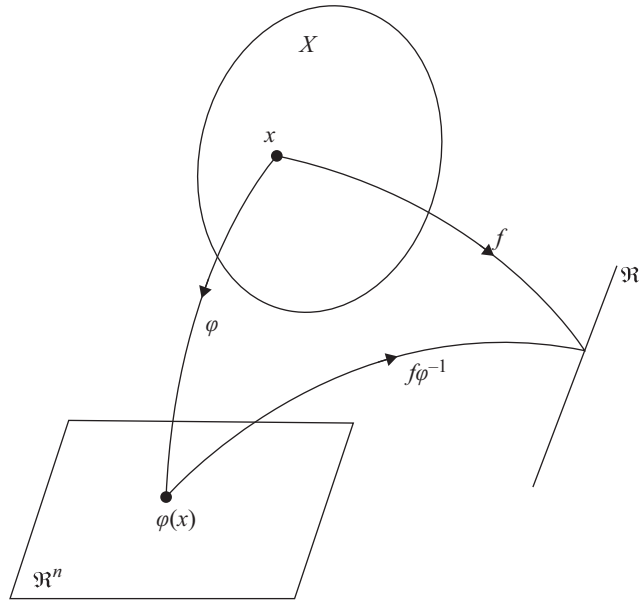


Figure 3.6.

Let us define φ by

$$\varphi_i \equiv \beta^i . \varphi,$$

so that we can write $\varphi = (\varphi^1, \ldots, \varphi^n)$ where the φ^i are the coordinate functions of φ in X. So φ^i maps X into $\Re^n$. That is, it takes the point $x \in X$ to the point $x^i \in \Re$,

$$\varphi^i(x) = x^i. \tag{3.30}$$

3.6 Tangent space

Let X be a manifold and $x \in X$. A curve on X is a mapping c: $I \to X$ of an open interval $I \subset \Re$ on X. If $t \in I$ have $c(t) \in X$ with $0 \in I$ and $c(0) = x$. Note that the 'curve' is defined to be the map itself, not the set of image points in X. If c is a differentiable curve and f is a function on X, differentiable at x, the composite mapping $f.c$ is a differentiable function at I at $t = 0$.

Since a curve in a manifold does not live in $\Re^n$, the tangent vector cannot be defined in a simple way using derivatives. We have to resort to a more abstract approach.

Let $C(X)$ be the space of a smooth, real-valued function defined on X, and $f \in C(U)$ a function defined on an open neighborhood U of x. The directional derivative of f along c at the point x is given by

$$v_x^c(f) \equiv \frac{d}{dt}(f . c)(t)\bigg|_{t=0}. \tag{3.31}$$

Now we define the mapping $f \to v_x^c(f)$ to be the *tangent vector* at point x.

A tangent vector v_x^c is completely determined if we know how functions vary along the curve c. Thus, to all intents and purposes the notion of a tangent vector can be equated with the operation of taking the directional derivative along c: the vector tangent to a manifold X at a point x is the operator which when applied to a function of real value f, analytic around x, leads to the real number $v_x^c(f)$.

We say that two curves c_1 and c_2 are tangent at x, or that they have the same tangent vector at x, if $c_1(f) = c_2(f)$ for all f. On the other hand, given a tangent vector v_x it is always possible to find a differentiable curve passing through x to which v_x is tangent. A tangent vector is an abstraction created to represent the structure common to a class of parameterized curves, all of them tangent at one point.

Given a map (U, φ), $U \subset X$, that is $\varphi(x) \to (x^1, \ldots, x^n)$, we can take the function $\phi^i(x) = x^i$ and define the components of the tangent vector v_x to the curve c by:

$$(v_x^c)^i \equiv v_x^c(\varphi^i) = \frac{d}{dt}(\varphi^i \cdot c)(t)\bigg|_{t=0}. \tag{3.32}$$

As $c(t) \in X$ we have $\varphi \cdot c(t) \in \Re^n$. We can therefore write:

$$\varphi \cdot c(t) = (c^1(t), \ldots, c^n(t)), \text{ or } \varphi \cdot c = (c^1, \ldots, c^n). \tag{3.33}$$

Using equation (3.30) we can write

$$\beta^i \cdot \varphi \cdot c(t) = \varphi^i \cdot c(t) = \beta^i(c^1(t), \ldots, c^n(t)) = c^i(t). \tag{3.34}$$

Then we can write equation (3.32) as

$$(v_x^c)^i = \frac{d}{dt}c^i(t)\bigg|_{t=0}, \tag{3.35}$$

and equation (3.31) as

$$v_x^c(f) = \frac{d}{dt}(f \cdot \varphi^{-1} \cdot \varphi \cdot c)(t)\bigg|_{t=0}. \tag{3.36}$$

Defining $\overline{f} = f.\ \varphi^{-1}$ we note that $\overline{f}$ it is a function of into $\Re$, that is, $\overline{f}$ is the expression for f on the map (U, φ). Equation (3.36) can then be written as

$$v_x^c(f) = \frac{d}{dt}[\overline{f}\ .\ (c^1(t), \ldots, c_n(t))]\bigg|_{t=0}. \tag{3.37}$$

Using the chain rule for compound functions and remembering that $c^i(0) = x^i$ we can write equation (3.37) as

$$v_x^c(f) = \frac{\partial \overline{f}}{\partial x^i}\bigg|_{\varphi(x)} \frac{dc^i}{dt}\bigg|_{t=0} = \frac{\partial \overline{f}}{\partial x^i}\bigg|_{\varphi(x)} (v_x^c)^i. \tag{3.38}$$

Using $v^i = (v_x^c)^i$ we have

$$v_x(f) = v^i \frac{\partial \overline{f}}{\partial x^i}\bigg|_{\varphi(x)}. \tag{3.39}$$

The representative of v_x can be written in a short way as: $v^i \partial/\partial x^i$, where $\partial/\partial x^i \in \Re^n$, although it spans a vector space. The vector space T_xX of tangent vectors to X at the point x, with the addition and multiplication by scalars defined by: $(au_x + bv_x)(f) = au_x(f) + bv_x(f)$ is called the *tangent (vector) space* at the point x.

From equations (3.31), (3.33) and (3.35) we can write

$$v_x^c(\varphi) = \frac{d}{dt}(\varphi \cdot c)(t)\Big|_{t=0} = (v^1, \ldots, v^n). \tag{3.40}$$

Let us define $\omega_k \in T_xX$ as

$$\omega_k = \frac{\partial}{\partial x^k}, \tag{3.41}$$

so that we can write

$$\omega_k(\varphi) = \frac{\partial \varphi(x)}{\partial x_k}. \tag{3.42}$$

Using equation (3.28) we get

$$\frac{\partial \varphi(x)}{\partial x^k} = (0, \ldots, 1, \ldots, 0), \tag{3.43}$$

(where 1 is at the kth position). This means that $\omega_k(\varphi) = e_k$, where e_k is the standard basis. The vectors of T_xX that are represented by $(\partial/\partial x^1, \ldots, \partial/\partial x^n)$ form a basis for the tangent vector space, called the natural basis. It follows that the dimension of T_xX is the same as the dimension of the manifold X. A map (U, φ) on X then induces an isomorphism of T_xX in $\Re^n$. The representative of v_x, which we shall call v, in relation to the natural basis, is then given by $v = (v^1, \ldots, v^n)$.

3.7 Cotangent space

The *cotangent space* in x of a manifold X is defined as the dual space of T_xX and is denoted by T_x^*X, that is, T_x^*X is the space of the linear mappings of $T_x X$ in $\Re$. So if $\omega_x \in T_x^*$ and $v_x \in T_xX$, we have $\omega_x(v_x) \in \Re$. However, $T_x^{**} = T_xX$, and so $v_x(\omega_x) = \omega_x(v_x)$. We call the elements of T_x^*X of cotangent vectors, covariant vectors, or differential forms. Thus a differential p-form ω at the point x of the manifold X is an external p-form in the tangent space T_xX at a point x in X.

We have seen that, given a base $\{e_i\}$ in a vector space (which may be T_xX) of dimension n we can construct the dual basis $\{\theta^i\}$ (in T_x^*X) as follows. The components v_x^i of a vector v_x with respect to the base $\{e_i\}$ constitute n linear forms defined in v_x. We define the form θ^i by

$$\theta^i(v_x) = v_x^i. \tag{3.44}$$

Since $v_x = v^j e_j$, we can write equation (3.44) as

$$\theta^i(v^j e_j) = v^j \theta^i(e_j) = v_x^i, \tag{3.45}$$

which leads to

$$\theta^i(e_j) \equiv \langle \theta^i, e_j \rangle = \delta^i_j. \tag{3.46}$$

Let us denote the dual of the natural basis $\{\partial/\partial x^i\}$ by $\{dx^i\}$, this is

$$\langle dx_i, \partial/\partial x^j \rangle = \delta^i_j. \tag{3.47}$$

If we denote the components of ω_x related to the basis $\{dx^i\}$ by ω_{x_i}we can write

$$\langle \omega_x, v_x \rangle = \langle \omega_{x_i} dx^i, v^j \partial/\partial x^j \rangle = \omega_{x_i} vxi. \tag{3.48}$$

Example. Suppose the form $\omega = 2x^1 dx^1 + 3x^2 dx^2$ and the vector field $v = 2\frac{\partial}{\partial x^1} - \frac{\partial}{\partial x^2}$.

We want to calculate $\omega(v)$ at the point (1, 2). We have

$$\omega(v) = (2x^1 dx^1 + 3x^2 dx^2)\left(2\frac{\partial}{\partial x^1} - \frac{\partial}{\partial x^2}\right) = 4x^1 - 3x^2, \quad \omega(v)\Big|_{(1,2)} = -2.$$

3.8 Push-forward

Let M and N be two manifolds (possibly of different dimensions), ϕ a map ϕ: $M \to N$, and f a function f: $N \to \Re$ (figure 3.7). We can compose ϕ with f to construct a map $(f.\ \phi)$: $M \to \Re$, which is simply a function on M. We define the pull back of f by ϕ, written as $\phi^* f$, by

$$\phi^* f = (f.\ \phi). \tag{3.49}$$

Functions can be pulled back, but they cannot be pushed forward. If one has a function g: $M \to \Re$, there is no way one can compose g with φ to create a function on N.

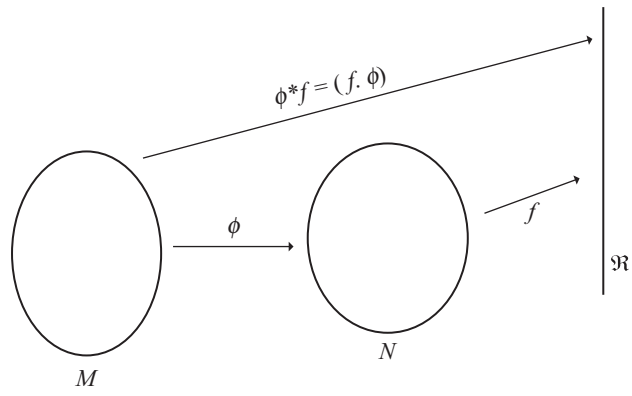


Figure 3.7.

But vectors can be considered as a derivative operator that maps smooth functions to real numbers, and this allows us to define the push-forward of a vector as we will discuss in the following.

To some extent, a tangent space can be regarded as a local linearization of the manifold, and it is important that a map f between two manifolds X and Y can be linearized with the aid of the tangent spaces and the vector space structure that they carry.

Let X be a manifold of dimension n and Y a manifold of dimension m. Let f: $X \to Y$ be a differentiable mapping of the manifold X into the manifold Y. (The mapping f is called differentiable when, in local coordinates in X and Y, it is determined by differentiable functions.) Let c be a curve on X and v_x^c the tangent vector to that curve at the point x, f maps the curve c on X into a curve $c' = f.c$ on Y and the point $x \in X$ into the point $y \in Y$. Let $w_y^{c'}$ be the tangent vector to the curve c' at the point y. The mapping taking v_x^c into $w_y^{c'}$ is called the *push-forward* derivative of f. Or more precisely: we call the push-forward of the mapping f at the point $x \in X$ the linear mapping of the tangent spaces

$$Df(x)\colon T_xX \to T_yY, \tag{3.50}$$

such that tangent curves are taken into tangent curves by the following definition

$$Df(x)v = w, \tag{3.51}$$

where $v \in T_xX$, $w \in T_yY$. We define w such that, for each differentiable function at $y = f(x)$, where $y \in Y$, we have

$$w(h) = v(h.f). \tag{3.52}$$

The vector w is called the image of v over f and the mapping $Df(x)$ is called the push-forward of f at x (Df is sometimes written as f_*).

Given two manifolds X and Y a map φ: $X \to Y$ can be used to pull objects back and push other objects forward. However, it generally does not work both ways because φ might not be invertible. If φ is invertible (and both φ and φ^{-1} are smooth), then it defines a diffeomorphism between X and Y, which is what happens if X and Y are the same manifold.

3.9 Fiber bundle

A fiber bundle is a generalization of the product of two manifolds M and N (Collinicci and Wyns 2006, Duffel 2017, Penrose 2007). One way to interpret a product manifold is to place a copy of M at each point of N. For instance in $\mathfrak{R}^2 = \mathfrak{R} \times \mathfrak{R}$ we take a line in $\mathfrak{R}$ as our base and place another line at each point of the base, forming a plane. If M is a circle S^1 and N a line, then $M \times N$ is now a cylinder. If both M and N are circles $M \times N$ is a torus.

In any local covering, a fiber bundle looks like $M \times N$, but globally it is not in general a product of manifolds. Let M be a manifold, which we call the base, and F a manifold that we call the fiber. A fiber bundle E over M with fiber F is a manifold that is locally a direct product $M \times N$. That is, if M is covered by a set of local coordinates in a covering $\{U_i\}$, then the bundle E is topologically described in each coverture U_i by the product manifold $U_i \times F$. However, the local direct-product

structure leaves a great deal of information about the global topology of E undetermined. For a complete specification of the bundle E, it is necessary to provide a set of transition functions $\{\phi_{ij}\}$ which tell us how the fiber manifolds match up in the overlap between two coverings $U_i \cap U_j$. Locally the topology of the bundle is trivial; however, the global topology, determined by $\{\phi_{ij}\}$ can be complicated due to the relative twisting of neighboring fibers. A trivial example happens when all the transition functions are the identity. In this case the global topology of the bundle is that of a direct product $E = M \times F$, and the bundle is called a trivial fiber bundle.

A simple and famous non-trivial example of a fiber bundle is the Möbius strip (figure 3.8). We start with $M = S^1$ and N as a line segment. At each point of the circle we attach a copy of the interval $[-1, 1]$, but instead of attaching a band of parallel intervals to the circle, the intervals perform a π twist as we go around. Locally the Möbius strip is indistinguishable from a piece of a cylinder. That is, the twist is not located at any particular point on the band; it is a global property of the manifold.

Suppose we parameterize the Möbius strip by a point θ in the circle, and a real number t on the interval $[-1, 1]$. By increasing θ we can transport a point around the strip. However because of the twist, the point will not coincide with its initial position in the strip after a 2π rotation in the circle. We see that the parameterization for F does not work globally. However, in the neighborhood of a point in E we can use a parameterization and when we go to another neighborhood we use a different parameterization. We show how this is done.

We can cover S^1 by two semicircular coverings $U_\pm$ as follows

$$\begin{aligned} U_+ &= \{\theta\colon -\varepsilon < \theta < \pi + \varepsilon\} \\ U_- &= \{\theta\colon \pi - \varepsilon < \theta < 2\pi + \varepsilon = 0 + \varepsilon\}' \end{aligned} \tag{3.53}$$

We take the fiber F to be an interval in the real line with coordinates $t \in [-1, 1]$, as was said before. The bundle then consists of the two local pieces

$$\begin{aligned} U_+ \times F &\text{ with coordinates } (\theta, t_+) \\ U_- \times F &\text{ with coordinates } (\theta, t_-) \end{aligned}$$

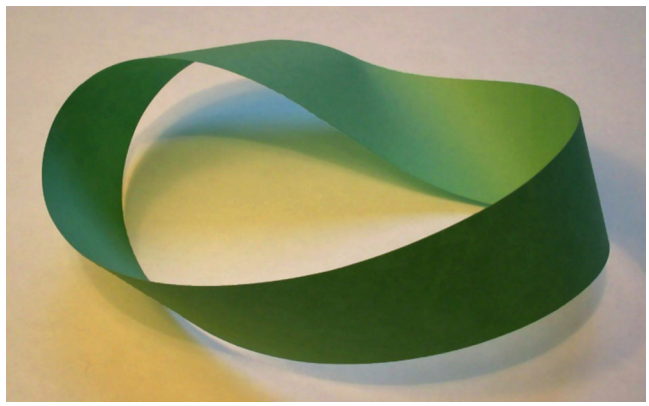


Figure 3.8. David Benbennick / Wikimedia Commons / https://creativecommons.org/licenses/by-sa/3.0/deed.en

and the transition functions relating t_+ to t_- in $U_+ \cap U_-$. This overlap consists of two regions I and II given by

$$\mathrm{I} = \{\theta\colon -\varepsilon < \theta < \varepsilon\}, \qquad \mathrm{II} = \{\theta\colon \pi - \varepsilon < \theta < \pi + \varepsilon\}. \tag{3.54}$$

We choose the transition functions to be $t_+ = t_-$ in region I, $t_+ = -t_-$ in region II.

Identifying t_+ with $-t$ in region II twists the bundle and gives it the non-trivial topology of the Möbius strip. If we set $t_+ = t$ in both regions we get a trivial bundle equal to the piece of a cylinder. The Möbius strip is the construction appropriate to the spinor behavior of a quantum mechanical two-level system.

To capture the fact that we are attaching a copy of the fiber F to each point x of the base space, we introduce a projection π which maps the fiber bundle E onto the base space M by shrinking each fiber to a point. If $x \in M$, $\pi^{-1}(x)$ is the fiber over x. We can now give a formal definition of a fiber bundle.

A differentiable fiber bundle (E, π, M, F, G) consists of the following elements

(1) A differentiable manifold E called the bundle space or total space.
(2) A differentiable manifold M called the base space.
(3) A surjection $\pi\colon E \to M$ called the projection map.
(4) A differentiable manifold called the typical fiber. If $x \in M$, the set F_x defined by $F_x = \pi^{-1}(x)$ is called the fiber of E at the point $x \in M$. Note that $\pi^{-1}(x)$ is the inverse image of the point x, that is, it consists of the elements of E that are taken by the transformation π in the element x of M. We require that each F_x be homeomorphic to F. The bundle space itself may be thought of as being completely made up of a whole family of fibers F. There are many copies of the fiber F in the bundle E, one entire copy standing above each point x of M. The copies are all disjoint, i.e. no two intersect, and together they make up the entire space E.

The general idea of a bundle is outlined in figure 3.9. For most purposes these items are sufficient to characterize a bundle. For the sake of consistency, we

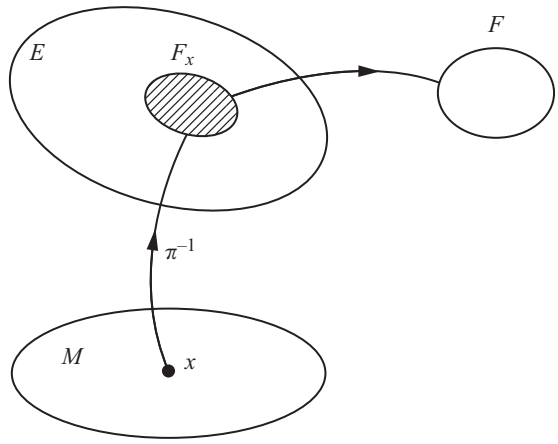


Figure 3.9.

present the complementary items that are included in the complete definition of a bundle.

(1) A (Lie) group G, called the structure group of the fiber bundle, of homeomorphism of F into itself.
(2) An open covering $\{U_i\}$ of M and a set of diffeomorphisms ϕ_i: $U_i \times F \to \pi^{-1}(U_i)$. Letting (x, f) denote a point of $U_i \times F$ we require that $\pi[\phi_i(x, f)] = x$ for consistency. The map ϕ_i is called a local trivialization.
(3) At each point $x \in M$, $\varphi_{i,x}(f) \equiv \varphi_i(x, f)$ is a diffeomorphism, $\phi_{i,x}$: $F \to F_x$. On each overlap $U_i \cap U_j \neq \{\varnothing\}$, we require $h_{ij} = \phi_{i,x}^{-1}\phi_{j,x}$: $F \to F$ to be an element of G, i.e. we have a smooth map h_{ij}: $U_i \cap U_j \to G$ such that $\varphi_j(x, f) = \varphi_i(x, h_{ij}(x)f)$.

The transition functions h_{ij} satisfy the conditions:

(a) h_{ii} = identity,
(b) $h_{ij}h_{jk} = h_{ik}$ for $x \in U_i \cap U_j \cap U_k$.

A bundle is completely determined by its transition functions.

A *vector bundle* over a space M is a continuous map π: $E \to M$ such that each fiber $\pi^{-1}(p)$, where $p \in M$, is a finite-dimensional vector space. A vector bundle is said to have dimension n if the fiber over every point of the base space is homeomorphic to $\Re^n$. A continuous map ϕ: $E_1 \to E_2$ between vector bundles over the same base space is an isomorphism if it takes each fiber $\pi_1^{-1}(x)$ to the corresponding fiber $\pi_2^{-1}(x)$ by an isomorphism.

A *line bundle* is a vector bundle with a one-dimensional vector space as fiber. It is a family of lines parameterized by the base space M.

The *tautological bundle* $S \to \Re P^n$ is a line bundle defined as follows: each point $l \in \Re P^n$ is a line in $\Re^{n+1}$; we let the fiber above l be that line.

In figure 3.10 I show an example, where the space M is a circle S^1, and the fiber F is a one-dimensional vector space, which we can consider topologically as a copy of the real line $\Re$ (with the origin marked). In figure 3.10(a) we have a line bundle over S^1, which as was mentioned in the beginning of this section, is the trivial case $M \times F$

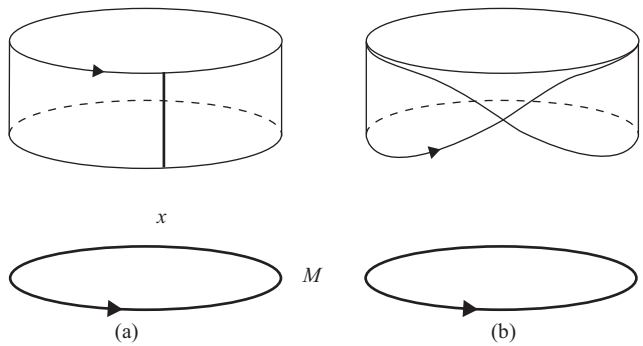


Figure 3.10. Reproduced from Frucart and Carpentier (2013). Copyright © 2013 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

which is the ordinary cylinder. In figure 3.10(b) we have a twisted case and we get the Möbius strip.

Definition 2. Let (E, π, M) with π: $E \to M$ be a bundle with base space M, and f: $M' \to M$ a continuous map from a manifold M' to the base space M. Then, the *pull back* of E to M' denoted $f^* E \to M'$ is the bundle whose fiber above $x' \in M'$ is $\pi^{-1}(f(x'))$.

We can understand the above definition using figure 3.11. E is a bundle over M with projection π: $E \to M$, f is a map from M' to M. Then we pull back the bundle over M to become a bundle over M' that we will call E'. We construct M' by taking a subset of $M' \times E$ composed of all pairs $(x', e) \in M' \times E$ such that $f(x') = x = \pi(e)$. We define a projection from E' to M' by $\pi'(x', e) = x'$. In this way we have pulled back the space E which fibers over M, to a space E' that fiber over M'.

If f, g are a pair of maps from M into M' that are homotopic, then we can show that $f^*(E)$ and $g^*(E)$ are isomorphic bundles over M.

If $B_F(M)$ denotes the set of isomorphism classes of F-bundles over M, any F-bundle ξ over M' generates a map

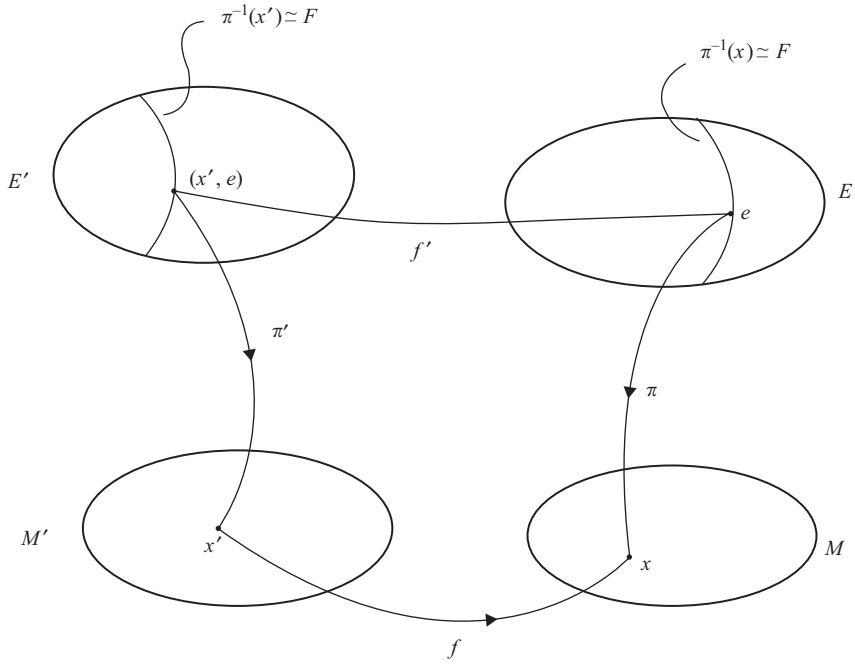


Figure 3.11.

$$\alpha_\xi : [M, M'] \to B_F(M),\ f \to f^*(\xi), \tag{3.55}$$

where $[M, M']$ denotes the set of homotopy classes of maps from M into M'. For a class of spaces M, which include all differential manifolds, there exist universal bundles: $UF = F \to EF \to BF$, with the property that the map

$$\alpha_{UF} : [M, BF] \to B_F(M) \tag{3.56}$$

is both one-to-one and onto. It can be shown that any F-bundle whose total space EF is contractible can serve as a model for a universal F-bundle.

If $f_1, f_2 : Y \to X$ are homotopic maps, then $f_1^* E$ and $f_2^* E$ are isomorphic. It follows that homotopy-equivalent spaces have the same isomorphism classes of vector bundles over them. In particular, any bundle over a contractible space is trivial.

If the fiber F is the same as the group G, and if the transitions functions act on $F = G$ by left translations, the fiber bundle is called a principal fiber bundle. This is of special interest because the mathematical structure of a non-interacting insulator is that of a principal fiber bundle.

For a given fiber bundle (E, M, G) we can ask how many different fiber bundles can be constructed and how they differ from the trivial bundle. The question of whether two different bundles are the same can be quite hard to answer. Two fiber bundles are considered different if they cannot be continuously deformed one into the other. A continuous deformation of a fiber bundle means a continuous deformation of the base space and of the fibers. For example, the surface of a cylinder can be continuously deformed into the surface of a cone which lies between two parallel planes cutting the cone (perpendicular to the cone axis), but it cannot be deformed into the Möbius strip.

Homotopy theory provides a way to measure the twisting of the fibers of a fiber bundle (the integer associated with the homotopy classes indicates the degree of twisting of the fiber). How many different principal bundles can be constructed for a given base manifold M and a group G can be found using homotopy theory. Other tools, which we will see in chapter 5, to measure the non-triviality of the twisting of a fiber bundle are characteristic classes. In this sense the Chern classes are obstructions which prevent a bundle from a trivial bundle.

Let $\pi : E \to M$ be a vector bundle over M. A vector bundle $\pi : F \to M$ is called a subbundle if F is a submanifold of E that, for each point $p \in M$, the fiber F_p is a vector subspace of the fiber E_p of E.

3.10 Magnetic monopole

We saw in chapter 1 that the vector potential of a magnetic monopole could not be globally defined on the entire surface of a sphere and should be singular at one point of the surface. As such the sphere cannot be covered by a global choice of a coordinate system. Here we show how we can treat the problem using a fiber bundle.

The magnetic field of a magnetic charge g at the origin of a coordinate system in $\mathfrak{R}_3$ is given by

$$\vec{B} = g\vec{r}/2r^3, \tag{3.57}$$

which gives $\vec{\nabla}.\ \vec{B} = 4\pi g\delta(r)$.

The flux of the magnetic field through a spherical surface S around the magnetic charge is given by, using Stoke's theorem

$$\int_S \vec{B} \cdot d\vec{\sigma} = \int_S \vec{\nabla} \times \vec{A} \cdot d\vec{\sigma} = \int_\lambda \vec{A} \cdot d\vec{l} = \varphi(S_1) - \varphi(S_2) = 0, \tag{3.58}$$

where γ is a closed curve in S, which we can take along the equator. S_1 is the top hemisphere and S_2 the lower hemisphere. We see that the magnetic flux ϕ through a closed surface around the magnetic monopole is zero. If the field of a magnetic monopole is to be given by a potential we have to do some modifications in the above procedure. Since the vector potential A is not globally defined, we use a non-trivial fiber bundle. We introduce two open covertures $U_\pm$ covering the regions $z > -\varepsilon$ and $z < +\varepsilon$ of $\Re^3 - \{0\}$ such that in $U_+ \cap U_-$ the potentials are related by a gauge transformation, which can be used to define the transition functions. In fact, $U_+ \cap U_-$ is equal to the xy plane in $z = 0$ without the origin. We have for the potential:

$$\begin{aligned} A_+ &= \frac{1}{2}(1 - \cos\theta)d\varphi \ \text{ in } \ U_+ \\ A_- &= \frac{1}{2}(-1 - \cos\theta)d\varphi \ \text{ in } \ U_- \end{aligned} \tag{3.59}$$

or

$$A_\pm = \frac{1}{2r}\frac{1}{z \pm r}(xdx - ydy), \tag{3.60}$$

where $r^2 = x^2 + y^2 + z^2$. A_+ and A_- have singularities at $\theta = \pi$ and $\theta = 0$, respectively. In the superposition region $\theta = \pi/2$, $r > 0$, both potentials are regulars. A_+ and A_- are related by the gauge transformation

$$A_+ = A_- + d[\tan^{-1}(y/x)] = A_- + d\varphi, \tag{3.61}$$

The magnetic field in $U_\pm$ is given by

$$F = \frac{1}{2r_3}(xdy \wedge dz + ydz \wedge dx + zdx \wedge dy). \tag{3.62}$$

In Dirac's formulation of the problem, coordinate patches were not used and this led to the appearance of a fictitious 'string singularity' on the $\pm z$ axis. In the differential geometry approach, $A_\pm$ are defined only in their respective coordinate patches $U_\pm$ and we avoid the problem of the singularity.

3.11 Tangent bundle

Let X be a differentiable manifold of dimension n. We denote by TX the space of pairs (x, v_x) for all $x \in X$ and $v_x \in T_xX$. This is

$$TX = \bigcup_{x \in X} T_x X. \tag{3.63}$$

TX is a vector fiber bundle called the tangent bundle (figure 3.12). The total space is TX, the base space is the manifold X and the fiber over x is T_xX. The projection is given by $\pi(p) = x$, where $p \in TX$, $p = (x, v_x)$, $x \in X$. The typical fiber is the space $\mathfrak{R}^n$.

The structure group G for the tangent bundle is the group GL(n, $\mathfrak{R}$) of isomorphism of $\mathfrak{R}^n$ into itself, whose matrix representation is the set of real invertible $n \times n$ matrices.

The tangent bundle can be divided in two subbundles. The *vertical bundle* consists of all vectors that are tangent to the fibers, while the *horizontal bundle* is then a particular choice of a subbundle of the tangent bundle which is complementary to the vertical bundle. Or in other words, if π: $E \to X$ is a smooth fiber bundle over a smooth manifold X and $e \in E$ with $\pi(e) = x \in X$, then the vertical space V_eE at e is the tangent space $T_e(E_x)$ to the fiber E_x containing e. That is, $V_eE = T_e(E_{\pi(e)})$. The horizontal space H_eE is then a choice of a subspace T_eE such that $T_eE = V_eE \oplus H_eE$.

3.12 Vector field

A cross section (or section) of the fiber bundle (E, π, X) is a smooth map S: $X \to E$ such that the image of each point $x \in X$ lies in the fiber $\pi^{-1}(x)$ over x (figure 3.13). That is: π. S = identity. (Or in other words, a cross section is a rule which assigns a preferred point $S(x)$ on each fiber to each point of the base manifold.) We call this a 'lift' of the base space X into the bundle E. For a trivial bundle the cross section can be interpreted simply as the continuous functions on the base space X which takes values in the space E.

A vector field v in a manifold X is a cross section of the tangent bundle TX. So, a vector field associates with each point $x \in X$ a tangent vector $v_x \in T_xX$.

A vector field v on an open subset U of X gives rise to a vector field, denoted v_U on any subset U of X. The vector field v_U is said to be the restriction of v to U. A tensor field of the type $\left[{}^{s}_{r}\right]$ in a manifold X is an association to each $x \in X$ of a tensor

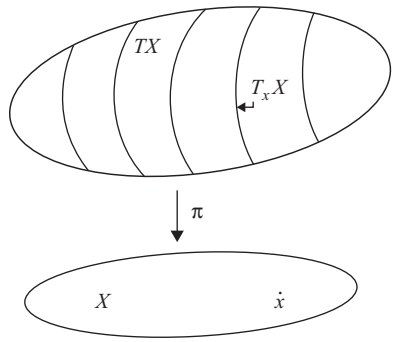


Figure 3.12. Tangent bundle.

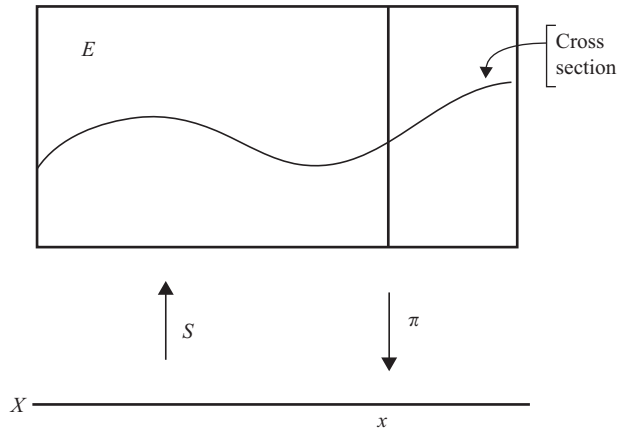


Figure 3.13. A cross section of a product bundle E over X is a continuous image of X in E which meets each individual fiber in a single point.

$T(x) \in (TX)^s_r$ such that if (U, ϕ) is a map, we have the following transformation of U into $\Re$

$$T^{n_1 \ldots n_s}_{m_1 \ldots m_r}(x) = T(x)\left(\frac{\partial}{\partial x^{m_1}}, \ldots, \frac{\partial}{\partial x^{m_r}}, dx^{n_1}, \ldots, dx^{n_s}\right). \tag{3.64}$$

For instance, if

$$T = T^k_{ij} dx^i \otimes dx^j \otimes \frac{\partial}{\partial x^k}, \tag{3.65}$$

in coordinates $\{x^i\}$ and

$$T = \tilde{T}^k_{ij} d\tilde{x}^i \otimes d\tilde{x}^j \otimes \frac{\partial}{\partial \tilde{x}^k}, \tag{3.66}$$

in coordinates $\{\tilde{x}^i\}$, we have

$$\begin{aligned} \tilde{T}^k_{ij} &= T\left(\frac{\partial}{\partial \tilde{x}^i}, \frac{\partial}{\partial \tilde{x}^j}, d\tilde{x}^k\right) = T\left(\frac{\partial n^n}{\partial \tilde{x}^i}\frac{\partial}{\partial x^n}, \frac{\partial x^m}{\partial \tilde{x}^j}\frac{\partial}{\partial x^m} - \frac{\partial \tilde{x}^k}{\partial x^p} dx^p\right) \\ &= \frac{\partial x^n}{\partial \tilde{x}^i}\frac{\partial x^m}{\partial \tilde{x}^j}\frac{\partial \tilde{x}^k}{\partial x^p} T^p_{nm} \end{aligned}. \tag{3.67}$$

Then the transformation law for the tensor components is

$$\tilde{T}^k_{ij} = \frac{\partial x^n}{\partial \tilde{x}^i}\frac{\partial x^m}{\partial \tilde{x}^j}\frac{\partial \tilde{x}^k}{\partial x^p} T^p_{nm}. \tag{3.68}$$

This equation can be generalized for other tensor fields.

In Riemannian geometry the vector spaces include the tangent space, the cotangent space, and the higher tensor spaces constructed from these. These spaces are intimately associated with the manifold itself, and are naturally defined once the

manifold is set up; the tangent space, as we have seen, is the space of directional derivatives at a point. This is what is of interest in general relativity, for instance, where the concept of fiber bundles is generally omitted. In gauge theories, on the other hand, we are concerned with 'internal' vector spaces. An internal vector space can be of any dimension we like, and has to be defined as an independent addition to the manifold. For instance, we can introduce an internal three-dimensional vector space, and sew the fibers together with ordinary rotations; the structure group of the fiber bundle is then SO(3). A field that lives in this bundle might be denoted by $\phi^i(x^\mu)$, where i runs from one to three; it is a three-vector (an internal one) for each point on the manifold. 'Physical quantities' should be left invariant under local SO (3) transformations. Such transformations are known as *gauge transformations.* In chapter 5, I will treat the gauge fields.

References and further reading

Altland A and Simons B 2010 *Condensed Matter Field Theory* (Cambridge: Cambridge University Press)

Carrol S M 2004 *An Introduction to General Relativity Spacetime and Geometry* (San Francisco: Addison Wesley)

Chouquet-Bruhat Y, De Witt-Morette C and Dillard-Blick M 1982 *Analysis, Manifolds and Physics* (Amsterdam: North-Holland)

Collinicci A and Wyns A 2006 *Topology of fiber bundles and global aspects of gauge theory* (arXiv: hep-th/0611201).

Curtis W D and Miller F R 1985 *Differential Manifolds and Theoretical Physics* (Orlando, FL: Academic)

Duffel P 2017 *Geometric Concepts in Physics* //duffel.org/chapter_4.pdf.

Eguchi T, Gilkey P B and Hanson A J 1980 Gravitation, gauge theories and differential geometry *Phys. Rep.* **66** 213

Frucart M and Carpentier D 2013 An introduction to topological insulators *Comptes Rendus Phys.* **14** 779–815

Isham C J 1999 *Modern Differential Geometry for Physicists* (Singapore: World Scientific)

Lee J 2003 *Introduction to Smooth Manifolds* (New York: Springer)

Manton N and Sutcliffe P 2004 *Topological Solitons* (London: Cambridge University Press)

Mathews J and Walker R L 1964 *Mathematical Methods for Physics* (Menlo Park, CA: Benjamin)

Mombelli S 2018 *Homotopy theory and characteristic classes* https://www.ethz.ch/content/dam/ethz/.../o2_Mombelli.pdf

Nash C and Sen S 1988 *Topology and Geometry for Physicists* (San Diego: Academic)

Penrose R 2007 *The Road to Reality* (New York: Vintage Books)

Warner F W 1983 *Foundations of Differentiable Manifolds and Lie Groups* (Heidelberg: Springer)

IOP Concise Physics

A Brief Introduction to Topology and Differential Geometry in Condensed Matter Physics

Antonio Sergio Teixeira Pires

Chapter 4

Metric and curvature

4.1 Metric in a vector space

Let V be a vector space of dimension n. Let g be a tensor $g \in V_2^0$, i.e. the space of all bilinear applications $V \times V \to \Re$. Thus $V_2^0 = V^* \otimes V^*$. If $\{e_n\}$ is a basis for V and if v and w lie in V, we can write $v = v^i e_i$, $w = w^j e_j$. Therefore, we have the expression

$$g(v, w) = g(v^i e_i, w^j e_j) = v^i w^j g(e_i, e_j) = g_{ij} v^i w^j, \tag{4.1}$$

where $g_{ij} = g(e_i, e_j)$ are the components of g relative to the base $\{e_n\}$.

The tensor $g \in V_2^0$ is called the *metric*, or *metric tensor*, if g is a symmetric and non-degenerate tensor. (Remember that non-degenerate means that $g(v, w) = 0$ for every $v \in V$ if and only if $w = 0$.) The tensor g gives to the vector space V a scalar product defined by $\langle v|\ w\rangle = g(v, w)$ for all vectors v and w lying in V.

If g is a metric in V, it can be shown (Curtis and Miller 1985) that there exists a base $\{e_n\}$ for V such that $g(e_i, e_j) = 0$ for $i \neq j$ and $g(e_i, e_j) = \pm 1$ for each i. Such a base is called the orthonormal basis for g. It is also usual to use the notation ds^2.

4.2 Metric in manifolds

A metric g on a manifold X is an association to each $x \in X$ of a metric $g(x)$ in $T_x X$, and such that if (U, φ) is a map in X and if we define g_{ij}: $U \to \Re$ by

$$g_{ij}(x) = g(x)\left(\left.\frac{\partial}{\partial x^i}\right|_x, \left.\frac{\partial}{\partial x^j}\right|_x\right), \tag{4.2}$$

g_{ij} should be C^∞ for all i and j. Note that the metric is defined in the tangent space and not in the manifold. A manifold X with a metric g is called a pseudo-Riemannian manifold. If the metric g is positive definite (that is, if for every

doi:10.1088/2053-2571/aaec8fch4

$x \in X$ we have $g(x)\ (v, v) > 0$ for all $v \in T_x X$ and $v \neq 0$) the manifold is called Riemannian. It can be shown that in every manifold X we can construct a definite positive metric (Curtis and Miller 1985). The index of a metric g is the number of vectors on an orthonormal basis for which $g(e_i, e_j) = -1$. If the metric is Riemannian the index is zero. The metric is called the Lorentz metric if its index is $n - 1$.

If $\{dx^i(x)\}$ is a basis on $T_x^* X$, $g_{ij}(x)$ is the matrix of the components of $g(x)$ with respect to that basis. Then we can write in U

$$g = g_{ij} dx^i \otimes dx^j. \tag{4.3}$$

Here I introduce the summation convention in which indices appearing repeated are summed over. Since $g_{ij} = g_{ji}$ we can also write

$$g = \frac{1}{2} g_{ij} (dx^i \otimes dx^j + dx^j \otimes dx^i), \tag{4.4}$$

The transformation rule for a change of coordinates can be calculated easily:

$$g_{kl} dx^k \otimes dx^l = g_{kl} \left(\frac{\partial x^k}{\partial y^i} dy_i \right) \otimes \left(\frac{\partial x^l}{\partial y^j} dy^j \right) = g'_{ij} dy^i \otimes dy^j, \tag{4.5}$$

and so

$$g'_{ij} = g_{kl} \frac{\partial x^k}{\partial y^i} \frac{\partial x^l}{\partial y^j}. \tag{4.6}$$

On a Riemannian manifold X with dimension n, a canonical n-form with coordinate-invariant definition exists. It is given by

$$\omega = \sqrt{g}\, dx^1 \wedge \ldots \wedge dx^n, \tag{4.7}$$

where $g = \det\{g_{ij}\}$ is the determinant of the metric tensor. If the metric is non-degenerate the determinant does not vanish. It can be shown that ω has the same representation in every coordinate system. This n-form is called the volume element of the manifold.

Example. In polar coordinates (r, θ, ϕ) in $\Re^3$ the metric is given by

$$g = dr \otimes dr + r^2 d\theta \otimes d\theta + r^2 \sin^2 d\phi \otimes d\phi. \tag{4.8}$$

It is quite easy to see that $\det g = r^4 \sin^2 \theta$, and then

$$\omega = r_2 \sin \theta dr \wedge d\theta \wedge d\phi. \tag{4.9}$$

We can define the inverse metric as g^{ij} via $g^{ij} g_{jk} = g_{nk} g^{ni} = \delta_k^i$. The metric and its inverse may be used to raise and lower indices on tensors.

4.3 Symplectic manifold

Let V be a vector space of dimension p and ω a 2-form in V:

$$\omega\colon V \times V \to \Re. \tag{4.10}$$

We say that ω is non-degenerate if $\omega(v_1, v_2) = 0$ for all $v_2 \in V$ implies $v_1 = 0$. (Remember that $\omega(v, u) = -\omega(u, v)$.)

Definition 1. Let X be a differentiable manifold of even dimension. A symplectic structure in X is a closed, differentiable, non-degenerate 2-form in X. Thus $\omega \in T^*X$ and $d\omega = 0$. The pair (X, ω) is called a symplectic manifold.

The symplectic form plays a role analogous to that of the metric tensor in Riemannian geometry. Where the metric tensor measures lengths, the symplectic form measures areas. However, there is no notion of curvature for a symplectic manifold, which might serve to distinguish one symplectic manifold from another, locally. Symplectic manifolds play an important role in classical mechanics.

Example. The space $\Re^{2n}$ with coordinates (x^i, x^j) and a 2-form $\omega = \sum_i dx^i \wedge dx^j$ is a symplectic manifold.

4.4 Exterior derivative

Let U be an open set in a manifold X of dimension n such that U is the domain of a map. The set of all smooth forms of degree p on U will be denoted by $\Lambda^p(U)$. In particular, $\Lambda^p(U)$ is the set of all smooth functions in U. It can be shown (Curtis and Miller 1985) that there exists an operator

$$d\colon \Lambda^p(U) \to \Lambda^{p+1}(U), \tag{4.11}$$

that takes each p-form ω into a $(p = 1)$-form $d\omega$, with the following properties

(1) $d(\omega + \eta) = d\omega + d\eta$, $d(a\omega) = ad\omega$.
(2) $d(\alpha \wedge \omega) = d\alpha \wedge \omega + (-1)^p \alpha \wedge d\omega$, where $\alpha \in \Lambda^p U$ and p is the degree of α.
(3) For each ω, $d(d\omega) = 0$.
(4) For each function f, df is the ordinary derivative of f, $df = \frac{\partial f}{\partial x^i} dx^i$.

Example. Let $X = \Re^3$, then

(a) For a 0-form f, we have

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz.$$

(b) For a 1-form $\omega = Pdx + Qdy + Rdz$, we have

$$d\omega = \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)dy \wedge dz + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)dz \wedge dx + \left(\frac{\partial Q}{\partial P} - \frac{\partial P}{\partial y}\right)dx \wedge dy.$$

(c) For a 2-form $\alpha = Ady \wedge dz + Bdz \wedge dx + Cdx \wedge dy$, we have

$$d\alpha = \left(\frac{\partial A}{\partial x} + \frac{\partial B}{\partial y} + \frac{\partial C}{\partial z}\right)dx \wedge dy \wedge dz.$$

4.5 The Hodge star operator

Let V be a vector space of dimension n with a basis $\{e_n\}$. As we have seen a p-form in V (with $p \leqslant n$) is an anti-symmetric element in Λ^p. This is a vector space of dimension $n![p!(n-p)!]^{-1}$. We have for the basis in Λ^p, and respective dimensions:

Λ^0	1	dim = 1
Λ^1	$e^1, \ldots, e^n$	dim = n
Λ^2	$e^i \wedge e^j$	dim = $n(n\text{-}1)/2!$
Λ^3	$e^i \wedge e^j \wedge e^k$	dim = $n(n\text{ -}1)(n\text{ -}2)/3!$
.		
.		
.		
Λ^{n-1}	$e^{i_1} \wedge \ldots \wedge e^{i_{n-1}}$	dim = n
Λ^n	$e^{i_1} \wedge \ldots \wedge e^{i_n}$	dim = 1

Thus we see that Λ^p and Λ^{n-p} have the same dimensions as vector spaces. We see also that $\Lambda^p = 0$, for $p > n$. The wedge product may be used to make $(p + q)$ forms out of a given p-form and a given q-form. But since one gets zero for $p + n > n$, the resulting forms always belong to the original set of spaces.

The number of independent forms in Λ^p is the same as in Λ^{n-p}: there exists a duality between the two spaces. We can then introduce an operator, called the Hodge*, which transforms p-forms into $(n - p)$ forms.

If $\beta \in \Lambda^p$, we have

$$\beta = \beta_{|j_1 \ldots j_p|} e^{j_1} \wedge \ldots \wedge e^{j_p}. \tag{4.12}$$

We define ${}^*\beta \in \Lambda^{n-p}$, written as

$${}^*\beta = {}^*\beta_{|s_1 \ldots s_{n-p}|} e^{s_1} \wedge \ldots \wedge e^{s_{n-p}}, \tag{4.13}$$

in the following way.

Let $v_1, \ldots, v_n$ be orthonormal oriented vectors in V. We define

$$\beta(v_1, \ldots, v_p) = {}^*\beta(v_{p+1}, \ldots, v_n) \tag{4.14}$$

From equations (4.12), (4.13) and (4.14) we get

$${}^*\beta_{s_1 \ldots s_{n-p}} = \beta^{|i_1 \ldots i_p|} \tau_{i_1 \ldots i_p s_1 \ldots s_{n-p}} \tag{4.15}$$

where τ is the volume element of the metric

$$\tau = \sqrt{|g|}\, v^1 \wedge \ldots \wedge v^n. \tag{4.16}$$

Suppose that in an orthonormal basis $\{\theta^n\}$ we have $\beta = \theta^1 \wedge \ldots \wedge \theta^p$, then ${}^*\beta = \theta^{p+1} \wedge \ldots \wedge \theta^n$. As the volume element is given by $\tau = \theta^1 \wedge \ldots \wedge \theta^n$ we can write $\tau = \beta \wedge {}^*\beta$. If α and β are p-forms in V we have

$$\tau \langle \alpha | \beta \rangle = \alpha \wedge {}^*\beta. \tag{4.17}$$

Example. If $\{e^i\}$ is the standard dual base in E^3 we have:

$${}^*e^1 = e^2 \wedge e^3, \quad {}^*e^2 = e^3 \wedge e^1, \quad {}^*e^3 = e^1 \wedge e^2.$$

4.6 The pull-back of a one-form

While the push-forward transfer objects(vectors) along a map in the same direction as the map points, the pull-back works in the opposite direction and transfers objects from the target manifold to the source manifold.

Let X be a differentiable manifold of dimension n and Y a differentiable manifold of dimension m. Let f: $X \to Y$ be a differentiable mapping of X into Y. If $x \in X$ and $y \in Y$, we write $y = f(x)$. Consider a differential p-form ω in Y and $v_1, \ldots, v_p$ tangent vectors belonging to $T_x X$. We know that if $v \in T_x X$ then $f_* v \in T_{f(x)} Y$ (where $f_* = Df$ is the push-forward of f). The mapping f induces, from ω, a p-form in X, written $f^*\omega$ [$f^*\omega \in \Lambda^p(X)$] and called the *pull-back* of ω, defined by

$$(f^*\omega)(v_1, \ldots, v_p) = \omega(f_* v_1, \ldots, f_* v_p), \tag{4.18}$$

that is, the value of the form $f^*\omega$ in vectors $v_1, \ldots, v_p$ is equal to the value of the form ω in the images of these vectors.

In local coordinates f is given by $y^j(x^1, \ldots, x^n)$ with $j = 1, \ldots, n$. Writing

$$\omega = \omega_{|r_1 \ldots r_p|} dy^{r_1} \wedge \ldots \wedge dy^{r_p}, \quad r_i = 1, \ldots, n \tag{4.19}$$

we have

$$f^*\omega = \omega_{|r_1 \ldots r_p|}(y(x) dy^{r_1} \wedge \ldots \wedge dy^{r_p}) = (f^*\omega)_{|i_1 \ldots i_p|} dx^{i_1} \wedge \ldots \wedge dx^{i_p}, \tag{4.20}$$

where $i_j = 1, \ldots, n$ and

$$(f^*\omega)_{i_1\ldots i_p} = \frac{\partial y^{r_1}}{\partial x^{i_1}}\ldots\frac{\partial y^{r_p}}{\partial x^{i_p}}\omega_{r_1\ldots r_p}. \tag{4.21}$$

Example. If $X = \Re^2$ and $Y = \Re$, let f: $\Re^2 \to \Re$ be given by $f(x^1, x^2) = (x^1)^2 + (x^2)^2$. Let us take the form $\omega = dy$ in $\Re$ as $y = f(x)$. We have $y = (x^1)^2 + (x^2)^2$ and $f^*\omega = 2x^1dx^1 + 2x^2dx^2$.

4.7 Orientation of a manifold

As we saw in section 2.11, for $x \in X$ we can choose an orientation for a vector space and in particular for the tangent space T_xX. We call an orientation of X a choice of orientation for T_xX for each $x \in X$, such that for $x_0 \in X$ there exists a neighborhood U of x_0 and a continuous vector field $v_1, \ldots, v_n$ in U such that for every $x \in U$, $(v_1(x), \ldots, v_n(x))$ has the same orientation chosen above. A manifold for which an orientation chosen in this way can be constructed is called orientable. If a manifold is orientable a reference frame transported along any trajectory in the tangent bundle of the manifold returns to the starting point with the same orientation. For example, the sphere S^2 is orientable, but the Mobius strip is not.

A boundary of a manifold X, if there is one, written as ∂X, is given by

$$\partial X = \{x \in X| \text{ there is a map } (U, \varphi) \text{ in } x \text{ such that } \varphi(U) \text{ is open in } H^n \text{ and } \varphi(\mathrm{x}) \in \partial \mathrm{H_n}\}$$

where $H^n = \{x \in |\Re^n|, x^n \geqslant 0\}$ and $\partial H^n = \{x \in |\Re^n|, x^n = 0\}$. If $\partial X = 0$ the manifold is closed.

A 0-manifold is a set of discrete points, and the orientation is given by assigning a (+) or a (−) sign to each point. For a curve (1-manifold) the orientation provides a direction along the curve. The boundary of a finite line segment consists of its two end points (one with a + sign and the other with a − sign). For a 2d-surface the orientation is given by a small circular arc with an arrow on it, indicating the positive direction (for example, clockwise direction). The boundary of a closed unit disc is the unit circle. The boundary of a circle is empty. The boundary of a ball is the sphere and the boundary of a spherical surface is empty. The orientation of a 3-manifold is given by a triad of independent vectors specified as 'right hand'.

4.8 Integration on manifolds

I start this section with a summary of some basic results. The *support* of a function f is the smallest closed set outside which f is identically zero. Suppose a continuous function with compact support f: $\Re^n \to \Re$. Then $\int f dx^1 \ldots dx^n$ is defined as the Riemann integral on any rectangle containing the support of f.

Now let X be an oriented manifold of dimension n and ω an n-form in X with compact support. We also assume that the sets of interest are compact sets. These assumptions are made only for the sake of mathematical rigor. For readers

unfamiliar with these terms, it suffices for them to be aware that in making a statement I will assume the existence of the integral.

I will first define the integral of the form ω in the domain U of a map in X with coordinates $(x^1, \ldots, x^n)$. Suppose that ω is zero outside a set contained in U. By definition, ω is integrable in X if its components $\omega_{1\ldots n}$ are integrable in $\Re^n$. The integral of ω in U, which is also the integral in X, is then:

$$\int_X \omega = \int_U \omega = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \omega_{1\ldots n} dx^1 \ldots dx^n. \tag{4.22}$$

This definition does not depend on the choice of the coordinates in U, as long as they are consistent with the chosen orientation.

Let X and Y be manifolds of dimensions n and f: $X \to Y$ a diffeomorphism preserving the orientation. Let ω be an n-form in Y. Using what was presented before we have

$$\int_Y \omega = \int_X f^* \omega. \tag{4.23}$$

This property is equivalent to the substitution of variables in the integral in $\Re^n$.

Note. Even in $\Re^n$ the use of the external product notation has some advantages. For example, if we use $dx \wedge dy$ the orientation of the integral is implied and if we make a change of variables $x = x(\eta, \xi)$, $y = y(\eta, \xi)$ we get

$$dx = \frac{\partial x}{\partial \eta} d\eta + \frac{\partial x}{\partial \xi} d\xi, \quad dy = \frac{\partial y}{\partial \eta} d\eta + \frac{\partial y}{\partial \xi} d\xi,$$

which leads to

$$dx \wedge dy = \left(\frac{\partial x}{\partial \eta} \frac{\partial y}{\partial \xi} - \frac{\partial x}{\partial \xi} \frac{\partial y}{\partial \eta} \right) d\eta \wedge d\xi,$$

and we get the Jacobian.

Let us now generalize the definition for integrals of p-forms into p-chains in a manifold X of dimension n. First we define a chain.

A p-rectangle D in space $\Re^p$ is a naturally oriented subset defined by

$$a^i \leqslant x^i \leqslant b_i, \; i = 1, \ldots, p. \tag{4.24}$$

The role of the 'integration path' is played by what we call the elementary p-chain σ in X, which is the triplet $\sigma = (D, f, \mathrm{Or})$ consisting of the following elements (figure 4.1):

(1) A rectangle D in $\Re^p$.
(2) A differentiable mapping f: $U \to X$, where $U \subset \Re^p$ and $D \subset U$. The domain of f, that is U, is an open neighborhood of D, such that f can be applied to the boundary of D.
(3) The orientation of $\Re^p$, symbolized by Or.

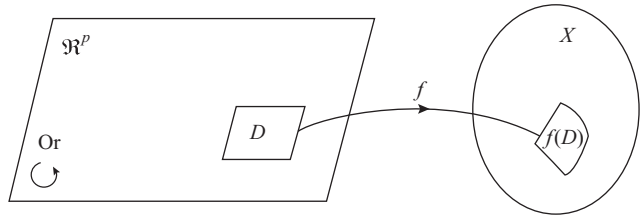


Figure 4.1.

The elementary p-chain σ is called an elementary p-domain of integration if f is a diffeomorphism of U into a differentiable submanifold X of dimension p. If ω is a p-form in X the integral of ω on the elementary p-domain of integration σ is, by definition, the integral of the corresponding form by the rectangle D

$$\int_\sigma \omega = \int_D f^* \omega. \tag{4.25}$$

Note that the set $f(D)$ is not necessarily a smooth submanifold of X (it may have mutual intersections or folds, for example).

We call a p-chain C in the manifold X a formal linear combination of p-elementary chains

$$C = \sum_i m_i \sigma_i, \tag{4.26}$$

where the coefficients m_i are integers. If the linear combination is locally finite, with $m_i = \pm 1$, and the σ_i are elementary integration domains, C is called an integration domain. Thus the integral of a p-form ω by the p-chain C, if it exists, is given by

$$\int_C \omega = \sum_i m_i \int_{\sigma_i} \omega. \tag{4.27}$$

Example. Given the form $\omega = xdy \wedge dz + ydx \wedge dy$ and the surface $x = u = v$, $y = u - v$, $z = uv$, with $0 \leqslant u \leqslant 1$, $0 \leqslant v \leqslant 1$, we want to calculate $\int_S \omega$.

We have:

$$\begin{aligned} f^*\omega &= (u+v)d(u-v) \wedge d(uv) + (u-v)d(u+v) \wedge d(u-v) \\ &= (u^2 - 2u + v^2 + 2v + 2uv)du \wedge dv \int_S \omega \\ &= \int_D f^*\omega = \int_0^1 \int_0^1 (u^2 - 2u + v^2 + 2v + 2uv)dudv = 7/6. \end{aligned}$$

The boundary of an oriented k-rectangle D on $\Re^k$ is the $(k-1)$ chain ∂D on $\Re^k$ defined as

$$\partial D = \sum_i \sigma_i, \tag{4.28}$$

where the elementary chains σ_i are the $(k-1)$-dimensional faces of D with orientation inherited from the orientation of $\Re^k$. One can easily extend this definition to the definition of the boundary of a elementary chain $\partial\sigma$ on the manifold X and then to the boundary of a chain as

$$\partial C_k = \sum_i m_i \partial\sigma_i. \tag{4.29}$$

∂C_k is a $(k-1)$ chain on X. We define a 0-chain as a collection of points with multiplicities and the boundary of an oriented interval $A\vec{B}$ as $B-A$. The boundary of a point is empty. The boundary of the boundary of an elementary chain is zero and therefore $\partial(\partial C_k) = 0$.

4.9 Stokes' theorem

Let ω be a n-form, which is at least C^1, in a manifold X. Let C be a p-chain and ∂C the boundary of C oriented coherently with C. We have:

$$\int_C d\omega = \int_{\partial C} \omega. \tag{4.30}$$

Demonstration: I will give the demonstration for a rectangle (see figure 4.2). The extension to the case of a chain is immediate. I will take, for simplicity, a 2-rectangle. Let $\omega = a(x, y)dx + b(x, y)dy$, then

$$d\omega = \left(\frac{\partial b}{\partial x} - \frac{\partial a}{\partial y}\right) dx \wedge dy. \tag{4.31}$$

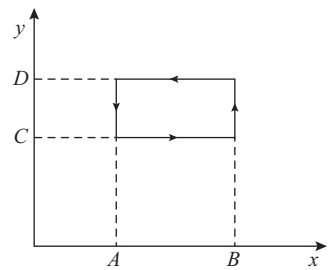


Figure 4.2.

Then

$$\begin{aligned}\iint_D d\omega &= \iint_D \frac{\partial b}{\partial x} dxdy - \iint_D \frac{\partial a}{\partial y} dxdy \\ &= \int_C^D [b(B, y) - b(A, y)]dy - \int_A^B [a(x, D) - a(x, C)]dx \\ &= \int_A^B a(x, C)dx + \int_C^D b(B, y)dy + \int_B^A a(x, D)dx + \int_D^C b(A, y)dy \\ &= \int_{\partial D} \omega,\end{aligned} \tag{4.32}$$

and the theorem is demonstrated.

Using equation (4.30) we have

$$\int_C d^2\omega = \int_{\partial C} d\omega = \int_{\partial^2 C} \omega = 0, \tag{4.33}$$

because $d^2\omega = 0$. This means that the boundary of a region has no boundary.

***Example* 1.** Given the 2-form in $\Re^3 - \{0, 0, 0\}$

$$\omega = \frac{xdy \wedge dz - ydx \wedge dz + zdx \wedge dy}{r^3}$$

where $r^2 = x^2 + y^2 + z^2$, we want to calculate the integral of ω over the sphere of radius R with center at the origin. On the surface of the sphere we can write $x = R \sin\theta \cos\varphi$, $y = R \sin\theta \sin\varphi$, $z = R \cos\theta$, where $0 \leqslant \theta \leqslant \pi$ and $0 \leqslant \varphi \leqslant 2\pi$. On the surface of the sphere we find by a direct calculation

$$\begin{aligned}dx \wedge dy &= R^2 \cos\theta \sin\theta d\theta \wedge d\varphi, \\ dx \wedge dz &= -R^2 \sin^2\theta \sin\varphi d\theta \wedge d\varphi, \\ dy \wedge dz &= R^2 \sin^2\theta \cos\varphi d\theta \wedge d\varphi,\end{aligned}$$

which leads to $\omega = \sin\theta d\theta \wedge d\varphi$. The integral is then given by

$$\int_{S^2} \omega = \int_{S^2} \sin\theta d\theta \wedge d\varphi = \int_0^{2\pi} d\varphi \int_0^{\pi} \sin\theta d\theta = 4\pi.$$

***Example* 2.** Given the 2-form $\omega = \frac{1}{2}(xdy \wedge dz + ydz \wedge dx + zdx \wedge dy)$ on $\Re^3$ we want to calculate the integral of ω over the sphere: $x^2 + y^2 + z^2 = R^2$ with an orientation given by the outward normal. We have $d\omega = dx \wedge dy \wedge dz$.

The sphere is the boundary of the ball B^3: $x^2 + y^2 + z^2 \leqslant R^2$. i.e. $S^2 = \partial B^3$. The form ω is defined on the whole B^3. Thus we can use Stokes' theorem:

$$\int_{S^2} \omega = \int_{B^3} dx \wedge dy \wedge dz = \frac{4}{3}\pi R^3.$$

4.10 Homology

As was mentioned in section 3.3, homotopy groups are quite difficult to calculate, even for the simplest structures. However, there are important relations between homotopy groups and homology of a topological space X, and homology can be used to distinguish topologically inequivalent manifolds.

A cycle on a manifold X is a chain whose boundary is equal to zero. A chain that is the boundary of another chain is called a boundary. All boundaries are cycles, however not all cycles are boundaries. For example, let us consider one of the cycles of a torus. Its boundary is zero, however it does not bound any chain on the torus. Let us consider two k-cycles a and b such that their difference is a boundary of a $(k+1)$ chain, i.e. $a - b = \partial C_{k+1}$. If $d\omega^k = 0$, we have from (4.33)

$$\int_a \omega^k = \int_b \omega^k, \tag{4.34}$$

that is, cycles can be replaced one by another. The quotient group

$$H_k(X) = \frac{(\text{cycles})}{(\text{boundaries})}, \tag{4.35}$$

is called the *kth homology group* of X. An element of this group is a class of cycles homologous to one another. We can think of representatives cycles in H_k as manifolds patched together to 'surround' a hole; we ignore cycles which can be 'filled in'.

There is a theorem known as *the Gurevich theorem* relating homotopy and homology groups that says:

If $\pi_k(X) = 0$ for all $k < n$, then $\pi_n(X) = H_n(X)$.

4.11 Cohomology

Let X be a m-dimensional manifold. A k-form $\omega \in \Lambda^k(X)$ is closed if $d\omega = 0$ and exact if $\omega = d\theta$ for some $\theta \in \Lambda^{k-1}(X)$. According to Poincaré's lemma any closed form is locally exact. The existence of locally but not globally exact closed forms is related to topological properties of X. A closed k-form is cohomologous to zero if it is exact. Closed k-forms ω and ω' on a manifold X are cohomologous if their difference is exact: $\omega - \omega' = d\theta$ for some $(k-1)$ smooth form θ on the whole X. This is an equivalence relation. Equivalence classes with respect to this relation are called cohomology classes. The set of all cohomology classes of degree k (also called dimension k) is denoted $H^k(X)$. $H^k(X)$ is a vector space for each $k = 0, 1, 2,..$ dual to $H_k(X)$. $H^k(X)$ is also an abelian group called the kth 'cohomology group'.

The set of closed k-forms is denoted by $Z^k(X)$ and the set of exact k-forms by $B^k(X)$. Since $d^2 = 0$, it follows that $Z^k(X) \supset B^k(X)$. $H^k(X)$ is also called the kth de Rahm cohomology group when written in the form

$$H^k(X) = \frac{Z^k(X)}{B^k(X)}. \tag{4.36}$$

The purpose of the de Rham cohomology is to classify the different types of closed forms on a manifold. There are other cohomology theories, but for manifolds they all give the same objects. A closed k-form is also said to be a k-cocycle and an exact k-form is said to be a k-coboundary for the de Rham cohomology.

Note that $\omega = d\theta$ implies $d\omega = 0$, but generally $d\omega = 0$ does not imply the existence of a θ such that $\omega = d\theta$. In $\Re^3$ all closed form are exact, but this is not true for general manifolds. For instance, in $U = \Re^2 - \{(0, 0)\}$ the form

$$\omega = -\frac{ydx + xdy}{x^2 + y^2}, \tag{4.37}$$

is closed, since $d\omega = 0$. However, although we can write $\omega = d[\arctan(y/x)]$, ω is not an exact form since $\arctan(y/x)$ is not differentiable in $\Re^2$ (integrating ω in a closed curve around the origin leads to a non-zero result). The dimension of H^k is called the kth Betti number of X and usually denoted $b_k(X)$.

To any smooth manifold X we can associate the graded vector space

$$H^*(X) = \bigoplus_{k \geqslant 0} H^k(X). \tag{4.38}$$

The Poincaré polynomial of X denoted by $P_X(t)$, is defined by

$$P_X(t) = \sum_k b_k(X) t^k. \tag{4.39}$$

The alternating sum

$$\chi(X) = \sum_k (-1)^k b_k(X), \tag{4.40}$$

is called the Euler characteristic of X. The Gauss–Bonnet theorem gives a formula for $\chi(X)$ in terms of curvature as we shall see later.

The de Rham cohomology of $\Re^n$ is trivial since any closed form can be expressed as the exterior derivative of a lower form in $\Re^n$. So $H^0(\Re^n) = 1$, dim $H^p(\Re^n) = 0$, $p > 0$. Non-trivial de Rham cohomology occurs only when the local coordinate neighborhoods are patched together in a globally non-trivial way. Only H^0 and H^n are non-zero for S^n and both have dimension 1 (note that in a circle S^1 the 1-form $d\theta$ is not exact since θ is not continuous or differentiable when considered as a function on the entire space S^1). H^0 consists of the space of constant functions and H^n consists of the constant multiples of the volume element. So we have

$$P_{S^n} = 1 + t^n. \tag{4.41}$$

For the torus $H^k(T^2) = \Re$ if $k = 0, 2$ and $\Re \oplus \Re$ if $k = 1$. If X is not connected but has n connected components, it is easy to see that dim $H^0(X)$ is equal to the number of connected components of X.

The k-forms form a cochain complex. In a cochain complex, the differential increases the index, while in a chain complex the boundary operator decreases the index. To be more explicit, if Λ^k is the vector space of k-forms, and C_k is the dual space of k-chains, then the action of the differential d and the boundary operator ∂ is as follows:

$$\rightarrow \Lambda^k \xrightarrow{d_{k+1}} \Lambda^{k+1} \xrightarrow{d_{k+2}} \Lambda^{k+2} \rightarrow$$

$$\leftarrow C_k \xleftarrow{\partial_{k+1}} C_{k+1} \xleftarrow{\partial_{k+2}} C_{k+2} \leftarrow.$$

Cochain complexes give cohomology and chain complexes give homology.

As we saw in equation (4.30), the Stokes' theorem relates the integral of a form over a boundary to the integral of the differential of the form over the interior of the boundary. In terms of the operator (ω, c) that evaluates the form ω on the chain c, we can write in a compact notation: $(\omega, \partial c) = (d\omega, c)$.

We can say that cohomology classes represent differential forms on a manifold and homology classes represent things we can integrate them over. We conclude that three types of invariant can be assigned to a topological space: homotopy, homology and cohomology groups

4.12 Degree of a map

Let $f\colon M \rightarrow N$ be a smooth map between two closed, connected and orientable manifolds M and N of the same dimension, and suppose that for $y \in N$, $x \in M$, with $f(x) = y$, the linear map $df_x : TM_x \rightarrow TN_y$ is non-singular, i.e. an isomorphism. If $\det(df_x) > 0$, we define sign $df_x = 1$, and if $\det(df_x) < 0$, let sign $df_x = -1$. We define the degree of f at y to be

$$\deg(f, y) = \sum_{x \in f^{-1}(y)} \text{sign } df_x. \tag{4.42}$$

Let M be a smooth oriented manifold of dimension m, and v a smooth vector field on M. If $v(x) = 0$ for some $x \in M$ we say that x is a zero of v. Now, suppose that x is an isolated zero of a vector field v, and we have a local coordinate system near x. Let D be a closed disk centered at x, so that x is the only zero of v in D. Then we define the index of x for v, written as $\text{index}_x(v)$, to be the degree of the map

$$u\colon \partial D \rightarrow S^{m-1}, \quad u(x) = \frac{v(x)}{|v(x)|}.$$

For instance, if M is a two-dimensional surface, the $\text{index}_x(v)$ is the winding number of the map $S^1 \rightarrow S^1$, that is, it is the change in the oriented angle $u(x)$ makes after

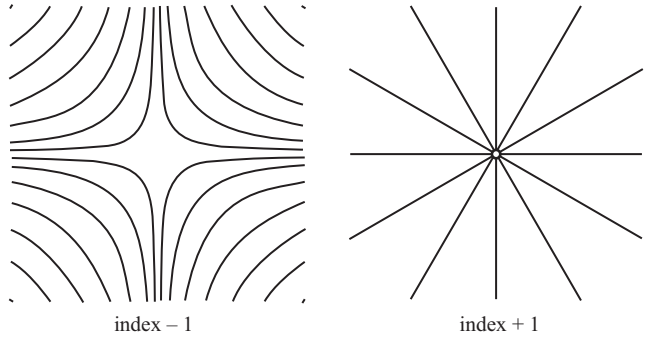


Figure 4.3.

traveling counterclockwise once around a small circle centered at x. Figure 4.3 shows a vector field in $\Re^2$ with index -1 and $+1$.

We can also define the degree of f as:

$$\deg(f) = \int_M f^*(\omega), \tag{4.43}$$

where ω is the volume form on N satisfying $\int_N \omega = 1$, and $f^*(\omega)$ is the pull-back of ω on M under f. In local coordinates, if

$$\omega = \beta(y)dy^1 \wedge dy^2 \wedge \ldots \wedge dy^d, \tag{4.44}$$

and f is represented by the function $y(x)$, then we have

$$\begin{aligned} f^*(\omega) &= \beta(y(x))\frac{\partial y^1}{\partial x_j}dx^j \wedge \frac{\partial y^2}{\partial x^k}dx^k \wedge \ldots \wedge \frac{\partial y^d}{\partial x^l}dx^l \\ &= \beta(y(x))\det\left(\frac{\partial y^i}{\partial x^j}\right)dx^1 \wedge dx^2 \wedge \ldots \wedge dx^d. \end{aligned} \tag{4.45}$$

It can be proved that equation (4.43) is identical to (4.42). The degree is always an integer, but may be positive or negative depending on the orientation. The topological degree is a homotopy invariant of f, because an integer cannot change under a continuous deformation. It is also independent of the choice of ω, because the difference of two normalized volume forms on N is a d-form whose integral is zero, and hence an exact form. The pull-back of the difference is therefore exact on M, and integrates to zero.

A simple example for the mapping f: $S^1 \to S^1$ was presented in equation (3.28) where the degree was equal to the winding number. Sometimes the degree of more general maps between higher-dimensional manifolds is called a winding number. We can think of the degree as the number of times that the domain manifold wraps around the range manifold under the mapping.

If X and Y are S^n. the degree of a map corresponds to the integer number associated with the homotopy class of the map.

4.13 Hopf–Poincaré theorem

The Hopf–Poincaré theorem can be stated as: Let M be a compact differential manifold and let v be a vector field on M with isolated zeros. Then we have

$$\sum_i \text{index}_{x_i}(v) = \chi(M), \tag{4.46}$$

where the sum of the indices is over all isolated zeros of v and $\chi(M)$ is the Euler characteristic of M.

For the sphere S^2, $\chi(M) = 2$, and so the total vorticity is 2. This is the well-known result that it is impossible to comb the hair on a 2-sphere without creating a vortex. In one combs the hair along the longitude (for instance), there are two +1 vortices at the north and south poles. That is, any continuous tangent vector field on the sphere must have a point where the vector is zero. The analogous statement for S^3 is not true.

4.14 Connection

In Euclidean space two vectors of different origins are compared to one another by the parallel translation of one or both of them to the same origin. On the other hand, in a curved space, such as the surface of a sphere, the translation is not well defined and the parallel transport depends on the trajectory along which the vector is displaced. Take as an example a vector located at a point on the equator. Move the vector by parallel transport following a meridian to the north pole, return again to the equator following another meridian and finally to the initial point along the equator. The displaced vector will not match the original vector (figure 4.4).

Two vectors can be compared in a natural way only if they are elements of the same tangent space. There is no standard way of generalizing these two related concepts, parallel transport and derivative transport, to vectors in an arbitrary differential manifold X. To speak of an equality of the components of two vectors, defined at different points p and q of the manifold X, we must have a means of assigning a coordinate frame defined in q uniquely given a frame in p, and to define a derivative we must somehow have a parallel transport of vectors. The definition of a connection makes possible the definition of parallel transport of a vector along a curve and the definition of the derivative of a vector, which in turn is a tensor. Before defining a connection we need to define a germ.

We say that two functions f and g defined on the manifold X, differentiable at the point $x \in X$ have the same germ at x, if there is a neighborhood of x where they coincide. The class of equivalence of functions differentiable at x, which have the same germ as a function f, is called the germ of f.

A linear *connection* in a smooth manifold X is a linear mapping $v \to \nabla v$ of the germs of vector fields in X into the germs of tensor fields of type $\binom{1}{1}$ in X such that

(1) $\nabla(v + u) = \nabla v + \nabla u$,
(2) $\nabla(fv) = df \otimes v + f\nabla v$,

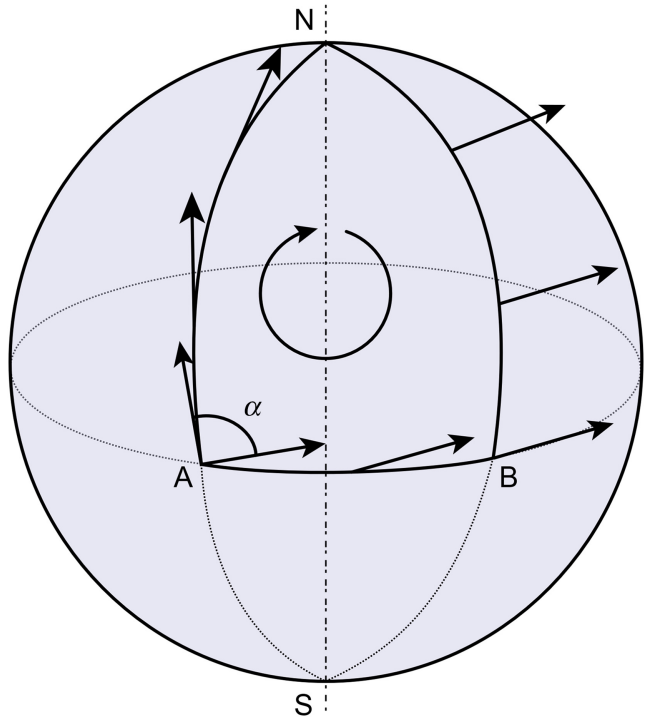


Figure 4.4. Parallel transport of a vector. Fred the Oyster / Wikimedia Commons / https://creativecommons.org/licenses/by-sa/4.0/deed.

where f is the germ of a differentiable function in X. The tensor ∇v is called the covariant derivative of v. The connection coefficients Γ^j_{ki} are defined by the relation

$$\nabla e_i = \Gamma^j_{ki}\theta^k \otimes e_j, \tag{4.47}$$

where $\{e_i\}$ and $\{\theta^i\}$ are dual basis. We have then

$$\nabla v = \nabla(v^i e_i) = dv^i \otimes e_i + v^i \nabla e_i = (dv^i + v^j \gamma^i_{kj}\theta^k) \otimes e_i. \tag{4.48}$$

The components of the tensor ∇v will be denoted by $\nabla_k v^i$ or $v^i_{;k}$. We have

$$\nabla v = \nabla_k v^i \theta^k \otimes e_i = v^i_{;k}\theta^k \otimes e_i. \tag{4.49}$$

We can also write in terms of $\omega_i{}^j = \gamma^j_{ki}\theta^k$

$$\nabla v = (dv^i + \omega_j{}^i v^j) \otimes e_i. \tag{4.50}$$

Using the natural basis

$$dv^i = \frac{\partial v^i}{\partial x^k}dx^k, \quad \theta^i = dx^i, \quad e_j = \frac{\partial}{\partial x^j}. \tag{4.51}$$

we can write

$$\nabla v = \left(\frac{\partial v_i}{\partial x^k} + \Gamma^i_{kn} v^n\right) dx^k \otimes \frac{\partial}{\partial x^i}, \tag{4.52}$$

where Γ^i_{kn} are the connection coefficients in the natural basis, also called Christofell symbols of second species.

Example. Let us consider the basis in spherical coordinates (e_r, e_θ, e_ϕ) written in terms of the Cartesian basis:

$$\begin{aligned} e_r &= \sin\theta\cos\phi e_x + \sin\theta\sin\phi e_y + \cos\theta e_z, \\ e_\theta &= r\cos\theta\cos\phi e_x + r\cos\theta\sin\phi e_y - r\sin\theta e_z, \\ e_\phi &= -r\sin\theta\sin\phi e_x + r\sin\theta\cos\phi e_y. \end{aligned}$$

Differentiating the coordinate basis and remembering that the Cartesian basis vectors are constant we find, for instance:

$$\frac{\partial e_r}{\partial r} = 0, \quad \frac{\partial e_r}{\partial \theta} = \cos\theta\cos\phi e_x + \cos\theta\sin\phi e_y - \sin\theta e_z = \frac{1}{r} e_\theta,$$
$$\frac{\partial e_r}{\partial \phi} = -\sin\theta\sin\phi e_x + \sin\theta\cos\phi e_y = \frac{1}{r} e_\phi, \quad \frac{\partial e_\phi}{\partial \theta} = \cot\theta e_\phi.$$

Calculating the other terms we find for the Γ coefficients

$$\Gamma^\theta_{r\theta} = \Gamma^\theta_{\theta r} = \frac{1}{r}, \quad \Gamma^r_{\theta\theta} = -r, \quad \Gamma^\phi_{r\phi} = \Gamma^\phi_{\phi r} = \frac{1}{r}, \quad \Gamma^\phi_{\phi\theta} = \Gamma^\phi_{\theta\phi} = \cot\theta$$
$$\Gamma^r_{\phi\phi} = -r\sin^2\theta, \quad \Gamma^\theta_{\phi\phi} = \sin\theta\cos\theta.$$

The other Γ vanish.

4.15 Covariant derivative

The covariant derivative $\nabla_u v$ of the vector v in the direction of the vector u is defined as

$$\nabla_u v = (\nabla v)(u), \quad \text{or} \quad \nabla_u v(.) = \nabla v(u, .). \tag{4.53}$$

We note that $\nabla_u v$ is linear in u, that is

$$\nabla_{fu+gw} v = f\nabla_u v + g\nabla_w v, \quad \text{where} \quad f, g: X \to \Re. \tag{4.54}$$

The covariant derivative of a vector in the basis along another vector in the basis is a vector and then can be written as a linear combination of the elements in the basis, that is

$$\nabla_{e_i} e_j = \Gamma^k_{ij} e_k. \tag{4.55}$$

For vector fields $v = v^i e_i$, $u = u^i e_i$ we have

$$\nabla_v u = \left(v^i u^j \Gamma^k_{ij} + v^i \frac{\partial u^k}{\partial x^i}\right) e_k. \tag{4.56}$$

The first term in the above equation is responsible for the 'rotation' of the coordinate frame in relation to the covariant derivative, and the second term is responsible for the component change of the vector field. In particular

$$\nabla_{e_j} u = \nabla_j u = \left(\frac{\partial u^i}{\partial x^j} + u^k \Gamma^i_{jk} \right) e_i. \tag{4.57}$$

We see that the covariant derivative is the standard derivative along the coordinates with correction terms that tell us how the coordinates change. Equation (4.57) is sometimes written as

$$\nabla_i u_k = \frac{\partial u^k}{\partial x^i} + \Gamma^k_{ij} u^j. \tag{4.58}$$

The covariant derivative of a covector α is given by

$$\nabla_v(\alpha(u)) = (\nabla_v \alpha)u + \alpha(\nabla_v u). \tag{4.59}$$

The concept of covariant derivative can be generalized for tensors of higher orders. For instance,

$$\nabla_i T^{jk} = \partial_i T^{jk} + \Gamma^j_{in} T^{nk} + \Gamma^k_{in} T^{jn}. \tag{4.60}$$

4.16 Curvature

Curvature measures the extent to which parallel transport is path dependent. The curvature tensor (also called Riemann tensor) in a connection ∇ in X is a $\binom{1}{3}$ tensor $T_x X \times T_x X \times T_x X \to T_x X$ for $x \in X$ defined by

$$R(u, v, w) = \nabla_u \nabla_v w - \nabla_v \nabla_u w - \nabla_{[u, v]} w, \tag{4.61}$$

where u, v, w are vector fields in X and $[u, v] = uv - vu$. In local coordinates we can write

$$R\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k} \right) = R^n_{ijk} \frac{\partial}{\partial n^n}, \tag{4.62}$$

with

$$R^n_{ijk} = \left(\frac{\partial \Gamma^n_{jk}}{\partial x^i} - \frac{\partial \Gamma^n_{ik}}{\partial x^j} \right) + (\Gamma^m_{jk} \Gamma^n_{im} - \Gamma^m_{ik} \Gamma^n_{jm}). \tag{4.63}$$

If $u = \partial/\partial x^i$, and $v = \partial/\partial x^j$ are coordinate vector fields then $[u, v] = 0$ and we can write equation (4.61) as $R(u, v, w) = [\nabla_u, \nabla_v]w$.

Given a local basis $\{e_i\}$ $1 \leqslant i \leqslant N$, we can define $N \times N$ matrices $\Gamma_j = \Gamma^n_{mj}$, and $\mathfrak{R}_{jk} = R^n_{mjk}$, and write equation (4.63) as

$$\mathfrak{R}_{jk} = \partial_j \Gamma_k - \partial_k \Gamma_j + [\Gamma_j, \Gamma_k]. \tag{4.64}$$

Note that $\mathfrak{R}_{jk}$ is anti-symmetric in j and k. Now we define a connection 1-form Γ and a curvature 2-form Ω by

$$\Gamma = \sum_j \Gamma_j dx_j, \quad \Omega = \frac{1}{2}\sum_{j,k} \mathfrak{R}_{jk} dx_j \wedge dx_k, \tag{4.65}$$

and equation (4.64) can be written in a shorter form as

$$\Omega = d\Gamma + \Gamma \wedge \Gamma. \tag{4.66}$$

Another important curvature tensor is defined by $R_{ijkl} = g_{in} R^{n}_{jkl}$, where g_{in} are the components of the metric tensor. The Ricci tensor is given by $R_{ij} = R^{k}_{ikj}$. The trace of the Ricci tensor is the curvature scalar $R = g^{ij} R_{ij}$. These tensors are important in the general relativity theory. In condensed matter we are more interested in two-dimensional surfaces in $\mathfrak{R}^3$. For a Riemannian manifold X of dimension 2, each component R_{ijkl} of the curvature tensor is either 0 or $\pm$ the quantity below

$$R_{1212} = R_{2121} = gk, \text{ where } g = \det(g_{jk}). \tag{4.67}$$

One calls k the Gauss curvature of X (when dim $X = 2$). We have also $k = R/2$. If the metric is given by $ds^2 = g_{ij} dx^i dx^j$ the Gauss curvature has the following expression as a function of the metric

$$\begin{aligned} k = \frac{1}{2g_{11}g_{22}}\Bigg[& -\frac{\partial^2 g_{11}}{\partial (x^2)^2} - \frac{\partial^2 g_{22}}{\partial (x^1)^2} + \frac{1}{2g_{11}}\left\{ \frac{\partial g_{11}}{\partial x^1}\frac{\partial g_{22}}{\partial x^1} + \left(\frac{\partial g_{11}}{\partial x^2}\right)^2 \right\} \\ & + \frac{1}{2g_{22}}\left\{ \frac{\partial g_{11}}{\partial x^2}\frac{\partial g_{22}}{\partial x^2} + \left(\frac{\partial g_{22}}{\partial x^1}\right)^2 \right\} \Bigg]. \end{aligned} \tag{4.68}$$

This is Gauss's *Theorema Egregium* (remarkable theorem).

The maximum and minimum values of the curvatures of a two-dimensional surface are called principal curvatures k_1 and k_2, respectively. The Gauss curvature is the product of the two principal curvatures. Another way to express the Gauss curvature in a simple way is as follows (Givental 2017, Moore 2014):

At a point in a curved two-dimensional surface, there is only one tangent plane, and consequently only one number is required to describe the Gauss curvature. We start with a vector at a point and then transport it parallel to itself around a closed loop. It arrives back at the starting point turned by some angle. We have

$$\text{Gaussian curvature} = \frac{\text{angle turned through}}{\text{area circunnavigated}}. \tag{4.69}$$

It can be shown that all definitions above are equivalent.

4.17 The Gauss–Bonnet theorem

Bending does not change the intrinsic geometry of a surface (that is, distances and angles), as can be seeing by drawing a curve in a cylinder and then unwrapping it in a plane. However, we cannot wrap a hemisphere with a sheet of paper. This is because the sphere is curved. Curvature prevents one surface being wrapped around another without stretching or wrinkling. It does not disappear under bending since it is an intrinsic property of geometry (Voronov 2011).

A smooth surface can be approximated with a polyhedron (i.e. a surface formed by flat polygons connected at their sides and vertices). The approximation can be made arbitrarily good by making the number of faces sufficiently great and the size of each face sufficiently small. We can study the geometric properties of surfaces by studying the geometric properties of polyhedra.

Polyhedron	V	E	F
Tetrahedron	4	6	4
Cube	8	12	6
Octahedron	6	12	8

The common use of triangles and the triangulation of a regular surface S (i.e. a surface differentiable in a neighborhood of each point) in two dimensions is defined as a finite collection of triangles $\{T_j\}_{j=1}^n$ such that $U_{j=1}^n T_j = S$ and the only possible intersections of T_i and T_j with $i \neq j$ is a common edge or a common vertex. Every regular surface S admits a triangulation.

If we cut out a piece of the surface of a polyhedra that contains no vertex, it is possible to unbend the piece along edges to make it lie flat on the plane. However, even a small piece around a vertex cannot be made to lie flat on the plane. We see that curvature is associated with vertices. For instance, let us consider part of a polyhedron made with four triangles. The triangles that belong to the vertex, laid out on a flat surface, fail to meet.

To be made flat, the corner near a vertex of a polyhedral must have the adjacent angles adding up to 2π. The difference 2π − sum of angles at a vertex is called the deficit angle of the vertex (let us call it *dif*). The deficit angle measures the amount of curvature concentrated at the vertex.

Let V, E, and F denote the number of vertices, edges and faces of a given polyhedron. Then we have that:

$$\textit{Total deficit of all vertices} = 2\pi \times V - \textit{sum of all angles at all vertices}.$$

The sum of all angles at all vertices is the same as the sum of all angles in all faces. Now we can see that the sum of angles in one face is equal to π multiplied by the number of sides of this face minus 2π. We arrive then at the following result

$$\textit{The sum of angles of all faces} = \pi \times \textit{total number of sides of all faces} - 2\pi \times F$$

Taking into account that each edge is a side to exactly two faces, we can rewrite this result as:

$$\textit{Sum of all angles at vertices} = 2\pi \times E - 2\pi \times F.$$

Using the first result we can get

$$\textit{The total angle deficit} = 2\pi \times (V - E - F).$$

The total angle deficit divided by 2π is called the Gauss number of the polyhedron P, and $\chi = V - E - F$ is the Euler number (or Euler characteristic) of the polyhedron P. The above result says that these two numbers are equal. In the table below I show χ for some polyhedra.

In all cases $\chi = 2$. χ is unchanged under continuous deformation of the surface, it remains unchanged so long as the topology remains the same.

The Euler number characterizes the combinatorics of P. Partitioning faces of a given polyhedron into smaller faces changes combinatorics of it, but it does not change the geometry. Therefore, the Gauss number does not change, and by the above result the Euler number remains the same too (for a more detailed discussion see Givental 2017). We arrive then at the following result, known as the Gauss–Bonnet theorem for polyhedra

The Gauss and Euler numbers of every polyhedron are equal to each other and depend only on the topology of the polyhedron.

If a polyhedron has the same topology as the sphere, then its Euler number is 2, and the total sum of its angle deficit is 4π.

Smooth surfaces can be approximated by polyhedra with lots of tiny faces and lots of vertices with very small angle deficits. The sums of angle deficits of those vertices that lie in a region of the surface characterize the curvature of this region. It is the Gaussian curvature of this region. By choosing better and better approximations of a surface and applying the Gauss–Bonnet theorem for polyhedra, we arrive at the following result:

The total Gauss curvature of a closed surface depends only on the topology of the surface and it is equal to 2π *times the Euler number of the surface.*

Note. The 2π appears due to the way Gauss measured the full angle.

All surfaces in two dimensions are classified up to homeomorphism by their genus, which is related to the Euler number. Essentially, the number of holes in a surface classifies it topologically.

The Gauss–Bonnet theorem bridges the gap between topology and differential geometry. Its importance lies in relating geometrical information of a surface to a purely topological characteristic. We can present the Euler–Bonnet theorem in its general form as:

If S is an orientable compact surface then

$$\int_S \kappa d\sigma = 2\pi\chi(S), \tag{4.70}$$

where κ is the Gaussian curvature, $\chi(S) = 2 - 2g$ and g is the genus of a surface, i.e. the number of holes in the surface. The total curvature depends exclusively on the topological characteristics, i.e. the genus. For instance, the sphere has $g = 0$ and then $\chi = 2$. For the torus $g = 1$, we have $\chi = 0$. Any two-dimensional closed orientable surface is classified by the number of handles, i.e. the genus g.

The Gauss–Bonnet theorem can be generalized to a smooth manifold of dimension $2k$, which is oriented and compact. However, for dimensions larger than 2, instead of the curvature we use $e(M)$, the Euler class associated to the tangent bundle of M. For any compact, orientable manifold of odd dimension, the Euler characteristic is always zero.

4.18 Surfaces

Considering that two-dimensional surfaces play an important role in condensed matter, we are going to discuss some details of them here.

Submanifolds of Riemannian manifolds inherit the Riemannian structure. The simplest example is provided by a surface X in $\Re^3$. A smooth surface in $\Re^3$ is a subset $X \subset \Re^3$ such that each point has a neighborhood $U \subset X$ and a map r: $V \to \Re^3$ from an open $V \subset \Re^3$ such that; (a) r: $V \to U$ is a homeomorphism, (b) $r(u, v) = (x(u, v), y(u, v), z(u, v))$ has derivatives of all orders. The induced Riemannian structure in X is called the *first quadratic form.* If the surface is given parametrically as $r(u, v) \in \Re^3$, then we have

$$g = Edu \otimes du + 2Fdu \otimes dv + Gdv \otimes dv. \tag{4.71}$$

The expression for the metric can be obtained from the Euclidian metric $dx \otimes dx + dy \otimes dy + dz \otimes dz$ by substituting the differentials of the coordinates functions $x(u, v)$, $y(u, v)$, $z(u, v)$ of the curve $r = (x, y, z)$ in terms of the parameters u, v on the surface.

***Example* 1.** The sphere can be parameterized as

$$r(\theta, \varphi) = (a \sin\theta \sin\varphi, a \cos\theta \sin\varphi, a \cos\theta), \text{ which leads to}$$
$$g = a^2 d\theta \otimes d\theta + a^2 \sin^2 d\varphi \otimes d\varphi.$$

***Example* 2.** For the cylinder one can consider the following parameterization of the surface:

$$r(h, \varphi) = \begin{cases} x = a \cos\varphi \\ y = a \sin\varphi \\ z = h \end{cases}$$

and we get $g = a^2 d\varphi \otimes d\varphi + dh \otimes dh$.

***Example* 3.** The saddle is given by the equation $z - xy = 0$, and we can use the parameterization: $x = u$, $y = v$, $z = uv$. We obtain

$$g = (1 + v^2)du \otimes du + 2uvdu \otimes dv + (1 + u^2)dv \otimes dv.$$

The area element σ is given by $\sigma = \pm\sqrt{|g|}\, du \wedge dv$, where $|g| = \det|g_{ij}|$. The sign $\pm$ is fixed by a choice of orientation of the surface. From equation (4.71) we have $\det g = EG - F^2$.

The quadratic form g can be represented as

$$g = u_1^2 + u_2^2, \tag{4.72}$$

where u_1 and u_2 are differential 1-forms. Locally, such reduction is always possible. For instance, if

$$g = dx^2 + 2x^2y^2dxdy + dy^2, \tag{4.73}$$

where by some error of notation we omit the symbol $\otimes$ in equation (4.73). We can write

$$g = (dx + x^2y^2dy)^2 + (1 - x^4y^4)dy^2, \tag{4.74}$$

and take $u_1 = dx + x^2y^2dy$, $u_2 = \sqrt{1 - x^4y^4}\, dy$.

We remark that $\sigma = u_1 \wedge u_2$ is the area form on X. There is no reason for u_1 and u_2 to be differentials of some functions, so that the differentials du^1 and du^2 could be non-vanishing. Now we define functions α_1 and as: $du_1 = \alpha_1\sigma$, $du_2 = \alpha_2\sigma$ and set $\theta = \alpha_1u_1 + \alpha_2u_2$. In the example above we have: $du_1 = 2xy^2dx \wedge dy$, $du_2 = (-2x^3y^4/\sqrt{1 - x^4y^4})dx \wedge dy$.

Differentiating, we find $d\theta = \kappa\sigma$, for some function κ. This function is called the curvature of the metric g. If the metric is the first quadratic form of a surface in $\mathfrak{R}^3$, the curvature κ coincides with the Gaussian curvature.

References and further reading

Curtis W D and Miller F R 1985 *Differential Manifolds and Theoretical Physics* (Orlando, FL: Academic)

Givental A 2017 Geometry of surfaces and the Gauss–Bonnet theorem http://math.berkeley.edu/~giventh/difgem.pdf

Moore J E 2014 Draft lectures V–VI notes for Les Houches 2014 https://pks.mpg.de/~topo14/slides/presentation_Moore_5+6.pdf

Voronov T 2011 Differential manifolds http://maths.manchester.ac.uk/~tv/manifolds.html

IOP Concise Physics

A Brief Introduction to Topology and Differential Geometry in Condensed Matter Physics

Antonio Sergio Teixeira Pires

Chapter 5

Dirac equation and gauge fields

5.1 The Dirac equation

Fermionic quasiparticles described by a pseudo-relativistic Dirac equation are common in condensed matter. In this section I will just present the general idea of Dirac, Weyl and Majorana fermions without going into detail.

The Dirac equation is a relativistic equation describing free particles of spin 1/2. It has the form

$$(i\gamma^{\mu}\partial_{\mu} - m)\psi = 0, \tag{5.1}$$

where m is the mass and $\psi(x)$ is a wave function with four components, and γ^{μ} are 4×4 matrices. In this section, I will use the standard particle physics units where $\hbar = c = 1$. We remember that in Minkowsky space–time for x^{μ} we have $x^0 = t$, $x^1 = x$, $x^2 = y$, $x^3 = z$, while for x_{μ}, $x_0 = x^0$, $x_1 = -x^1$, $x_2 = -x^2$, $x_3 = -x^3$.

A convenient representation of the γ^{μ} matrices is

$$\gamma^0 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}, \qquad \gamma^k = \begin{pmatrix} 0 & \sigma^k \\ -\sigma^k & 0 \end{pmatrix}, \qquad k = 1, 2, 3. \tag{5.2}$$

where $I = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ and σ^k are the Pauli matrices. The γ^{μ} matrices satisfy the anti-commutation relation: $\{\gamma^{\mu}, \gamma^{\nu}\} = 2g^{\mu\nu}$ (where $g^{\mu\nu} = \mathrm{diag}(+---)$). The fermion wave function can be written as

$$\psi(p) = e^{-ix_{\mu}p^{\mu}}\, u(p),$$

where $p_{\mu} = (E, -p_x, -p_y, -p_z)$ and $x^{\mu} = t, x, y, z$ and so $-ix_{\mu}p^{\mu} = -i(Et - \vec{x}.\,\vec{p}\,)$. For a particle at rest there are four independent solutions

doi:10.1088/2053-2571/aaec8fch5

$$u_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \qquad u_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \qquad u_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \qquad u_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \tag{5.3}$$

with eigenvalues m, m, $-m$ and $-m$, respectively. The first two solutions are interpreted as positive energy particle solutions with spin up and spin down. The solutions with negative energies can be interpreted (in a second quantization procedure) as describing antiparticles with positive energies. Then we have the following correspondence

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} \rightarrow \begin{pmatrix} \text{particle spin up} \\ \text{particle spin down} \\ \text{antiparticle spin up} \\ \text{antiparticle spin down} \end{pmatrix}. \tag{5.4}$$

To study a particle in motion we write the wave function as

$$\psi = \begin{pmatrix} \psi_A \\ \psi_B \end{pmatrix}, \quad \text{where } \psi_A = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}, \quad \text{and } \psi_B = \begin{pmatrix} \psi_3 \\ \psi_4 \end{pmatrix}. \tag{5.5}$$

Taking this into the Dirac equation and replacing ∂_μ by p_μ we obtain the following equation for $u(p)$

$$(\gamma^\mu p_\mu - m)u(p) = 0, \tag{5.6}$$

with $u = \begin{pmatrix} u_A \\ u_B \end{pmatrix}$.

Writing equation (5.6) in a matrix form we get

$$\begin{pmatrix} E - m & -\vec{\sigma}.\,\vec{p} \\ \vec{\sigma}.\,\vec{p} & -E - m \end{pmatrix} \begin{pmatrix} u_A \\ u_B \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{5.7}$$

where we have used

$$\gamma^\mu p_\mu = \gamma^0 p_0 - \vec{\gamma}.\,\vec{p} = p_0 \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} - \vec{p} \cdot \begin{pmatrix} 0 & \vec{\sigma} \\ -\vec{\sigma} & 0 \end{pmatrix} = \begin{pmatrix} p_0 & -\vec{p}\,.\,\vec{\sigma} \\ \vec{p}\,.\,\vec{\sigma} & -p_0 \end{pmatrix}. \tag{5.8}$$

From equation (5.8) we have

$$(\vec{\sigma}.\,\vec{p}\,)u_B = (E - m)u_A, \qquad (\vec{\sigma}.\,\vec{p}\,)u_A(E + m)u_B. \tag{5.9}$$

Explicit solutions for u_A and u_B can be found in Griffiths (2008).

Defining

$$\varphi_R = u_A + u_B, \qquad \text{and } \varphi_L = u_A - u_B, \tag{5.10}$$

we may rewrite equation (5.8) as

$$\begin{pmatrix} -m & E+\vec{\sigma}.\vec{p} \\ E-\vec{\sigma}.\vec{p} & -m \end{pmatrix}\begin{pmatrix} \phi_R \\ \phi_L \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{5.11}$$

If the mass of the fermion is zero the two equations are decoupled and we get

$$(p^0 - \vec{\sigma}.\vec{p})\phi_R = 0, \qquad (p^0 + \vec{\sigma}.\vec{p})\phi_L = 0. \tag{5.12}$$

These are known as Weyl equations and ϕ_L and ϕ_R are called Weyl spinors. Since for a massless particle $p^0 = |\vec{p}|$, these equations become

$$\frac{\vec{\sigma}.\vec{p}}{|\vec{p}|}\phi_R = \phi_R, \qquad \frac{\vec{\sigma}.\vec{p}}{|\vec{p}|}\phi_L = -\phi_L. \tag{5.13}$$

For $m = 0$, the states ϕ_L and ϕ_R do not get mixed up by the equation of motion. We have thus two kinds of massless particles named left-handed described by ϕ_L and right-handed described by ϕ_R.

The projection of $\vec{\sigma}$ along the direction of motion is known as the helicity:

$$\hat{h} = \frac{\vec{\sigma}.\vec{p}}{|\vec{p}|}. \tag{5.14}$$

Thus Weyl spinors are eigenstates of the helicity, the left-handed (right-handed) spinor having negative (positive) helicity. In the absence of a mass term, the helicity operator commutes with the Dirac Hamiltonian, and the helicity is, therefore, a good quantum number.

The chiral representation, also called Weyl representation is given by

$$\gamma^0 = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}, \qquad \gamma^k = \begin{pmatrix} 0 & -\sigma^k \\ \sigma^k & 0 \end{pmatrix}. \tag{5.15}$$

Writing

$$\psi = \begin{pmatrix} \varphi_R \\ \varphi_L \end{pmatrix}, \qquad \gamma^5 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}, \tag{5.16}$$

we have

$$\gamma^5\psi = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}\begin{pmatrix} \varphi_R \\ \varphi_L \end{pmatrix} = \begin{pmatrix} \varphi_R \\ -\varphi_L \end{pmatrix}. \tag{5.17}$$

Defining

$$\psi_L = \begin{pmatrix} 0 \\ \varphi_L \end{pmatrix}, \qquad \psi_R = \begin{pmatrix} \varphi_R \\ 0 \end{pmatrix}, \tag{5.18}$$

we can write

$$\psi_L = \frac{1}{2}\left[\begin{pmatrix} \varphi_R \\ \varphi_L \end{pmatrix} - \begin{pmatrix} \varphi_R \\ -\varphi_L \end{pmatrix}\right], \qquad \psi_R = \frac{1}{2}\left[\begin{pmatrix} \varphi_R \\ \varphi_L \end{pmatrix} + \begin{pmatrix} \varphi_R \\ -\varphi_L \end{pmatrix}\right]. \tag{5.19}$$

Thus

$$\psi_L = \left(\frac{1-\gamma^5}{2}\right)\psi, \qquad \psi_R = \left(\frac{1+\gamma^5}{2}\right)\psi. \tag{5.20}$$

The matrix γ^5 is called the chirality operator. ψ_R is an eigenstate of γ^5 with eigenvalue $+1$, while ψ_L is an eigenstate with eigenvalue -1.

In the chiral representation we have the correspondence

$$\psi = \begin{pmatrix} \text{particle right hand} \\ \text{antiparticle right hand} \\ \text{particle left hand} \\ \text{antiparticle left hand} \end{pmatrix}. \tag{5.21}$$

The chiral states are mixtures of spin states 1/2 and $-1/2$.

In 1937 Majorana wrote a Dirac equation which had real solutions. To do that he used the following representation of purely imaginary matrices

$$\tilde{\gamma}^0 = \begin{pmatrix} 0 & \sigma^2 \\ \sigma^2 & 0 \end{pmatrix}, \quad \tilde{\gamma}^1 = \begin{pmatrix} i\sigma^1 & 0 \\ 0 & i\sigma^1 \end{pmatrix}, \quad \tilde{\gamma}^2 = \begin{pmatrix} 0 & \sigma^2 \\ -\sigma^2 & 0 \end{pmatrix}, \quad \tilde{\gamma}^3 = \begin{pmatrix} -i\sigma^3 & 0 \\ 0 & i\sigma^3 \end{pmatrix}. \tag{5.22}$$

Note that we have $(\tilde{\gamma}^\mu)^* = -\tilde{\gamma}^\mu$. The Dirac equation can then be written as

$$(i\tilde{\gamma}^\mu \partial_\mu - m)\tilde{\psi} = 0. \tag{5.23}$$

This equation describe neutral spin 1/2 fermions that are their own antiparticles.

The components of $\tilde{\psi}$ are all real. We then only need the two upper components. There is evidence that particle excitations in some condensed matter systems are Majorana fermions (Zhao 2013, He *et al* 2017).

5.2 Two-dimensional Dirac equation

Any mathematical objects that satisfy the so called Clifford algebra

$$\{\gamma^\mu, \gamma^\mu\} = 2, \qquad \{\gamma^\mu, \gamma^\nu\} = 0 \text{ for } \mu \neq \nu \text{ and } (\gamma^0)^2 = 1, \tag{5.24}$$

are a good representation for the Dirac equation. In three spatial dimensions, as we saw, γ^μ are 4×4 matrices. However, in the case of two space dimensions there are only three matrices and they can be given a two-dimensional representation. The Pauli matrices σ^z, σ^x, σ^y satisfy the Clifford algebra, and we can write the 2D Dirac Hamiltonian as

$$H_{2D} = c\vec{\sigma}.\,\vec{p} + mc^2\sigma^z, \tag{5.25}$$

where $\vec{\sigma} = (\sigma^x, \sigma^y)$. As before, the massive Dirac Hamiltonian has non-degenerate energy levels $E = \pm\sqrt{p^2c^2 + m^2c^4}$. As we will see in chapter 8, identifying c with the Fermi velocity v_F, the Hamiltonian (5.25) for $m = 0$, can be used to describe the low-energy properties of electrons in graphene.

The Dirac equation in (4 + 1) dimensions has been used to study topological insulators in (4 + 1) dimensions (Qi *et al* 2008).

5.3 Electrodynamics

Let us start with the Dirac Lagrangian

$$L = i\overline{\psi}\gamma^{\mu}\partial_{\mu}\psi - m\overline{\psi}\psi, \tag{5.26}$$

where $\overline{\psi} = \psi^{+}\gamma^{0}$ and we have set $\hbar = 1$, $c = 1$. To obtain a $U(1)$ gauge theory, one requires that L is invariant under the local phase transformation

$$\psi \to e^{i\theta(x)}\psi, \tag{5.27}$$

(where x means x^{μ}). We have

$$\partial_{\mu}(e^{i\theta}\psi) = i(\partial_{\mu}\theta)e^{i\theta}\psi + e^{i\theta}\partial_{\mu}\psi, \tag{5.28}$$

and therefore, the Lagrangian is not invariant under the transformation (5.27), since

$$L \to L - (\partial_{\mu}\theta)\overline{\psi}\gamma^{\mu}\psi. \tag{5.29}$$

It is convenient to define

$$\lambda(x) = -\frac{1}{q}\theta(x) \tag{5.30}$$

and then

$$L \to L + (q\overline{\psi}\gamma^{\mu}\psi)\partial_{\mu}\lambda, \tag{5.31}$$

under the transformation $\psi \to e^{-iq\lambda(x)}\psi$. If we demand that L be invariant under the local phase transformation (5.27), we have to add a gauge field A_{μ} that transforms as

$$A_{\mu} \to A_{\mu} + \partial_{\mu}\lambda, \tag{5.32}$$

and define a covariant derivative as

$$D_{\mu} \equiv \partial_{\mu} + iqA_{\mu}. \tag{5.33}$$

We can verify that

$$D_{\mu}\psi \to e^{-iq\lambda}D\psi, \tag{5.34}$$

and L becomes invariant under the local phase transformation.

For the fields A_{μ} to be dynamical, we need to introduce terms involving their derivatives in the Lagrangian. We do that using the field tensor

$$F_{\mu\nu} \equiv \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}. \tag{5.35}$$

This field is gauge invariant, since under the gauge transformation (5.32)

$$F_{\mu\nu} \to \partial_\mu(A_\nu + \partial_\nu\lambda) - \partial_\nu(A_\mu + \partial_\mu\lambda) = \partial_\mu A_\nu - \partial_\nu A_\mu = F_{\mu\nu}. \tag{5.36}$$

A Lorentz invariant Lagrangian for the field $F_{\mu\nu}$ is given by (Griffiths 2008):

$$L = -\frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu}, \tag{5.37}$$

(where the term $-1/16\pi$ is just a convention). The Euler–Lagrange equation yields

$$\partial_\mu F^{\mu\nu} = 0, \tag{5.38}$$

which are Maxwell's equations for empty space. The final result is

$$L = i\overline{\psi}\gamma^\mu\partial_\mu\psi - m\overline{\psi}\psi - \frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu} - (q\overline{\psi}\gamma^\mu\psi)A_\mu. \tag{5.39}$$

The first term describes free fermions, the second term the photon and the third term gives the interaction between fermions and photons. Note that the electric charge appears only in this term as a coupling constant.

The current density is given by

$$J^\mu = q(\overline{\psi}\gamma^\mu\psi). \tag{5.40}$$

The current conservation

$$\partial_\mu J^\mu = 0, \tag{5.41}$$

follows from the anti-symmetry of $F^{\mu\nu}$.

So, the requirement of local phase invariance, applied to the Dirac Lagrangian, generates all of electrodynamics and gives the current produced by Dirac particles with charge q. It is simple to verify that

$$\partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} + \partial_\lambda F_{\mu\nu} = 0. \tag{5.42}$$

This equation can be expressed as closure of the two-form $F_{\mu\nu}$:

$$dF = 0. \tag{5.43}$$

We can also show that $F = dA$, where A_μ is the one-form vector potential. Gauge invariance under $A \to A + d\lambda$ is evident since $d^2\lambda = 0$.

There is one more term that we could add to the Lagrangian consistent with gauge invariance given by (Schwartz 2014):

$$L_\theta = \theta\varepsilon^{\mu\nu\alpha\beta}F_{\mu\nu}F_{\alpha\beta} = 2\theta\partial_\mu(\varepsilon^{\mu\nu\alpha\beta}A_\nu F_{\alpha\beta}), \tag{5.44}$$

where θ is a number and $\varepsilon^{\mu\nu\alpha\beta}$ is the fully asymmetric 4D Levi-Civita tensor. The term in equation (5.44) is known as the *axion* field term, and it has been introduced in particle physics in order to solve a strong charge-parity violation problem. The term (5.44) has the following contribution to the action of the electromagnetic field

$$S_\theta = \frac{\theta\alpha}{4\pi^2}\int d^3x dt \vec{E}\ .\ \vec{B}, \tag{5.45}$$

where $\alpha = e^2/\hbar c$. S_θ is a topological term—it depends only on the topology of the underlying space, not on the geometry. Since the $\vec{E}$ field is invariant under time reversal, whereas the $\vec{B}$ field changes sign, S_θ breaks time reversal symmetry. For a periodic system, however, there are two values of θ, namely $\theta = 0$ and $\theta = \pi$, that preserve the time reversal symmetry. When θ is constant it plays no role in electrodynamics. However, if we consider θ (x, t), the presence of the axion field can have profound consequences at surfaces and interfaces, when gradients of $\theta(x)$ appear. This term has been used to study topological magnetoelectric effects in topological insulators (Hasan and Kane 2010).

5.4 Time reversal

The time reversal operator is defined by

$$T\colon (t, \vec{x}) \to (-t, \vec{x}). \tag{5.46}$$

The kinetic part in the Lagrangian (5.26) should be invariant under (5.46). To do this, we need $\overline{\psi}\gamma^\mu\partial_\mu\psi$ to transform as a 4-vector under T, so that (Schwartz 2014)

$$i\overline{\psi}\gamma^\mu\partial_\mu\psi(t, \vec{x}) \to i\overline{\psi}\gamma^\mu\partial_\mu\psi(-t, \vec{x}). \tag{5.47}$$

In particular, we need the 0-component to transform as $\overline{\psi}\gamma^0\psi \to -\overline{\psi}\gamma^0\psi$, which implies $\psi^+\psi \to -\psi^+\psi$. Thus we require T to take i into $-i$ in the whole Lagrangian in addition to acting on fields. This makes T an anti-linear operator. We can also verify that T flips the spins of particles, but does not turn particles into antiparticles.

5.5 Gauge field as a connection

In this section I will show another way to treat a gauge field. Let us consider a complex field $\phi(x)$. The physics of the problem does not change by the transformation $\phi(x) \to e^{i\alpha(x)}\phi(x)$. We can choose $\alpha(x)$ at a point x and $\alpha(y)$ at a point y, but then it is impossible to compare fields at different points. To solve this problem we introduce a new field $W(x, y)$ that transform as

$$W(x, y) \to e^{i\alpha(x)}W(x, y)e^{-i\alpha(y)}. \tag{5.48}$$

Then

$$\begin{aligned} W(x, y)\varphi(y) - \varphi(x) &\to e^{i\alpha(x)}W(x, y)e^{-i\alpha(y)}e^{i\alpha(y)}\varphi(y) - e^{i\alpha(x)}\varphi(x) \\ &= e^{i\alpha(x)}[W(x, y)\varphi(y) - \varphi(x)]. \end{aligned} \tag{5.49}$$

We see that $[W(x, y)\phi(y) - \phi(x)]$ is independent of the choice of a local phase. Writing $y^\mu = x^\mu + \delta x^\mu$, dividing by δx^μ and taking $\delta x^\mu \to 0$ we get

$$D_\mu\varphi(x) \equiv \lim_{\delta\mu\to 0}\frac{W(x, x+\delta x)\varphi(x+\delta x) - \varphi(x)}{\delta x^\mu}. \tag{5.50}$$

$$D_\mu\varphi(x) \to e^{i\alpha(x)}D_\mu\varphi(x). \tag{5.51}$$

Imposing the condition $W(x, y) = 1$, we can write for δx^μ small

$$W(x, x + \delta x) = 1 - ie\delta x^\mu A_\mu(x) + O(\delta x^2), \tag{5.52}$$

where e is an arbitrary constant. Using equation (5.48) we get

$$A_\mu(x) \to A_\mu(x) + \frac{1}{e}\partial_\mu\alpha(x), \tag{5.53}$$

and from equation (5.51)

$$D_\mu\varphi(x) = \partial_\mu\varphi(x) - ieA_\mu(x), \tag{5.54}$$

Which is equation (5.33) with $e = -q$. Here, the gauge field was introduced as a connection, allowing us to compare fields at different points, despite their arbitrary phases. We can see the similarity of equation (5.54) with equation (4.58).

As we saw in chapter 4, the tangent space is directly associated with a manifold and has the same dimension as the manifold. The tangent bundle TM in a manifold M is defined as $TM = \cup_{p\in M} T_pM$. We also mentioned that an internal vector space in a fiber bundle can be of any dimension and is defined as an independent addition to the manifold. As an example, let us consider (for simplicity) scalar electrodynamics, where a wave function ψ takes its values at the complex place C, and take as the base manifold the space–time $\Re^4$. We can picture ψ supposing that there is a copy E_p (a fiber) of the complex plane C associated with each point $p \in \Re^4$, and ψ assigns to p a point in the space E_p. The collection of the spaces $\{E_p\}_{p\in\Re^4}$ forms a vector bundle, and the map ψ associating p to a vector in E_p is a cross section of the bundle. Just like the tangent bundle of a manifold, we have a collection of vector spaces that vary smoothly.

We saw above that to compare ψ at different points (i.e. in different E_p) we had to introduce a connection, which we take now on the vector bundle, and a covariant derivative, using an analogy to what was done when we considered Riemannian geometry. The electromagnetic potential can then be geometrically interpreted as a connection on the vector bundle in which the field ψ takes its values. Just as with connections on a manifold, we can study the parallel transport of vectors in a bundle around a closed path and use this to define a notion of curvature for the bundle. In analogy to what was done in chapter 4, we define a curvature as the field strength tensor. Note that the curvature defined in this form for the vector bundle has nothing to do with the intrinsic Riemannian curvature of the base manifold.

We can study gauge theory (as will be done in the following) independent of any physical interpretation.

5.6 Chern classes

In chapter 3, we saw that homotopy theory provides a way to measure the twisting of the fibers of a fiber bundle. Another tool which allows us to measure the

non-triviality of a fiber bundle is characteristic classes (Mombelli 2018, Rahman 2017). They are invariant under vector bundle isomorphism, and thus capture information about the isomorphism class of a vector bundle. Chern and Weil showed that one can construct such characteristic classes in the case of complex bundles using connections and their curvatures. These bundles appear often in physics where the connection is a gauge field. To present this concept, I first need to introduce invariant polynomials. Having in mind that in gauge theory, fiber bundles are in general complex spaces I will use here complex vector spaces. (In appendix B, I present a brief discussion of complex manifolds.)

Let $M(n, C)$ denote the space of $n \times n$ complex matrices. An invariant polynomial on $M(n, C)$ is a function

$$P\colon M(n,\ C) \to C, \tag{5.55}$$

which is basis invariant, i.e. which satisfies $P(TAT^{-1}) = P(A)$ for every nonsingular matrix T. Both the determinant and the trace are invariant polynomials.

The *Chern–Weil* theorem is as follows: If P is an invariant polynomial of the 2-form curvature Ω defined on a manifold M, then $P(\Omega)$ satisfies:

(a) $dP(\Omega) = 0$.
(b) If Ω and Ω' are curvature two-forms corresponding to different connections, then: $d(\Omega') - d\Omega$ is exact.

We see from this theorem that we can associate to an invariant polynomial P a cohomology class of M, which does not depend on the connection used. Topologically equivalent fiber bundles can be associated to the same cohomology class of M, represented by P. The cohomology class of M corresponding to P is called *characteristic class* of the fiber bundle.

Characteristic classes are topological invariants of the fiber bundle, in the sense that topologically equivalent bundles have the same characteristic classes. There are several kinds of characteristic classes, and here I am going to present what is called *Chern classes*.

Definition 1. Let E be a fiber bundle on a manifold M whose fiber is the complex space C^k, and Ω_E the curvature 2-form. We define the total Chern class $C(\Omega_E)$ as

$$C(E) \equiv \det\left(1 + \frac{it}{2\pi}\Omega_E\right) = 1 + C_1 t + C_2 t^2 + \ldots \tag{5.56}$$

C_k is called the kth Chern class and each C_k is a cohomology class, i.e. $C_k(E) \in H^{2k}(M)$. Chern classes were originally defined to be elements of $H^*(M)$ via algebraic topology. Chern proved its equivalence to Chern classes defined by the curvature form. Let E and E' be two complex vector bundles over M, then we have $C(E \oplus E') = C(E)\ .\ C(E')$.

Now we use the result that if an $n \times n$ matrix A is diagonalized by a similarity transformation

$$SAS^{-1} = diag(x_1, x_2, \ldots, x_N), \tag{5.57}$$

and considering that the determinant and the trace are invariant under such a transformation, we can write

$$\begin{aligned}\det(1 + A) = \prod_{j=1}^{N} (1 + x_j) &= 1 + (x_1 + \ldots + x_N) + (x_1x_2 + \ldots + x_{N-1}x_N) \\ &\quad + \ldots + x_1x_2\ldots x_N \\ &= 1 + trA + \frac{1}{2}[(trA^2) - trA^2] + \ldots + \det A.\end{aligned} \tag{5.58}$$

Using equation (5.58) in (5.56) we get

$$C_0 = 1, \qquad C_1 = \frac{i}{2\pi} tr\Omega_E. \tag{5.59}$$

The Chern class C_k is real and the curvature Ω_E is purely imaginary. We write $\Omega_E = i\Omega$.

If the dimension of the manifold is 2, only C_1 is non-zero. The Chern numbers are given by

$$c_l = \int_M C_l. \tag{5.60}$$

The Chern numbers are independent of the choice of the connections. They are integers, invariant under a continuous deformation of the manifold. Thus, they reveal the topology of the fiber bundle. Chern showed that there are n topological invariants associated with every fiber bundle constructed upon a $2n$ dimensional orientable manifold. Different Chern numbers imply different topologies, but the reverse is not true. We also note that there is no Chern number for a base manifold with odd dimension. The Chern classes are defined for complex vector bundles. The Pontryagin classes are the analog for real vector bundles. The kth Pontryagin class p_k are given in term of Chern classes as

$$p_k(M) = (-1)^k C_{2k}(M). \tag{5.61}$$

If E is a real vector bundle, the Euler class of E is given by

$$C(E) = \sqrt{\det(\Omega/2\pi)} = Pf(\Omega/2\pi).$$

5.7 Abelian gauge fields

Using the similarity with electromagnetism, we define an abelian gauge potential, a, in a manifold of dimension n, as a differential 1-form (Manton and Sutcliffe 2004)

$$a = a_i dx^i. \qquad (i = 1, \ldots, n) \tag{5.62}$$

Let us introduce the field 2-form

$$f = da = \sum_{i<j} (\partial_i a_j - \partial_j a_i) dx^i \wedge dx^j. \tag{5.63}$$

We have $df = d(da) = 0$, so f is closed. Under a gauge transformation $a \to a + d\alpha$, f is gauge invariant since $d(a + d\alpha) = da + d(d\alpha) = da$. If $da = 0$, a is closed and the form a is exact and there exists g such that $a = dg$. This happens in a simply-connected space. If the 1-form a is given on a manifold X which is not simply connected $\pi_1(X) \neq 0$, it is possible to have a closed but not exact 1-form a, as we will see in the *Aharonov–Bohm effect* (chapter 6).

The field f plays the role of a curvature for the space where the gauge field lies, and so we can use the results of section 5.6 to construct Chern classes for a gauge theory. The first Chern class of an abelian gauge field is defined to be the

$$C_1 = \frac{1}{2\pi} f. \tag{5.64}$$

For a gauge field with base manifold the plane $\Re^2$, the first Chern number c_1 is defined by

$$c_1 = \frac{1}{2\pi} \int_{\Re^2} f. \tag{5.65}$$

If f is smooth and decays to zero as $|\vec{x}| \to \infty$ more rapidly than $|\vec{x}|^{-2}$, c_1 is finite (we should remember once more that this curvature does not come from the intrinsic geometry of the base manifold $\Re^2$). In Cartesian coordinates we have

$$c_1 = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\partial_x a_y - \partial_y a_x) dx dy. \tag{5.66}$$

Since we can consider f_{xy} as a magnetic field in the plane, c_1 is the total magnetic flux through the plane, divided by 2π. If the plane is embedded in $\Re^3$, $-f_{xy}$ is the component of the magnetic field in the z-direction.

Since $f = da$ locally, f is a closed form, satisfying $df = 0$. Because a is not globally well defined, f is not necessarily an exact 2-form. For a compact Riemann surface X without boundary, one obtains the constraint on the first Chern number c_1 = an arbitrary integer.

As we have seen before, 2D surfaces can be topologically classified by their genus, g, which counts the number of holes. The Gauss–Bonnet theorem states that the integral of the Gaussian curvature over a closed surface is related to g. The Chern number defined above is an integral of a related curvature.

Chern forms can be locally expressed as exact forms. The expression whose exterior derivative gives the Chern form is called a Chern–Simons form. For instance, for the abelian case studied here we have from equation (5.64) $C_1 = da/2\pi$, so the Chern–Simons 1-form is given by $Y_1 = a/2\pi$.

5.8 Non-abelian gauge fields

I will now briefly present the extension of the gauge theory formalism to the non-abelian case. This case is natural for a Lagrangian whose global symmetries include more than just a simple phase rotation.

Let G be a finite dimensional non-abelian Lie Group, and $\{t^a\}$ a base for the Lie algebra. We will assume that $\{t^a\}$ is orthonormalized satisfying the condition $Tr(t^a t^b) = C\delta^{ab}$, where C is a negative constant. The normalization condition implies that the structure constants f^{abc} are totally antisymmetric in their indices.

We will be concerned with a field theory described by the field $\Phi = (\Phi_1, \ldots, \Phi_n)$, where Φ_i are complex scalar fields. The action of the group G can be expressed as

$$\Phi \to g\Phi, \qquad g \in G, \tag{5.67}$$

and the indices have been omitted. We are interested in a theory invariant under space–time dependent gauge transformations

$$\Phi(x) \to g(x)\Phi(x). \tag{5.68}$$

As in the abelian case, we have to introduce a covariant derivative of Φ and a Lie-algebra-valued gauge field $A_\mu(x)$. In terms of the basis $\{t^a\}$, A_μ can be written as

$$A_\mu = A_\mu^a t^a. \tag{5.69}$$

We write the covariant derivative as

$$D_\mu\Phi = \partial_\mu\Phi + A_\mu\Phi, \tag{5.70}$$

and assume that the gauge transformation is given by

$$A_\mu \to gA_\mu g^{-1} - \partial_\mu g g^{-1}. \tag{5.71}$$

Then

$$\begin{aligned} D_\mu\Phi &\to \partial_\mu(g\Phi) + (gA_\mu g^{-1} - \partial_\mu g g^{-1})g\Phi \\ &= \partial_\mu g\Phi + g\partial_\mu\Phi + gA_\mu\Phi - \partial_\mu g\Phi = gD_\mu\Phi, \end{aligned} \tag{5.72}$$

showing that $D_\mu\Phi$ transforms like Φ.

The Yang–Mills field tensor $F_{\mu\nu}$ is defined by the following expression

$$F_{\mu\nu}\Phi = [D_\mu, D_\nu]\Phi. \tag{5.73}$$

or

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]. \tag{5.74}$$

We have under a gauge transformation

$$\begin{aligned} &F_{\mu\nu} \to \partial_\mu(gA_\nu g^{-1} - \partial_\nu g g^{-1}) - \partial_\nu(gA_\mu g^{-1} - \partial_\mu g g^{-1}) \\ &+ [gA_\mu g^{-1} - \partial_\mu g g^{-1}, gA_\nu g^{-1} - \partial_\nu g g^{-1}] = gF_{\mu\nu}g^{-1}, \end{aligned} \tag{5.75}$$

(where we have used $\partial_\mu g^{-1} = -g^{-1}\partial_\mu g g^{-1}$, which leads to $\partial_\mu(g^{-1}g) = \partial_\mu I = 0$).

F is a two 2-form (see the similarity with equation (4.66)):

$$F = DA + A \wedge A. \tag{5.76}$$

It is not closed, however, it satisfies the Biachi identity

$$DF = dF + A \wedge F - F \wedge A = 0. \tag{5.77}$$

We can say, in a loose way, that gauge theory is the study of vector bundles over manifolds, provided with connections satisfying some gauge invariant curvature condition. I would like to stress once more that curvature of space–time and gauge field strength is not the same. They are objects of a similar type, but the curvature of space–time is the curvature of the Levi-Civita connections on the tangent bundle, and the field strength is proportional to the curvature of a connection on a principal $SU(N)$ bundle.

5.9 Chern numbers for non-abelian gauge fields

We can use the results of section 5.6 with Ω replaced by F, and for a $U(n)$ gauge theory to write

$$C_1 = \frac{1}{2\pi} trF. \tag{5.78}$$

We can show that the Chern classes are invariant under the gauge transformation

$$C_l(gFg^{-1}) = C_l(F). \tag{5.79}$$

This implies that c_l is a closed form: $dc_l = 0$.

The second Chern class C_2 is defined as

$$C_2 = \frac{1}{8\pi^2}[tr(F \wedge F) - trF \wedge trF]. \tag{5.80}$$

If F has no $U(1)$ part, then only the term $tr(F \wedge F)$ contributes (Manton and Sutcliffe 2004). Using $F = DA + A \wedge A$ we can write $tr(F \wedge F) = dY_3$, where Y_3 is a Chern–Simons 3-form

$$Y_3 = tr\left(dA \wedge A + \frac{2}{3} A \wedge A \wedge A\right). \tag{5.81}$$

Thus, C_2 is a closed 4-form since $dC_2 = 0$. The second Chern number $c_2 = \int C_2$ is characterized by the homotopy $\pi_4(S^4)$ and it requires a four-dimensional space. We can define a Chern–Simmons number y_3 as $y_3 = \int_{\Re^3} Y_3$, and it is precisely the Qi *et al* (2008) formula for the 'θ' term in equation (5.44).

For a general k we have:

$$C_k = \left(\frac{1}{2\pi}\right)^k \det(F). \tag{5.82}$$

In the abelian case the above definitions reduces to

$$C_0 = 1, \quad C_1 = \frac{1}{2\pi} f, \quad C_2 = \frac{1}{8\pi^2} f \wedge f, \ \cdots C_k = \left(\frac{1}{2\pi}\right)^k \underbrace{f \wedge f \wedge \ \ldots \ \wedge f}_{k} \tag{5.83}$$

For an abelian field on a closed 4-manifold X, c_2 can be non-zero if X has topologically non-trivial closed two-dimensional submanifolds, and is an integer.

The Chern numbers are independent of the choice of either the gauge, or the connections. They are integers, invariant under a continuous deformation of the manifold. As was mentioned above, they reveal the topology of the fiber bundle.

Any gauge theory which does not depend on the metric of the underlying space–time manifold must have a Lagrangian which depends only on topological terms. These terms are precisely the characteristic classes of the relevant bundles.

5.10 Maxwell equations using differential forms

In this section I will treat the Maxwell equations in the context of differential forms. The contravariant antisymmetric electromagnetic tensor $F^{\mu\nu}$ is given by

$$F^{\mu\nu} = \begin{vmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{vmatrix}. \tag{5.84}$$

The covariant form $F_{\mu\nu} = \eta_{\mu\alpha} F^{\alpha\beta} \eta_{\beta\nu}$, (where $\eta_{\alpha\varphi} = \text{diag}(+---)$) written as

$$F_{\mu\nu} = \begin{vmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{vmatrix}, \tag{5.85}$$

can be interpreted as a 2-form F: $V \times V \to \Re$. Using the basis (dx, dy, dz, dt) we can write

$$\begin{aligned} F = {} & E_x dx \wedge dt + E_y dy \wedge dt + E_z dz \wedge dt + B_x dy \wedge dz \\ & + B_y dz \wedge dx + B_z dx \wedge dy. \end{aligned} \tag{5.86}$$

Taking the derivative we find

$$\begin{aligned} dF = {} & \left(\frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z}\right) dx \wedge dy \wedge dz + \left(\frac{\partial B_x}{\partial t} + \frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z}\right) dt \wedge dy \wedge dz \\ & \times \left(\frac{\partial B_y}{\partial t} + \frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x}\right) dt \wedge dz \wedge dx + \left(\frac{\partial B_z}{\partial t} + \frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y}\right) dt \wedge dx \wedge dy. \end{aligned} \tag{5.87}$$

Using Maxwell's equations

$$\vec{\nabla} \cdot \vec{B} = 0, \qquad \vec{\nabla} \times E = -\frac{\partial \vec{B}}{\partial t}, \tag{5.88}$$

we get $dF = 0$. Thus F is a closed 2-form. We can write F as $F = dA$, where A is a 1-form, the 4-potential. From equation (5.88) we get

$$\begin{aligned} {}^*F = {} & E_x dy \wedge dz + E_y dz \wedge dx + E_z dx \wedge dy + B_x dt \wedge dx \\ & + B_y dt \wedge dy + B_z dt \wedge dz. \end{aligned} \tag{5.89}$$

Following the same procedures that we arrived at equation (5.87) we find

$$d^*F = \left(\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} \right) dx \wedge dy \wedge dz + \ldots. \tag{5.90}$$

Now, using the Maxwell equations

$$\vec{\nabla} \cdot \vec{E} = 4\pi\rho, \qquad \frac{\partial \vec{E}}{\partial t} - \nabla \times \vec{B} = -4\pi \vec{J}, \tag{5.91}$$

we find

$$\begin{aligned} d^*F = {} & 4\pi(\rho dx \wedge dy \wedge dz - J_x dt \wedge dy \wedge dz - J_y dt \wedge dz \wedge dx \\ & - J_z dt \wedge dx \wedge dt) = 4\pi {}^*J \end{aligned} \tag{5.92}$$

where (ρ, J_x, J_y, J_z) are the components of the 4-current, and ${}^*J$ is a 3-form. The equations $dF = 0$ and $d^*F = 4\pi {}^*J$ are the Maxwell equations in the language of differential geometry independent of the coordinate system and without mentioning the metric. From equation (5.92) we get the charge conservation law

$$d^*J = 0. \tag{5.93}$$

References and further reading

Griffiths D 2008 *Introduction to Elementary Particles* (Morlenbach: Wiley)

Hasan M Z and Kane C L 2010 Topological insulators *Rev. Mod. Phys.* **82** 3045

He Q L *et al* 2017 Chiral Majorana fermions modes in a quantum anomalous hall insulator–superconductor structure *Science* **21** 294

Manton N and Sutcliffe P 2004 *Topological Solitons* (London: Cambridge University Press)

Mombelli S 2018 *Homotopy Theory and Characteristic Classes* https://ethz.ch/content/dam/ethz/…/o2_Mombelli.pdf

Qi X L, Hughes T L and Zang S C 2008 Topological field theory of time reversal invariant insulators *Phys. Rev.* B **78** 195424

Rahman A 2017 *Chern–Weil theory* https://math.uchicago.edu/~may/REU2017/REUpapers/Rahman.pdf

Schwartz M D 2014 *Quantum Field Theory and the Standard Model* (Cambridge: Cambridge University Press)

Zhao C 2013 *The topological properties in topological insulators and superconductors* (arXiv:1309.2056)

IOP Concise Physics

A Brief Introduction to Topology and Differential Geometry in Condensed Matter Physics

Antonio Sergio Teixeira Pires

Chapter 6

Berry connection and particle moving in a magnetic field

6.1 Introduction

In chapter 3 we saw that the simplest example of a fiber bundle was that of a tangent bundle: a Riemannian manifold with a tangent vector space attached at each point. The metric defined on the tangent plane, varying smoothly from point to point, allowed us to define parallel transport and curvature. In quantum mechanics, we have instead of a tangent vector space, a Hilbert space, with a Hermitian inner product, that varies smoothly from point to point in a parameter space. The structure is in some ways similar.

Let us consider a quantum system described by a Hamiltonian H that depends on parameters $\lambda = (\lambda^1, \lambda^2, \ldots, \lambda^n)$ $\lambda \in M$,where M, the manifold of external parameters, is our base space. The complex vector fiber is the Hilbert space H of the problem, independent of λ. The evolution with time t of the parameters corresponds to a curve $\lambda(t)$, $t \in [t_i, t_f]$ on the base space. A vector $|\psi(\lambda)\rangle$ continuously defined as a function of λ is a cross-section of the vector bundle. In the present case we can write $M \times H$ for the fiber bundle. For simplicity, I am going to use just a single parameter. For instance, if we have a particle in a box, λ is the side of the box. Now we vary λ slowly. The adiabatic theorem says that if the time evolution is slow compared with the characteristic time scale of the system and if the system is in a non-degenerate eingestate, the system will remain in the same state. The energy gap between adjacent eigenstates sets the time scale of the system. There are no transitions between eigenstates during the evolution. (The theorem can apply to degenerate eigenvalues, in which case the states corresponding to the eigenvalues can mix among themselves.) As an example, let us consider a particle in the ground state of an infinite square well in one dimension with λ the size of the box. If $\lambda = a$ at $t = 0$ the initial state is

doi:10.1088/2053-2571/aaec8fch6

$$\psi_i = \sqrt{\frac{2}{a}} \sin\left(\frac{\pi x}{a}\right). \tag{6.1}$$

If we move the right wall slowly, the particle will remain in the ground state with wave function

$$\psi_i = \sqrt{\frac{2}{\lambda}} \sin\left(\frac{\pi x}{\lambda}\right). \tag{6.2}$$

However, if the wall is moved suddenly at $t = 0$, the particle will not instantly adapt to the change and the resulting state will remain ψ_i, which is no longer an eingestate of the new Hamiltonian.

6.2 Berry phase

We start with the Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} |\, \psi(t)\rangle = H(\lambda(t))|\, \psi(t)\rangle. \tag{6.3}$$

we can choose a basis of eigenstates $|n(\lambda)\rangle$ such that

$$H(\lambda)|\, n(\lambda)\rangle = E_n(\lambda)|\, n(\lambda)\rangle. \tag{6.4}$$

We suppose that the spectre of H is discrete, the eigenvalues are non-degenerate and that level crossing does not occur during the evolution. We suppose also that in $t = 0$ the system is in the nth state, i.e.

$$|\, \psi(0)\rangle = |\, n(\lambda(0))\rangle. \tag{6.5}$$

Although the system remains in the same state it acquires a phase that does not affect the physics. We can write

$$|\, \psi(t)\rangle = e^{i\phi_n(t)} |\, n(\lambda)\rangle, \tag{6.6}$$

where $\phi_n(t) = \theta_n(t) + \gamma_n(t)$, $\theta_n(t)$ is the dynamic phase

$$\theta_n(t) = -\frac{1}{\hbar}\int_0^t E_n(\lambda(t'))dt', \tag{6.7}$$

and $\gamma_n(t)$ is the geometric phase. Taking the time derivative of equation (6.6)

$$\frac{\partial |\, \psi(t)\rangle}{\partial t} = e^{i\phi_n(t)} \frac{\partial}{\partial t} |\, n(\lambda)\rangle + i\left(\frac{\partial \theta_n(t)}{\partial t} + \frac{\partial \gamma_n(t)}{\partial t}\right) e^{i\phi_n(t)} |\, n(\lambda)\rangle, \tag{6.8}$$

where

$$\frac{\partial \theta_n(t)}{\partial t} = -\frac{1}{\hbar} E_n \tag{6.9}$$

Taking equations (6.5), (6.6) and (6.7) into (6.8) we get

$$\frac{\partial \mid n(\lambda)\rangle}{\partial t} + i\frac{\partial \gamma_n(t)}{\partial t} \mid n(\lambda)\rangle = 0. \tag{6.10}$$

Multiplying equation (6.10) by $\langle n(\lambda)|$ we obtain

$$\frac{d\gamma_n(t)}{dt} = i\langle n(\lambda)| \frac{d}{dt} \mid n(\lambda)\rangle, \tag{6.11}$$

which leads to

$$\gamma_n(t) = i\int_0^t \langle n(\lambda(t))| \frac{d}{dt'} \mid n(\lambda(t))\rangle dt'. \tag{6.12}$$

Differentiating $\langle n(\lambda(t))||n(\lambda(t))\rangle = 1$ we get

$$\frac{\partial}{\partial t}(\langle n(\lambda(t))|) \mid n(\lambda(t))\rangle + \langle n(\lambda(t))| \left(\frac{\partial}{\partial t} \mid n(\lambda(t))\rangle\right) = 0, \tag{6.13}$$

or

$$\langle n(\lambda(t))| \left(\frac{\partial}{\partial t} \mid n(\lambda(t))\rangle\right) = -\langle n(\lambda(t))| \left(\frac{\partial}{\partial t} \mid n(\lambda(t))\rangle\right)^*, \tag{6.14}$$

which shows that the phase is real. Coming back to λ^j we can write equation (6.12) as

$$\gamma_n(t) = i\int_{\lambda(0)}^{\lambda(t)} \langle n(\lambda(t'))| \frac{\partial}{\partial \lambda^j} \mid n(\lambda(t'))\rangle d\lambda^j = \int_{\lambda(0)}^{\lambda(t)} A_{n,j}(\lambda) d\lambda, \tag{6.15}$$

where

$$A_{n,j} = i\langle n(\lambda)| \frac{\partial}{\partial \lambda^j} \mid n(\lambda)\rangle, \tag{6.16}$$

is the Berry vector potential, also called the Berry connection (Berry 1984). In the language of differential forms

$$\gamma_n(t) = \int_C A_n, \tag{6.17}$$

where the Berry one-form is

$$A_n = A_{n,j} d\lambda^j = i\langle n(\lambda)| \frac{\partial}{\partial \lambda^j} \mid n(\lambda)\rangle d\lambda^j = i\langle n(\lambda)| \, d \mid n(\lambda)\rangle, \tag{6.18}$$

and d is the exterior derivative.

As we saw in chapter 4, a connection is a way to compare vector spaces that are attached to different points of a manifold. In our case, there is a one-dimensional complex vector space attached at each point in the parameter space, spanned by the local eigenstate $|n(\lambda)\rangle$. The Berry connection allows us to compare two vectors from slightly different points.

Differentiating $H|n(\lambda)\rangle = E_n(\lambda)|n(\lambda)\rangle$ we get

$$dH \mid n(\lambda)\rangle + Hd \mid n(\lambda)\rangle = dE_n(\lambda)\mid n(\lambda)\rangle + E_n(\lambda)d \mid n(\lambda)\rangle. \tag{6.19}$$

Multiplying equation (6.19) by $\langle n(\lambda)|$ we find

$$\langle n(\lambda)\mid dH \mid n(\lambda)\rangle = dE_n(\lambda). \tag{6.20}$$

Multiplying equation (6.19) by $\langle m(\lambda)|$ with $m \neq n$ we get

$$\langle m(\lambda)\mid d \mid n(\lambda)\rangle = \frac{\langle m(\lambda)\mid dH(\lambda)\mid n(\lambda)\rangle}{E_n(\lambda) - E_m(\lambda)}. \tag{6.21}$$

The states $|n(\lambda)\rangle$ are not uniquely defined, there is a gauge freedom

$$\mid n(\lambda)\rangle \rightarrow \mid n(\lambda)\rangle' = e^{i\omega_n(\lambda)} \mid n(\lambda)\rangle. \tag{6.22}$$

Note that this is not possible if M has a non-trivial topology. We have now

$$A'_{n,j} = A_{n,j} + \frac{\partial \omega_n}{\partial \lambda^j} = A_{n,j} + d\omega_n. \tag{6.23}$$

The gauge transformation of the Berry vector potential is

$$\gamma_n(t) \rightarrow \gamma'_n(t) = \int_{\lambda(0)}^{\lambda(t)} A'_n.\, d\lambda = \gamma_n(t) - \omega_n(\lambda(t)) + \omega_n(\lambda(0)). \tag{6.24}$$

If the path in the parameter space is a closed loop, $\lambda(T) = \lambda(0)$, and then

$$\gamma_n(t) \rightarrow \gamma'_n(T) = \gamma_n(T) + 2\pi \times \text{integer}.$$

The geometric phase around a closed loop is the Berry phase. It is gauge invariant and cannot be removed.

As an example, let us consider a particle of mass m in a two-dimensional box with sides a and b. The energy levels are given by

$$E = \frac{\pi^2\hbar^2}{2m}\left(\frac{n_x^2}{a^2} + \frac{n_y^2}{b^2}\right). \tag{6.25}$$

Here $\lambda^1 = a$, $\lambda^2 = b$. We can take n as the state with $n_x = n_y = 1$, and starting with some initial values for a and b, vary them slowly and return to the same values keeping $a \neq b$.

As we saw in chapter 5, in electromagnetism the gauge invariance $A_\mu \rightarrow A'_\mu + \partial_\mu \omega$ leads to the fields

$$F_{\mu\nu} = \frac{\partial A_\mu}{\partial x^\nu} - \frac{\partial A_\nu}{\partial x_\mu}. \tag{6.26}$$

Doing an analogy, we can define the fields

$$\mathfrak{F}_{n,ij}(\lambda) = \frac{\partial A_{n,i}}{\partial \lambda^j} - \frac{\partial A_{n,j}}{\partial \lambda^i}, \tag{6.27}$$

where $\mathfrak{F}_{n,ij}$ is the connection defined in the parameter space.

The corresponding two-form

$$F_n = \frac{1}{2}\mathfrak{F}_{n,jk} d\lambda^j \wedge d\lambda^k = dA_n, \tag{6.28}$$

is the Berry curvature. The Berry curvature does not come from the intrinsic geometry of the manifold of parameters but rather with how the attached Hilbert space evolves as parameters change. The Berry curvature is gauge invariant

$$F_n \to F'_n = dA'_n = dA_n - d^2\omega_n = dA_n = F_n. \tag{6.29}$$

Using Stokes' theorem we can write

$$e^{i\gamma} = \exp\left(-i\oint_C A_{n,j}(\lambda) d\lambda^j\right) = \exp\left(-i\int_S \mathfrak{I}_{n,ij} dS^{ij}\right), \tag{6.30}$$

where S is a two-dimensional surface in the parameter space bounded by the curve C.

The Berry phase is observable. We can detect interference effects when one system is taken around a closed path in parameter space, while another system initially identical to the first is taken around a different path.

6.3 The Aharonov–Bohm effect

Let us consider a particle with charge q moving in a region where there is a magnetic potential $\vec{A}$. The plane wave solution of the Schrödinger equation is $\psi \propto e^{i\vec{p}\,.\,\vec{r}/\hbar}$, where $\vec{p} = m\vec{v} - q\vec{A}$. Moving the particle slowly around a contractible loop will generate a local phase ϕ given by

$$\phi = \exp\left(i\frac{q}{\hbar}\oint \vec{A}\,.\,d\vec{r}\right) = \exp\left(i\frac{q}{\hbar}\int \vec{B}\,.\,d\vec{\sigma}\right). \tag{6.31}$$

If the local phases for all contractible loops vanish, the magnetic field is zero. The Berry phase is an example of a local phase.

The global phase is related to non-contractible loops. To understand this phase let us consider the double-slit experiment shown in figure 6.1. A solenoid between the slits and the screen, aligned along the z-direction, gives rise to a magnetic vector potential in the exterior regions inside and outside the solenoid, but the magnetic field is zero in the outside region. A classical particle excluded from the solenoid will not change its motion because $B = 0$. However, a quantum particle, as mentioned above, acquires an extra phase factor ϕ. If we perform a gauge transformation $\vec{A} \to \vec{A} + \vec{\nabla}\omega$, then

$$\phi = \frac{q}{\hbar}\int (\vec{A} + \vec{\nabla}\omega)\,.\,d\vec{r}, \tag{6.32}$$

and so we can choose ϕ as we want, all we need to do is to pick a particular form for ω. However, we are interested in the phase difference

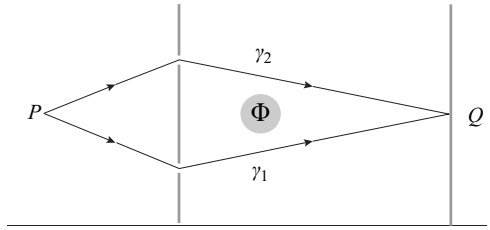


Figure 6.1. The Aharonov–Bohm effect. Reproduced from Buniy *et al* (2013).

$$\Delta\phi = \phi_1 - \phi_2 = \frac{q}{\hbar}\int_{P_1} \vec{A} \,.\, d\vec{r} + \int_{P_2} \vec{A} \,.\, d\vec{r}, \tag{6.33}$$

where P_1 and P_2 are the two paths. We can write equation (6.33) as

$$\Delta\phi = \frac{q}{\hbar}\oint_{C=P_1-P_2} \vec{A} \,.\, d\vec{r}. \tag{6.34}$$

Using Stokes' theorem we find

$$\Delta\phi = \frac{q}{\hbar}\int \vec{\nabla} \times \vec{A} \,.\, d\vec{\sigma} = \frac{q}{\hbar}\int \vec{B} \,.\, d\vec{\sigma} = \frac{q}{\hbar}\Phi, \tag{6.35}$$

where Φ is the flux through the solenoid. This leads to a shift in the interference pattern on the screen that can be observed experimentally. Gauge transformation does not affect the result because here $\oint \vec{\nabla}\omega .\, d\vec{r} = 0$. The effect is topological. This happens because the wave function is defined in the plane $\Re^2$ minus the origin, i.e. a sheet with a hole in it. The result is the same if we place an infinite thin flux tube of flux Φ at $r = 0$. Such a phase is called the global phase (Wen 2004). Electromagnetism has a $U(1)$ symmetry, which has the same topology as the circle S^1. The definition of a phase is equivalent to mapping S^1 to a path around a hole, this is $\pi_1(S^1) = Z$. As we saw in chapter 3, the mapping is characterized by an integer winding number. Global phases exist only when the space is not simply connected. They do not affect the classical equations of motion, but will affect the quantum properties of the particle.

6.4 Non-abelian Berry connections

To complete the study of Berry phases we will present here the non-abelian case that is used in the *non-abelian quantum Hall states*. We consider here the case where the ground state is N-fold degenerate and remains in this situation for all values of the parameter λ. We will change the Hamiltonian in a way that the degeneracy is not broken. If we start with the system in one of the N degenerate ground states, and then vary the parameters in a closed path, the final state could be any one of the degenerate states in the N-dimensional space. For simplicity, we will assume that the ground state has energy $E = 0$ for all values of λ.

As before we have the Schrödinger equation

$$i\frac{\partial \mid \psi\rangle}{\partial t} = H(\lambda(t))\mid \psi\rangle = 0. \tag{6.36}$$

For any choice of the parameters λ, we use an N-dimensional basis of ground states

$$\mid n_i(\lambda)\rangle,\ i = 1, \ldots, N. \tag{6.37}$$

Note that we can use any other choice of bases for each λ. The time evolution of the basis (6.37) through the Schrödinger equation is given by

$$\mid \psi_i(t)\rangle = U_{ij}(t)\Big| n_j(t)\Big\rangle, \tag{6.38}$$

where $U(t)$ is a time-dependent unitary matrix. Taking equation (6.38) into (6.36) we find

$$\frac{d \mid \psi_i\rangle}{dt} = \frac{d}{dt}U_{ij}\mid n_j\rangle + U_{ij}\frac{d}{dt}\mid n_j\rangle, \tag{6.39}$$

which can be written as

$$U_{ik}^{+}\frac{d}{dt}U_{ij} = -\langle n_i \mid \frac{d}{dt}\mid n_j\rangle = -\langle n_i \mid \frac{\partial}{\partial\lambda^{\alpha}}\mid n_j\rangle\frac{d\lambda^{\alpha}}{dt}. \tag{6.40}$$

Now we define a non-abelian Berry connection as

$$(A_{\alpha})_{ji} = -i\langle n_i \mid \frac{\partial}{\partial\lambda^{\alpha}}\mid n_j\rangle. \tag{6.41}$$

A_{α} is a $N \times N$ matrix and is a gauge connection over the space of parameters. But the system is invariant under a rotation of the bases at each point

$$\mid n'_i(\lambda)\rangle = \Omega_{ij}(\lambda)\Big| n_j(\lambda)\Big\rangle, \tag{6.42}$$

where $\Omega(\lambda)$ is a unitary rotation. The Berry connection is now given by

$$A'_{\alpha} = \Omega A_{\alpha}\Omega^{+} + i\frac{\partial\Omega}{\partial\lambda^{\alpha}}\Omega^{+}. \tag{6.43}$$

But this is the gauge connection studied in section 5.7. As we saw the field strength is given by

$$F_{\alpha\beta} = \frac{\partial A_{\alpha}}{\partial\lambda^{\beta}} - \frac{\partial A_{\beta}}{\partial A^{\alpha}} - i[A_{\alpha}, A_{\beta}]. \tag{6.44}$$

References and further reading

Berry M V 1984 Quantum phase factors accompanying adiabatic changes *Proc. R. Soc.* A **392** 45
Buniy R V *et al* 2013 *J. Phys.: Conf. Ser.* **462** 012025
Wen X G 2004 *Quantum Field Theory of Many-Body Systems* (Oxford: Oxford University Press)

IOP Concise Physics

A Brief Introduction to Topology and Differential Geometry in Condensed Matter Physics

Antonio Sergio Teixeira Pires

Chapter 7

Quantum Hall effect

7.1 Integer quantum Hall effect

In 1980, von Klitzing *et al* (1980) discovered experimentally that in a two-dimensional electron gas produced at a semiconductor hereto-junction subject to a strong magnetic field, the longitudinal conductance vanished while quantum plateaus appeared in the Hall conductance at values $\nu e^2/h$, where ν is an integer. To understand the integer quantum Hall effect we start by studying the motion of a particle in a magnetic field. So, we take a uniform magnetic field B in the z direction. Using $\vec{B} = \vec{\nabla} \times \vec{A}$ we can write

$$B_z = \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}. \tag{7.1}$$

We can take $\vec{A} = Bx\hat{j}$, that is

$$A_y = Bx. \tag{7.2}$$

Another choice is

$$A_y = \frac{Bx}{2}, \quad A_x = -\frac{By}{2}. \tag{7.3}$$

The Hamiltonian for a particle with mass m and charge e moving in the x–y plane in the presence of the potential $\vec{A} = Bx\hat{j}$ is

$$H = \frac{1}{2m}(p_x^2 + p_y^2). \tag{7.4}$$

Using equation (7.2) we have for the momentum $p_y \to p_y - \frac{ie}{c}A_y$, so that we can write the Schrödinger equation as

doi:10.1088/2053-2571/aaec8fch7

$$-\frac{\hbar^2}{2m}\left[\frac{\partial^2}{\partial x^2}+\left(\frac{\partial}{\partial y}-\frac{ieBx}{\hbar c}\right)^2\right]\psi(x,y)=E\psi(x,y). \tag{7.5}$$

The translational invariance in the y direction allows us to write

$$\psi_{n,k}(x,y)=e^{iky}\phi_n(x). \tag{7.6}$$

Taking equation (7.6) into (7.5) we get

$$\frac{\hbar\omega_B}{2}\left[-l^2\frac{\partial^2}{\partial x_2^2}+\left(\frac{x}{l}-lk\right)^2\right]\phi_n(x)=\varepsilon_n\phi_n(x), \tag{7.7}$$

where

$$\omega_B\equiv\frac{eB}{mc},\quad l\equiv\sqrt{\frac{\hbar c}{eB}}. \tag{7.8}$$

As we can see, equation (7.7) is the Schrödinger equation for a harmonic oscillator in the variable $x'=x-l^2k$. We can thus write the solution for this equation as

$$\psi_{n,k}(x,y)=e^{iky}H_n(x/l-lk)\exp\left[-\frac{(x-x_k)^2}{2l^2}\right], \tag{7.9}$$

where H_n are the Hermitian polynomials, $x_k=l^2k$, and

$$\varepsilon_n=\hbar\omega_B\left(n+\frac{1}{2}\right). \tag{7.10}$$

The wave function (7.9) depends on the quantum numbers n and k, but the energy (7.10) depends only on n. The levels of the above oscillator, called the Landau levels, are infinitely degenerate. Landau levels are one of the simplest band structures in condensed matter physics, they define a complex vector bundle over the momentum space. Here we have an example where from a bulk Hamiltonian which has translational invariance along one direction (the y direction in our case), we obtain an edge Hamiltonian by a partial Fourier transform. The momenta k in the y direction can then be treated as a parameter. If we plot the energy as a function of k, we find the dispersion relation. The dispersion contains a bunch of horizontal lines, which we take as the energy bands. All the quantum states in one band (the same n but different k) have the same energy. To get a finite number of levels, we consider a rectangular region of sides L_x and L_y, and then calculate the number of states in this region. Using the periodic boundary condition we have

$$e^{i(y+L_y)k}=e^{iky}\rightarrow e^{ikL_y}=1. \tag{7.11}$$

Thus, k is quantized in units $2\pi/L_y$. The situation with the x variable is a little more complicated since the gauge we have chosen is not translational invariant in the x direction. But let us use a simple argument. The wave functions are exponentially

localized around $x_k = kl^2$. For a region $0 \leqslant x \leqslant L_x$, we expect that the allowed values for k are in the range $L_x/l^2 \leqslant k \leqslant 0$. Thus the total number of states is given by

$$N = \frac{L_y}{2\pi}.\frac{L_x}{l^2} = \frac{eBA}{2\pi\hbar c} = \frac{eBA}{hc}, \tag{7.12}$$

where $A = L_xL_y$. The magnetic flux is given by the magnetic field B multiplied by the area A. So we can introduce the flux quantum $\phi_0 = 2\pi\hbar c/e$, and write

$$N = \frac{BA}{\phi_0}. \tag{7.13}$$

We note that N is the number of electrons in each level n, $n_B = eB/hc$ is the number of states per unit of area. If the number of electrons per unit of area is n_e, we define the filling factor ν by

$$\nu = \frac{n_e}{n_B} = \frac{n_e hc}{Be}. \tag{7.14}$$

Whenever the magnetic field takes a value

$$B = \frac{n_e hc}{ep}, \tag{7.15}$$

with p an integer, the filling factor will be an integer and hence ν Landau levels will be completely filled.

We can arrive at equation (7.12) following another procedure as follows. We start with free electrons in a two-dimensional box of side L. The energy levels are given by

$$E = \frac{\hbar^2\pi^2}{2mL^2}(n_x^2 + n_y^2). \tag{7.16}$$

Let us write

$$r^2 = n_x^2 + n_y^2 = (2mE)L^2/(\pi\hbar)^2. \tag{7.17}$$

The number of states with energy less than E is given by

$$\varphi(E) = \frac{1}{4}\pi r^2 = \frac{L^2 2mE}{4\pi\hbar^2}, \tag{7.18}$$

since we use only positive values for n_x and n_y. The number of states with energy lying in the interval between E and $E + \delta E$ is

$$G(E) = \frac{d\varphi}{dE}\delta E = \frac{L^2 m}{2\pi\hbar^2}\delta E. \tag{7.19}$$

The density of states is

$$g(E) = \frac{Am}{2\pi\hbar^2}, \tag{7.20}$$

where $A = L^2$ is the area. If the particle has spin 1/2 we multiply equation (7.20) by 2. In the presence of a magnetic field the total number of states does not change, but the states of free electrons collapse to certain allowed discrete levels satisfying the condition (7.10). That is, the allowed values for the energy become discrete with a spacing between levels of $\hbar\omega_B$. The number of states in the interval $\hbar\omega_B$ is then given by

$$g(E)\hbar\omega_B = \frac{Am\hbar eB}{\pi\hbar^2 mc} = \frac{2eAB}{hc}, \tag{7.21}$$

which is equation (7.12) (the factor 2 is due to the two values of spin). The electrons are strongly correlated because all the states in a given Landau level are completely degenerate in kinetic energy.

Our next step is to study the motion of charged particles. The classical equation of motion for a particle moving in the presence of an electric and a magnetic field is

$$m\frac{d\vec{v}}{dt} = -e\vec{E} - e\vec{v} \times \vec{B} - \frac{m\vec{v}}{\tau}, \tag{7.22}$$

where τ, the damping term, represent the average collision time. In equilibrium $d\vec{v}/dt = 0$, and we can write

$$\vec{v} + \frac{e\tau}{m}\vec{v} \times \vec{B} = -\frac{e\tau}{m}\vec{E}. \tag{7.23}$$

The current density is given by $\vec{j} = -n_e e\vec{v}$, where $n_e e$ is the charge density. Equation (7.23) can be written as

$$v_x + e\omega_B v_y = -(e\tau/m)E_x, \quad v_y - e\omega_B v_x = -(e\tau/m)E_y, \tag{7.24}$$

which can be written (multiplying equation (7.24) by $-n_e e$):

$$\begin{pmatrix} 1 & \omega_B\tau \\ -\omega_B\tau & 1 \end{pmatrix}\begin{pmatrix} j_x \\ j_y \end{pmatrix} = \frac{e^2 n_e \tau}{m}\begin{pmatrix} E_x \\ E_y \end{pmatrix}. \tag{7.25}$$

Using the Ohm law $\vec{j} = \sigma\vec{E}$, with

$$\sigma = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} \\ -\sigma_{xy} & \sigma_{xx} \end{pmatrix}, \tag{7.26}$$

we get, using equation (7.24)

$$\sigma = \frac{\sigma_{DC}}{1 + \omega_B^2\tau}\begin{pmatrix} 1 & -\omega_B\tau \\ \omega_B\tau & 1 \end{pmatrix}, \tag{7.27}$$

where $\sigma_{DC} = n_e e^2\tau/m$ is the direct conductivity. The resistivity $\rho = \sigma^{-1}$ is then

$$\rho = \frac{1}{\sigma_{DC}}\begin{pmatrix} 1 & \omega_B\tau \\ -\omega_B\tau & 1 \end{pmatrix}. \tag{7.28}$$

We can write

$$\rho_{xx} = \frac{m}{n_e e^2 \tau}, \quad \rho_{xy} = \frac{\omega_B \tau}{\sigma_{DC}} = \frac{hc}{e^2 \nu}. \tag{7.29}$$

For a conventional system we expect ρ_{xx} to be a constant and ρ_{xy} to increase linearly with field. However, as was said before, von Klitzing and co-workers found something quite different. They found plateaus in ρ_{xy} occurring at values given by equation (7.29) with ν integer, and dramatic drops in ρ_{xx} which takes very small values when we have a plateau in ρ_{xy}. By varying the strength of the B field, the system turns from one insulator to a metal, and so on. Each insulation state corresponds to a plateau of ρ_{xy} and the step between two neighboring plateaux is a metallic state. Let us now analyze the integer quantum Hall effect.

A metal layer in the xy-plane in the presence of an electric field in the x direction supports a current I in the same direction (figure 7.1). By Ohm's law we have $j_x = \sigma_0 E_x$, where σ_0 is the conductivity. In the presence of a magnetic field oriented upward along the z direction, the electrons will be deflected in the y direction by the Lorentz force. However, since the layer has a finite width L_y, the deflected electrons will run in the edge of the layer. The accumulation of electrons at the edges produces an electric field E_y (called the Hall field) in the y direction (thus, perpendicular to the current). The accumulation ceases when the force eE_y cancels the Lorentz force, that is,

$$E_y = v_x B/c. \tag{7.30}$$

We have $I = L_y j_x$, $V_H = L_y E_y$. The Hall resistance is given by

$$R_H = \frac{V_H}{I} = \frac{L_y E_y}{L_y j_x} = \frac{E_y}{j_x}, \tag{7.31}$$

where V_H is the Hall voltage. Using equation (7.30) and remembering that $j_x = e n_e v_x$ we get

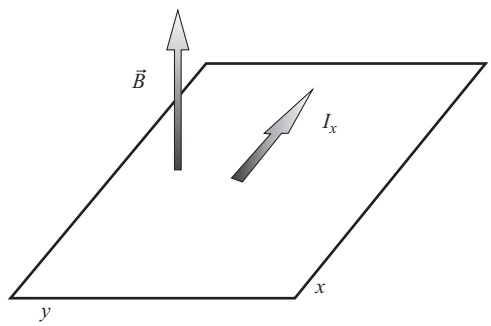


Figure 7.1. Reproduced from Bieri and Frohlich (2011). Copyright © 2011 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

$$R_H = \frac{v_x B}{c j_x} = \frac{B}{e c n_e}. \tag{7.32}$$

The Hall conductance is

$$\sigma_H = \sigma_{xy} = \frac{1}{R_H} = \frac{e c n_e}{B} = \frac{e c \nu n_B}{B} = \frac{\nu e^2}{hc}. \tag{7.33}$$

When the magnetic field is very strong, and the temperature is low enough, ν is quantized and takes on positive integer values. Equation (7.33) is compatible with equation (7.29). In a Hall system an energy gap separates the occupied and empty states just like in an insulator. Unlike an insulator, though, an electric field causes the 'cyclotron orbits to drift' leading to a Hall current characterized by the quantized Hall conductivity (7.33).

7.2 Currents at the edge

A charge e in a magnetic field moves in a circular orbit with radius r under the action of the force evB. Using the expression for the centripetal force we can write

$$\frac{mv^2}{r} = evB \;\rightarrow\; r = \frac{mv}{eB}. \tag{7.34}$$

The energy in a Landau level is

$$\varepsilon_n = \left(n + \frac{1}{2}\right)\frac{\hbar e B}{m}. \tag{7.35}$$

In a semi-classical treatment we can put ε_n equal to the kinetic energy and obtain

$$mv = \sqrt{(2n+1)\hbar e B}, \tag{7.36}$$

and so

$$r_n = \sqrt{(2n+1)\frac{\hbar}{eB}}. \tag{7.37}$$

In the bulk of a sample, clockwise and counterclockwise pieces of neighboring orbits overlap, and the current in the bulk vanishes. At the edge, the orbits are truncated in response to the confining potential created by the boundary. Once an electron is reflected by the boundary, it still attempts to move in a circular orbit. This induces a skipping-type motion of an electron at the boundary of the sample, as shown in figure 7.2 (Philips 2003). Such a motion generates an edge current that flows to the right in the upper edge and to the left in the lower edge, for a magnetic field oriented in the positive z-direction. The electronic states responsible for this motion are chiral in the sense that they propagate in one direction only along the edge. These states are insensitive to disorder because there are no states available for backscattering and the charge carriers have a ballistic motion. Each Landau level generates an edge channel.

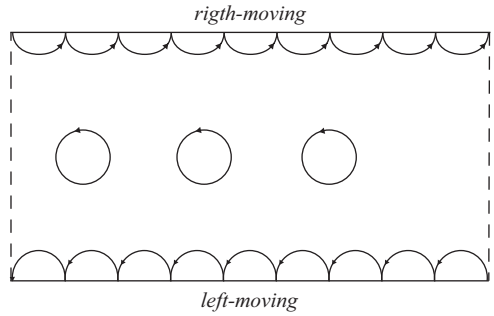


Figure 7.2. Adapted from https://topocondmat.org/w3_pump_QHE/QHEedgestates.html.

The edge of the sample can be modeled by a confining potential $V(x)$ that varies slowly in the bulk but rises steeply near the edge. We write

$$\hbar\omega_k = \left(n + \frac{1}{2}\right)\hbar\omega_c + V(kl^2). \tag{7.38}$$

The group velocity (which represents the electron velocity) is given by

$$v = \frac{\partial\omega_k}{\partial k} = \frac{1}{\hbar}\frac{\partial V}{\partial k} = \frac{1}{\hbar}\frac{\partial V}{\partial r}\frac{\partial r}{\partial k}, \tag{7.39}$$

where $r = kl^2$. We have $v = cE/B$, where $E = -\partial V(r)/\partial r$ is the electric field in the edge originating from the confining potential. We take $E = 0$ in the bulk. Writing $\Delta = E(k) - E_F$, and $\tilde{k} = k - k_F$, we have for $E \approx E_F$,

$$\Delta = \hbar v\tilde{k}. \tag{7.40}$$

The fact that the current arises from states at the edge is a novel feature of quantum Hall systems, and the integer Hall effect can be understood simply as a quantization of the edge currents. Because the Landau levels are discrete, each Landau level will generate one edge channel. Consequently, the number of filled Landau levels, i.e. the filling factor, determines the quantized Hall conductance. The metallic edges are 'chiral' quantum wires (each wire gives one conductance e^2/h). The topological invariant of the bulk 2D material just gives how many wires there have to be at the boundaries of the system.

We can imagine an interface where a crystal slowly interpolates as a function of the distance x between a quantum Hall state and a trivial insulator. Somewhere along the way the energy gap has to vanish, because otherwise it is impossible for the topological invariant to change. If the system remains insulating, the invariant does not change. Therefore the system must not remain insulating. If a bulk Hamiltonian has non-trivial topology, then there should be gapless states localized at the boundary which reflect the bulk topological invariant. The same idea will apply to other topological phases. That is: a topological invariant based on the Berry phase, will lead to a non-trivial edge or surface state at any boundary to an ordinary

insulator or vacuum. However, the physical origin of the phenomenon will be different. I will come back to this subject later.

Our discussion up to now only shows that ρ_{xy} is quantized at those values of B at which an integer number of Landau levels are completely filled, but does not explain the occurrence of the plateau nor the vanishing of ρ_{xx}. The Fermi level jumps from one Landau level to the next as the levels get filled up. Thus, we should expect σ_{xy} to be a monotonically increasing function of the electron density. The explanation of the observability of the steps in σ_{xy} involves both the presence of impurities and of states at the edges of the sample. Disorder changes both the spatial extent and the energy of electronic states. Hence, the degenerate band of states comprising each Landau level can be thought of as being broadened into a band of states. The quantum Hall effect occurs only in imperfect samples. If the samples were ideal, the effect would disappear. I will not go into these details here.

7.3 Kubo formula

It is well known from perturbation theory that in the interaction picture the temporal evolution of a state $|\psi(t)\rangle$ is given by

$$|\psi(t)\rangle = U(t, t_0)|\psi(t_0)\rangle, \tag{7.41}$$

where the time-displacement operator $U(t, t_0)$ is given by

$$U(t, t_0) = T \exp\left[-\frac{i}{\hbar}\int_{t_0}^{t} H_I(t_1)dt_1\right], \tag{7.42}$$

H_I is the interaction Hamiltonian and T the time-ordering operator. Following Tong (2016) we suppose that for $t_0 \to -\infty$ the system is in the unperturbed ground state $|0\rangle$, and write $U(t) = U(t, t_0 \to -\infty)$. From the equation of the canonical momentum operator $\vec{p} = m\vec{v} + q\vec{A}/c$, we have

$$p^2/2m = mv^2/2 + q\vec{v}.\,\vec{A} + q^2A^2/2m, \tag{7.43}$$

where I have taken $c = 1$. Neglecting the term in A^2 we can write (considering that the electron has negative charge)

$$H_I = -\vec{J}.\,\vec{A}. \tag{7.44}$$

In the absence of a scalar potential we have $\vec{E} = -\partial\vec{A}/\partial t$ and supposing $\vec{E}(t) = \vec{E}e^{-i\omega t}$ we get $\vec{A} = (\vec{E}/i\omega)e^{-i\omega t}$.

Now we are going to calculate $<\vec{J}(t)> = \langle\psi(t)|\vec{J}(t)|\psi(t)\rangle$.

We have

$$<\vec{J}(t)> = \langle 0|\, U^{-1}(t)\vec{J}(t)U(t)|\, 0\rangle. \tag{7.45}$$

Expanding equation (7.42) we obtain

$$U(t) \approx 1 - i\int_{-\infty}^{t} dt' H_I(t'), \tag{7.46}$$

and then we can write equation (7.45) as

$$<\vec{J}(t)> = \langle 0 \mid \left(\vec{J}(t) + \frac{i}{\hbar}\int_{-\infty}^{t} dt'[H_I(t'), \vec{J}(t)]\right) \mid 0\rangle. \tag{7.47}$$

The term $\langle 0|\vec{J}(t)|0\rangle$ is the current for $E = 0$ and we can take it as zero. Using equation (7.44) and $\vec{A} = (\vec{E}/i\omega)e^{-i\omega t}$, we get

$$<J_i(t)> = \frac{1}{\hbar\omega}\int_{-\infty}^{t} dt'\langle 0 \mid [J_j(t'), J_i(t)] \mid 0\rangle E_j e^{-i\omega t'}. \tag{7.48}$$

Making a change of variables $t'' = t - t'$ (since the system is translational invariant in time) we can write

$$<J_i(t)> = \frac{1}{\hbar\omega}\left(\int_0^{\infty} dt'' e^{i\omega t''}\langle 0 \mid [J_j(0), J_i(t'')] \mid 0\rangle\right) E_j e^{-i\omega t}. \tag{7.49}$$

It follows from equation (7.49) that the Hall conductance σ_{xy} can be written in the form

$$\sigma_{xy} = \frac{1}{\hbar\omega}\int_0^{\infty} dt e^{i\omega t}\langle 0 \mid [J_y(0), J_x(t)] \mid 0\rangle. \tag{7.50}$$

Inserting $\sum_n |n\rangle\langle n| = 1$ into equation (7.50) and using $J(t) = e^{iH_0 t}J(0)e^{-iH_0 t}$ we get

$$\sigma_{xy} = \frac{1}{\hbar\omega}\int_0^{\infty} dt e^{i\omega t}\sum_n [\langle 0|J_y|n\rangle\langle n|J_x|0\rangle e^{i(E_n - E_0)t} - \langle 0|J_x|n\rangle\langle n|J_y|0\rangle e^{i(E_0 - E_n)t}]. \tag{7.51}$$

Performing the integral, expanding the denominator $(\omega + E_n - E_0)^{-1}$ for small ω and noting that due to the conservation of the current the contribution from the first term vanishes and considering that we should have used the current in equation (7.44), instead of the current density (Tong 2016) we get the final result

$$\sigma_{xy} = i\hbar L_x L_y \sum_{n\neq 0} \frac{\langle 0 \mid J_y \mid n\rangle\langle n \mid J_x \mid 0\rangle - \langle 0 \mid J_x \mid n\rangle\langle n \mid J_y \mid 0\rangle}{(E_n - E_0)^2}. \tag{7.52}$$

This expression is known as the Kubo formula for the Hall conductance σ_{xy}. It is important to mention that the states $|n\rangle$ are the exact eigenstates of the Hamiltonian and E_n are the exact energy levels (Fradkin 1991).

7.4 The quantum Hall state on a lattice

As we saw in chapter 2, a two-dimensional lattice with sides L_x and L_y and periodic boundary conditions can be mapped on the torus T^2. Problems involving a lattice in solid state physics make use of translational symmetry. The translation operator for a translation $\vec{d}$ is given by $T_d = e^{-i\vec{d}\cdot\vec{p}}$, where $\vec{p}$ is the momentum operator. For a periodic system T_d commutes with the Hamiltonian. However, in the presence of a magnetic field T_d does not commute with H. This is because H has become a

function of the magnetic vector potential $\vec{A}$ and this changes as we translate the Hamiltonian. Now we have a translation symmetry that is invariant under a combination of a translation with a gauge transformation.

If we apply a magnetic field perpendicular to the plane, the wave function at the edges should be related by a gauge transformation. To see what happens let us start with the magnetic translator operator

$$T_d = e^{-\vec{d}.\vec{p}} = \exp\left[-i\vec{d}.\,(i\vec{\nabla} + e\vec{A})\right], \tag{7.53}$$

where I have taken $\hbar = 1$. This operator translates a state around a cycle of the torus back to itself. Let us choose the gauge $_{Ax} = 0$, $B_y = Bx$. Then we have, applying T_d to the wave function and translating it around a cycle of the torus

$$\begin{aligned} T_x\psi(x, y) &= \psi(x + L_x, y) = \psi(x, y) \\ T_yT_x\psi(x, y) &= T_y\psi(x + L_x, y) = e^{-ieBL_y(x+L_x)}\psi(x + L_x, y + L_y) \\ T_y\psi(x, y) &= e^{-ieBL_yx}\psi(x, y + L_y) \\ T_xT_y\psi(x, y) &= e^{-ieBL_yx}\psi(x + L_x, y + L_y), \end{aligned} \tag{7.54}$$

which leads to

$$T_yT_x = e^{-ieBL_xL_y}T_xT_y. \tag{7.55}$$

Imposing the condition that we should get the same result we obtain

$$BL_xL_y = \frac{2\pi}{e}n,\ \ n \in Z. \tag{7.56}$$

Since generators of translations do not commute with one another in a magnetic field, electronic states cannot be labeled with momentum. However, if a unit cell with area $2\pi/eB$ enclosing a flux quantum is defined, then lattice translations do commute, so Bloch's theorem allows states to be labeled by 2D crystal momentum $\vec{k}$. In the absence of a periodic potential, the energy levels are simply the $\vec{k}$ independent Landau levels. In the presence of a periodic potential with the same lattice periodicity, the energy levels will disperse with $\vec{k}$.

Now let us see what happens when we apply fluxes Φ_x and Φ_y in the x and y directions of the rectangle. This is equivalent to a flux Φ_x perpendicular to the plane of the torus and a flux Φ_y along the torus. Since flux $= B \times$ area, and if we take $A_x = By$, we get $\Phi_x = Bxy = A_xx$, we can write

$$A_x = \frac{\Phi_x}{L_x} \quad A_y = \frac{\Phi_y}{L_y} + Bx. \tag{7.57}$$

The perturbed Hamiltonian becomes now

$$H_I = \sum_{i=x,y} \frac{J_i\Phi_i}{L_i}. \tag{7.58}$$

We will suppose that the ground state $|\psi_0\rangle$ is non-degenerate and that there is a gap to the first excited state. The perturbed ground state $|\psi_0\rangle'$ can be calculated in first order using perturbation theory. We find

$$|\psi_0\rangle' = |\psi_0\rangle + \sum_{n\neq\psi_0} \frac{\langle n|H_I|n\rangle}{E_n - E_0}|n\rangle. \tag{7.59}$$

If we now take Φ_i as an infinitesimal amount, we get

$$\lim_{\Phi_i\to 0} \frac{|\psi_0\rangle' - |\psi_0\rangle}{\Phi_i} = \left|\frac{\partial\psi_0}{\partial\Phi_i}\right\rangle = -\frac{1}{L_i}\sum_{n\neq\psi_0} \frac{\langle n|J_i|\psi_0\rangle}{E_n - E_0}|n\rangle. \tag{7.60}$$

Taking equation (7.60) into the Kubo formula (7.52) we get

$$\sigma_{xy} = i\left[\left\langle \frac{\partial\psi_0}{\partial\Phi_y}\middle|\frac{\partial\psi_0}{\partial\Phi_x}\right\rangle - \left\langle \frac{\partial\psi_0}{\partial\Phi_x}\middle|\frac{\partial\psi_0}{\partial\Phi_y}\right\rangle\right], \tag{7.61}$$

or

$$\sigma_{xy} = i\left[\frac{\partial}{\partial\Phi_y}\left\langle\psi_0\middle|\frac{\partial\psi_0}{\partial\Phi_x}\right\rangle - \frac{\partial}{\partial\Phi_x}\left\langle\psi_0\middle|\frac{\partial\psi_0}{\partial\Phi_y}\right\rangle\right]. \tag{7.62}$$

The flux Φ_i is periodic and the space of these parameters is a torus T^2_Φ, that should be distinguished from the initial torus.

The calculation of the Hall conductivity of the lattice model for a particle in a magnetic field is more complicated and I refer the reader to Tong (2016) where the subject is very well treated.

7.5 Particle on a lattice

Let us consider a d-dimensional crystal. The Bloch theorem states that the solutions $\psi(\vec{r})$ for the equation

$$H\psi(\vec{r}) = \left(\frac{p^2}{2m} + V(\vec{r})\right)\psi(\vec{r}) = E\psi(\vec{r}), \tag{7.63}$$

where $V(\vec{r})$ is periodic with the periodicity of the lattice are of the form

$$\psi_{n\vec{k}}(\vec{r}) = e^{i\vec{k}.\vec{r}}u_{n\vec{k}}(\vec{r}), \tag{7.64}$$

where $u_{n\vec{k}}(\vec{r})$ are lattice-periodic and are eigenstates of the Bloch Hamiltonian

$$H(\vec{k})\left|u_{n\vec{k}}\right\rangle = E_n(\vec{k})\left|u_{n\vec{k}}\right\rangle, \tag{7.65}$$

where $\vec{k}$ is a crystal momentum and n labels the bands. The Bloch Hamiltonian is given by

$$H(\vec{k}) = e^{-i\vec{k}.\vec{r}} H e^{i\vec{k}.\vec{r}}. \tag{7.66}$$

The symmetry of the lattice implies

$$H(\vec{k} + \vec{G}) = H(\vec{k}). \tag{7.67}$$

If we neglect interaction effects we can use a single electron Hamiltonian. As the quasi-momentum $\vec{k}$ is restricted to the first Brillouin zone, and is defined up to a reciprocal lattice vector G, the Brillouin zone has the topology of a d-dimensional torus T^d. For all $\vec{k} \in T^d$ the Hamiltonian $H(k)$ can be represented by a $N \times N$ matrix with real eigenvalues

$$E_1(k) < E_2(k) < \ldots < E_{N-1}(k) < E_N(k). \tag{7.68}$$

We assume that the $E_n(k)$ are non-degenerate everywhere (i.e. no band crossing).

The evolution of each $E_n(\vec{k})$ as $\vec{k}$ evolves in the Brillouin torus defines a band, and we have N bands. If a band is partially filled one has a metal. If there is a gap in energy between the empty bands above the gap and the filled bands (valence bands) below, one has an insulator. Note that all conventional insulators are equivalent. The Bloch Hamiltonian $H(k)$ defines for each $\vec{k}$, Hermitian operators on the effective Hilbert space $\mathcal{H}_k$. The collection of all spaces of $\mathcal{H}_k$ forms a topological space E which we consider as the total space of a fiber bundle on the base space T^d, which happens to be trivial (due to the vanishing of the total Berry curvature, as will be discussed in the next section) (Fruchart and Carpentier 2013). We have

$$\pi: E \to T^d. \tag{7.69}$$

The fiber over $\vec{k} \in T^d$ is a Hilbert space of dimension N (where, as was said before, N is number of bands) given by all $u_n(\vec{k})$ with $\vec{k}$ fixed. A cross section is a map $T^d \to \mathcal{H}_k$ that assigns to each point $\vec{k} \in T^d$ a specific state $u_n(\vec{k})$ in $\mathcal{H}_k$. Note that here we have a complex Hilbert vector space. But it is of finite dimension and the theory developed before can be used with very few modifications (as was mentioned in chapter 5). For details of complex manifolds see appendix B. If no further symmetry conditions are imposed, the structural group of the bundle is $U(N)$, where N is the dimension of $\mathcal{H}_k$.

If two different Hamiltonians can be deformed into each other continuously they share physical properties that are homotopy invariant. Any two Hamiltonians acting on the same bundle of states are topologically equivalent. There are at least two well-defined subbundles in an insulator: The valence bands bundle (filled bands) and the conduction bands bundle, which corresponds to all empty bands over the energy gap. The valence subbundle underlies the ground state properties of the insulator. The question of whether an insulator is topologically non-trivial or not is equivalent to asking if the total space of the valence fiber subbundle is non-trivial or not. Since the complete Bloch bundle is always trivial, the topology of the filled band will reflect the topological properties of the empty one.

7.6 The TKNN invariant

As we have seen above, a 2D band structure consists of a mapping from the crystal momentum $\vec{k}$, defined on a torus, to the Bloch Hamiltonian $H(k)$. Here we will consider an insulator. If there are no accidental degeneracies, when $\vec{k}$ is transported around a close loop, the state $|u_k\rangle$ acquires a well-defined Berry phase given by the line integral of the Berry connection over T^2, defined by

$$A_m(\vec{k}) = -i\left\langle u_k \left| \frac{\partial}{\partial k_m} \right| u_k \right\rangle, \quad m = x, y. \tag{7.70}$$

In section 6.2 the connection was defined on the space of parameters of the Hamiltonian; here the connection is defined on the space of states itself. The field associated with A_m is given by

$$\mathfrak{F}_{xy} = \frac{\partial A_x}{\partial k_y} - \frac{\partial A_y}{\partial k_x} = -i\left\langle \frac{\partial u_k}{\partial k_y} \middle| \frac{\partial u_k}{\partial k_x} \right\rangle + i\left\langle \frac{\partial u_k}{\partial k_x} \middle| \frac{\partial u_k}{\partial k_y} \right\rangle. \tag{7.71}$$

The Chern invariant is the total Berry flux in the Brillouin zone

$$C = -\frac{1}{2\pi}\int_{T^2} d^2k \mathfrak{F}_{xy}. \tag{7.72}$$

C is also called the TKNN (from Thouless *et al* 1982) invariant. As we saw before, the Chern number is always an integer, and we associate a C_α to each band α.

Using the Kubo formula (7.52) we can write

$$\begin{aligned}\sigma_{xy} = i\hbar \sum \int_{T^2} \frac{d^2k}{(2\pi)^2} \\ \times \frac{\langle u_k^\alpha \mid J_y \mid u_k^\beta\rangle\langle u_k^\beta \mid J_x \mid u_k^\alpha\rangle - \langle u_k^\alpha \mid J_x \mid u_k^\beta\rangle\langle u_k^\beta \mid J_y \mid u_k^\alpha\rangle}{(E_\beta(\vec{k}) - E_\alpha(\vec{k}))^2},\end{aligned} \tag{7.73}$$

where the sum is over $E_\alpha < E_F < E_\beta$, and the index α runs over the filled bands and β runs over the empty bands. This formula takes into account only off diagonal matrix elements (α states are below the Fermi energy, whereas β states are above). We have to integrate in each band. Following Tong (2016) we define the generalized current as

$$\vec{J} = \frac{e}{\hbar}\frac{\partial H}{\partial \vec{k}}, \tag{7.74}$$

where $H(\vec{k})$ is given by equation(7.66). Equation(7.73) can then be written as

$$\begin{aligned}\sigma_{xy} = \frac{ie^2}{\hbar} \sum \int_{T^2} \frac{d^2k}{(2\pi)^2} \\ \times \frac{\langle u_k^\alpha \mid \partial_y H \mid u_k^\beta\rangle\langle u_k^\beta \mid \partial_x H \mid u_k^\alpha\rangle - \langle u_k^\alpha \mid \partial_x H \mid u_k^\beta\rangle\langle u_k^\beta \mid \partial_y H \mid u_k^\alpha\rangle}{(E_\beta(\vec{k}) - E_\alpha(\vec{k}))^2}\end{aligned} \tag{7.75}$$

where $\partial_m = \partial/\partial k_m$, $m = x,y$, and we always sum with the condition $E^\alpha < E_F < E^\beta$. We have

$$\begin{aligned}\langle u_k^\alpha \mid \partial_m H \mid u_k^\beta\rangle &= \langle u_k^\alpha \mid \partial_m(H \mid u_k^\beta\rangle) - \langle u_k^\alpha \mid H \mid \partial_m u_k^\beta\rangle \\ &= (E_\beta(k) - E_\alpha(k))\langle u_k^\alpha|\partial_m u_k^\beta\rangle \\ &= -(E_\beta(k) - E_\alpha(k))\langle \partial_m u_k^\alpha|u_k^\beta\rangle.\end{aligned} \tag{7.76}$$

We do not have the term proportional to $\partial_m E_\beta$ because α and β are distinct bands. Taking equation (7.76) into (7.75) we obtain

$$\sigma_{xy} = \frac{ie^2}{\hbar}\sum\int_{T^2}\frac{d^2k}{(2\pi)^2}\Big(\big\langle \partial_y u_k^\alpha|u_k^\beta\big\rangle\langle u_k^\beta|\partial_x u_k^\alpha\rangle - \langle \partial_x u_k^\alpha|u_k^\beta\rangle\big\langle u_k^\beta|\partial_y u_k^\alpha\big\rangle\Big). \tag{7.77}$$

As we assumed the Fermi energy is in a gap between two bands, we have

$$\sum\Big(|u_k^\alpha\rangle\langle u_k^\alpha| + |u_k^\beta\rangle\langle u_k^\beta|\Big) = 1. \tag{7.78}$$

The sum over the empty bands can then be written as

$$\sum_\beta |\, u_k^\beta\rangle\langle u_k^\beta\,| = 1 - \sum_\alpha |\, u_k^\alpha\rangle\langle u_k^\alpha\,|. \tag{7.79}$$

The second term vanishes by symmetry and so we can write

$$\sigma_{xy} = \frac{ie^2}{\hbar}\sum_\alpha\int_{T^2}\frac{d^2k}{(2\pi)^2}\Big(\big\langle \partial_y u_k^\alpha\big|\partial_x u_k^\alpha\big\rangle - \big\langle \partial_x u_k^\alpha\big|\partial_y u_k^\alpha\big\rangle\Big). \tag{7.80}$$

The sum in equation (7.80) is only over the filled bands α. Using equation (7.71), we have the final result

$$\sigma_{xy} = -\frac{e^2}{2\pi\hbar}\sum_\alpha C_\alpha. \tag{7.81}$$

The Chern number (and therefore the Hall conductivity) is a topological invariant in the sense that it cannot change when the Hamiltonian varies smoothly. This helps to explain the robust quantization of the Hall conductivity. A non-zero Chern number may also be interpreted as a topological obstruction to globally defining a basis of valence Bloch eigenvalues for H. We have seen in chapter 4 that the total Gaussian curvature is quantized only when we consider a closed 2D manifold without edges. For Berry curvature, the same is true.

To see if a system is a topological insulator we have to calculate C for a given model. In the next chapter I will present a simple example.

For the case of a single non-degenerate band, we can define the projection operator $P_k = |u_k\rangle\langle u_k|$ at each point of the Brillouin zone. The projection operator is invariant under $U(1)$transformations of u_k. Equation (7.80) can then be written as

$$\sigma_{xy} = \frac{ie^2}{\hbar}\int_{T^2}\frac{d^2k}{(2\pi)^2}tr[dP_k \wedge dP_k P_k], \tag{7.82}$$

where

$$dP_k = \partial_x P_k dk_x + \partial_y P_k dk_y, \tag{7.83}$$

is a differential form where the coefficients are operators (the wedge product acts only on dk_x and dk_y).

The work of TKNN was performed in a system without edges. The topological role of the edge states was unclear since the TKNN integers are written in terms of the bulk wave functions. On the other side, if one uses Stokes' theorem in equation (7.72), we get zero since there is no boundary in the Brillouin zone in T^2. However, this procedure, as shown by Hatsugai (1993), is incorrect. The phase of the wave function is not well-defined globally over the Brillouin zone (BZ). There is no choice of a global gauge where the energy eigenstate $|u_k\rangle$ is a smooth function of $\overrightarrow{k}$ everywhere. Therefore, one cannot apply Stokes' theorem here. Dividing T^2 into several regions, requiring that the wave functions is an analytical function of k_x and k_y on the BZ, and using a geometry with edges, Hatsugai (1993) has shown that the Chern invariant in the case of an infinite sample, with no edges, is equivalent to the number of edge states for finite samples.

7.7 Quantum spin Hall effect

The quantum spin Hall effect was first studied in graphene, and later in HgTe quantum wells with stronger spin–orbit coupling. The spin–orbit term $H_{SO} = \lambda\vec{L}.\vec{S}$ leads to a momentum dependent force on the electron, somewhat like a magnetic field. To study this effect we start with the Lorentz transformation for electromagnetic fields

$$\vec{B}'_{\perp} = \frac{\vec{B}_{\perp} - \vec{v}\times\vec{E}/c^2}{\sqrt{1 - v^2/c^2}}. \tag{7.84}$$

In a reference frame moving with velocity $\vec{v}$ relative to an electric field $\vec{E}$ (caused by the atoms in the crystal), we have a magnetic field $\vec{B} = \vec{v}\times\vec{E}/c$, neglecting higher order terms. The interaction with a spin $\vec{\sigma}$ is given by

$$\vec{\sigma}.\vec{B} \to \vec{\sigma}.(\vec{p}\times\vec{E}) \to \vec{\sigma}.(\vec{k}\times\vec{E}). \tag{7.85}$$

Each spin component sees an opposite magnetic field. However, the spin dependence means that the time reversal symmetry is different from a real magnetic field.

Let us consider now a uniformly charged cylinder in the presence of an electric field $\vec{E} = E(x\hat{i} + y\hat{j})$. Then we have

$$(\vec{E}\times\vec{k}).\vec{\sigma} = E\sigma_z(k_y x - k_x y). \tag{7.86}$$

The above expression is called spin–orbit interaction, because the deduction is similar to the one for the standard spin–orbit interaction.

The quantum hall state requires an external magnetic field which explicitly breaks the time reversal symmetry. Quantum spin Hall (QSH) states, in contrast, are time reversal (TR) invariant and do not require an external magnetic field. Strong spin–orbit coupling acts like an internal magnetic field without violating the TR symmetry. An applied electrical field causes oppositely directed Hall currents of up and down spins. The charge current is zero, but the spin current is non-zero. Within such materials, we can have the spin-up forward mover and the spin-down backward mover on the upper edge. In the bottom edge, the spin and associated momentum direction are reversed (figure 7.3). To have backscattering the spin of the carriers has to be flipped. Such a spin-flip scattering process is forbidden for non-magnetic impurity (Saha and Jayannavar 2016).

The quantum spin Hall effect is the prototype of a topological insulator with time reversal symmetry. A topological phase is insulating, but always has metallic edges/surfaces, where the conduction occurs, when put next to vacuum or an ordinary phase. Topological insulators will be studied in chapter 8.

7.8 Chern–Simons action

Here we are going to study a topological Lagrangian, that is, a Lagrangian that does not make use of the metric tensor g_{ij}. The Chern–Simons Lagrangian in (2 + 1) dimensional space–time is given by

$$L = -\frac{1}{2}\kappa\varepsilon^{\mu\nu\lambda}a_\mu\partial_\nu a_\lambda, \tag{7.87}$$

where a_μ is a $U(1)$ gauge field, κ is a constant and $\varepsilon^{\mu\nu\lambda}$ is the antisymmetric symbol defined by

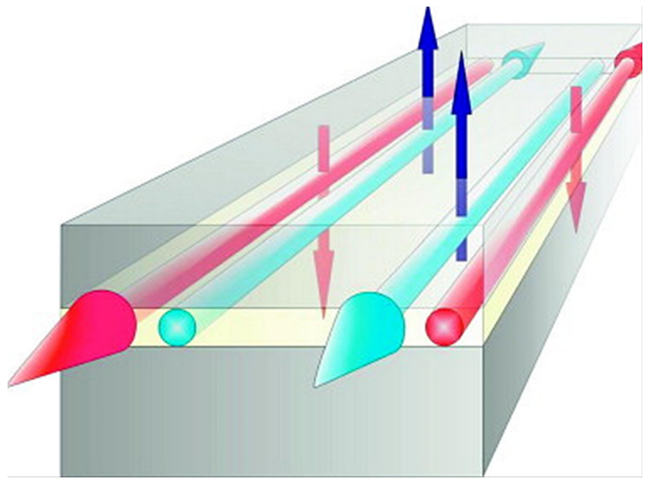


Figure 7.3. Reproduced from Brüne *et al* (2013).

$$\varepsilon^{012} = \varepsilon^{201} = \varepsilon^{120} = 1, \quad \varepsilon^{021} = \varepsilon^{210} = \varepsilon^{102} = -1, \tag{7.88}$$

the other components are zero. In a (2 + 1) dimensional topological theory the contraction of indices is carried out through the use of the antisymmetric symbol $\varepsilon^{\mu\nu\lambda}$.

The topological invariant Chern–Simons action

$$S_{CS}[a] = \int d^3xL, \tag{7.89}$$

(not the Lagrangian density) is gauge invariant. The Chern–Simons action is sometimes called a Hopf invariant (see section B.3).

Writing the field a as a 1-form $a = a_\mu dx_\mu$ and using equation (7.87), equation (7.89) assumes the compact form

$$S_{CS}[a] = -\frac{\kappa}{2}\int a \wedge da. \tag{7.90}$$

We can make the gauge transformation $a_\mu \to a_\mu + \partial_\mu\omega$, and get

$$\varepsilon^{\mu\nu\lambda}a_\mu\partial_\nu a_\lambda \to \varepsilon^{\mu\nu\lambda}a_\mu\partial_\nu a_\lambda + \varepsilon^{\mu\nu\lambda}\partial_\mu\omega\partial_\nu a_\lambda. \tag{7.91}$$

The resulting change in the action is then

$$\delta S = -\frac{\kappa}{2}\int d^3x\varepsilon^{\mu\nu\lambda}\partial_\mu(\omega\partial_\nu a_\lambda). \tag{7.92}$$

The integral is gauge invariant up to a total derivative. It vanishes under the assumption that we may drop boundary terms. But we should be careful. If ω is not single valued, the proof of gauge-invariance is not quite correct. There are some situations where the total derivative does not vanish.

The Chern–Simons term (7.87) respects rotational invariance, but breaks both parity and time reversal. In $d = 2 + 1$ dimensions, parity is defined as

$$x^0 \to x^0, \quad x^1 \to -x^1, \quad x^2 \to x^2, \tag{7.93}$$

and correspondingly

$$a_0 \to a_0, \quad a_1 \to -a_1, \quad a_2 \to a_2. \tag{7.94}$$

The measure $\int dx^3$ is invariant under parity (although $x^1 \to -x^1$, the limits of the integral also change). Thus the Chern–Simons action can only arise in systems that break parity.

Using the Euler–Lagrange equation, we obtain the following equation for the a_μ fields

$$\kappa\varepsilon^{\mu\nu\lambda}\partial_\nu a_\lambda = 0. \tag{7.95}$$

If we define the tensor $f_{\mu\nu} = \partial_\mu a_\nu - \partial_\nu a_\mu$, we get from equation (7.95)

$$f_{\mu\nu} = 0. \tag{7.96}$$

We can couple the fields a_μ to a source J^μ, which is the conserved current of some other field, writing

$$L = -\frac{1}{2}\kappa\varepsilon^{\mu\nu\lambda}a_\mu\partial_\nu a_\lambda + a_\mu J^\mu. \tag{7.97}$$

The equation of motion is now

$$J^\mu = \kappa\varepsilon^{\mu\nu\lambda}\partial_\nu a_\lambda = \frac{1}{2}\kappa\varepsilon^{\mu\nu\lambda}f_{\nu\lambda}. \tag{7.98}$$

We can show that the current is conserved: $\partial_\mu J^\mu = 0$.

Analogous with electromagnetism we can define $b = \partial_2 a_1 - \partial_1 a_2$ as a magnetic field, and $\partial_0 a_i - \partial_i a_0$ the ith component of the electric field e^i. Writing $J^\mu = (\rho, J)$ we get from equation (7.98)

$$\rho = \kappa(\partial_1 a_2 - \partial_2 a_1) = -\kappa b,\ \ J^1 = \kappa(-\partial_0 a_2 + \partial_2 a_0),\ \ J^2 = \kappa(\partial_0 a_1 - \partial_1 a_0), \tag{7.99}$$

or $J^i = -\kappa\varepsilon^{ij}e^j$.

We see that the charge density is locally proportional to the magnetic field. Thus the effect of a Chern–Simmons field is to tie magnetic flux to the 'electric' charge.

Integrating over the two-dimensional space we get the charge of the source field

$$q = -\kappa\int b dx^2. \tag{7.100}$$

We remark that the integral gives the flux of the a_μ field.

The presence of the $\varepsilon^{\mu\nu\lambda}$ symbol means that the action in Euclidean space picks up an extra factor of i. We know that the electron wave function transform as $e^{ie\omega/\hbar}$, and this term should be single valued and not ω. Due to the S^1 geometry of τ, if the gauge transformations wind around the circle with $\omega = 2\pi\hbar\tau/e\beta$ (where β is the period of τ) the exponential $e^{ie\omega/\hbar}$ remains invariant.

As was mentioned above, the Chern–Simons action is invariant under gauge transformations that vanish at the boundary, i.e. $\omega = 0$ on the boundary. Now let us see what happens at the edge. We choose $a_0 = 0$ and from the equation of motion we get $\varepsilon^{ij}a_j = 0$. We can, therefore write $a_i = \partial_i\phi$, where ϕ is a function. Substituting this into equation (7.89) we get

$$\begin{aligned} S &= \frac{\kappa}{2}\int\varepsilon^{ij}a_i\partial_0 a_j d^2x dt = \frac{\kappa}{2}\int(\partial_x\phi\partial_0\partial_y\phi - \partial_y\phi\partial_0\partial_x\phi)d^2x dt \\ &= \frac{\kappa}{2}\int[\partial_x(\phi\partial_0\partial_y\phi) - \partial_y(\phi\partial_0\partial_x\phi)]d^2x dt \\ &= \frac{\kappa}{2}\int(\vec{\nabla}\times\vec{u})d^2x dt = \frac{\kappa}{2}\int\vec{u}.d\vec{l}\,dt \end{aligned} \tag{7.101}$$

where $\vec{u} = (\varphi\partial_0\partial_x\varphi,\ \varphi\partial_0\partial_y\varphi)$. Therefore, at the boundary (assuming runs along x) we get the following term, after integration by parts

$$S_{\text{edge}} = -\frac{\kappa}{2}\int \partial_t\phi\partial_x\phi dxdt. \tag{7.102}$$

This is a topological term. However, as shown by Moore (2014), to obtain an accurate physical description of the theory we need to include non-topological terms. For details the reader is referred to Moore (2014).

7.9 The fractional quantum Hall effect

In 1982, Tsui *et al* observed that using high magnetic fields, quantum plateaus appeared at filling factors ν with rational fractions ($\nu = 1/3, 2/3, 1/5, 2/5, 3/5, 12/5,\ldots$). Known as the fractional Hall effect, this effect originates fundamentally on the electron–electron interaction as well as the Landau level quantization. When $0 < \nu < 1$, the zeroth Landau level is only partially filled. This gives huge ground state degeneracy for free electrons and therefore electron–electron interactions should be responsible for this effect. Fractional quantum Hall (FQH) states constitute a new state of matter which is described by a topological field theory, and one way to study these states is by means of a low-energy effective theory which captures the response of the quantum Hall ground state to low-energy perturbations. We will be interested in general properties of the model, without worrying about the microscopic details. We will be interested only in the case with $1/\nu$ equal to an odd integer. (For other cases see the references.)

Let us rewrite equation (7.33) as

$$\sigma_{xy} = \frac{J_x}{E_y} = \frac{\nu e^2}{2\pi\hbar} \quad \rightarrow \quad J_x = \left(\frac{e^2\nu}{2\pi\hbar}\right)E_y. \tag{7.103}$$

The density of charge can be calculated as: $(-e)\times$ (number of states per Landau level) $\times$ (filling factor)/area. That is

$$\rho = -eJ^0 = -\frac{e^2\nu}{2\pi\hbar}B. \tag{7.104}$$

We take $B < 0$, so that $(-e)B > 0$. Introducing the potential vector A in (2+1) dimensions we can write

$$B = \partial_x A^y - \partial_y A^x, \quad E^i = \partial_t A^i - \partial_i A^t.$$

From equations (7.103) and (7.104) we can write

$$-eJ^\mu = -\left(\frac{e^2\nu}{2\pi\hbar}\right)\varepsilon^{\mu\nu\lambda}\partial_\nu A^\lambda. \tag{7.105}$$

Now we look for an effective Lagrangian that produces equation (7.105). To do this we introduce another $U(1)$ gauge field a_μ which couples to the electromagnetic field A^μ and write

$$J^{\mu} = \frac{1}{2\pi}\varepsilon^{\mu\nu\lambda}\partial_{\nu}a_{\lambda}. \tag{7.106}$$

This current automatically satisfies the conservation law. Using the Euler–Lagrange equation we can verify that the simplest effective Lagrangian that leads to equation (7.106) is given by

$$L = -\frac{s}{4\pi}\varepsilon^{\mu\nu\lambda}a_{\mu}\partial_{\nu}a_{\lambda} + \frac{e}{2\pi}\varepsilon^{\mu\nu\lambda}A_{\mu}\partial_{\nu}a_{\lambda} + j^{\mu}a_{\mu}, \tag{7.107}$$

with $s = 1/\nu$ and $j^{\mu} = 0$. The first term on the right is the Chern–Simons Lagrangian. The source j^{μ} will create excitations with a charge l (constrained to be an integer). Note that we have two gauge fields. One field of electromagnetism A^{μ} which gives rise to the electric field $\vec{E}$ and magnetic field B, whose excitations couple to the charge $-e$, and a topological field a_{μ}, whose excitations couple to the charge l. So, the FQH quasiparticles carry charge e and charge l.

Let us localize the quasiparticle source term at $\vec{x}_0$. We have then

$$j^{\mu}a_{\mu} = la_0\delta^{(2)}(\vec{x} - \vec{x}_0). \tag{7.108}$$

From the equation of motion $\delta L/\delta a_0 = 0$ we get

$$eJ^0 = \frac{1}{2\pi}\varepsilon_{ij}\partial_i a_j = -\frac{e^2}{2\pi s}B + \frac{el}{s}\delta^{(2)}(\vec{x} - \vec{x}_0). \tag{7.109}$$

The first term on the right-hand side is the electric charge density expected from the FQH ground state and indicates that the filling fraction $\nu = (2\pi J^0/ - eB)$ is indeed $\nu = 1/s$. The second term corresponds to the increase in the electron density associated with excitations. We see then that the source term creates FQH quasiparticles with electric charge

$$Q = -le/s. \tag{7.110}$$

We also see that the excitations created by the source term (7.109) is associated with l/s units of the a_{μ} flux. Thus, if we have two quasiparticle excitations carrying a_{μ} charges of l_1 and l_2, and move one around the other in a circle, we obtain a phase of $2\pi \times$ (number of a_{μ} flux quanta) $\times$(a_{μ} charge), namely

$$\Delta\phi = 2\pi\frac{l_1}{s}l_2. \tag{7.111}$$

If we set $l_1 = l_2 = l$, the two quasiparticle excitations will be identical. Interchanging them will induce half of the phase in equation (7.111), that is

$$\theta = \pi l^2/s. \tag{7.112}$$

The angle θ will determine the statistics of the quasiparticles. Therefore, we conclude that the excitations of the FQH ground state are quasiparticles with electric charge $-el/s$ and a_{μ} charge of l. From experiments we know that $k = 1/s$ has values such as 1/3 and thus this means that the quasiparticles carry a fraction of the electron charge

and possess fractional statistics. The FQH quasiparticle excitations can be called quasielectrons and quasiholes. A single electron has electric charge $-el/s = -e$, and thus $l/s = 1$. Therefore, it carries $l = s$ units of a_μ charge. From equation (7.112) we get $\theta = \pi s$, and considering that electrons are fermions we conclude that s is an odd integer, which is what we wanted to deduce. Electrons with $l = s = 3$ units of a_μ-charge form a fractional quantum Hall condensate with $\nu = 1/3$. The quasielectrons in this condensate carry fractional charge $-e/3$ and a_μ-flux of $2\pi/3$, because of their strong coulomb interaction. The fractional quantum Hall effect can be regarded as the integer quantum Hall effect of these quasiparticles.

For more details of the above discussion see Wen (2004) and Lancaster and Blundell (2015). The original explanation of the FQH effect was made by Laughlin using a trial wave function approach, but here I was concerned with the topological aspects of the problem.

References and further reading

Bieri S and Frohlich J 2011 Physical principles underlying the quantum Hall effect *Comtes Rendus Phys.* **12** 333–46

Brüne C, Buhmann H and Molenkamp L W 2013 Quantum spin Hall state in HgTe Contemporary *Concepts of Condensed Matter Science* vol 6 (Elsevier) pp 125–42

Fradkin E 1991 *Field Theories of Condensed Matter Systems* (Redwood City: Addison-Wesley)

Fruchart M and Carpentier D 2013 An introduction to topological insulators *C. R. Acad. Sci.—Ser. IV Phys.* **14** 779

Hatsugai Y 1993 Chern number and edge states in the integer quantum Hall effect *Phys. Rev. Lett.* **71** 3697

Lancaster T and Blundell S J 2015 *Quantum Field Theory for Gifted Amateur* (Oxford: Oxford University Press)

Moore J E 2014 *Draft lectures V–VI notes for Les Houches 2014* https://pks.mpg.de/~topo14/slides/presentation_Moore_5+6.pdf

Philips P 2003 *Advanced Solid State Physics* (Cambridge: Westview Press)

Saha A and Jayannavar A M 2016 Emerging trends in topological insulator and tological superconductor (arXiv:1611.04252)

Thouless D J, Kohmoto M, Nightingale M P and den Nijs M 1982 Quantized Hall conductance in a two-dimensional periodic potential *Phys. Rev. Lett.* **49** 405

Tong D 2016 *The integer quantum Hall effect* (arXiv:16006.06687)

Tsui D C, Storm H L and Gossard A C 1982 Two-dimensional magnetotransport in the extreme quantum limit *Phys. Rev. Lett.* **48** 1559

von Klitzing K, Dorda G and Pepper M 1980 New method for high accuracy determination of the fine-structure constant based on quantized Hall resistance *Phys. Rev. Lett.* **45** 494

Wen X G 2004 *Quantum Field Theory of Many-Body Systems* (Oxford: Oxford University Press)

IOP Concise Physics

A Brief Introduction to Topology and Differential Geometry in Condensed Matter Physics

Antonio Sergio Teixeira Pires

Chapter 8

Topological insulators

8.1 Two bands insulator

In this chapter I will first consider lattice models with non-vanishing Chern numbers without a magnetic field (and therefore no Landau levels), but with time reversal breaking symmetry. Later I will discuss systems with time reversal symmetry. A topological insulator, like an ordinary insulator, has a bulk energy gap separating the highest occupied electronic band from the lowest empty band. The edge in two dimensions of a topological insulator, however, necessarily has gapless states.

The simplest example of a topological insulator is a particle in a two-dimensional lattice with two internal degrees of freedom. The translational motion is described by wavectors k_x and k_y. We have thus two bands, one above and one below the band gap. We have a two-dimensional Hilbert space at each point of the Brillouin torus. If the lowest band is filled one has the so-called Chern insulator.

Since the one particle system has two levels, the Hamiltonian can be written as a 2×2 Hermitian matrix. We can use as a basis the three Pauli matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \ \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{8.1}$$

and the unitary matrix $\sigma_0 = I$. The most general two band Hamiltonian is then written as (Tong 2016)

$$H(\vec{k}) = h_x\sigma_x + h_y\sigma_y + h_z\sigma_z + h_0\sigma_0, \tag{8.2}$$

or

$$H(\vec{k}) = \begin{pmatrix} h_0 + h_z & h_x - ih_y \\ h_x + ih_y & h_0 - h_z \end{pmatrix}, \tag{8.3}$$

doi:10.1088/2053-2571/aaec8fch8

and we suppose that the terms h_μ are periodic on the torus. (We can generalize the procedure to an N-band model, where now σ_i are spin $S = (N-1)/2$ matrices.) In the single particle Hamiltonian, the vector $\vec{h}(\vec{k})$ acts as a 'Zeeman field' applied to a 'pseudo-spin' σ_i of a two level system. Solving the eigenvalues equation

$$H(\vec{k})u_{1,2}(\vec{k}) = \omega_{1,2}(\vec{k})u_{1,2}(\vec{k}), \tag{8.4}$$

we find

$$\omega_{1,2}(\vec{k}) = h_0(\vec{k}) \pm \sqrt{h_x^2(\vec{k}) + h_y^2(\vec{k}) + h_z^2(\vec{k})}, \tag{8.5}$$

and

$$u_{1,2}(\vec{k}) = \left(1 + \frac{h_z + \omega_{1,2}^2}{h_x^2 + h_y^2}\right)^{-1/2} \begin{pmatrix} \dfrac{\omega_{1,2}}{h_x + ih_y} \\ 1 \end{pmatrix}. \tag{8.6}$$

A shift in the energy of both levels has no effect on the topological properties, provided the system remains insulating. Thus, we can take $h_0 = 0$. The system is then insulating provided $h(\vec{k}) = \sqrt{h_x^2(\vec{k}) + h_y^2(\vec{k}) + h_z^2(\vec{k})}$ never vanishes on the whole Brillouin torus.

Let us consider the eigenvalues $u_1(\vec{k})$ corresponding to the filled band. We can construct a map that assigns $u_1(\vec{k})$ to each point of the Brillouin torus, defining a one-dimensional vector bundle on the torus.

The Hamiltonian (8.2) can be parameterized by a vector $\vec{h}$, written in spherical coordinates as

$$\vec{h} = h \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix}. \tag{8.7}$$

The parameter space is a 2-sphere S^2. We can write u_1, the eigenstate of the filled band, as

$$u_1 = \begin{pmatrix} \sin(\theta/2)e^{-i\varphi} \\ -\cos(\theta/2) \end{pmatrix}. \tag{8.8}$$

At the north pole we have

$$u_1 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \tag{8.9}$$

and at the south pole

$$u_1 = \begin{pmatrix} e^{-i\varphi} \\ 0 \end{pmatrix},$$

which is not well defined. We can otherwise write

$$\tilde{u} = e^{i\varphi}u_1 = \begin{pmatrix} \sin(\theta/2) \\ -\cos(\theta/2)e^{i\varphi} \end{pmatrix}, \tag{8.10}$$

which is not well defined at the north pole. It is not possible to have a coherent phase convention for all points of the sphere, i.e. it is not possible to get rid of this singularity. The base manifold cannot be covered by one open set. There is always an obstruction to define a continuous eigenvector on the whole sphere. The topology of a vector bundle on the sphere is not trivial. The transition functions from trivializations on the superior hemisphere U_N to the trivialization on the inferior hemisphere U_S is thus a phase change on the equator.

Let us introduce the unitary vector

$$\vec{n}(k) = \frac{\vec{h}(\vec{k})}{|\vec{h}(\vec{k})|}, \tag{8.11}$$

that describes a point on a sphere S^2. When we move on the Brillouin torus, $\vec{n}(\vec{k})$ leads to a map $T^2 \to S^2$. The Chern number for this system is given by

$$C_1 = \frac{1}{4\pi}\int_{T^2} \vec{n}.\left(\frac{\partial \vec{n}}{\partial k_x} \times \frac{\partial \vec{n}}{\partial k_y}\right) dk_x \wedge dk_y. \tag{8.12}$$

This is the pull-back, via $\vec{n}$, of the 2-sphere S^2 winding number (i.e. it is the index of the mapping from the Brillouin torus to S^2, counting the winding of this map around the sphere).

8.2 Nielsen–Ninomya theorem

Let $S[\psi]$ be the Euclidian action describing fermions on a regular lattice of even dimensions with periodic boundary conditions, and suppose that S is local, hermitian and translation invariant. Then the theory describes as many-left handed as right-handed states. Equivalently, the theorem implies that there are as many states of chirality $+1$ as of chirality -1. This means that in a lattice of even dimension, Dirac points (massless fermions) always appear in pairs, where each pair comprises opposite chirality (Nielsen and Ninomiya 1981).

8.3 Haldane model

A simple example of a two-band model displaying a topological insulating phase was proposed by Haldane (1988). This model is of interest because it provides not only a simple two-band description of the quantum Hall effect, but also a stepping stone to the two-dimensional quantum spin Hall insulator (Kane 2013). As there is no pure magnetic field, the quantum Hall conductance originates from the electron band structure for the lattice instead of the discrete Landau levels created in a strong magnetic field.

This model describes graphene in a periodic magnetic field, which is, on average, zero. Graphene is a 2D form of carbon that has been widely studied. The interesting fact in graphene is that the conduction band and valence band touch each other at two distinct points in the Brillouin zone. Near those points the electronic dispersion resembles the linear dispersion of massless relativistic particles, described by the Dirac equation.

The simplest description starts with a tight-binding model of spinless electrons on a two-dimensional honeycomb lattice. The honeycomb lattice is a bipartite but not a Bravais lattice. It can be viewed as composed of two interlacing triangular sublattices, A and B. Every site has three nearest neighbors on the other sublattice, and six next-nearest-neighbors on the same sublattice. The lattice parameter (the shortest distance between nearest neighbors) will be set as unity.

The base vectors of the Bravais lattice can be chosen as

$$\vec{b}_1 = \begin{pmatrix} -\sqrt{3/2} \\ 3/2 \end{pmatrix}, \vec{b}_2 = \begin{pmatrix} -\sqrt{3/2} \\ -3/2 \end{pmatrix}. \tag{8.13}$$

The Hamiltonian of Haldane's model is written as

$$H = t \sum_{<i,j>} |i\rangle\langle j| + t_2 \sum_{\ll i,j \gg} |i\rangle\langle j| + M\left[\sum_{i\in A} |i\rangle\langle i| - \sum_{j\in B} |j\rangle\langle j|\right], \tag{8.14}$$

where $|i\rangle$ is an electronic state localized at site i, $<,>$ means nearest neighbors and $\ll,\gg$ next-nearest-neighbors. The parameters t and t_2 are hopping terms. The last term is a sublattice symmetry breaking term with on-site energies M for sites on sublattice A, and $-M$ for sites on sublattice B, which breaks inversion symmetry. A local magnetic field is added in such a way that the flux through a unit cell vanishes. This can be done by assuming that the nearest neighbor hopping coefficient is real and the nearest neighbor hopping coefficient acquires a phase: $t_2 \to t_2 e^{i\phi}$, where the phase ϕ due to local magnetic flux is taken as a parameter of the model. Fourier transforming the Hamiltonian (8.14) we obtain the Bloch Hamiltonian

$$H(\vec{k}) = h^{\mu}(\vec{k})\sigma_{\mu}, \tag{8.15}$$

where σ_{μ} are the Pauli matrices and

$$h_0 = 2t_2 \cos\varphi \sum_{i=1}^{3} \cos(\vec{k}.\,\vec{b}_i),\ h_z = M - 2t_2 \sin\varphi \sum_{i=1}^{3} \sin(\vec{k}.\,\vec{b}_i), \tag{8.16}$$

$$h_x = t[1 + \cos(\vec{k}.\,\vec{b}_1) + \cos(\vec{k}.\,\vec{b}_2)],\ h_y = t[\sin(\vec{k}.\,\vec{b}_1) - \sin(\vec{k}.\,\vec{b}_2)]. \tag{8.17}$$

The two energy bands touch each other when $\|\vec{h}\| = 0$. For $h_z = 0$ this occurs at the corners of the Brillouin zone:

$$\vec{K} = \frac{1}{2}\left(\vec{b}_1^{\,*} + \vec{b}_2^{\,*}\right) \text{ and } \vec{K}' = -\vec{K} \tag{8.18}$$

where

$$\vec{b}_1^* = 2\pi\begin{pmatrix} -1/\sqrt{3} \\ 1/3 \end{pmatrix}, \vec{b}_2^* = 2\pi\begin{pmatrix} -1/\sqrt{3} \\ -1/3 \end{pmatrix}. \tag{8.19}$$

At a generic point this degeneracy is lifted, and the system is an insulator, except when $M = \pm 3\sqrt{3}\, t_2 \sin\varphi$ (Fruchart and Carpentier 2013). At the other points the system is an insulator. If we are interested in topological invariants when this condition is not satisfied, we can change the Hamiltonian as long as we do not close the gap and we can take $|\vec{h}|$ everywhere. This is the case studied in section 8.1, and the Chern number can be calculated using equation (8.12).

In figure 8.1 we show the phase diagram for the Haldane model with the Chern numbers (as calculated in Fruchart and Carpentier 2013) in each region. In the critical lines which separate insulating phases with different Chern numbers, a phase transition takes place and the system is not insulating: it is a semi-metal.

In Haldane's model, the time reversal symmetry is broken, and the gap at the Dirac points opens because $h_z(\vec{K}) \neq 0$, but $h_x(\vec{K}) = h_y(\vec{K}) = 0$. We can linearize (8.15) around a Dirac point $\vec{K}$ with $\vec{k} = \vec{K} + \vec{q}$ and get a massive Dirac Hamiltonian with mass $m = h_z(\vec{k})$

$$H = \hbar v_F(q_x\sigma_x + q_y\sigma_y) + m\sigma_z. \tag{8.20}$$

The dispersion $E(\vec{q}) = \pm\sqrt{|\hbar v_F\vec{q}|^2 + m^2}$ has an energy gap $2|m|$.

It can be shown (Fruchart and Carpentier 2013) that the masses $m = h_z(\vec{K})$ and $m' = h_z(\vec{K}')$ have the same sign in the trivial case, and the opposite signs in the topological case.

Graphene is time reversal invariant and parity-invariant. When both symmetries are present h_z is zero. The Nielsen–Ninomiya theorem implies that Dirac points come in pairs in a time reversal invariant system. Hence, the simplest case is one with two Dirac points $\vec{K}$ and $\vec{K}'$. For small $\vec{q} \equiv \vec{k} - \vec{K}$, we can write $\vec{h}(\vec{q}) = \hbar v_F\vec{q}$,

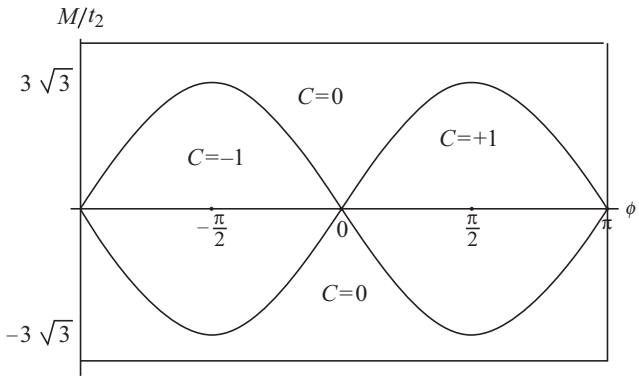


Figure 8.1. Adapted from www.researchgate.net/figure/Phase-diagram-of-the-Haldane-model-33-as-a-function-of-the-Aharonov-Bohm-flux-ph_fig5_264624015.

where v_F is a velocity. So, $H(\vec{q}) = \hbar v_F \vec{q}$. $\vec{\sigma}$ has the form of a 2D massless Dirac Hamiltonian.

Another simple model with a nontrivial Chern number was proposed by Qi *et al* (2006) investigating the quantum Hall effect in a two-dimensional paramagnetic semiconductor. The Hamiltonian is given by

$$H(\vec{k}) = \sin k_x \sigma_x + \sin k_y \sigma_y + (m + \cos k_x + \cos k_y)\sigma_z, \tag{8.21}$$

with energies

$$E_{\pm}(k_x, k_y) = \pm\sqrt{\sin^2 k_x + \sin^2 k_y + (m + \cos k_x + \cos k_y)^2}\,. \tag{8.22}$$

Physically, this model describes the quantum anomalous Hall effect realized with both strong spin–orbit coupling (σ_x and σ_y terms) and ferromagnetic polarization (σ_z term).

For general values of m, the system is an insulator with a gap. However, the gap vanishes (and the bands touch each other) if $m = 0$, or $m = \pm 2$. Near the point $k_x = k_y = 0$, equation (8.22) can be written as

$$E_{\pm} = \pm\sqrt{k_x^2 + k_y^2 + (m+2)}\,. \tag{8.23}$$

If $m = -2$, the gap vanishes at $k_x = k_y = 0$, and $E_{\pm} = \pm|\vec{k}|$. In the neighborhood of this point (a Dirac point) the dispersion relation takes a form of a cone (Dirac cone). For $m = 2$, we have a Dirac cone at $k_x = k_y = \pm\pi$. For $m = 0$, we have two cones. One is located at $k_x = 0$, $k_y = \pm\pi$, and the other at $k_x = \pm\pi$, $k_y = 0$. The Chern number of this system is

$$c_1 = \begin{cases} 1 & \text{for } 0 < m < 2 \\ -1 & \text{for } -2 < m < 0 \\ 0 & \text{otherwise} \end{cases}. \tag{8.24}$$

For $m = -2$, the gap occurs at $k_x = k_y = 0$ (Γ point); for $m = 0$, at $k_x = 0$, $k_y = \pi$ and $k_x = \pi$, $k_y = 0$ (two inequivalent X points); for $m = 2$, at $k_x = \pi$, $k_y = \pi$ (M point). Note that $k_x = \pm\pi$, $k_y = \pm\pi$ are all equivalent. For all other values the spectrum is gapped.

8.4 States at the edge

Let us consider an interface in the two band Dirac model where the mass m at one of the Dirac points changes sign as a function of y (Hasan and Kane 2010). We will take $m(y) > 0$ for the trivial insulator for $y > 0$, and $m(y) < 0$ for the topological insulator for $y < 0$. We can set $m(0) = 0$. The Hamiltonian obtained by replacing $\vec{q}$ by $-i\vec{\nabla}$ in equation (8.20) noting that we have translation symmetry in the x direction is (Fruchart and Carpentier 2013)

$$H = \begin{pmatrix} m(y) & -i\partial_x - \partial_y \\ -i\partial_x + \partial_y & -m(y) \end{pmatrix}. \tag{8.25}$$

We can rotate the basis using the unitary matrix

$$U = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \tag{8.26}$$

to get

$$\begin{pmatrix} -i\partial_x & \partial_y + m(y) \\ -\partial_y + m(y) & i\partial_x \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = E\begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \tag{8.27}$$

or

$$\begin{aligned} (-i\partial_x - E)\alpha = -(\partial_y + m(y))\beta = 0 \\ (i\partial_x - E)\beta = (\partial_y - m(y))\alpha = 0 \end{aligned}. \tag{8.28}$$

The exact solution is

$$\psi_{q_x}(x, y) \propto e^{iq_x x} \exp\left[-\int_0^y \frac{dy'}{v_F} m(y')\right]\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \tag{8.29}$$

with energy $E(q_x) = \hbar v_F q_x$.

This band of states intersects the Fermi energy E_F at $q_x = 0$, with a positive group velocity $\hbar v_F$ and thus corresponds to a right moving chiral edge mode.

Solving the Haldane model in a semi-infinite geometry with an edge at $y = 0$ we obtain the energy levels as a function of the momentum k_x along the edge as shown in figure 8.2.

The difference $N_R - N_L$ between the number of right moving and left moving modes, is an integer topological invariant characterizing the interface. We have the

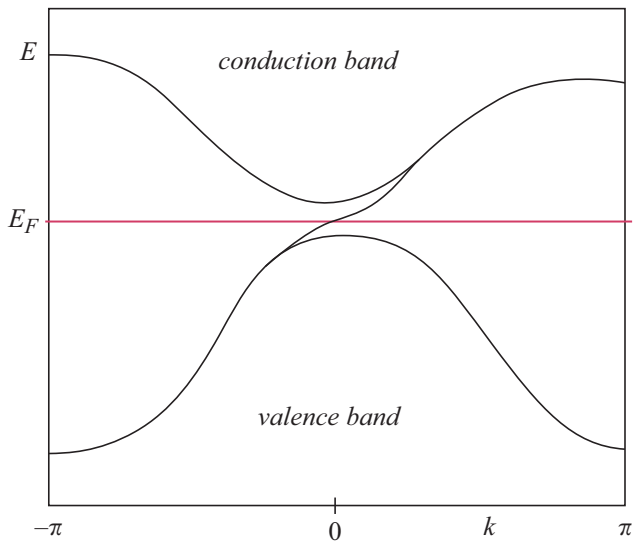


Figure 8.2.

bulk-boundary correspondence $N_R - N_L = \Delta n$, where Δn is the difference in the Chern number across the interface.

In a trivial insulator the surface states have an even number of Fermi level crossings. In a topological insulator the surface states are gapless and cross the Fermi energy an odd number of times.

8.5 Z_2 topological invariants

Here I will comment briefly on the subject, for a more detailed exposition see the excellent text by Kane (2013). For a recent review and references see Shankar (2018).

Since the Chern invariant is odd under time reversal, the topologically nontrivial Chern states can only occur when time reversal symmetry is broken either by an external magnetic field or by magnetic order. A consequence of the time reversal symmetry and inversion symmetry is the fact that the Berry curvature must vanish when this happens. In a material with only spin–orbit the Chern numbers always vanish. Kane and Mele (2005) found a new topological invariant in time reversal invariant systems of fermions. But it is not an integer. It is a parity 'odd' or 'even', that is a 'Z_2 invariant'. Systems in the 'odd' class are 2D topological insulators.

To proceed, I will present briefly the concept of time reversal symmetry. A physical system has time reversal symmetry if it is invariant under the transformation T: $t \to -t$. In the Brillouin zone T changes $\vec{k}$ into $-\vec{k}$. Kramers showed that the square of the time reversal operator is connected to a 2π rotation, which implies that $T^2 = (-1)^{2s}$, where S is the total spin quantum number of a state. Thus, time reversal symmetry takes the complex conjugate of the wave function and rotates the spin. So, the operation of time reversal symmetry T for spin 1/2 particles is given by the anti-unitary operator $\Theta = \exp(i\pi\sigma_y)K$, where K is the complex conjugation. Note that $\Theta^2 = -1$. Anti-unitary operators act as follows on any two generic states $|\phi\rangle$ and $|\psi\rangle$:

$$\langle \Theta\phi|\Theta\psi\rangle = \langle\psi|\varphi\rangle = \langle\varphi|\psi\rangle^*. \tag{8.30}$$

A consequence of time reversal symmetry is the Kramers' theorem that says that all einenstates of a time reversal invariant Hamiltonian are at least two-fold degenerate. The proof is as follows. If a non-degenerate state $|\psi\rangle$ existed, then $\Theta|\psi\rangle = c|\psi\rangle$ for some constant c. This leads to $\Theta^2|\psi\rangle = |c|^2|\psi\rangle$, which is not possible because $|c|^2 \neq -1$. Kramers degeneracy, in the absence of spin–orbit interactions, is simply the degeneracy between up and down spins. The two partners of a Kramers pair live in the same fiber. With the additional structure given by the time reversal symmetry, the Bloch bundle modeling the band structure becomes a quaternionic vector bundle.

If a Bloch Hamiltonian is T invariant, we should have

$$\Theta H(\vec{k})\Theta^{-1} = H(-\vec{k}). \tag{8.31}$$

To obtain a Z_2 invariant we start by defining a unitary matrix

$$w_{mn}(\vec{k}) = \langle u_m(\vec{k})|\,\Theta\,|u_n(-\vec{k})\rangle, \tag{8.32}$$

called the 'sewing matrix' where $|u_m(\vec{k})\rangle$ are occupied Bloch states. Using the fact that Θ is anti-unitary and $\Theta^2 = -1$, leads to

$$w^T(\vec{k}) = -w(-\vec{k}). \tag{8.33}$$

That relation implies that the matrix w is anti-symmetric at points of the Brillouin zone T^d where $\vec{K} = -\vec{K}$. There are 2^d such points in T^d. So, in $2d$ there are four special points $\vec{K}_a$ in the bulk of the Brillouin zone where this happens. At these time reversal invariant points, the filled band fiber of a Z_2 topological insulator is equipped with a quaternionic structure. As we saw in section 2.12 the determinant of an anti-symmetric matrix is the square of its Pfaffian. We define

$$\delta_a = \frac{Pf[w(\vec{K}_a)]}{\sqrt{\det[w(\vec{K}_a)]}} = \pm 1. \tag{8.34}$$

It is always possible to choose $|u_m(\vec{k})\rangle$ continuously throughout the Brillouin zone, so the branch of the square root can be specified globally, and the Z_2 invariant is

$$(-1)^{\nu} = \prod_{a=1}^{4} \delta_a. \tag{8.35}$$

This formulation can be generalized to 3D topological insulators (Fruchart and Carpentier 2013). The theory did not take into account the geometric framework behind and was mathematically established only in 2016 (Fiorenza *et al* 2016).

The association $\vec{k} \to H(\vec{k})$ is a map H: $T^d \to H_{\text{her}}$, where H_{her} is the space of Hermitians $n \times n$ matrices. Taking into account the action of T on T^d, then to specify a map with time reversal symmetry it is enough to know the restriction of H onto a fundamental domain of the T action (8.31). (As the time reversal maps the fibers at $\vec{k}$ and $-\vec{k}$, there is a redundancy in the description of the system on the whole Brillouin torus.) Such a choice of fundamental domain is called the *effective Brillouin zone* (EBZ). This EBZ consists of half of the Brillouin torus, keeping only one member of each Kramers pair $(\vec{k}, -\vec{k})$, except at the boundary. In the 2D-case the EBZ is given by the region $[0, \pi] \times [-\pi, \pi]$, which we can regarded as a cylinder $C = S^1 \times I$. The boundary of this EBZ (denoted ∂EBZ) still has a nontrivial action on T and we can consider a fundamental domain of the boundary of the EBZ, which we call R^{d-1}. In $2d$, we have $R^1 = [0, \pi]$.

Using this idea, Fu and Kane (2006) proposed another form of the Z_2 invariant given by

$$\nu = \frac{1}{2\pi}\left[\oint_{\partial EBZ} A - \int_{EBZ} F\right] (\text{mod. } 2) \tag{8.36}$$

where A is the total Berry connection (the sum of the Berry connections of all bands) constructed from Kramers pairs, and $F = dA$ is the Berry curvature. In the trivial case, $\nu = 0$, while $\nu = 1$ for the nontrivial case. They also showed the equivalence between this expression and the previous one. In 2005 Kane and Mele (2005)

generalized the spinless Haldane model to a graphene lattice of spin 1/2 electrons with spin–orbit coupling. The strong spin–orbit coupling was introduced to replace the periodic magnetic flux in the Haldane model. We will not consider the Kane–Mele model here.

Unlike Chern insulators, Z_2 topological insulators also exist in three dimensions. In three-dimensional topological insulators, the topological invariants are parameterized by four binary $(\nu_0, \nu_1, \nu_2, \nu_3)$ in $(Z_2)^2$. The hallmark of the Z_2 topological order in 3D is the existence of surface states with a linear dispersion and obeying the Dirac equation. The dispersion for surface states arise around a single, or an odd number, of Dirac points in the Brillouin zone (in two-dimensional materials these Dirac points occur only in pairs, in agreement with the Nielsen–Ninomya theorem). In three dimensions, the Z_2 invariant (8.36) can be shown indirectly to be identical to the Chern–Simons invariant y_3 presented in section 5.8.

References and further reading

Fiorenza D, Monaco D and Panati G 2016 Z_2 invariants of topological insulators as geometric obstructions *Commun. Math. Phys.* **343** 1115

Fruchart M and Carpentier D 2013 An introduction to topological insulators *C. R. Acad. Sci. Ser. IV* **14** 779

Fu L and Kane C L 2006 Time reversal polarization and a Z_2 adiabatic spin pump *Phys. Rev.* B **74** 195312

Haldane F D M 1988 Model for a quantum Hall effect without Landau levels: condensed-matter realization of the 'parity anomaly' *Phys. Rev. Lett.* **61** 2015

Hasan M Z and Kane C L 2010 Topological insulators *Rev. Mod. Phys.* **82** 3045

Kane C L and Mele E J 2005 Z_2 topological order and the quantum spin Hall effect *Phys. Rev. Lett.* **95** 146802

Kane C L 2013 Topological band theory and the Z2 invariant *Contemporary Concepts of Condensed Matter Science* **vol 6** p 3

Nielsen H B and Ninomiya M 1981 Absence of neutrinos on a lattice: (I) proof by homotopy theory *Nucl. Phys.* B **185** 20

Qi X L, Wu Y S and Zhang S C 2006 Topological quantization of the spin Hall effect in two-dimensional paramagnetic semiconductors *Phys. Rev.* B **74** 085308

Shankar R 2018 *Topological insulators—A review* (arXiv: 1804.06471)

Tong D 2016 *The integer quantum Hall effect* (arXiv: 16006.06687)

IOP Concise Physics

A Brief Introduction to Topology and Differential Geometry in Condensed Matter Physics

Antonio Sergio Teixeira Pires

Chapter 9

Magnetic models

In this chapter I will treat the one- and two-dimensional antiferromagnetic Heisenberg model. I will be interested only in the topological properties. The thermodynamic and dynamical properties of these models have been presented in a very clear and extensive way in the book by Sachdev (2011).

9.1 One-dimensional antiferromagnetic model

First I will consider the topological term given by equation (1.28) written as

$$\omega[\vec{n}] = \frac{1}{8\pi}\int_0^1 d\rho \int_0^\beta dt \varepsilon^{\mu\nu}\vec{n}.\,[\partial_\mu\vec{n}\times\partial_\nu\vec{n}], \tag{9.1}$$

where I have assumed periodic boundary conditions in time $\vec{n}(\beta) = \vec{n}(0)$, ρ is an auxiliary coordinate $\rho \in [0, 1]$ and the $\vec{n}$ field is extended to $\vec{n}(t, \rho)$ in such way that $\vec{n}(t, 0) = (0, 0, 1)$ and $\vec{n}(t, 1) = \vec{n}(t)$. The indices μ, ν take values t, ρ. The term $\omega[\vec{n}]$ is an oriented area in the sphere swept out by the curve $\vec{n}(t)$.

In an antiferromagnetic spin chain, the neighboring spins prefer antiparallel alignment and we have $\vec{n}_i \to (-1)^i\vec{n}_i$. Then the topological term can be written as

$$S_{\text{top}}[\vec{n}] = S\sum_i (-1)^i\omega[\vec{n}_i]. \tag{9.2}$$

The contribution of two neighboring configurations $\vec{n}_{i+1}$ and $\vec{n}_i$ evaluate, respectively, the areas bounded by the curves $\omega[\vec{n}_{i+1}]$ and $\omega[\vec{n}_i]$. The area difference is given approximately by

$$\begin{aligned}\delta\omega &= \frac{1}{4\pi}\int_0^1 d\rho\int_0^\beta dt\varepsilon^{\mu\nu}\vec{n}.[\partial_\mu\delta\vec{n}\times\partial_\nu\vec{n}]\\ &= \frac{1}{4\pi}\int_0^1 d\rho\int_0^\beta dt\partial_\mu\{\varepsilon^{\mu\nu}\vec{n}.\,[\delta\vec{n}\times\partial_\nu\vec{n}]\} = \frac{1}{4\pi}\int_0^\beta dt\delta\vec{n}.\left(\frac{d\vec{n}}{dt}\times\vec{n}\right),\end{aligned} \tag{9.3}$$

doi:10.1088/2053-2571/aaec8fch9

where I have used $\delta\vec{n}.\,[\partial_\mu\delta\vec{n}\times\partial_\nu\vec{n}]=0$ because all three vectors $\delta\vec{n}$, $\partial_\mu\vec{n}$, $\partial_\nu\vec{n}$ lie in the same plane (tangent to the two-dimensional sphere $\vec{n}^2=1$).

The term

$$\int dt\delta\vec{n}.\left(\frac{d\vec{n}}{dt}\times\vec{n}\right), \tag{9.4}$$

is odd with respect to spin inversion, i.e.

$$\omega[-\vec{n}(t)]=(4\pi-\omega[\vec{n}(t)])=-\omega[\vec{n}(t)]. \tag{9.5}$$

Converting differences into derivatives when the continuum limit is taken and further using the fact that derivatives are contributed by every other link on the chain (resulting in a factor of 1/2) we have for the topological Euclidian action

$$S_{\text{top}}=\frac{iS}{2}\int dxd\tau\partial_x\omega[\vec{n}(x,\tau)], \tag{9.6}$$

or

$$S_{\text{top}}=\frac{iS}{2}\int d\tau dx\vec{n}.\left(\frac{\partial\vec{n}}{\partial x}\right)\times\left(\frac{\partial\vec{n}}{\partial\tau}\right)\equiv i\theta q. \tag{9.7}$$

Terms of this kind are generically called Wess–Zumino terms. The treatment of the Heisenberg model using the path-integral method is well presented in the book by Fradkin (1991). He showed that the Wess–Zumino term distinguishes ferromagnets from antiferromagnets. For the antiferromagnet he obtained a non-linear sigma model, but the action in the path integral had a contribution due to the topological term equal to $2\pi Sq$, where S is the spin and q the winding number. The topological term then gives a contribution

$$e^{i2\pi Sq}=(-1)^{2Sq}. \tag{9.8}$$

Thus if S is an integer, the spin chain is described, at low energies, by the non-linear sigma model. For half-integer S, each topological class contributes with a sign which is positive (negative) if the winding number q is even (odd). The integer and half-integer spin chains fall in different universality classes. An important result of this difference is that the integer spin chains have a gap, while the half-integer chains do not.

Now, I will treat the antiferromagnetic chain with nearest-neighbor exchange using another procedure proposed by Affleck (1989). The Hamiltonian is given by

$$H=J\sum_i\vec{S}_i\cdot\vec{S}_{i+1}. \tag{9.9}$$

I will consider a bipartite lattice and since we expect that at least the short-order should have Néel character we can split the spin field into two pieces, the order parameter $\vec{n}$, and a small varying part $\vec{l}$:

$$\vec{S}_i\approx\pm S\vec{n}_i+\vec{l}_i, \tag{9.10}$$

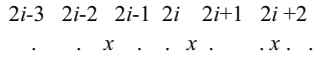


Figure 9.1.

where $\vec{n}$ and $\vec{l}$ vary slowly on the scale of the lattice. Following Affleck (1989) we combine each spin on the even site, $2i$, with the spin to its right, at $(2i + 1)$ (figure 9.1).

$$\vec{l}(2i + 1/2) = (\vec{S}_{2i} + \vec{S}_{2i+1})/2, \quad \vec{n}(2i + 1/2) = (\vec{S}_{2i+1} - \vec{S}_{2i})/2S. \tag{9.11}$$

Assuming that $\vec{l}$ and $\vec{n}$ are slowly varying variables we can write

$$\begin{aligned}
\vec{S}_{2i}.\ \vec{S}_{2i+1} &= 2\vec{l}(2i + 1/2)^2 + \text{constant} \\
\vec{S}_{2i}.\ \vec{S}_{2i-1} &= -S^2\vec{n}(2i + 1/2).\ \vec{n}(2i - 3/2) \\
&\quad + S[\vec{l}(2i - 3/2).\ \vec{n}(2i + 1/2) - \vec{n}(2i - 3/2).\ \vec{l}(2i + 1/2)] \\
&\quad + \vec{l}(2i + 1/2).\ \vec{l}(2i - 3/2) \\
&\approx 2S^2\left(\frac{\partial\vec{n}}{\partial x}\right)^2 - S\left(\vec{l}.\ \frac{\partial\vec{n}}{\partial x} + \frac{\partial\vec{n}}{\partial x}.\ \vec{l}\right) + 2\vec{l}^{\,2}.
\end{aligned} \tag{9.12}$$

The Hamiltonian can then be written as

$$H = J\int\frac{dx}{2}\left[4\vec{l}^{\,2} + 2S^2\left(\frac{\partial\vec{n}}{\partial x}\right)^2 - S\left(\vec{l}.\ \frac{\partial\vec{n}}{\partial x} + \frac{\partial\vec{n}}{\partial x}.\ \vec{l}\right)\right], \tag{9.13}$$

or

$$H = \frac{\mathrm{v}}{2}\int dx\left\{g^2\left[\vec{l} - \frac{\theta}{4\pi}\frac{\partial\vec{n}}{\partial x}\right]^2 + \frac{1}{g^2}\left(\frac{\partial\vec{n}}{\partial x}\right)^2\right\}, \tag{9.14}$$

where $\mathrm{v} = 2JS$, $g = 2/S$ and $\theta = 2\pi S$. This Hamiltonian can be obtained from the following Lagrangian density

$$L = \frac{1}{2g}\partial_\mu\vec{n}.\ \partial^\mu\vec{n} + \left(\frac{\theta}{8\pi}\right)\varepsilon^{\mu\nu}\vec{n}.\ (\partial_\mu\vec{n}\times\partial_\nu\vec{n}), \tag{9.15}$$

where the first term is the Lagrangian for the one-dimensional non-linear sigma model, with the constraint $\vec{n}^2 = 1$, and $\mu = x, t$. The topological term is equal to the one obtained by Fradkin (1991).

9.2 Two-dimensional non-linear sigma model

Let us start with the following Lagrangian in two dimensions

$$L = \frac{1}{2g}(\partial_\mu\vec{n}).\ (\partial_\mu\vec{n}) \tag{9.16}$$

with $\mu = x, y, it$ and subject to the constraint $\vec{n}.\,\vec{n} = 1$. In section 9.1 it was shown how the Heisenberg antiferromagnet for integer spins becomes the O(3) non-linear sigma model in the continuum limit and of course equation (9.16) for $g = 2/S$ describes the Néel phase of the 2D antiferromagnet. However equation (9.16) is more general. Varying the parameter g we get a quantum phase transition to a paramagnetic phase (which can be obtained when we add terms to equation (9.9) for the 2D case such as next-nearest-neighbor interaction). The topological term vanishes on the 2D square lattice for all smooth spacetime configurations of $\vec{n}$ in the Néel state (Sachdev 2011). However, there are important singular configurations on $\vec{n}(x, y, y)$ that do yield a non-vanishing contribution to the topological term in the quantum paramagnetic phase. A simple solution to equation (9.16) is the linear (harmonic) solution that corresponds to the spin wave (or magnons) excitation. However, here I will be interested in more general topological solutions known as skyrmions.

The action for the Lagrangian (9.16) can be written as

$$S = \frac{1}{2g}\int dtdxdy\left[\left(\frac{\partial\vec{n}}{\partial t}\right)^2 - (\nabla\vec{n})^2\right]. \tag{9.17}$$

Writing $\vec{n} = (n_1, n_2, n_3)$ and considering that the allowed values of $\vec{n}$ form the surface of a sphere S^2_{int}, we can use stereographic projection of this sphere onto a plane with Cartesian coordinates w_1 and w_2 to write,

$$w_1 = \frac{n_1}{1+n_3}, \quad w_2 = \frac{n_2}{1+n_3}. \tag{9.18}$$

Introducing the complex variables $z = x + iy$, and $w = w_1 + iw_3$, equation (9.18) can be rewritten as

$$w(z) = \frac{n_1 + in_2}{1+n^3}. \tag{9.19}$$

From equations (9.18) and (9.19) we get

$$n_1 = \frac{2\,\text{Re}\,w}{1+|w|^2}, \quad n_2 = \frac{2\text{Im}w}{1+|w|^2}, \quad n_3 = \frac{1-|w|^2}{1+|w|^2}. \tag{9.20}$$

Now we derive some relations:

$$n_3^2 = 1 - n_1^2 - n_2^2 \rightarrow\ n_3\partial_\mu n_3 = -n_1\partial_\mu n_1 - n_2\partial_\mu n_2, \tag{9.21}$$

$$n_3\partial_\mu n_3 = -n_a\partial_\mu n_a, \quad \text{with } a = 1, 2 \tag{9.22}$$

(remember that we sum over repeated indices),

$$\partial_\mu n_3\partial_\mu n_3 = \frac{(n_3\partial_\mu n_3)(n_3\partial_\mu n_3)}{n_3^2} = \frac{1}{n_3^2}(n_a\partial_\mu n_a)(n_b\partial_\mu n_b). \tag{9.23}$$

Using these expressions we can write the Lagrangian as

$$L = \frac{1}{2g}\left(\partial_\mu n_a \partial_\mu n_a + \frac{n_a n_b}{n_3^2}\partial_\mu n_a \partial_\mu n_b\right). \tag{9.24}$$

We also have

$$\partial_\mu n_a = (1 + n_3)\partial_\mu w_a + w_a \partial_\mu n_3, \tag{9.25}$$

$$w\overline{w} = w_1^2 + w_2^2 = \frac{n_1^2 + n_2^2}{(1 + n_3)^2} = \frac{1 - n_3^2}{(1 + n_3)^2} = \frac{1 - n_3}{1 + n_3}, \tag{9.26}$$

$$w_a \partial_\mu w_a = -\frac{\partial_\mu n_3}{(1 + n_3)^2}. \tag{9.27}$$

Taking equation (9.27) into (9.25) we obtain

$$\partial_\mu n_a = (1 + n_3)[\partial_\mu w_a + w_a(1 + n_3)w_b \partial_\mu w_b]. \tag{9.28}$$

Using the above results, we get after some calculations

$$L = \frac{1}{2g}(1 + n_3)^2 \partial_\mu w_a \partial_\mu w_a. \tag{9.29}$$

Using now

$$(1 + n_3)^2 = \frac{4}{(1 + w\overline{w})^2}, \tag{9.30}$$

we can write equation (9.24) as

$$S = \frac{2i}{g}\int \frac{dtdzd\overline{z}}{(1 + |w|^2)^2}\left\{\frac{1}{2c^2}\left|\frac{\partial w}{\partial t}\right|^2 - \left|\frac{\partial w}{\partial z}\right|^2 - \left|\frac{\partial w}{\partial \overline{z}}\right|^2\right\}. \tag{9.31}$$

Using

$$\frac{\partial w}{\partial x}\frac{\partial \overline{w}}{\partial x} = \left|\frac{\partial w_1}{\partial x}\right|^2 + \left|\frac{\partial w_2}{\partial x}\right|^2, \tag{9.32}$$

together with the Cauchy–Riemann condition

$$\frac{\partial w_1}{\partial x} = \frac{\partial w_2}{\partial y}, \quad \frac{\partial w_1}{\partial y} = -\frac{\partial w_2}{\partial y}, \tag{9.33}$$

we find

$$\left|\frac{\partial w}{\partial x}\right|^2 + \left|\frac{\partial w}{\partial y}\right|^2 = 2\left\{\left(\frac{\partial w}{\partial z}\right)\left(\frac{\partial \overline{w}}{\partial \overline{z}}\right) + \left(\frac{\partial w}{\partial \overline{z}}\right)\left(\frac{\partial \overline{w}}{\partial z}\right)\right\}, \tag{9.34}$$

which allows us to write

$$S = \frac{8i}{g}\int \frac{dtdzd\bar{z}}{(1+|w|^2)}\left(\frac{1}{4c^2}\left|\frac{\partial w}{\partial t}\right|^2 - \left|\frac{\partial w}{\partial z}\right|^2\right) + \frac{4\pi}{g}q, \tag{9.35}$$

where

$$q = \frac{1}{\pi}\int \frac{dxdy}{(1+|w|^2)}\left(\frac{\partial w}{\partial z}\frac{\partial \overline{w}}{\partial \bar{z}} - \frac{\partial w}{\partial \bar{z}}\frac{\partial \overline{w}}{\partial z}\right). \tag{9.36}$$

Using the Euler–Lagrange equations we obtain

$$\frac{1}{2}\frac{\partial^2 \overline{w}}{\partial t^2} - 2\frac{\partial^2 \overline{w}}{\partial z \partial \bar{z}} + \frac{w}{1+|w|_2}\left[4\frac{\partial \overline{w}}{\partial z}\frac{\partial \overline{w}}{\partial \bar{z}} - \left(\frac{\partial \overline{w}}{\partial t}\right)^2\right] = 0. \tag{9.37}$$

Any analytic function $w(z)$ or $w(\bar{z})$ automatically solves equation (9.37) for $t = 0$. A general solution can be written as

$$w(z) = \prod_{i=1}^{N}\left(\frac{z-a_i}{z-b_i}\right), \tag{9.38}$$

where a_i and b_i are complex numbers.

Writing $\vec{n} = (\sin\theta\cos\varphi,\ \sin\theta\sin\varphi,\ \cos\theta)$, we get from equation (9.19)

$$w = \frac{\sin\theta e^{i\phi}}{1+\cos\theta} = \tan\left(\frac{\theta}{2}\right)e^{i\phi}. \tag{9.39}$$

Let us now write $z = re^{i\varphi}$ and consider the simple solution

$$w(z) = \lambda/\bar{z}. \tag{9.40}$$

We obtain $\varphi = \phi$ and $\tan(\frac{\theta}{2}) = \frac{\lambda}{r}$.

We see that

$$r = 0 \to \theta(0) = \pi \ \text{ and } \ r = \infty \to \theta(\infty) = 0.$$

Static solutions with non-zero but finite energy are called skyrmions in condensed matter (see figure 9.2). (They are not strictly solitons since they have an instability associated with changes in their scale.) The solutions must satisfy (Manton and Sutcliffe 2004, Rajaraman 1987)

$$\lim_{r\to\infty}\vec{n}(\vec{r}) = \vec{n}^{(0)}, \quad r\,\|\,\mathrm{grad}\vec{n}\,\| \to 0, \quad r\to\infty \tag{9.41}$$

Note that as we tend to infinity in coordinate space in different directions, $\vec{n}(\vec{r})$ must approach the same limit $\vec{n}^{(0)}$, otherwise $\vec{n}(\vec{r})$ will depend on the coordinate angle even at $r = \infty$ and the angular component of the gradient will not satisfy equation (9.41). This boundary condition spontaneously breaks the $O(3)$ symmetry to an $O(2)$ symmetry. We may identify the points at infinity as a single point, i.e. the space can be identified with a sphere S^2_{phys}. Since the $\vec{n}$ field lives on a sphere of unit radius S^2_{int}, any finite-energy static configuration $\vec{n}(\vec{r})$ is a mapping of S^2_{phys} into S^2_{int}. We saw in

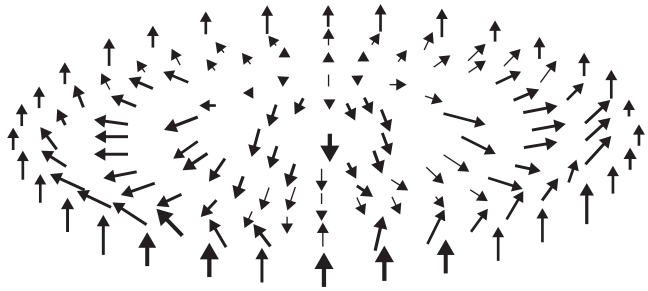


Figure 9.2. In this example of a skyrmion, the field at infinite points in positive z direction, while near the origin it points in the $-z$ direction. It has topological charge of +1.

section 3.3 that all non-singular mapping of one sphere S^2 into another can be classified into homotopy sectors: $\pi_2(S^2) = Z$. So, each field configuration is characterized by an integer, that here we call the topological charge q. The charge q can be interpreted as the number of lumps (skyrmions) in the field configuration.

Calculating q using the expression

$$q = \frac{1}{8\pi}\int \varepsilon_{\mu\nu}\vec{n}.\,(\partial_\mu\vec{n} \times \partial_\nu\vec{n})dxdy, \tag{9.42}$$

we can show that this equation is equivalent to equation (9.28). Calculating q for the solution

$$w(z) = \left(\frac{z - z_0}{\lambda}\right)^n, \tag{9.43}$$

where n is a positive integer, λ a real number and z_0 a complex number, we find $q = n$.

Skyrmions plays an important role in the study of low temperature properties of two-dimensional magnetic systems.

Static solutions to the two-dimensional sigma model correspond to Euclidian solutions to the one-dimensional model. On the other side, topological time dependent solutions for the Euclidian action of the two-dimensional non-linear sigma model which corresponds to static solutions for the model in three dimensions are called *instantons* (map of S^3 into S^2). There are no smooth solutions with non-trivial winding numbers in three dimensions. We have a singular solution where $\vec{n}$ points form a center. Such a solution is called a *hedgehog*.

The topological action for skyrmions in a ferromagnet is given by (Tanaka and Takayoshi 2015)

$$S_{\text{top}} = \int dtdxdy\rho\omega[\vec{n}(t, x, y)], \tag{9.44}$$

where ρ is the density of skyrmions. We assume that the time dependence takes the form $\vec{n} = \vec{n}(\vec{r} - \vec{R}(t))$, where $\vec{R}(t)$ stands for the center of the skyrmion. Using the fact that the variation of the surface angle is given by

$$\delta\omega = \int dt\vec{n}.\ \delta\vec{n} \times \partial_t\vec{n}, \tag{9.45}$$

we obtain for the force acting on a skyrmion

$$\vec{F} = -\frac{\delta S_{\text{top}}}{\delta \vec{R}} = 4\pi Sq\rho\vec{v} \times \hat{z}. \tag{9.46}$$

Thus a skyrmion in motion behaves like a charged particle in the presence of a magnetic field proportional to $\rho\hat{z}$.

Skyrmions have also been used to study quantum Hall ferromagnets (Girvin 2000).

9.3 *XY* model

In this section I will consider the two-dimensional XY model. I start with the quantum model described by the following Hamiltonian

$$H = -J \sum_{<i,j>} \left(S_i^x S_j^x + S_i^y S_j^y\right), \tag{9.47}$$

where $<i, j>$ means next-nearest-neighbor interactions. Although the spins are constrained to lie in a plane, the quantum spin $\vec{S}$ has three components. Using the Villain representation (Villain 1974)

$$\begin{aligned} S_i^+ &= e^{i\varphi_i}\sqrt{(S+1/2)^2 - (S_i^z + 1/2)^2}, \\ S_i^- &= \sqrt{(S+1/2)^2 - (S_i^z + 1/2)^2}\, e^{-i\varphi_i}, \end{aligned} \tag{9.48}$$

where φ_i is the angle the ith spin makes with some arbitrary axis we can write equation (9.47) as (Pires 2007)

$$H = J \sum_{<i,j>} [(S_i^z)^2 - \tilde{S}^2 \cos(\phi_i - \phi_j)], \tag{9.49}$$

where $\tilde{S}^2 = S(S+1)$. I will consider only slowly varying configurations, that is, those with nearly equal adjacent angles. Expanding up to quadratic terms and using

$$\frac{\partial\varphi_n}{\partial t} = -i[\varphi_n, H] = -i2JS_n^z, \tag{9.50}$$

we get in the continuum limit

$$H = \int dxdy \left\{ \frac{1}{2}\left(\frac{\partial\varphi}{\partial t}\right)^2 + J\tilde{S}^2\left[\left(\frac{\partial\varphi}{\partial x}\right)^2 + \left(\frac{\partial\varphi}{\partial y}\right)^2\right]\right\}, \tag{9.51}$$

neglecting a constant term. The effective low energy Hamiltonian of a superfluid or a superconductor has a form similar to equation (9.51). The Lagrangian for Hamiltonian (9.51) is given by

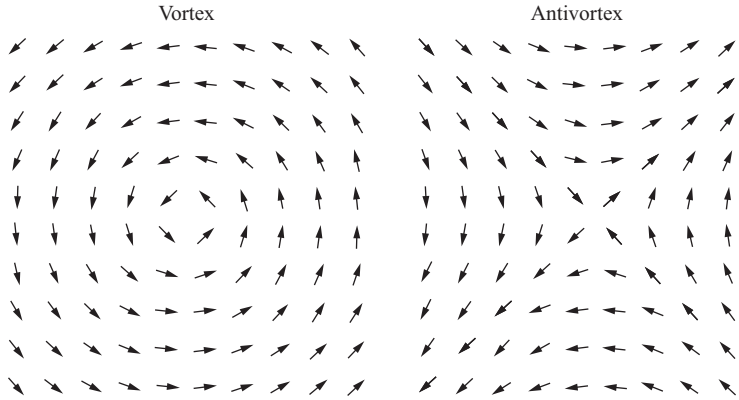


Figure 9.3. A vortex with vorticity +1 and a antivortex with vorticity −1.

$$L = \int dxdy \frac{\partial \varphi}{\partial t} + H \tag{9.52}$$

where the first term in the right-hand side of equation (9.52) is a topological term. From now on I will consider the classical case where we have a system of rotors.

The equation of motion for the classical model is

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} - \frac{1}{c^2}\frac{\partial^2 \phi}{\partial t^2} = 0. \tag{9.53}$$

This equation admits spin wave solutions; however, I will be interested in topological solutions called vortices. A static vortex solution is given by

$$\varphi = \tan^{-1}(y/x). \tag{9.54}$$

If we consider a vortex configuration, as we go around some closed path containing the center of the vortex, φ will change by 2π for each revolution of the spin. We have thus a mapping $S^1 \rightarrow S^1$, i.e. $\pi_1(S^1)$. The vorticity q is given by (figure 9.3)

$$\frac{1}{2\pi}\oint d\phi(r) = q, \quad \text{where } q = 0, \pm 1, \pm 2, \ldots. \tag{9.55}$$

For an isolated vortex the energy is

$$E \approx 2\pi J \ln(R/a), \tag{9.56}$$

where R is the radius of the system and a the lattice spacing. The energy increases logarithmically with the size of the system. Thermal generation of vortices produces an even number of vortices and antivortices. Bound pairs have lower energies than free vortices, but have low entropy. In order to minimize the free energy the system undergoes a topological transition at a critical temperature T_{BKT}, called the Berezinskii, Kosterlitz, Thouless temperature resulting from the unbinding of vortex–antivortex pairs. There is no local order parameter below T_{BKT}, and the correlation functions decay algebraically.

9.4 Theta terms

Let us address again the effect of topology in the path integral formalism using what was presented in section 4.16. In general we have (Altland and Simons 2010)

$$Z = \sum_{W \in G} \int D\phi_W e^{-S[\phi]}, \tag{9.57}$$

where G is the homotopy group and $\int D\phi_W$ is the integration over a 'topological sector' (a homotopy class defined by an element $W \in G$) of the theory. As we saw in chapter 1, there are cases when we have

$$S[\phi] = S_0[\phi] + S_{\text{top}}[\phi], \tag{9.58}$$

where $S_{\text{top}}[\phi]$ is a topological action term, that depends only on the topological class of the field ϕ. If this happens we can write

$$Z = \sum_{W \in G} e^{-F(W)} \int D\phi_W e^{-S_0[\phi]}, \tag{9.59}$$

where $F(W) \equiv S_{\text{top}}[\phi]$. We can suppose that the action is linear in the field and therefore linear in the topological index so that we can write

$$F[W] = i\theta W, \tag{9.60}$$

up to a constant. So the term $e^{-i\theta W}$ weighing the different sectors assumes the form of a phase. The topological action is sometimes called a θ-term. Theta terms are topological terms of a particular type. Essentially, they are just complex weights of different topological sectors in the path integration.

We saw in chapter 4, that the degree of a map is a topological invariant. Then we can define a general coordinate invariant representation of the θ-term as

$$S_{\text{top}}[\phi] = i\theta \int_M \phi^* \omega. \tag{9.61}$$

The reader is also referred to Altland and Simons (2010) where several examples of the θ-term are presented. Equation (9.61) is a particular example where Gurevich theorem (section 4.10) applies: there is a correspondence between homotopies (degrees of mapping) to cohomologies (integral of a form).

Up to now we have been studying maps $S^d \to S^d$. Let us increase the target space to S^{d+1}. Then all configurations become contractible $\pi_d(S^{d+1}) = 0$. It does not mean, however, that a geometric phase vanishes. It exists due a non-zero homotopy group $\tilde{\pi}_{d+1}(S^{d+1}) = Z$. To see that we extend the mapping π: $S^d \to S^{d+1}$ continuously to $\tilde{\pi}$: $D^{d+1} \to S^{d+1}$ where the base space is the boundary of the disk $S^d = \partial D^{d+1}$. We write

$$W_D[\pi] = \int_{D^{d+1}} \tilde{\pi}^* \omega^{d+1}, \tag{9.62}$$

where $\tilde{\pi}^*\omega^{d+1}$ is the pull-back of the volume form on the target space to the auxiliary disk D^{d+1}. Equation (9.62) is multiply-valued and W_D is defined modulo integer. The term

$$S_{WZ}[\pi] = i2\pi k W_D, \tag{9.63}$$

with k an integer is called a Wess–Zumino term. Under the reduction back to S^d this term converts into a θ term (Abanov 2009, Abanov and Wiegmann 2000).

References and further reading

Abanov A G 2009 http://felix.physics.sunysb.edu/~abanov/Teaching/Spring2009/Notes/abanov-cp08-upload.pdf

Abanov A G and Wiegmann P B 2000 Theta-terms in non-linear sigma models *Nucl. Phys.* B **570** 685

Affleck I 1989 Quantum spin chains and the Haldane gap *J. Phys. Condens. Matter* **1** 3047

Altland A and Simons B 2010 *Condensed Matter Field Theory* (Cambridge: Cambridge University Press)

Fradkin E 1991 *Field Theories of Condensed Matter Physics* (Redwood City: Addison-Wesley)

Girvin S M 2000 Spin and isospin: exotic order in quantum Hall ferromagnets *Phys. Today* **53** 39

Manton N and Sutcliffe P 2004 *Topological Solitons* (London: Cambridge University Press)

Pires A S T 2007 Quantum- phase transition in a XY model *Phys.* A **373** 387

Rajaraman R 1987 *Solitons and Instantons* (Amsterdam: North-Holland)

Sachdev S 2011 *Quantum Phase Transit* (Cambridge: Cambridge University Press)

Tanaka A and Takayoshi S 2015 A short guide to topological terms in the effective theories of condensed matter *Sci. Technol. Adv. Mater.* **16** 0144404

Villain J 1974 Quantum theory of one- and two-dimensional ferro and antiferromagnet *J. Phys. France* **35** 27

IOP Concise Physics

A Brief Introduction to Topology and Differential Geometry in Condensed Matter Physics

Antonio Sergio Teixeira Pires

Appendix A

Lie derivative

A.1 Integral curve

As we saw before, a vector field v on a smooth manifold M is an assignment to every point $p \in M$ of a vector $v(p) \in T_pM$. In a local coordinate system we have (Curtis and Miller 1985)

$$v(p) = v^i(p)\frac{\partial}{\partial x^i}\bigg|_p. \tag{A.1}$$

Given two vector fields u and w we can define a new vector field $[u, w]$, called the commutator of u and v, by

$$[u, w](f) = u[w(f)] - w[u(f)]. \tag{A.2}$$

In terms of components we have

$$[u, w]^i = u^j\frac{\partial w^i}{\partial x^j} - w_j\frac{\partial u^i}{\partial x^j}. \tag{A.3}$$

A *curve* on a manifold M (as we saw before) is the smooth mapping

$$\sigma: I \to M,$$

where $I \subset \Re$, and if $t \in I$ we have $\sigma(t) \in M$. The 'curve' is defined to be the map itself, not the set of image points in M (see figure A.1).

doi:10.1088/2053-2571/aaec8fch10

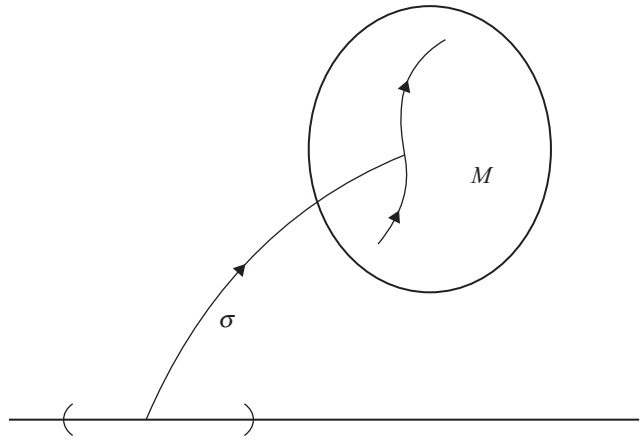


Figure A.1. A curve in the manifold M.

An *integral curve* σ of a vector field v is a curve in the manifold M whose tangent at $p = \sigma(t)$ is the vector $v(p)$ with $p \in M$. This is

$$\frac{d\sigma(t)}{dt} = v(\sigma(t)) = v^i(\sigma)\frac{\partial}{\partial x^i}. \tag{A.4}$$

Example. Let us consider in $\Re^2$ the vector field $v = y\partial/\partial x - (y + x)\partial/\partial y$ and the curve $\sigma(t) = (x(t), y(t))$. The integral curves of the vector field v satisfy the equation $(\frac{dx}{dt}, \frac{dy}{dt}) = v$. We have then $\frac{dx}{dt} = y$, $\frac{dy}{dt} = -(y + x)$, and so

$$\frac{dx^2}{dt^2} + \frac{dx}{dt} + x = 0,$$

which is the equation for the damped harmonic oscillator.

I am going to skip some theorems since they are too mathematical for our purposes. For any vector field v on the manifold M there exists a smooth map $\sigma_\cdot\colon \Re \times M \to M$ called the *flow* of v, written as $\sigma_t(p)$ with $t \in \Re$, and $p \in M$, such that $\sigma_0(p) = p$, $\sigma_t(\sigma_s(p)) = \sigma_{t+s}(p)$, and

$$\frac{d\sigma_t}{dt} = v(\sigma_t(p)). \tag{A.5}$$

The flow is a diffeomorphism with inverse σ_{-t}. Geometrically, σ_t sends each $p \in M$ to the point obtained by moving along the integral curve of v through p for a time t. Since σ_t is a map $M \to M$ we can think of the component σ_t^i with respect to the local coordinates x^i. Using equation (A.5) we get for small t

$$\sigma_t^i(p) = x^i(p) + tv^i(p) + O(t^2). \tag{A.6}$$

The set of points (t, p) with $t \in \Re$ and $p \in M$ is an open set of the space $\Re \times M$ and then a smooth manifold of dimension $n + 1$. Here I am supposing that M is a compact manifold

A.2 The Lie derivative

In general, geometric objects can be compared only if they are defined at the same point in the manifold. The geometric operation that provides the measure of the rate of change of a map is called the Lie derivative (Ebrain 2010, Friedman 2017, Harmark 2008). The simplest case is that of a function. The derivative $f: M \to \Re$ with respect to a vector field u, quantifies how much f changes along the flow of u. We use the difference between f in the point p and f in the translated point $\sigma_t(p)$:

$$\pounds_u f(p) = \lim_{t \to 0} \left[\frac{f(\sigma_t(p)) - f(p)}{t} \right], \tag{A.7}$$

which becomes

$$\pounds_u f(p) = \frac{d}{dt}\bigg|_{t=0} f(\sigma_t(p)) = u(f)\bigg|_p . \tag{A.8}$$

Let us now turn to the case of vectors. But until we provide additional information, the concept of a vector field derivative is not well defined. We want to measure the rate of change of a vector field as we move from one point to another in the manifold, which means that we are implicitly comparing tangent vectors defined at different points p and q of M. There is no unique way to do this, since tangent vectors in p are in the tangent space T_pM and tangent vectors in q are in a different space T_qM. There are several ways to define mappings between these two spaces, but there is no special or natural mapping. As we saw before, choosing a particular mapping between tangent spaces in the manifold imposes an additional structure called a connection. The derivative that uses a connection defined in M is the covariant derivative. The Lie derivative provides an alternative method, which does not require a connection, to derive vector fields and therefore apply in a general unstructured smooth manifold.

If u and v are vector fields in a variety of M the Lie derivative, which gives the measure of the rate of change of v in the direction of u, can be defined as follows.

Let $q \in M$ be defined by $q = \sigma_t(p) \equiv p(t)$, and v a tangent vector $v(p) \in T_pM$.

Note that both vectors u and v are tangent vectors at p, but only u is tangent to the curve described by the point p (t). Let M and N be smooth manifolds and f a mapping $f: M \to N$, with $p \in M$, $q \in N$. If c is a curve in M with $dc(0)/dt = u$, then $f.c$ is a curve in N. We saw in chapter 3 that $Df: T_pM \to T_qN$ takes a tangent vector in M into a tangent vector in N. Now, if $M = N$, $f: M \to M$, with $f(p) = q$ where $p, q \in M$, takes a tangent vector at p into a tangent vector at q.

In figure A.2, u is a tangent vector to the curve p (t) (not shown in the figure) and v is the vector we want to calculate the variation along the curve. We know that σ_t

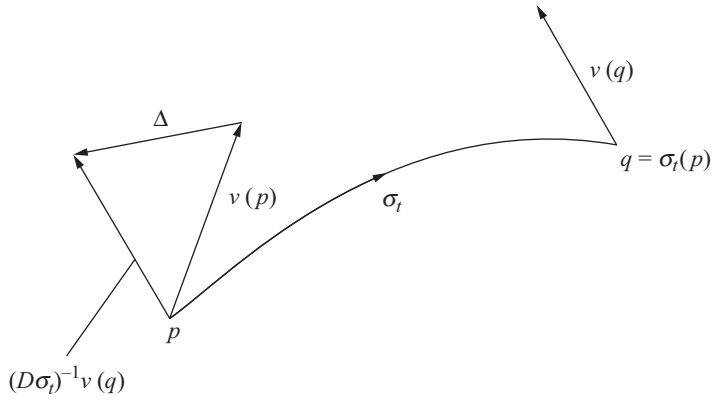


Figure A.2. Lie derivative of a vector.

takes the point p at $t = 0$ in the point $p\ (t) = q$. We have that $v(q) = v(\sigma_t(p))$ is the vector v at the point q.

We also know that $D\sigma_t$: $T_pM \to T_qN$ takes a tangent vector in p in the tangent vector in q and therefore the inverse mapping $(D\sigma_t)^{-1}$ takes the tangent vector in q in the tangent vector in p (pulls back the vector). Note that $v\ (p)$ is tangent to a curve $\tilde{c}(t)$ that passes through p, but this curve is not the curve $p(t) = \sigma_t(p)$ which u is tangent. The difference between $v\ (q)$ and $v\ (p)$ in the point p is then given by

$$\Delta = (D\sigma_t)^{-1}v(q) - v(p), \tag{A.9}$$

which leads to the following definition of the Lie derivative

$$\pounds_u v(p) = \lim_{t=0} \frac{1}{t}[(D\sigma_t)^{-1}v(\sigma_t(p)) - v(p)], \tag{A.10}$$

or

$$\pounds_u v(p) = \frac{d}{dt}\bigg|_{t=0} (D\sigma_t)^{-1}.\ v.\ \sigma_t(p). \tag{A.11}$$

The derivative $\pounds_u v(p)$ measures how $v(p)$ changes, as compared with what would happen were it simply 'dragged along' by the vector field u. Writing the Lie derivative in terms of coordinates it is easy to show that

$$\pounds_u w = [u,\ w]. \tag{A.12}$$

The Lie derivative has the following properties

(a) $\pounds_u(v + w) = \pounds_u v + \pounds_u w$.
(b) $\pounds_u(fv) = f\pounds_u v_+\ (\pounds_u f)v$.
(c) $\pounds_{[v,u]} = [\pounds_v,\ \pounds_u]$.

The Lie derivative of a 1-form α can be obtained taking a vector field v and the function $f = \alpha(v)$. We have

$$\pounds_u f = (\pounds_u \alpha).\ v + \alpha.\ \pounds_u v, \tag{A.13}$$

which gives

$$(\pounds_u \alpha).\ v = d(\alpha.\ v)u - \alpha[u, v]. \tag{A.14}$$

The Lie derivative can be generalized to tensors. In particular a Killing vector on a (pseudo-) Riemannian manifold M is a vector field u which has the property that Lie differentiation with respect to it annihilates the metric:

$$\pounds_u g = 0. \tag{A.15}$$

A.3 Interior product

Let M be a manifold of dimension n, $\omega \in \Lambda^{k+1}(M)$ a $k+1$ form and v is a vector field. We define the interior product i_v: $\Lambda^{k+1}(M) \to \Lambda^k(M)$ by

$$i_v\omega(v_1, v_2, \ldots, v_k) = \omega(v, v_1, \ldots, v_k).$$

So $i_v\omega$ is a k-form. It can be shown that

(a) $di_v + i_v d = \pounds_v$.

(b) $[\pounds_v, i_u] = i_{[v,u]}$.

References and further reading

Curtis W D and Miller F R 1985 *Differential Manifolds and Theoretical Physics* (Orlando: Academic)

Ebrain E 2010 A self-contained introduction to Lie derivatives http://web.math.ucsb.edu/~ebrahim/liederivs_tame.pdf

Friedman J L 2017 Lie derivatives, forms, densities, and integration http://ictp-saifr.org/schoolgr/Lecture0friedman.pdf

Harmark T 2008 Vector fields, flows and Lie derivatives http://nbi.dk/~harmark/killingvectors.pdf

IOP Concise Physics

Antonio Sergio Teixeira Pires

Appendix B

Complex vector spaces

The definition of a vector space presented in chapter 1, can be extended to the complex case taking the scalars as complex numbers. The scalar product is now substituted by the Hermitian product (Groecheneg 2016, Kobayashi 1987).

Let V be a vector space over the complex numbers. A Hermitian product on V is a rule which to any pair of elements v, u of V associates a complex number, denoted $\langle v, u\rangle$, satisfying the following conditions:

(a) We have $\langle v, u\rangle = \overline{\langle u, v\rangle}$ for all $v, u \in V$ (the bar denotes complex conjugate).
(b) If, u, v, $w \in V$, then $\langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle$.
(c) If $\alpha \in C$, then $\langle u, \alpha v\rangle = \alpha\langle u, v\rangle$, $\langle \alpha u, v\rangle = \overline{\alpha}\langle u, v\rangle$.

The Hermitian product is called *positive definite* if $\langle v, v\rangle \geqslant 0$ for all $v \in V$, and $\langle v, v\rangle > 0$ if $v \neq 0$.

B.1 Complex manifolds

At first it seems that the definition of a complex manifold is similar to that of a smooth manifold, replacing open subsets of $\Re^n$ by open subsets of C^n, and smooth functions by holomorphic functions. However, the complex analogous often yields more restrictions.

Let z denote any point of some neighborhood of a fixed point z_0, where the neighborhood is within the domain of definition of a function f. The derivative of f at z_0 is defined by the equation (Churchill 1960)

$$f'(z_0) = \lim_{\Delta z \to 0} \frac{f(z_0 + \Delta z) - f(z_0)}{\Delta z}, \tag{B.1}$$

where $\Delta z = z - z_0$.

doi:10.1088/2053-2571/aaec8fch11

We say the f is complex differentiable, or holomorphic, if and only if the limit exists for every $x \in U$ ($U \subset C$). The derivative of a complex differentiable function is always continuous. Every holomorphic function is infinitely many times complex differentiable.

Let M be a bounded compact n-dimensional manifold with boundary ∂M. For every smooth $(n-1)$ form ω we have the Stokes' theorem

$$\int_M d\omega = \int_{\partial M} \omega. \tag{B.2}$$

The function $f\colon U \to C$ (where $f(x, y) = u(x, y) + iv(x, y)$) is holomorphic if and only if u and v are differentiable and satisfy the Cauchy–Riemann condition:

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \qquad \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}. \tag{B.3}$$

We introduce the complex 1-form $dz = dx + idy$, and define a second complex 1-form as $\omega = f(z)dz = (u + iv)(dx + idy) = (udx - vdy) + i(udy + vdx)$. Note that dz and $d\bar{z} = dx - idy$ are linearly independent.

A complex valued function *f,* such that u and v are differentiable, satisfies the Cauchy–Riemann condition only if $\omega = f(z)dz$ satisfy $d\omega = 0$.

Proof:

$$\begin{aligned} &d[(udx - vdy) + i(udy + vdx)] \\ &= \frac{\partial u}{\partial y} dy \wedge dx - \frac{\partial v}{\partial x} dx \wedge dy + i\frac{\partial u}{\partial x} dx \wedge dy + i\frac{\partial v}{\partial y} dy \wedge dx. \end{aligned}$$

Using $dx \wedge dy = -dy \wedge dx$ we find

$$d\omega = \left[\left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right) + i\left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}\right)\right] dx \wedge dy. \tag{B.4}$$

So $d\omega = 0$ implies the vanishing of both the real and imaginary part, which corresponds to the Cauchy–Riemann condition.

Definition 1. Let X be a topological space together with an open covering $\{U_i\}$ with $i \in I$, and homeomorphisms: $\phi_i\colon U_i \to U'_i$, where $U'_i \subset C^n$ is an open subset. If for every pair (i, j) the induced mapping

$$\varphi_j \cdot \varphi_i^{-1}\colon \varphi_i(U_i \cap U_j) \to \varphi_j(U_i \cap U_j), \tag{B.5}$$

is holomorphic, we say that X is endowed with the structure of a complex manifold of complex dimension n. The pair (U_i, ϕ_i) are called charts.

The covering $X = \bigcup_i U_i$ introduces a system of locally defined complex coordinates.

Definition 2. Let X be a continuous manifold and f: $X \to C$ a continuous map. We say that f is holomorphic if for every chart (U, ϕ) as above, the composition $f \cdot \phi^{-1}$: $U' \to C$ is a holomorphic function on U'.

If X is a compact, connected, complex manifold, then every holomorphic function f: $X \to C$ is constant. This implies that a compact complex manifold cannot be embedded into any C^n.

We can think of a vector bundle $E \to X$ as a family of complex vector spaces over X. To each point $x \in X$ we associate a vector space $\pi^{-1}(X)$. The complex manifold E is the total space, X is the base, and π the structure map.

A *complex line bundle* L over a manifold M is a manifold L and a smooth mapping π: $L \to M$, such that (Murray 2016)

(1) Each fiber $\pi^{-1}(m) = L_m$ is a complex one-dimensional vector space.
(2) Every $m \in M$ has an open neighborhood $U \in M$ for which there is a diffeomorphism ϕ: $\pi^{-1}(U) \to U \times C$ such that $\phi(L_m) \in \{m\} \times C$ for every m, and the map $\phi|_{L_m}$: $L_m \to \{m\} \times C$ is an isomorphism.

A *Hermitian manifold* is a complex manifold with a smoothly varying Hermitian inner product on each (holomorphic) tangent space. We can also define a Hermitian manifold as a real manifold with a Riemannian metric that preserves a complex structure.

The *Grassmannian* Gr(m, n), which is the space of m-dimensional subspaces of an n-dimensional complex vector space C^n (with $m \leqslant n$) is relevant to the classification of topological phases of condensed matter because it is associated with n-dimensional quantum system with m occupied levels.

Let V be a real vector space. The complexification of V is defined by taking the tensor product of V with the complex numbers (thought of as a two-dimensional vector space over the real numbers):

$$V^C = V \otimes C. \tag{B.6}$$

Every vector v in V^C can be written in the form

$$v = v_1 \otimes 1 + v_2 \otimes i, \tag{B.7}$$

where $v_1, v_2 \in V$. It is common to write: $v = v_1 + v_2$.

Let V be a vector space with dimension n over the complex numbers. If $\{e_n\}$ is a basis for V and $v \in V$ we can write $v = x_1e_1 + x_2e_2 + \ldots + x_ne_n$, where x_i are complex numbers. Writing $x_i = a_i + ib_i$, where a_i and b_i are real numbers we have $v = a_1e_1 + b_1ie_1 + \ldots + a_ne_n + b_nie_n$. Then $\{e_n, ie_n\}$ is a basis for the underlying real vector space V_R of dimension $2n$.

If η is a complex vector bundle, then the underlying real bundle η_R has a canonical orientation.

Proof: Let V be a finite dimensional complex vector space with basis $\{e_n\}$ over C. As was demonstrated above the set $\{e_n, ie_n\}$ gives a real basis for V_R. This ordered basis

determines the required orientation for V_R, since if $\{e'_n\}$ is another complex basis of V, then there is a $n \times n$ complex matrix A (with $\det A \neq 0$) which transform the first basis into the second. This transformation does not alter the orientation of the real vector space, since A is the coordinate change matrix. Then the underlying $2n \times 2n$ real matrix A_R has: $\det A_R = |\det A|^2 > 0$. Hence A_R preserves the orientation of the underlying real vector space. We may apply this construction to each fiber of η to obtain the required orientation of η_R.

By the above discussion, we conclude that every complex manifold is oriented, since an orientation of the tangent bundle of a manifold induces an orientation of the manifold itself.

Let us consider a closed curve γ: $S^1 \to C/\{0\}$ in the complex plane without the zero. The winding number of γ can be expressed as the complex integral

$$w[\gamma] = \deg[\lambda] = \frac{1}{2\pi i}\int_\gamma \frac{dz}{z} = \frac{1}{2\pi i}\int_\lambda d\,\log z, \tag{B.8}$$

and $w[\gamma] \in Z$ is an integer.

B.2 Complex projective space

A complex projective space CP^n (C) is the set of lines in C^{n+1} passing through the origin. If $z = (z_0, \ldots, z_n) \neq 0$, the z determines the same line if $z = cz'$ for some complex $c \neq 0$ and they are called equivalent. CP^n (C) is a complex manifold. It can be shown that CP^1 is diffeomorphic to the sphere S^2.

B.3 Hopf bundle

The Hopf bundle describes a 3-sphere in terms of circles and ordinary spheres. Hopf found a many-to-one continuous map from the 3-sphere onto 2-sphere such that each distinct point of the 2-sphere comes from a distinct circle of the 3-sphere (Penrose 2007).

The unit 3-sphere $x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1$ in $\Re^4$ can be thought of as a 3-sphere in C^2, defined by the equation $|w|^2 + |z|^2 = 1$, where $w = x_1 + ix_2$, $z = x_3 + ix_4$. Let us consider now the space CP^1 of complex straight lines in C^2 passing through the origin. Each line is given by an equation of the form $aw + bz = 0$, where a and b are complex numbers (not both zero). This line is a copy of a complex plane, and it meets S^3 in a circle S^1, which we can think of as a unit circle in that plane. These circles are the fibers of the bundle. The different lines can meet only at the origin, so no two distinct S^1s can have a point in common. Thus, this family of S^1s constitute fibers giving S^3 a bundle structure. We can multiply a and b by the same non-zero complex number and get the same line: it is the ratio a/b that distinguishes one line from another. The space of such ratios is a Riemann sphere S^2, which we identify as the base space of the bundle. Thus the 3-sphere is realized as a disjoint union of circular fibers. The Hopf bundle is the simplest non-trivial vector bundle over the sphere S^2.

If we write $S^2 = \{(x, y, z) \in \Re^3, x^2 + y^2 + z^2 = 1\}$, $S^3 = \{(a, b, c, d) \in \Re_4$, $a^2 + b^2 + c^2 + d^2 = 1\}$, the Hopf map π: $S^3 \to S^2$ is given by

$$\pi(a, b, c, d) = [(a^2 + b^2 - c^2 - d^2), 2(bc + ad), 2(bd - ac)i]. \tag{B.9}$$

Since $\pi_3(S^2) = Z$, there is an associated integer called the Hopf invariant. This invariant cannot be the degree of the mapping, since the domain and target spaces have different dimensions. We can define the mapping as follows. Let ω denote the area 2-form on the target S^2 and let $f = \vec{n}^*\omega$ be its pull-back under $\vec{n}$ to the domain S^3 (here $\vec{n}$ is a 3D unit vector). Since ω is closed, f is also closed. The triviality of the second cohomology group of 3-spheres $H^2(S^3) = 0$ demands its pull-back to be an exact 2-form, which we write as $f = da$. The Hopf invariant is then given by integrating the Chern–Simmons 3-form over S^3

$$W = \frac{1}{4\pi^2}\int_{S^3} f \wedge a. \tag{B.10}$$

This integral is independent of the choice of a, because if $a \to a + d\alpha$, we have

$$\Delta W = \frac{1}{4\pi^2}\int_{S^3} (d(f\alpha) - (df)\alpha) = 0, \tag{B.11}$$

because $df = 0$, and by Stokes' theorem the integral of $d(f\alpha)$ is zero over a closed three-manifold.

If we consider a two-band insulator in 3D and take the Brillouin zone as a sphere S^3, we can use the same Hamiltonian used in 2D and from the mapping $\pi_3(S^2) = Z$ conclude that there are many different non-trivial phases. However, by adding a few trivial bands and using what is called the K-theory it can be shown that the topological phases disappear (Thiang 2017).

References and further reading

Churchill R V 1960 *Complex Variables and Applications* (Tokyo: McGraw-Hill)

Groecheneg M 2016 Complex manifolds http://page.mi.fu-berlin.de/groemich/complex.pdf

Kobayashi S 1987 *Differential Geometry of Complex Vector Bundles* (Princeton: Princeton University Press)

Murray M 2016 Line Bundles http://maths.adelaide.edu.au/michael.murray/line_bundles.pdf

Penrose R 2007 *The Road to Reality* (New York: Vintage Books)

Thiang G C 2017 Lectures notes on symmetries, topological phases and K-theory http://adelaide.edu.au/directory/guochuan.thiang

IOP Concise Physics

A Brief Introduction to Topology and Differential Geometry in Condensed Matter Physics

Antonio Sergio Teixeira Pires

Appendix C

Fubini–Study metric and quaternions

C.1 Fubini–Study metric

Let us consider a Hamiltonian $H(\lambda)$ that depends smoothly on the set of parameter $\lambda = (\lambda_1, \ldots, \lambda_N)$ (Cheng 2013). If $\psi(\lambda)$ is a parameter-dependent wave function we can try to define a quantum distance upon an infinitesimal variation of the parameter λ by

$$\begin{aligned} ds^2 &= \| \psi(\lambda + d\lambda) - \psi(\lambda)\|^2 = \langle \delta\psi | \delta\psi \rangle = \left\langle \partial_\mu \psi | \partial_\nu \psi \right\rangle d\lambda^\mu d\lambda^\nu \\ &= (\gamma_{\mu\nu} + i\sigma_{\mu\nu}) d\lambda^\mu d\lambda^\nu, \end{aligned} \tag{C.1}$$

where $\gamma_{\mu\nu}$ is the real and $\sigma_{\mu\nu}$ the imaginary parts of ds^2. Using the property that the inner product is Hermitian we have

$$\gamma_{\mu\nu} + i\sigma_{\mu\nu} = \gamma_{\nu\mu} - i\sigma_{\mu\nu}, \tag{C.2}$$

which gives $\gamma_{\mu\nu} = \gamma_{\nu\mu}$, and $\sigma_{\mu\nu} = -\sigma_{\nu\mu}$. Therefore, $\sigma_{\mu\nu} d\lambda^\mu d\lambda^\nu$ vanishes due to the antisymmetry of $\sigma_{\mu\nu}$ and symmetry of $d\lambda^\mu d\lambda^\nu$. We can then write

$$ds^2 = \gamma_{\mu\nu} d\lambda^\mu d\lambda^\nu. \tag{C.3}$$

However, the above expression is not gauge invariant as we can see by taking

$$|\psi'(\lambda)\rangle = e^{i\alpha(\lambda)} |\psi(\lambda)\rangle, \tag{C.4}$$

and defining

$$\left\langle \partial_\mu \psi' | \partial_\nu \psi' \right\rangle = \gamma'_{\mu\nu} + i\sigma'_{\mu\nu}. \tag{C.5}$$

doi:10.1088/2053-2571/aaec8fch12

We find

$$\gamma'_{\mu\nu} = \gamma_{\mu\nu} - \beta_\mu \partial_\nu \alpha - \beta_\nu \partial_\mu \alpha + \partial_\mu \alpha \partial_\nu \alpha, \quad \sigma'_{\mu\nu} = \sigma_{\mu\nu}, \tag{C.6}$$

where $\beta_\mu(\lambda) = i\langle\psi(\lambda)|\partial_\mu\psi(\lambda)\rangle$ is the Berry connection, which is purely real due to the normalization $\langle\psi(\lambda)|\psi(\lambda)\rangle = 1$. The Berry connection upon the above gauge transformation changes as $\beta'_\mu = \beta_\mu + \partial_\mu\alpha$. We define a gauge invariant metric using the following expression

$$g_{\mu\nu} = \gamma_{\mu\nu} - \beta_\mu(\lambda)\beta_\nu(\lambda). \tag{C.7}$$

It is easy to show that under the gauge transformation we get $g'_{\mu\nu}(\lambda) = g_{\mu\nu}(\lambda)$.

We can verify also that the covariant derivative

$$\left| D_\mu\psi \right\rangle = \left| \partial_\mu\psi \right\rangle - \left| \psi \right\rangle \left\langle \psi \middle| \partial_\mu\psi \right\rangle, \tag{C.8}$$

transforms as $|\psi\rangle$. The last term projects out parts of $|\partial_\mu\psi\rangle$ not orthogonal to $|\psi\rangle$.

The Fubini–Study metric is defined as

$$Q_{\mu\nu}(\lambda) = \left\langle \partial_\mu\psi(\lambda) \middle| \partial_\nu\psi(\lambda) \right\rangle - \left\langle \partial_\mu\psi(\lambda) \middle| \psi(\lambda) \right\rangle \left\langle \psi(\lambda) \middle| \partial_\nu\psi(\lambda) \right\rangle. \tag{C.9}$$

We define $g_{\mu\nu} = \mathrm{Re}\, Q_{\mu\nu}$, $\sigma_{\mu\nu} = \mathrm{Im}\, Q_{\mu\nu}$.

Taking the inner product of $|\psi(\lambda)\rangle$ and $|\psi(\lambda + d\lambda)\rangle$, and expanding in a Taylor series we obtain

$$\langle\psi(\lambda)|\psi(\lambda + d\lambda)\rangle = 1 + i\beta_\mu(\lambda)d\lambda^\mu + \frac{1}{2}\left\langle \psi(\lambda) \middle| \partial_\mu\partial_\nu\psi(\lambda) \right\rangle d\lambda^\mu d\lambda^\nu + \ldots \tag{C.10}$$

Using the fact that $\langle\psi|\partial\psi\rangle$ is purely imaginary we find that $\langle\partial_\mu\psi|\partial_\nu\psi\rangle + \langle\psi|\partial_\mu\partial_\nu\psi\rangle$ is also purely imaginary. We then get

$$\mathrm{Re}\left\langle \psi \middle| \partial_\mu\partial_\nu\psi \right\rangle = -\mathrm{Re}\left\langle \partial_\mu\psi \middle| \partial_\nu\psi \right\rangle = -\gamma_{\mu\nu}. \tag{C.11}$$

Equation (C.10) can then be written

$$\left|\langle\psi(\lambda)|\psi(\lambda + d\lambda)\rangle\right| = 1 - \frac{1}{2}(\gamma_{\mu\nu}(\lambda) - \beta_\mu(\lambda))d\lambda^\mu d\lambda^\nu = 1 - \frac{1}{2}g_{\mu\nu}(\lambda)d\lambda^\mu d\lambda^\nu. \tag{C.12}$$

The quantum distance between quantum states labeled by λ_I and λ_F can be written as

$$|\langle\psi(\lambda_F)|\psi(\lambda_I\rangle| = 1 - \frac{1}{2}\int_{\lambda_I}^{\lambda_F} g_{\mu\nu}(\lambda)d\lambda^\mu d\lambda^\nu \tag{C.13}$$

The last term is called the geometric quantum distance.

As an example let us consider a two-level quantum system living in C^2, with wave function

$$\psi(x) = \begin{pmatrix} \cos(\theta/2)e^{i\varphi} \\ \sin(\theta/2) \end{pmatrix}. \tag{C.14}$$

We have

$$\langle\psi|\partial_\phi\psi\rangle = i\cos^2(\theta/2), \qquad \langle\psi|\partial_\theta\psi\rangle = 0, \tag{C.15}$$

$$\langle\partial_\phi\psi|\partial_\phi\psi\rangle = \cos^2(\theta/2), \qquad \langle\partial_\phi\psi|\partial_\theta\psi\rangle = -i\frac{1}{4}\sin\theta. \tag{C.16}$$

$$\langle\partial_\theta\psi|\partial_\theta\psi\rangle = 1/4. \tag{C.17}$$

The components of the Fubini–Study metric are then given by

$$g_{\theta\theta} = \frac{1}{4}, \quad g_{\theta\varphi} = 0, \quad g_{\varphi\varphi} = \frac{1}{4}\sin^2\theta. \tag{C.18}$$

We see that the metric agrees with the standard metric on a sphere of radius 1/4.

The field associated to the connection β_μ, is given by

$$\mathfrak{F}_{\mu\nu} = \partial_\mu\beta_\nu - \partial_\nu\beta_\mu = i\big(\partial_\mu\langle\psi|\partial_\nu\psi\rangle - \partial_\nu\langle\psi|\partial_\mu\psi\rangle\big). \tag{C.19}$$

From the normalization condition $\langle\psi|\psi\rangle = 1$, we get

$$\langle\psi|\partial_\mu\psi\rangle = -\langle\partial_\mu\psi|\psi\rangle. \tag{C.20}$$

Using equation (C.20) in (C.19) and comparing with equation (C.9) we find

$$\mathfrak{F}_{\mu\nu} = i(Q_{\mu\nu} - Q_{\nu\mu}) = -2\mathrm{Im}Q_{\mu\nu} = -2\sigma_{\mu\nu}. \tag{C.21}$$

We can also write

$$Q_{\mu\nu} = g_{\mu\nu} - \frac{i}{2}\mathfrak{F}_{\mu\nu}. \tag{C.22}$$

Suppose now that that there is a large gap between the ground state $|\phi_0(\lambda)\rangle$ and the first excited state, such that transitions can be ignored. We have

$$H(\lambda)\big|\,\phi_0(\lambda)\big\rangle = E_0(\lambda)\big|\,\phi_0(\lambda)\big\rangle, \tag{C.23}$$

with

$$\langle\phi_n(\lambda)|\phi_0(\lambda)\rangle = \delta_{n0}. \tag{C.24}$$

Taking the derivative of equation (C.23) we find

$$(\partial_\mu H)\big|\,\phi_0\big\rangle + H\partial_\mu\,\big|\,\phi_0\big\rangle = E_0\partial_\mu\,\big|\,\phi_0\big\rangle. \tag{C.25}$$

Using equation (C.20) we find

$$\begin{aligned}\langle\phi_n|\phi_0\rangle &= \frac{\langle\phi_n\,|\,\partial_\mu H\,|\,\phi_0\rangle}{E_0 - E_n}, \quad \text{if } n \neq 0,\\ \langle\phi_0\,|\,\partial_\mu H\,|\,\phi_0\rangle &= \partial_\mu E_0.\end{aligned} \tag{C.26}$$

On the ground state we have

$$\begin{aligned} Q_{\mu\nu} &= \langle \partial_\mu \phi_0 \left| \left(1 - \left| \phi_0 \right\rangle \langle \phi_0 \right| \right) \right| \partial_\mu \phi_0 \rangle = \sum_{n \neq 0} \langle \partial_\mu \phi_0 | \phi_n \rangle \langle \phi_n | \partial_\nu \phi_0 \rangle \\ &= \sum_{n \neq 0} \frac{\langle \phi_0 \left| \partial_\mu H \right| \phi_n \rangle \langle \phi_n \left| \partial_\nu H \right| \phi_0 \rangle}{(E_0 - E_n)^2}. \end{aligned} \tag{C.27}$$

Note. Sometimes the distance between two states $|\psi\rangle$ and $|\phi\rangle$ (not normalized) is presented in the literature as

$$\gamma(\psi, \phi) = \arccos \sqrt{\frac{\langle \psi | \phi \rangle \langle \phi | \psi \rangle}{\langle \psi | \psi \rangle \langle \phi | \phi \rangle}}. \tag{C.28}$$

Taking $\phi = \psi + \delta\psi$, and using the expansion $\cos \sqrt{ds^2} = 1 - \frac{1}{2} ds^2$, we get the former result for the Fubini–Study metric.

C.2 Quaternions

Quaternions have being used in the study of Landau levels in topological insulators in three dimensions (Li and Wu 2013). It is appropriate, therefore, that I comment on this subject briefly here.

In the 19th-century, the Irish mathematician William Rowan Hamilton (1805–65) generalized the complex numbers to a four-dimensional space, with the imaginary basis-vectors extended from one (i) to three (i, j, k) with the following property (Penrose 2007)

$$\begin{gathered} i^2 = j^2 = k^2 = ijk = -1 \\ ij = -ji = k, \quad jk = -kj = i, \quad ki = -ik = j. \end{gathered} \tag{C.29}$$

A quaternion q can be written as

$$q = q_0 + iq_1 + jq_2 + kq_3 \tag{C.30}$$

where q_i are real numbers.

Quaternions satisfy the commutative and associative laws of addition and the distributive laws of multiplication over addition, namely

$$\begin{gathered} a + b = b + a, \qquad a + (b + c) = (a + b) + c \\ a(bc) = (ab)c, \qquad a(b + c) = ab + ac, \qquad (a + b)c = ac + BC \end{gathered}$$

together with the existence of additive and multiplicative 'identity elements' 0, and 1, such that

$$a + 0 = a, \qquad 1a = a1 = a$$

we define $q^* = q_0 - iq_1 - jq_2 - kq_3$ and the norm of q is given by

$$|q| = \sqrt{qq^*} = \sqrt{q_0^2 + q_1^2 + q_2^2 + q_3^2}. \tag{C.31}$$

For each non-zero quaternion q, there is an inverse q^{-1} that satisfies

$$q^{-1}q = qq^{-1} = 1.$$

We have

$$q^{-1} = q^{*}(qq^{*})^{-1}.$$

The real number qq^{*} cannot vanish unless $q = 0$.

If $q_1 = (x_1, y_1, z_1, w_1)$ and $q_2 = (x_2, y_2, z_2, w_2)$, we have

$q_1q_2 = x_1x_2 - y_1y_2 - z_1z_2 - w_1w_2 + x_1y_2 + y_1x_2 + z_1w_2 - w_1z_2$
$+ x_1z_2 + z_1x_2 + w_1y_2 - y_1w_2 + x_1w_2 + w_1x_2 + y_1z_2 - z_1y_2.$

We can define vector spaces over quaternions, however, there is no satisfactory quaternionic analogous of the notion of a holomorphic function.

References and further reading

Cheng R 2013 *Quantum geometric tensor (Fubini–Study metric) in simple quantum system: a pedagogical introduction* (arXiv:1012.1337v2)

Li W and Wu C 2013 High-dimensional topological insulators with quaternionic analytic Landau levels *Phys. Rev. Lett.* **110** 216802

Penrose R 2007 *The Road to Reality* (New York: Vintage Books)

IOP Concise Physics

A Brief Introduction to Topology and Differential Geometry in Condensed Matter Physics

Antonio Sergio Teixeira Pires

Appendix D

K-theory

Before presenting the *K*-theory I will introduce a few concepts.

D.1 Rings

A ring is a set S, together with two operations $(x, y) \to x + y$ and $(x, y) \to xy$ satisfying

(1) S is a commutative group under the operation $(x, y) \to x + y$. (This is, S is a commutative group under addition.)
(2) $(xy)z = x(yz)$.
(3) $x(y + z) = xy + xz$; $(y + z)x = yx + zx$.

If $xy = yx$ for all x and y in S, we say that the ring S is commutative. If there is an element 1 in S such that $1x = x1$ for each x, we say that S is a ring with identity, and 1 is called the identity for S.

The set of integers, with the usual operations, is a commutative ring with identity.

D.2 Equivalence relations

An equivalence relation on a set X is a relation that is reflexive, symmetric and transitive. That is, for all a, b and c in X we have

(1) $a \sim a$,
(2) $a \sim b$ if and only if $b \sim a$,
(3) if $a \sim b$ and $b \sim c$ then $a \sim c$.

The equivalence class of a under $\sim$, denoted $[a]$, is defined as $[a] = \{b \in X | a \sim b\}$.

doi:10.1088/2053-2571/aaec8fch13

Examples:
- 'equal to' on the set of numbers,
- 'has the same absolute value' on the set of real numbers,
- 'has the same birthday as' on the set of all people.

D.3 Sum of vector bundles

Given two vectors bundles π_1: $E_1 \to X$, π_2: $E_2 \to X$, over the same base space X, we can define a third vector bundle over X whose fiber over each point of X is the direct sum of the fibers of E_1 and E_2 over this point. Thus, we define the direct sum of E_1 and E_2 as the space

$$E_1 \oplus E_2 = \{(v_1, v_2) \in E_1 \times E_2 | \pi_1(v_1) = \pi_2(v_2)\}. \tag{D.1}$$

There is then a projection $E_1 \oplus E_2 \to X$ sending (v_1, v_2) to the point $\pi_1(v_1) = \pi_2(v_2)$. The fibers of this projection are the direct sum of the fibers of E_1 and E_2.

D.4 *K*-theory

The classification of all the different vector bundles over a given base space with fibers of a given dimension is in general very difficult. In the absence of a full classification, there are two directions we can take to make some partial progress on the problem. One way is to look for invariants which distinguish at least some of the different vector bundles. I have discussed this topic in the text. Another procedure is to look for a rough classification using a weaker equivalence relation instead of the notion of isomorphism. In this section, I will present a brief discussion of this approach. I will follow mainly Eguchi *et al* (1980), Karoubi (1978) and Hatcher (2017) and the reader is referred to these and other references mentioned here.

The so-called K-theory is a powerful tool, in some ways more powerful than ordinary cohomology to compare vector bundles. Here, I will treat only the complex K-theory which is somewhat simpler than the real K-theory and suppose always that the base space is compact. In the following, I will suppose that the fibers of a vector bundle π: $E \to X$ are vector spaces of different dimensions, but assume local trivializations

$$\pi^{-1}(U) \to U \times \mathbb{C}^n,$$

where U is an open set in the base space X. The dimensions of the fibers are locally constant (however, it is not so globally if X is disconnected).

Suppose now that we have the following relation

$$E_1 \oplus F \approx E_2 \oplus F. \tag{D.2}$$

At first sight we can think that we could introduce a formal difference operation which would allow us to cancel F from both sides and get $E_1 \approx E_2$. However, the cancellation does not hold for vector bundles in general. For instance, if $E = TS^2$ and $F = S^2 \times \Re^2$, F is trivial, whereas E, the tangent bundle of the sphere S^2, is non-trivial. Adding NS^2, the bundle of normal vectors to S^2, to both bundles E and F, we obtain the same trivial bundle $S^2 \times \Re^3$. That is

$$E \oplus NS^2 = S^2 \times \Re^3, \;\; F \oplus NS^2 = S^2 \times \Re^3, \;\; \text{so } E \oplus NS^2 = F \oplus NS^2. \quad \text{(D.3)}$$

If we could perform the formal cancellation we would get $E = F$, which is false.

We can resolve the problem with the formal differences of vector bundle isomorphism using the notion of stable equivalence. In the following, I will write the trivial n-dimensional vector bundle over a fixed base space X as I^n. If E_1 and E_2 are two vector bundles over X (not necessarily of the same dimensions) we say that E_1 and E_2 are stably equivalent (or stably isomorphic) and write $E_1 \underset{s}{\approx} E_2$ if $E_1 \oplus I^n \approx E_2 \oplus I^n$ for some n. In the same way we write $E_1 \underset{s}{\sim} E_2$ if $E_1 \oplus I^n \approx E_2 \oplus I^m$ for some m and n. Both $\underset{s}{\approx}$ and $\underset{s}{\sim}$ are equivalence relations. Equivalence classes of either sort of operation of direct sum are well-defined, commutative and associative. A zero element is the class of I°. Taking the sum with trivial bundles serves to eliminate pathologies arising from low fiber dimensions. It can be shown that:

If X is compact, then the set of $\sim$-equivalence classes of vector bundles over X forms an abelian group with respect to $\oplus$. This group is called $\tilde{K}(X)$.

If E is a vector bundle, we can always find a complementary bundle F such that $E \oplus F \approx I^m$ is trivial for some m. For the direct sum operation $\underset{s}{\approx}$ equivalence classes, only the class I°, the zero element, can have an inverse since

$$E_1 \oplus E_2 \underset{s}{\approx} I^0 \rightarrow E_1 \oplus E_2 \oplus I^n \approx I^n, \quad \text{(D.4)}$$

for some n, which happens only if E_1 and E_2 has 0-dimensions. However, although there is no inverse, we have the cancellation property:

$$E_1 \oplus E_2 \underset{s}{\approx} E_1 \oplus E_3 \rightarrow E_2 \underset{s}{\approx} E_3, \quad \text{(D.5)}$$

since we can add to both sides of $E_1 \oplus E_2 \underset{s}{\approx} E_1 \oplus E_3$ a bundle E'_1 such that $E_1 \oplus E'_1 \approx I^n$ for some n.

For a compact base manifold X, we can construct an abelian group $K(X)$ consisting of formal differences $E - E'$ of vector bundles E and E' over X, with the equivalence relation

$$E_1 - E'_1 = E_2 - E'_2, \quad \text{(D.6)}$$

if and only if $E_1 \oplus E'_2 \underset{s}{\approx} E_2 \oplus E'_1$. The addition rule is

$$(E_1 - E'_1) + (E_2 - E'_2) = E_1 \oplus E_2 - (E'_1 \oplus E'_2). \quad \text{(D.7)}$$

The zero element is the equivalence class of $E - E$ for any E, and the inverse of $E - E'$ is $E' - E$. There is a homomorphism $K(X) \rightarrow \tilde{K}(X)$ sending $E - I^n$ to the $\sim$ class of E. This is well-defined since if $E - I^n = E' - I^m$ in $K(X)$. Then

$$E \oplus I^m \underset{s}{\approx} E' \oplus I^n, \quad \text{(D.8)}$$

which leads to $E \sim E'$.

Using the tensor product, we can define a natural multiplication for vector bundles. If E_1 and E_2 represent elements of $K(X)$, then products in $K(X)$ will be represented by the bundle $E_1 \otimes E_2$. Thus, for arbitrary elements of $K(X)$ represented by differences of vector bundles, we define their product in $K(X)$ as

$$(E_1 - E'_1)(E_2 - E'_2) = E_1 \otimes E_2 - E_1 \otimes E'_2 - E'_1 \otimes E_2 + E'_1 \otimes E'_2.$$

This makes $K(X)$ a commutative ring with identity I^1.

To summarize: we have obtained a weaker notion of isomorphism of vector bundles by defining two vector bundles over the same base space X to be stably isomorphic if they become isomorphic after direct sum with I^m for some m (where m can be different for the two bundles).

If X are spheres, the so-called Bott Periodicity asserts that $\tilde{K}(S^m)$ is Z for n even, and 0 for n odd, so we have a period two.

(*Note*: For the real case, the group is named $K\tilde{O}(X)$ and $K\tilde{O}(S^n)$ has period eight.)

The K-theory has been used to classify topological insulators (Zhao 2013), since adding an extra flat band to an insulator does not affect its topological classification.

References and further reading

Eguchi T, Gilkey P B and Hanson A J 1980 Gravitation, gauge theories and differential geometry *Phys. Rep.* **66** 213

Hatcher A 2017 Vector bundles and *K*-theory http://pi.math.cornell.edu/~hatcher/VBKT/VB.pdf

Karoubi M 1978 *K-Theory* (Heidelberg: Springer)

Zhao C 2013 *The topological properties in topological insulators and superconductors* (arXiv:1309.2056)