Session 2: Topologically non-trivial gauge fields

Gerard Valentí-Rojas[1,2 ⋆†]

1 Naquidis Center, Institut d'Optique Graduate School (IOGS), 91127, Palaiseau, France
2 Institut de Mathématiques de Bordeaux (IMB), UMR 5251, Université de Bordeaux, France

⋆ gerard.valenti-i-rojas@institutoptique.fr , † gvalenti.physics@gmail.com

Abstract

First draft of the notes for Session 2. Any feedback is appreciated as well as reporting typos. Good references for this Session are [1–8]. The material here provides the very first steps of incredibly vast topics in Physics and Mathematics. I have not attempted to be exhaustive, but comprehensive. Thus, the effort is put in linking key ideas that might have broad applications and generalisation.

Contents

The goal of this session is to study slightly non-trivial gauge fields and their associated theories. Mostly, we deal with non-trivial topology and its effects on Physics. We introduce the concept of geometric phases and observe a topological instance of a Berry phase in the Aharonov-Bohm effect. We also explore the natural appearance of Chern-Simons theory and its link with statistical transmutation.

1 Crash Course on Topology in Physics

Topology is often introduced in Physics as a branch of Mathematics dealing not with shape, but with the overall structure of spaces, their deformation and the immutability of some properties. This is often illustrated by the famous example of a coffee cup being continuously deformed into a doughnut, which makes the two objects topologically equivalent. Sometimes one is left with the feeling of not understanding the implications of this example. Hence, let me start rather differently.

> *Geometry is local and quantitative.*
> *Topology is global and qualitative.*

Topology in Physics is very much related to counting in integers. There are only a few things that do not vary continuously, the number of holes in a surface is one, knots in a string is another. Similarly, we can count the number of singularities, twists, branch-cuts, or discontinuities of a function in a given space. Let us refer to these collectively as *defects* when discussing Physics and *holes* when discussing Mathematics. We can classify objects according to the number and type of defects they have. We then say that the coffee cup and the doughnut are of the same type, while the circle and a disk are not. Furthermore, we can define closed paths (i.e. loops) enclosing these defects and count the number of times one can circulate around them. Altogether allows us to classify objects according to a set of *topological invariant* quantities and define equivalence classes. Two families of these invariants are: (i) Homotopy groups $\pi_n(X)$, classifying topologically equivalent *mappings* and; (ii) Homology $H_p(x)$ and Cohomology $H^p(X)$ groups, classifying topologically equivalent *manifolds*. Homotopy deals with the connectedness of paths, their continuous deformation and their relation to holes. Homology and Cohomology deal with the definition of those holes as well as the connection of spaces and their boundaries. In the following we will make these notions in (i) precise. Despite their importance, for conciseness, we will not cover (ii) in this piece of work. Once again, we refer the interested reader to the canonical texts [1–3].

> **Topological Equivalence**
>
> Homeomorphisms define equivalence relations between topological spaces $X_1 \sim X_2$.
> If connected by an homeomorphism, two spaces are topologically equivalent. Topological invariants are quantities that do not change under homeomorphisms. Three fundamental topological invariants are *(i) dimensionality*, *(ii) compactness* and *(iii) connectedness*.

A *manifold* X is a topological space that locally looks Euclidean, meaning it is *locally homeomorphic* to $\mathbb{R}^d$ or $\mathbb{C}^d$, but not necessarily globally. In more familiar terms, manifolds are the generalisation of the idea of curves and surfaces to arbitrary dimensional objects. In the vicinity of each point in X one can define a local coordinate system, so that a point is represented by a set of numbers $x = (x^1, \dots, x^d)$. An *homeomorphism* defines the mathematical equivalence of “being able to deform an object into another continuously”. That is, there is a map $f : X \longrightarrow Y$ that has an inverse $f^{-1} : Y \longrightarrow X$ and both are *continuous*. In this way, the coffee cup is homeomorphic (e.g. equivalent) to the doughnut. This allows defining equivalence classes connected by homeomorphisms. Then, *topological invariants* are quantities allowing to identify non-homeomorphic spaces, and are conserved under homeomorphisms. A *diffeomorphism* is a similar map but in which the requirement is *not* for it to be continuous, but *differentiable*. A *homotopy type* is similar to an homeomorphism but without the need for an inverse.

Imagine now a a punctured plane and two loops α and β defined as closed paths on its surface. If α encircles the hole and β does not, we say α and β are homotopically different because we cannot continuously deform one into the other, however β is homotopic (e.g. deformable) to a point as it can be *retracted* to it (see Figure 1a). A *Möbius strip* and a normal closed strip are homotopic, as they both can be deformed into a circle, but they are not homeomorphic. A circle S^1 and a punctured plane $\mathbb{R}^2 - \{0\}$ — and more generally S^{d-1} and $\mathbb{R}^d - \{0\}$ — are homotopically equivalent.

An equivalence class of loops $[\alpha]$ is the homotopy class of α. The set of homotopy classes of loops at $x \in X$ is denoted by $\pi_1(X, x)$ and known as the *fundamental* or *first homotopy group* of X at x. The basic idea of the fundamental group is to test and assess connectedness by considering loops in space up to continuous deformation. Hence, a space X is *simply connected* if it has no holes, meaning it is path-connected for every $x \in X$, and thus $\pi_1(X, x) = 0$. We then, see that $\pi_1(\mathbb{R}^d) = \pi_1(\mathbb{C}^d) = \pi_1(D^d) = 0$ since they have no holes but $\pi_1(S^1) = \mathbb{Z}$ since the circle has a hole and one can wrap a line around it an integer number of times; think of wrapping a rubber band on a finger. For a torus, we find $\pi_1(T^1) = \pi_1(S^1 \times S^1) = \mathbb{Z} \oplus \mathbb{Z}$ as one can find two independent types of non-contractible circles on its surface.

One can generalise this scheme by increasing the dimensionality of the "wrapping object", so $\pi_n(X, x)$ known as the *n-th homotopy group* is to be thought as the number of times one can non-trivially wrap an n-dimensional object on a d-dimensional space. Wrapping a circle on a sphere is trivial $\pi_1(S^2) = 0$, similar to wrapping a sphere on a circle $\pi_2(S^1) = 0$. However, wrapping a sphere on a sphere is $\pi_2(S^2) = \mathbb{Z}$. Finally, wrapping a 3-sphere on a 2-sphere is highly non-trivially $\pi_3(S^2) = \mathbb{Z}$ and is known as the Hopf invariant, while the corresponding mapping is the *Hopf fibration*. In general, $\pi_n(S^d) = \mathbb{Z}$ for $n = d$, and the associated topological invariant in a Physical setup is, given a map $\vec{\pi} : S^d \longrightarrow S^d$, where field $\vec{\pi}(\mathbf{x}) = (\pi^0, \pi^1, \ldots, \pi^d) \in S^d$, so $|\vec{\pi}|^2 = 1$. We find the topological charge to be

$$Q = \frac{1}{\Omega_d} \int_{S^d} d^d\mathbf{x}\, \epsilon^{\mu_1 \ldots \mu_n}\, \pi^0(\mathbf{x})\, \partial_{\mu_1} \pi^1(\mathbf{x}) \ldots \partial_{\mu_d} \pi^d(\mathbf{x})\,, \tag{1}$$

where Ω_d is a normalisation constant related to the d-dimensional volume element. Homotopy can also be combined with our knowledge of fibre bundles, so that now not only can we determine local geometric properties of a gauge connection, but also global topological properties of the bundle by computing the relevant homotopy invariant of the corresponding Lie group, for instance $\pi_3\big(SU(2)\big) = \mathbb{Z}$, which signals the degree of *twisting* in the fibres.

1.1 My First Topological Defect

Let us consider a velocity field of some fluid expressed as the gradient of a scalar field $\mathbf{v} = \boldsymbol{\nabla}\theta(\mathbf{x})$ or, in mathematical language, as a mapping $\mathbf{v} : X \longrightarrow M$. Assume there is a defect in the fluid, in particular a vortex line. We can define a circle $X = S^1$ around the defect. The order parameter space is $M = \mathrm{U}(1)$, so the associated homotopy group is $\pi_1(M) = \mathbb{Z}$. In a way, what we are saying is that if we take the circulation of the velocity field around the defect

$$\omega := \oint_{S^1} \mathbf{v}(\mathbf{x}) = \oint_{S^1} \boldsymbol{\nabla}\theta(r, \varphi) = 2\pi\big[\theta(\varphi = 2\pi) - \theta(\varphi = 0)\big] = 2\pi N \tag{2}$$

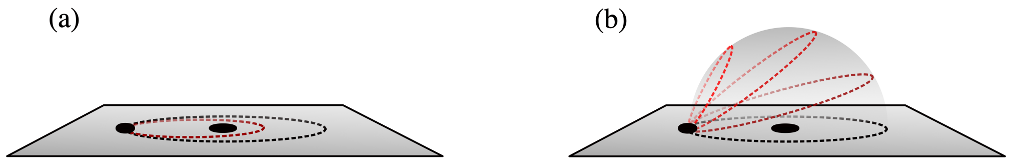

Figure 1: Deformation retraction of a closed contour to a point for two topologically distinct scenarios.

its values are quantised in terms of an integer known as the *winding number* $N \in \mathbb{Z}$, which counts the number of wrappings in field space. This is equivalent to saying that $\theta(\mathbf{x})$ is multi-valued, and thus, that it is not irrotational $\nabla \times \nabla\theta \neq 0$. In fact, it is ill-defined at $\mathbf{x} = 0$.

The idea discussed above can be applied to gauge fields, or more generally, gauge forms with group structure U(1). If we consider a gauge connection 1-form $A = A_i\, dx^i$ and define the curvature 2-form $F = \mathrm{d}A$, we can integrate the 2-form over a 2-chain $\int_{\mathbb{R}^2} F(x) = 2\pi c_1$. This roughly corresponds to the magnetic flux. Now, we can imagine a particular instance of a *locally flat* gauge connection $F = 0$, such as in the Aharonov-Bohm effect. By Poincaré's lemma we know we can write $A = \mathrm{d}\theta$, so that $c_1 = \mathbb{Z}$ is a topological invariant known as the *first Chern number*. See Appendix A for more details.

> The physics connection comes from understanding that "integrating" is, to some extent, "wrapping". Integrating matter or gauge fields that contain defects at a given location x often leads to *topological quantisation*. Furthermore, one can classify types of defects, field configurations and even phases of matter based on these criteria [7, 9, 10].

1.2 To Remove or Not To Remove... Gauge Fields

A gauge transformation connects locally different, but equivalent, gauge fields $A_\mu(x)$ and $A'_\mu(x)$ giving rise to the same curvature or field strength $F_{\mu\nu}(x)$. In practice, this amounts to adding a total derivative to the gauge field to find a gauge-transformed configuration without altering the Physics. Certain configurations or gauge orbits are equivalent — or can be deformed — to the identity. For an Abelian 1-form gauge field in $D = d + 1$ spacetime dimensions, a conventional gauge transformation is given by

$$A_\mu(x) \quad\longrightarrow\quad A'_\mu(x) = \mathcal{U}(x)A_\mu(x)\mathcal{U}^{-1}(x) - \frac{i}{g}[\partial_\mu \mathcal{U}(x)]\,\mathcal{U}^{-1}(x) = A_\mu(x) + \frac{1}{g}\,\partial_\mu \xi(x) \tag{3}$$

where $\mu = 0, \ldots, d$. We find the field strength $F_{\mu\nu}(x) = \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x) = F'_{\mu\nu}(x)$ to be gauge-invariant. Furthermore, matter couples to gauge fields via a gauge-covariant derivative, which transforms like

$$D_\mu \Psi(x) \quad\longrightarrow\quad D'_\mu \Psi'(x) = \left[\partial_\mu - igA'_\mu(x)\right]\Psi'(x) = \mathcal{U}(x)\left[\partial_\mu - igA_\mu(x)\right]\Psi(x) \tag{4}$$

where the matter field transforms like $\Psi'(x) = \mathcal{U}(x)\Psi(x)$, while for the conjugate field

$$\bar{\Psi}'(x) = \bar{\Psi}(x)\mathcal{U}^\dagger(x)\,. \tag{5}$$

Now, since the gauge transformation is given by a unitary matrix $\mathcal{U}^{-1}(x) = \mathcal{U}^\dagger(x)$, kinetic terms in the associated Lagrangians and Hamiltonians are invariant to gauge transformations.

If the gauge field can be expressed in the form[1]

$$A_\mu(x) = -i\left[\partial_\mu \tilde{\mathcal{U}}(x)\right]\tilde{\mathcal{U}}^{-1}(x) = \partial_\mu \phi(x) \tag{6}$$

where $\phi(x)$ is some well-defined scalar field, A_μ is referred to as a topologically trivial *pure gauge* configuration. The associated curvature is $F_{\mu\nu} = 0$, so it does not give rise to field strengths — e.g. electromagnetic fields — and the connection is said to be flat. In addition, it can be gauge transformed to $A'_\mu(x) = 0$. That is, it can be *removed* or *trivialised* via a gauge transformation.

Imagine now that the field $\phi(x)$ is not well-defined everywhere. It, for instance, can have a singularity at some spacetime point $x = x_0$, which we signal with the notation $\phi_s(x; x_0)$. Then, the above mechanism still holds everywhere except for precisely at x_0 as the gauge field and transformation become ill-defined. Here, we say that the gauge field is locally, but not globally, a pure gauge, or that the connection is non-trivially flat. The gauge field becomes "topological" as a consequence of this ill-defined point being a seed for topological invariants, i.e. countable objects to wrap around. More generally, defects cannot be naïvely lifted, so they provide *obstructions* to the removal of the gauge field. Fully removing these gauge fields requires a *large gauge transformation* parametrised by $\mathcal{W}(x)$, which does not connect physically equivalent states, but configurations with different homotopic properties yielding the same observable consequences. This means that matter fields transform like

$$\Psi'(x) = \mathcal{W}(x)\Psi(x)\,, \tag{7}$$

but $\mathcal{W}(x) \equiv \mathcal{W}(x; x_0) = e^{i\phi_s(x;x_0)}$ now carries defects. This has important implications as these features are transferred to the transformed matter field $\Psi'(x) \equiv \Psi'(x, x_0)$, implying that such a gauge transformation has the ability to add (or remove) defects. Hence, even though we believe we remove topologically non-trivial pure gauge fields from a system, this comes at a cost of altering the matter field, so they are not truly *removed* but *hidden*. This redefinition of the matter field can also be thought of as a *dressing* process. This type of transformation is recurrently found in literature, not as a gauge transformation, but as a mere identity or mapping of matter fields. In this context, $\mathcal{W}(x)$ might be referred to as a *disorder field* [11]. For quantised theories, it is promoted to a *disorder operator*. Depending on the specific type of defect it contains, it is typically known as a soliton, kink, vortex or monopole operator, amongst others. If $\Psi(x)$ is an order parameter or operator, we say that the mapping has an order-disorder operator structure [12–14]. This is a recurrent structure in Bose-Fermi dualities [15]. The reason is relatively simple and rarely spelled out in literature. If we consider the equal-time (anti)commutation relations $[\bullet,\bullet]_\mp$ of initially bosonic (fermionic) matter fields, we can apply transformation (7) to obtain

$$\left[\hat{\Psi}'(x), \hat{\Psi}'(x')\right]_\mp = \hat{\mathcal{W}}(x)\left[\hat{\Psi}(x), \hat{\mathcal{W}}(x')\right]_- \hat{\Psi}(x') + \hat{\mathcal{W}}(x)\hat{\mathcal{W}}(x')\left[\hat{\Psi}(x), \hat{\Psi}(x')\right]_- \tag{8}$$

$$+\left[\hat{\mathcal{W}}(x), \hat{\mathcal{W}}(x')\right]_- \hat{\Psi}(x')\hat{\Psi}(x) + \hat{\mathcal{W}}(x')\left[\hat{\mathcal{W}}(x), \hat{\Psi}(x')\right]_\mp \hat{\Psi}(x) \tag{9}$$

and

$$\left[\hat{\Psi}'(x), \hat{\Psi}'^{\dagger}(x')\right]_\mp = \hat{\mathcal{W}}(x)\left[\hat{\Psi}(x), \hat{\Psi}^{\dagger}(x')\right]_- \hat{\mathcal{W}}^\dagger(x') + \hat{\mathcal{W}}(x)\hat{\Psi}^{\dagger}(x')\left[\hat{\Psi}(x), \hat{\mathcal{W}}^\dagger(x')\right]_- \tag{10}$$

$$+\left[\hat{\mathcal{W}}(x), \hat{\Psi}^{\dagger}(x')\right]_- \hat{\mathcal{W}}^\dagger(x')\hat{\Psi}(x) + \hat{\Psi}^{\dagger}(x')\left[\hat{\mathcal{W}}(x), \hat{\mathcal{W}}^\dagger(x')\right]_\mp \hat{\Psi}(x)\,. \tag{11}$$

Therefore, we observe that the statistics of the original fields are transformed and depend on the statistics of the disorder operator. We also notice that these reduce to the usual relations

[1] Here $\tilde{\mathcal{U}}$ is simply used to indicate that it is a particular case of the general case $\mathcal{U}$. This is merely a stylistic point.

when $\hat{\mathcal{W}}(x) = \hat{\mathbb{I}}$. If the disorder operators trivially commute with themselves and with the matter fields, the original commutation relations are preserved up to re-scaling of the matter fields.

1.3 Bose-Fermi Statistical Transmutation

Inspired by our knowledge on disorder operators, we notice that certain canonical transformations can alter the statistical properties of quantum fields. In particular, these transformations can be large gauge transformations but, for now, we consider them as some mathematical mapping between fields. Let us compute a particular example for two particles in first quantised formulation. This amounts to finding a transformation $\mathcal{W}(x_1, x_2) = \exp[i\phi(x_1, x_2)]$, so that the transformed two-body wave function reads

$$\Psi'(x_1, x_2) = e^{i\phi(x_1,x_2)}\Psi(x_1, x_2) = \pm e^{i\phi(x_1,x_2)}\Psi(x_2, x_1)\,. \tag{12}$$

Now, upon exchange of particle position labels we might consider some properties for the scalar field; for instance, symmetry under exchange $\phi(x_1, x_2) = \phi(x_2, x_1)$. The previous expression then becomes

$$\Psi'(x_1, x_2) = \pm e^{i\phi(x_2,x_1)}\Psi(x_2, x_1) = \pm\Psi'(x_2, x_1)\,, \tag{13}$$

so we recover the original symmetry under exchange of particles. Nevertheless, we could have considered a slightly more exotic property for the scalar field such as $\phi(x_1, x_2) = \alpha + \phi(x_2, x_1)$ with $\alpha \in [0, 2\pi)$. We then see that Eq. (12) becomes

$$\Psi'(x_1, x_2) = \pm e^{i[\alpha+\phi(x_2,x_1)]}\Psi(x_2, x_1) = \pm e^{i\alpha}\Psi'(x_2, x_1)\,, \tag{14}$$

where the transformed field picks up an additional phase α with respect to the original symmetry properties under exchange of the two-body wavefunction. In particular, for $\alpha = \pi$ this reverses the statistical properties, so an originally bosonic two-body state transforms into a two-body fermionic state and *vice versa*. More profoundly, a bare field transforms into a composite one with altered — e.g. anyonic — statistics. A subclass of the possible values that the transformed or composite field can acquire are the standard bosonic and fermionic ones.

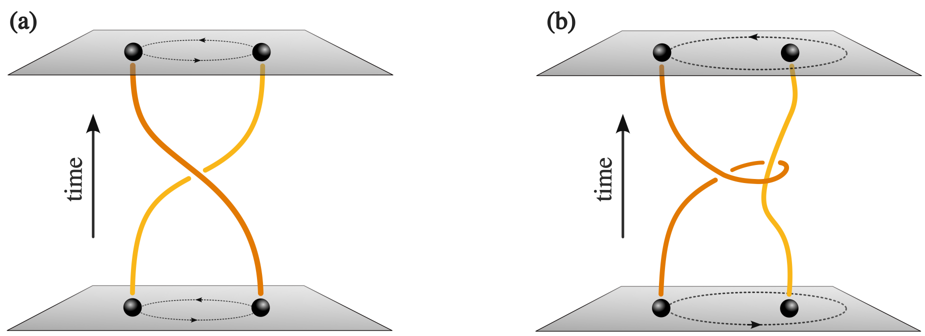

Figure 2: Worldline brading of particles. Encircling (b) is equivalent to a double exchange 2× (a).

For consistency, we might want to check the exchange properties when $x_1 = x_2$ such that, under exchange of labels, the field returns the identity. For a property of the form $\phi(x_1, x_2) = \alpha + \phi(x_2, x_1)$ when $x_1 = x_2$, the associated phase factor is the same if $\alpha = 2\pi N$ for $N \in \mathbb{Z}$. More generally, ϕ can be ill-defined at this point while retaining the desired structure elsewhere, i.e. when $x_1 \neq x_2$. This implies a delicate treatment should be taken in order to determine whether or not there is a Pauli exclusion principle. Examples of fields satisfying these conditions are, once again, topological defects with singularities at $x_1 = x_2$. For instance, in 2d, a non-trivial function can be a vortex $\phi(x_1, x_2) = \arg(\mathbf{x}_1 - \mathbf{x}_2)$. This corresponds to an Aharonov-Bohm-type vector potential, which is not surprising taking into account it gives rise to flux attachment and statistical transmutation. In the present context we see that this just comes in the explicit functional form of a scalar field $\phi(x)$, while the rest of the discussion is rather general. At this point, we believe to be on firm ground to motivate a framework that brings the above ideas together.

2 Geometric Phases & Non-integrable Phase Factors

An *holonomy* is the failure of a parametrised system to return to its initial state after moving in a closed circuit in parameter space. This failure can be caused by a non-trivial geometry (i.e. a curved manifold) and is regarded as a geometrical holonomy, by a non-trivial topology (e.g. manifolds with twists or singularities) and is then regarded as a topological holonomy, or a combination of both. This phenomenon manifests itself in the form of phases, generally known as *geometric phases*[2] which may or may not be topological. Common examples are the parallel transport of some vector in a loop over a curved parameter space, or a closed path encircling a hole. Let us not forget that topology can be related to geometry, for instance in the Gauss-Bonnet theorem. It is a thing of beauty to see direct observable manifestations of curved geometry. Geometric phases are fundamentally tied to mathematics, and it is not surprising that this feeds into fibre bundles, as curvature is defined in terms of connections which, in plain language, means *effective gauge fields*. This is a rather spectacular realisation! We now understand that gauge fields are not just a convenient way to rewrite electric and magnetic fields. Instead,

> Gauge fields are non-measurable local probes (of the properties) of curved spaces in which natural phenomena take place. In turn, geometric phases constitute measurable counterpart of that.

Geometric phases come in all sorts of physical incarnations. Depending on the context they receive different names. Those which are topological are often associated with the notion of a winding and a corresponding quantisation. In the following we introduce a class of these phases which plays a fundamental role in Quantum Matter. This is the *Berry phase*, understood as an umbrella term. Some well-known scenarios in Quantum Mechanics can be understood as a subclass of this type of quantum geometric phase, such as the case of the Aharonov-Bohm effect.

2.1 The Berry Phase

Physical states are *rays* and not unit vectors in Hilbert space. Hence, states are defined up to a global phase ambiguity. This ambiguity was neglected since the early days of quantum mechanics as it was not believed to have physical meaning, but this is wrong. In particular, a global phase of a quantum state might not return to its initial value after a cyclic evolution in parameter space. In other words, there is a “global change without a local change”. More

[2] As opposed to dynamical phases.

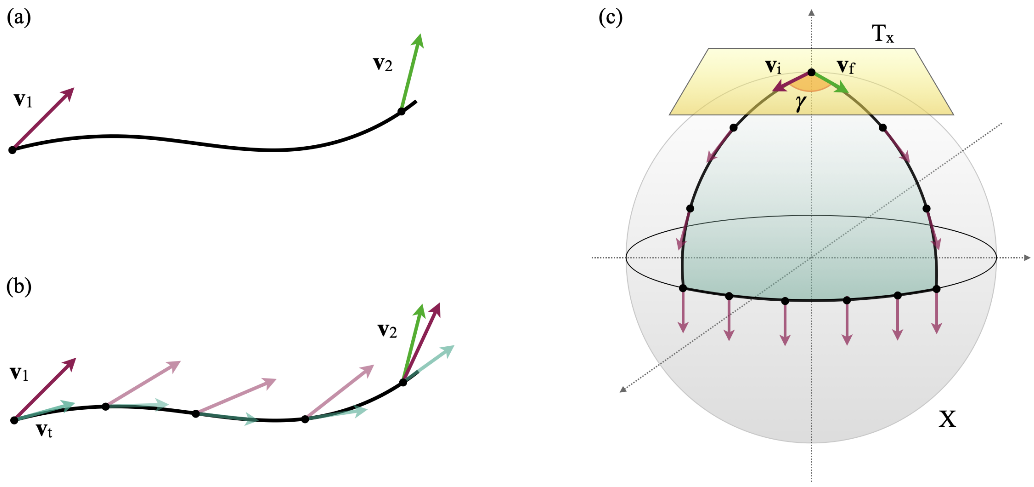

Figure 3: Parallel transport of a vector. (a) Given two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ located at different points A and B of a manifold, we want to see whether they are parallel or not. (b) We define a vector $\mathbf{v}_t$ tangent to the manifold, so that it forms an initial angle α with the vector $\mathbf{v}_1$. We then *parallel transport* the vector $\mathbf{v}_1$ rigidly from point A to point B, where we can compare it with $\mathbf{v}_2$ and verify that they are not parallel. (c) We consider the parallel transport of a vector on the surface of the curved manifold X. At each location $x \in X$, we can define a tangent space T_x. Starting from the initial vector $\mathbf{v}_\mathrm{i}$, we rigidly transport the vector over a closed circuit so that it returns to the initial position. The parallel transported vector $\mathbf{v}_\mathrm{f}$ differs from the initial vector by a phase γ, this is known as a geometrical holonomy or geometric phase.

precisely, it is the manifestation of a non-trivial holonomy, i.e. the parallel transport of a vector on a curved geometry. The Berry phase is nothing but the result of parallel transport of quantum states in Hilbert space. Let us see this more concretely.

Consider an $N-$dimensional set of parameters $\mathbf{R}$ which can evolve in time, so that $\mathbf{R} \equiv \mathbf{R}(t)$. Now consider a Hamiltonian that depends on this family of parameters $H(\mathbf{R}) \equiv H[\mathbf{R}(t)]$. For each $\mathbf{R}$ there exists a set of instantaneous orthonormal eigenstates of $H(\mathbf{R})$ such that

$$H(\mathbf{R})\,|n(\mathbf{R})\rangle = \epsilon_n(\mathbf{R})\,|n(\mathbf{R})\rangle \ . \tag{15}$$

We assume that this spectrum is *discrete* and *non-degenerate* in $\mathbf{R}-$space.

Then, the *adiabatic theorem* states that if a system is initially prepared in the n^{th} eigenstate $|\Psi_n(t=0)\rangle = |n\big(\mathbf{R}(t=0)\big)\rangle$, then it will evolve into the same eigenstate at later times, as long as the time-variation of H is *sufficiently slow*[3]. This means that in the adiabatic regime, we find

$$|\Psi_n(t)\rangle = C_n(t)\,|n\big(\mathbf{R}(t)\big)\rangle \ , \tag{16}$$

where $C_n(t)$ is a pure phase, since the time-evolution under the Hamiltonian $H(\mathbf{R})$ is given by a unitary transformation. Recall that for at time-dependent Hamiltonian H, we find that $C_n(t) = \exp\big(-\frac{i}{\hbar}\epsilon_n t\big)$. However, for a general time-dependent Hamiltonian such as $H(\mathbf{R})$, we have

$$C_n(t) = e^{i\gamma_n(t)}\, e^{-\frac{i}{\hbar}\int_0^t dt'\,\epsilon_n(t')} \ , \tag{17}$$

[3]Sufficiently slow meaning (roughly) less than the inverse of the spectral gap to other eigenstates.

where $\gamma_n(t)$ is some unspecified global phase shift, where the second term is know as the dynamical phase term. We might substitute this in the time-dependent Schrödinger equation to find

$$i\hbar \frac{\partial}{\partial t} |\Psi_n(t)\rangle = H[\mathbf{R}(t)] |\Psi_n(t)\rangle \ . \tag{18}$$

After substitution of Eq. (16) and taking the inner product with $\langle \Psi_n(t)|$ on both sides of the equation, we find the identity

$$\dot{\gamma}_n(t) = i \langle n(\mathbf{R})|\dot{n}(\mathbf{R})\rangle = i \langle n\big(\mathbf{R}(t)\big)| \frac{\partial}{\partial t} |n\big(\mathbf{R}(t)\big)\rangle \ . \tag{19}$$

Integrating over time yields

$$\gamma_n(t) = i \int_0^t dt' \ \langle n\big(\mathbf{R}(t')\big)| \frac{\partial}{\partial t'} |n\big(\mathbf{R}(t')\big)\rangle = i \int_{\mathcal{C}} d\mathbf{R} \cdot \langle n(\mathbf{R})| \nabla_{\mathbf{R}} |n(\mathbf{R})\rangle = \gamma_n[\mathcal{C}] \,, \tag{20}$$

which is (parameter) path-dependent but not explicitly time-dependent, implying that the phase γ_n is not *dynamical* but *geometric* in origin. We might define the vector field

$$\boldsymbol{\mathcal{A}}(\mathbf{R}) \equiv i \langle n(\mathbf{R})| \nabla_{\mathbf{R}} |n(\mathbf{R})\rangle \tag{21}$$

as the so-called *Berry connection*.

Exercise. Given $\langle n(\mathbf{R})|n(\mathbf{R})\rangle = 1$ and, thus, $\nabla_{\mathbf{R}} \langle n(\mathbf{R})|n(\mathbf{R})\rangle = 0$. Show that γ_n and $\boldsymbol{\mathcal{A}}$ are real-valued.

Now, note that so far we have not specified anything about the path $\mathcal{C}$. In such a general scenario, we can define a local transformation that leaves Eq.(15) invariant, for instance

$$|n(\mathbf{R})\rangle \longrightarrow |n'(\mathbf{R})\rangle = e^{-i\xi(\mathbf{R})} |n(\mathbf{R})\rangle \ . \tag{22}$$

From Eq.(21) we observe that the Berry connection transforms like a gauge field, namely

$$\boldsymbol{\mathcal{A}}(\mathbf{R}) \longrightarrow \boldsymbol{\mathcal{A}}'(\mathbf{R}) = \boldsymbol{\mathcal{A}}(\mathbf{R}) + \nabla_{\mathbf{R}}\, \xi(\mathbf{R}) \,, \tag{23}$$

and the transformation is, in fact, a gauge transformation. The global phase also transforms accordingly

$$\gamma_n \longrightarrow \gamma_n' = \gamma_n + \xi[\mathbf{R}(t)] - \xi[\mathbf{R}(0)] \,. \tag{24}$$

The new terms appearing in the phase will not, in general vanish, which implies that the geometric phase is gauge-dependent. For many years, it was assumed that, because of that, such a global phase had no measurable consequences and thus, no physical meaning.

Now, we observe that when the circuit $\mathcal{C}$ describes a closed path or loop in parameter space, meaning that after a time $t = t_f$, we find $\mathbf{R}(0) = \mathbf{R}(t_f)$, the phase γ_n becomes gauge invariant. This means that the phase can potentially have measurable physical consequences. This was realised by Michael Berry [16], and is thus known as the *Berry phase*, namely

$$\gamma_n[\mathcal{C}] = \oint_{\mathcal{C}} d\mathbf{R} \cdot \boldsymbol{\mathcal{A}}(\mathbf{R}) \,. \tag{25}$$

Having identified the Berry connection as a 1−form gauge connection, we can construct its 2−form curvature. In other words, we can consider the more general gauge field $\mathcal{A}_\mu(R)$ and the field strength

$$\mathcal{F}_{\mu\nu}(R^\alpha) = \frac{\partial}{\partial R_\mu} \mathcal{A}_\nu(R^\alpha) - \frac{\partial}{\partial R_\nu} \mathcal{A}_\mu(R^\alpha) \tag{26}$$

also known as the *Berry curvature*. From this we can define the spatial part of the Berry curvature as an "artificial" or "synthetic" magnetic field in parameter space

$$\boldsymbol{\mathcal{B}}(\mathbf{R}) = \nabla_{\mathbf{R}} \times \boldsymbol{\mathcal{A}}(\mathbf{R}) \tag{27}$$

and reinterpret the Berry phase as a *Berry flux*

$$\gamma_n = \oint_{\partial\Sigma} d\mathbf{R} \cdot \boldsymbol{\mathcal{A}}(\mathbf{R}) = \int_{\Sigma} d\mathbf{S} \cdot \boldsymbol{\mathcal{B}}(\mathbf{R}) \,. \tag{28}$$

over some region Σ with boundary $\partial\Sigma$ of the $N-$dimensional parameter space. Even though we talk about synthetic gauge and magnetic fields, the Berry connection and curvature are as physical and measurable as their conventional electromagnetic counterparts, just in parameter space.

The previous results are completely general with respect to the parameters $\mathbf{R}$. Hence, they include the particular case in which the parameter space is the real (position) space and $\mathbf{R} = \mathbf{x}$ become real-space coordinates. Similarly, they also include the particular case in which parameter space is reciprocal (momentum) space and $\mathbf{R} = \mathbf{k}$ are wave-vectors.

2.2 The Aharonov-Bohm Effect

The Aharonov-Bohm effect is one of the deepest and most unsettling modern results of quantum mechanics. There are, at least, five major and interdependent insights that we can obtain from the Aharonov-Bohm effect. We will discuss some in detail, but let us state them for clarity.

1. Gauge potentials can have observable consequences despite being in a region of null electromagnetic fields. Thus, gauge fields are more fundamental than electromagnetic fields. Alternatively, one can give up on the interpretation in terms of gauge fields at the expense of locality, i.e. accept quantum non-locality or the existence of an action-at-a-distance, which has implications for causality.
2. A gauge connection can be flat or pure gauge, and yet topologically non-trivial. This means that the vector potential, and the gauge transformation to remove it, are singular. This leads to either, multivaluedness of the wave function or single-valuedness plus quantisation.
3. The Aharonov-Bohm phase is an example of a geometric Berry phase effect. In this particular case it is, in addition, topological.
4. It shows the importance of connectedness of paths in parameter space and is a topological obstruction.
5. The Aharonov-Bohm effect can be understood as a change in spin and statistics.

The gedankenexperiment. We imagine a two-slit interference experiment for charged quantum particles. A very long, thin and ideal solenoid of radius R is placed transverse to the plane of the experiment and in between the two slits and the screen. It defines an inaccessible region for the particles and it confines the magnetic field to its core so that in all the accessible regions $|\mathbf{B}| = 0$. The experiment shows appreciable shifts in the interference fringes that are correlated with changes in the magnetic flux through the solenoid, even though particles are not in presence of a magnetic field at any point. In an ideal infinite solenoid, the magnetic field is constant and along the axis, meaning $\mathbf{B} = (0, 0, B_z) = B\,\hat{e}_z$ for $r < R$ but it is $\mathbf{B} = \mathbf{0}$ for $r > R$. Correspondingly, we can define the inner and outer magnetic fluxes $\Phi_B(r < R) = B\pi r^2$ and $\Phi_B(r \geq R) = B\pi R^2$. Since the problem is axis-symmetric, the only surviving components of the vector potential are $\mathbf{A} = (0, A_\varphi(r), 0)$ with $A_\varphi(r < R) = Br/2$ and $A_\varphi(r \geq R) = BR^2/(2r)$.

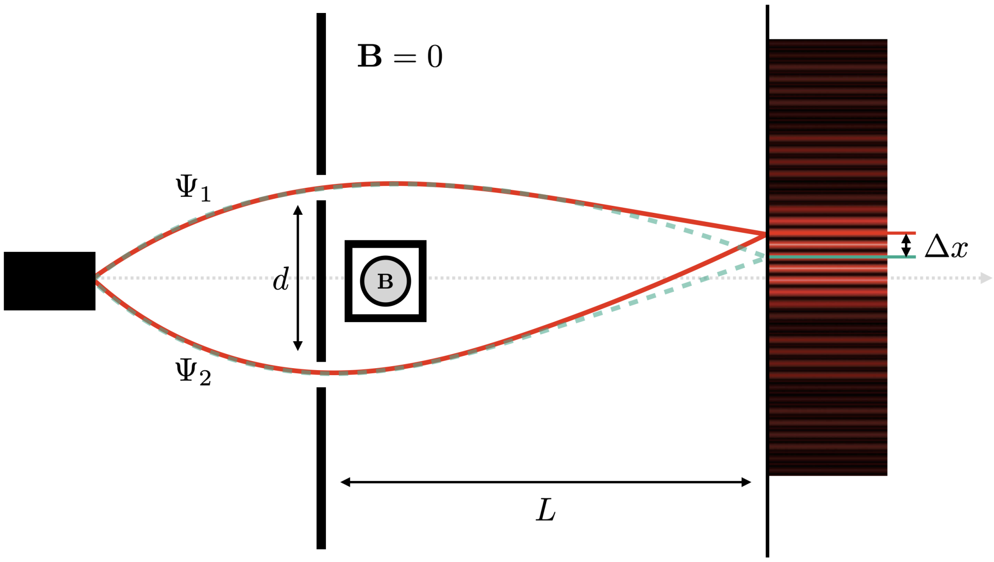

Figure 4: The Aharonov-Bohm interference experiment. (From left to right) We observe an electron gun, two slits, an impenetrable region containing a solenoid placed transversely to the plane, and a screen where an interference pattern is appreciated. The magnetic field vanishes in every point charged particles have access to. Two possibilities for the double path are depicted in solid red and dashed green respectively. Dashed green corresponds to the paths when there is no intensity running through the solenoid. Solid red depict the change in the paths and corresponding interference shift.

We imagine preparing the system with no flux on the solenoid. This allows us to define the wavefunctions $\Psi_a^{(0)}(t, \mathbf{x})$ for $a = 1, 2$ corresponding to the two possible paths the particles can take. Ramping up the flux introduces a change in the wavefunctions of the form

$$\Psi_a(t, \mathbf{x}) = e^{i\gamma_a(\mathbf{x})}\Psi_a^{(0)}(t, \mathbf{x})\,. \tag{29}$$

Such a change can be incorporated in the equation of motion for a charged quantum particle in the presence of a non-null vector potential

$$i\hbar\partial_t\Psi_a = \frac{1}{2m}\Big[-i\hbar\nabla - q\mathbf{A}(\mathbf{x})\Big]^2\Psi_a = \frac{e^{i\gamma_a(\mathbf{x})}}{2m}\Big[-i\hbar\nabla - q\mathbf{A}(\mathbf{x}) + \hbar\nabla\gamma_a(\mathbf{x})\Big]^2\Psi_a^{(0)} \tag{30}$$

so that for $q\mathbf{A} = \hbar\nabla\gamma_a$ we remove the gauge potential from the equation via gauge transformation. This gauge transformation is, thus, given by $\gamma_a(\mathbf{x}) = \frac{q}{\hbar}\int^{\mathbf{x}} d\mathbf{x}' \cdot \mathbf{A}(\mathbf{x}')$, which adds to the local phase of the wavefunction for the charged particle. Now the interference fringes of the experiment will depend on the phase difference of the wave function $\Psi_a = |\Psi_a|\, e^{i\theta_a}$, namely

$$\Delta\theta = \theta_1(\mathbf{x}) - \theta_2(\mathbf{x}) = \theta_1^{(0)}(\mathbf{x}) + \gamma_1(\mathbf{x}) - \theta_2^{(0)}(\mathbf{x}) - \gamma_1(\mathbf{x}) \tag{31}$$

$$= \Delta\theta^{(0)}(\mathbf{x}) + \frac{q}{\hbar}\left(\int_{P_1} d\mathbf{x}' \cdot \mathbf{A}(\mathbf{x}') - \int_{P_2} d\mathbf{x}' \cdot \mathbf{A}(\mathbf{x}')\right) = \Delta\theta^{(0)}(\mathbf{x}) + \gamma \tag{32}$$

where $\gamma \equiv \frac{q}{\hbar}\oint_P d\mathbf{x}' \cdot \mathbf{A}(\mathbf{x}')$ is the *Aharonov-Bohm phase* defined as the circulation of the vector potential over the positively-oriented closed circuit defined by the two trajectories starting at the source and ending in the screen $P = P_1 \cup P_2$. We can use Stokes' theorem and find $\gamma = \frac{q}{\hbar}\Phi_B$. We can now define the overlap of the two wavefunctions

$$\langle\Psi_1|\Psi_2\rangle \equiv \int d^3\mathbf{x}\, \Psi_1^*(\mathbf{x})\, \Psi_2(\mathbf{x}) = e^{i\gamma}\, \langle\Psi_1^{(0)}|\Psi_2^{(0)}\rangle \ , \tag{33}$$

from where we see that for certain values of the flux $\Phi_B = nh/q$ for $n \in \mathbb{Z}$, the solenoid has no effect on the overlap. Hence, the magnetic field in the solenoid induces a shift in the particles' interference pattern given by

$$\Delta x = \frac{\lambda_{th}}{2\pi}\frac{L}{d}\gamma = \frac{\lambda_{th} L}{2\pi d}\frac{q}{\hbar}\Phi_B \tag{34}$$

despite the fact that they experience a vanishing magnetic field. Thought in classical terms this is rather counter-intuitive, as quantum charged particles are deflected by an angle $\alpha \approx \Delta x/L$ without experiencing a Lorentz force, yet the result is analogous to classical charged particles being deflected by a magnetic field. This leads to the usual statement that, in quantum mechanics gauge fields are more "fundamental" [4] than electromagnetic fields and have measurable consequences, despite being unobservable themselves.

The modern setup We now take another look at the previous effect in a more schematic scenario. We observe that the Aharonov-Bohm bound-state problem, the particle on a ring or in a punctured plane, threaded with magnetic flux, and the particle orbiting a flux-tube, are topologically equivalent.

Let us consider a point particle of charge q propagating in the x–y plane under the influence of an infinitesimally thin magnetic flux tube or vortex piercing the plane along positive z-axis. The quantum-mechanical Hamiltonian density of the particle moving in 2d is given by

$$H_0 = -\frac{\hbar^2}{2m}\left[\boldsymbol{\nabla} - i\frac{q}{\hbar}\mathbf{A}(\mathbf{x})\right]^2 , \tag{35}$$

where the functional form of the vector potential is determined from

$$\mathbf{B} = B(t,\mathbf{x})\,\hat{\mathrm{e}}_z = \boldsymbol{\nabla} \times \mathbf{A}(t,\mathbf{x}) = \Phi_B\, \delta^{(2)}\big(\mathbf{x} - \mathbf{x}(t)\big)\,\hat{\mathrm{e}}_z \tag{36}$$

with the magnetic flux enclosed being

$$\Phi_B = \int_\Sigma d\mathbf{S} \cdot \mathbf{B} = \oint_{\partial\Sigma} d\mathbf{x} \cdot \mathbf{A} = \text{const.} \tag{37}$$

and $\mathbf{x}(t)$ defines the location of the flux tube. Fixing the Coulomb gauge $\boldsymbol{\nabla} \cdot \mathbf{A} = 0$ and expressing $A^i = -\epsilon^{ij}\partial_j\omega$, the magnetic field is written as $B\,\hat{\mathrm{e}}_z = \boldsymbol{\nabla}^2\omega\,\hat{\mathrm{e}}_z$. Using the solution of

[4] In the sense that they contain more information and, therefore, provide a more complete description of Nature.

the Poisson equation in two dimensions $\nabla^2 G(\mathbf{x}) = \delta^{(2)}(\mathbf{x})$, we obtain the form of the Green's function and thus, the form of the vector potential

$$G(\mathbf{x}) = \frac{1}{2\pi}\ln|\mathbf{x}-\mathbf{x}(t)| + \text{const.} \qquad \text{and} \qquad A^i(t,\mathbf{x}) = -\frac{\Phi_B}{2\pi}\epsilon^{ij}\frac{x_j - R_j(t)}{|\mathbf{x}-\mathbf{R}(t)|^2}\,. \tag{38}$$

One could also write the gauge potential as a pure gauge or flat connection of the form

$$\mathbf{A}(t,\mathbf{x}) = \frac{\Phi_B}{2\pi}\nabla\varphi\big[\mathbf{x}-\mathbf{x}(t)\big] = \frac{\Phi_B}{2\pi}\nabla\arg\big[\mathbf{x}-\mathbf{x}(t)\big] = \frac{\Phi_B}{2\pi}\tan^{-1}\left[\frac{x_2 - R_2(t)}{x_1 - R_1(t)}\right], \tag{39}$$

but this comes at the expense of it being singular, since the argument function is multivalued, which prevents derivatives from commuting in $\mathbf{B} = \frac{\Phi_B}{2\pi}\nabla\times\nabla\varphi \neq 0$. Now, we are in the position of solving the eigenvalue problem $H_0\Psi = E\Psi$ in polar coordinates $\mathbf{x} = (r,\varphi)$ with $\mathbf{A}(\mathbf{x}) = \frac{\Phi_B}{2\pi r}\,\hat{\mathrm{e}}_\varphi$. Let us now assume that the charged particle is orbiting the flux tube at a fixed radius so that the dynamics is only present in the polar coordinate. Hence, we might drop the radial component of the kinetic term and be left with

$$H_0\Psi = -\frac{\hbar^2}{2mr^2}\left(\frac{\partial}{\partial\varphi} - i\frac{q}{\hbar}\frac{\Phi_B}{2\pi}\right)^2\Psi = E\Psi \qquad \text{and} \qquad \Psi(\mathbf{x}) = \frac{1}{\sqrt{2\pi r}}e^{iN\varphi} \tag{40}$$

with $N\in\mathbb{Z}$, so that the allowed energies of the system are

$$E = \frac{1}{2mr^2}\left(\hbar N - \frac{q}{2\pi}\Phi_B\right)^2\,. \tag{41}$$

We observe that the spectrum is shifted according to the value of the flux. However, for multiples of $\Phi_0 = h/q$, it is left invariant, i.e. the flux tube has no effect. This can be seen differently by considering a gauge transformation $A_\mu \to A_\mu + \partial_\mu\xi$ that removes the gauge field, provided it is flat. This is

$$H_0\Psi \longrightarrow \tilde{H}_0\tilde{\Psi} = -\frac{\hbar^2}{2m}\nabla^2\tilde{\Psi} \tag{42}$$

$$\mathbf{A} \longrightarrow \tilde{\mathbf{A}} = \mathbf{A} - \frac{\Phi_B}{2\pi}\nabla\varphi\,, \tag{43}$$

$$\Psi \longrightarrow \tilde{\Psi} = \exp\left(-i\frac{q}{\hbar}\frac{\Phi_B}{2\pi}\varphi\right)\Psi\,. \tag{44}$$

Without further constraints, the transformed wavefunction $\tilde{\Psi}$ is ill-defined as it is multivalued. Therefore, the only gauge transformations allowed are those for which $\Phi_B = 2\pi\hbar l/q$ with $l\in\mathbb{Z}$, which keep the wavefunction single-valued. This topological quantisation coincides with the invariance of the spectrum for certain values of the flux.

Small *vs.* Large gauge transformations

Small or *conventional gauge transformations* are a manifestation of gauge redundancy and are physically equivalent configurations expressed in a different way, very much like changing a coordinate system. *Large gauge transformations*, on the other hand are topological. They typically relate homotopically distinct states with the same physical properties. It is often said that small gauge transformations are homotopic to the identity or deform to the identity at spatial infinity, while large gauge transformations do not.

> A simple example is that of the Aharonov-Bohm effect. The spectrum of the theory depends on the value of the flux traversing a solenoid. However, there is a periodicity on the spectrum labelled by a winding number. While the Hamiltonian at $\Phi_B = 0$ and at $\Phi_B = \Phi_0$ give the exact same energy spectrum, both configurations are topologically inequivalent. These two physical states, which are indistinguishable from the point of view of the particle are connected by a large gauge transformation. In transforming, we are moving between equivalence classes with different winding number but the same observable properties.

Berry phase interpretation Let us now take our flux tube and put it at the origin of $\mathbb{R}^3$. A test charge particle will be found at position $\mathbf{x}$, but it will be contained in a cubic box. A reference vertex of such a box is found at $\mathbf{X}_0$, let us imagine it is the bottom leftmost corner of the box. The hard walls of the box provide a confining potential so the Hamiltonian for the charged particle is now $H = H_0 + V(\mathbf{x}-\mathbf{X}_0)$, where H_0 is given in Eq. (35). In the absence of a flux tube (i.e. for $\mathbf{A}=\mathbf{0}$), the eigenvalue problem reads

$$H_{\mathbf{A}=\mathbf{0}}(\mathbf{x}-\mathbf{X}_0)\,\Psi^{(0)}(\mathbf{x}-\mathbf{X}_0) = E_n\,\Psi^{(0)}(\mathbf{x}-\mathbf{X}_0) \tag{45}$$

As soon as we find non-zero magnetic flux in the tube, the problem becomes $H\Psi = E_n\Psi$ and the original wave function acquires a phase

$$\Psi(\mathbf{x};\mathbf{X}_0) = e^{i\gamma_{\mathbf{X}_0}(\mathbf{x})}\,\Psi^{(0)}(\mathbf{x}-\mathbf{X}_0) = e^{i\frac{q}{\hbar}\int_{\mathbf{X}_0}^{\mathbf{x}} d\mathbf{x}'\cdot\mathbf{A}(\mathbf{x}')}\,\Psi^{(0)}(\mathbf{x}-\mathbf{X}_0)\,, \tag{46}$$

where the integral is taken over some path inside the box. This is no different than in the original discussion of the gedankenexperiment. Now however, we might want to compute the *Berry connection* in parameter space $\mathbf{X}_0$, which is defined as the quantity

$$\boldsymbol{\mathcal{A}}(\mathbf{X}_0) \equiv i\,\langle\Psi|\,\nabla_{\mathbf{X}_0}|\Psi\rangle = i\int d^3\mathbf{x}\,\Psi^*(\mathbf{x};\mathbf{X}_0)\,\nabla_{\mathbf{X}_0}\Psi(\mathbf{x};\mathbf{X}_0)\,. \tag{47}$$

Substitution of Eq. (46) and careful computation of $\nabla_{\mathbf{X}_0}\int_{\mathbf{X}_0}^{\mathbf{x}} d\mathbf{x}'\cdot\mathbf{A}(\mathbf{x}')$ with an adequate change of variables $\mathbf{x}' = \mathbf{X}_0 + \mathbf{x}''$ yields

$$\boldsymbol{\mathcal{A}}(\mathbf{X}_0) = i\int d^3\mathbf{x}\left[\Psi^{(0)}(\mathbf{x}-\mathbf{X}_0)\right]^*\Big(-i\frac{q}{\hbar}\mathbf{A}(\mathbf{X}_0)\Psi^{(0)}(\mathbf{x}-\mathbf{X}_0)+\nabla_{\mathbf{X}_0}\Psi^{(0)}(\mathbf{x}-\mathbf{X}_0)\Big) = \frac{q}{\hbar}\mathbf{A}(\mathbf{X}_0)\,. \tag{48}$$

So the Berry connection is directly related to the Aharonov-Bohm vector potential evaluated at $\mathbf{x} = \mathbf{X}_0$. Notice that the last term in brackets in Eq. (48) vanishes due to normalisation of $\Psi^{(0)}$. Transporting the box, without changing its orientation, around a closed circuit $\mathcal{C}$ containing the flux tube, allows us to compute the *Berry phase* defined as the circulation of the Berry connection

$$\gamma(\mathcal{C}) \equiv \oint_{\mathcal{C}} d\mathbf{X}_0\cdot\boldsymbol{\mathcal{A}}(\mathbf{X}_0) = \frac{q}{\hbar}\Phi_B\,. \tag{49}$$

This confirms that the Aharonov-Bohm phase can be thought of as a geometric (Berry) phase.

2.3 Many-body Aharonov-Bohm Effect

Let us consider the Physics of the many-body Aharonov-Bohm effect.

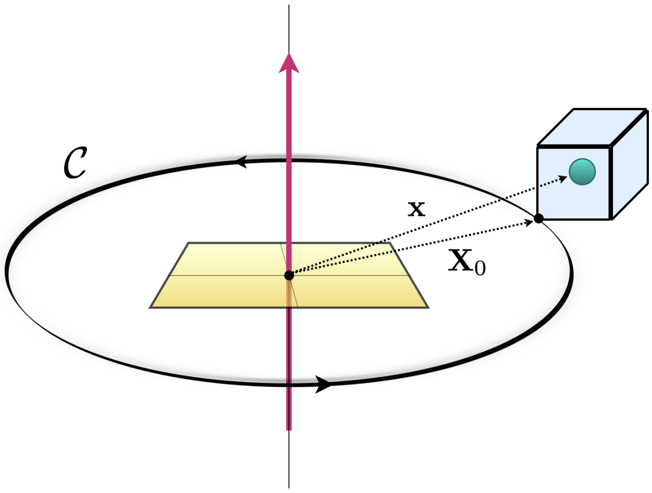

Figure 5: A test particle in a box is adiabatically circulated around a flux tube.

Flux Attachment. Let us consider a gas of N identical charge-flux-tube complexes, each carrying a magnetic flux Φ_B. From a distance, we can think of them as point objects located at position $\mathbf{x} = \mathbf{x}_i(t)$. The magnetic field experienced by each object is

$$\boldsymbol{b}(t,\mathbf{x}_i) = \boldsymbol{\nabla}_{\mathbf{x}_i} \times \boldsymbol{a}(t,\mathbf{x}_i) = \sum_{j\neq i} \Phi_B\, \delta^{(2)}(\mathbf{x}_i - \mathbf{x}_j)\,\hat{\mathrm{e}}_z\,. \tag{50}$$

In the Coulomb gauge, the corresponding vector potential is

$$\boldsymbol{a}(t,\mathbf{x}_i) = \frac{\Phi_B}{2\pi}\sum_{j\neq i}\boldsymbol{\nabla}_{\mathbf{x}_i}\varphi(\mathbf{x}_i - \mathbf{x}_j) = \frac{\Phi_B}{2\pi}\sum_{j\neq i}\frac{\hat{\mathrm{e}}_z \times (\mathbf{x}_i - \mathbf{x}_j)}{|\mathbf{x}_i - \mathbf{x}_j|^2}\,, \tag{51}$$

where φ is the polar angle. Defining the number density of point particles as $n(t,\mathbf{x}) = \sum_{i=1}^{N}\delta^{(2)}(\mathbf{x}-\mathbf{x}_i)$, we can re-express the initial magnetic field as $\boldsymbol{b}(t,\mathbf{x}_i) = \Phi_B\, n(t,\mathbf{x}_i)$. This links the magnetic field felt by one particle with the local number density of particles. Such a relation corresponds to a many-body version of the Aharonov-Bohm effect. Replacing the local point-particle number density $n(t,\mathbf{x}_i)$

Statistical Transmutation. The natural angular variable is the polar angle, which can take values in S^1 and has a singularity at $\mathbf{x} = 0$, so the relevant homotopy group is $\pi_1(S^1) = \mathbb{Z}$. This can be formalised by means of the argument function $\phi_{ml} \equiv \arg(\tilde{\mathbf{x}}_{ab}) = \arg(\mathbf{x}_a - \mathbf{x}_b\,;\hat{\mathrm{e}}_{\mathbf{x}})$, where the angle is taken with respect to some arbitrary reference, in this case, the x-axis. The exchange property for this function reads $\phi_{ab} = \pm\pi + \phi_{ba}$. We can now compute the gauge potential as

$$\boldsymbol{a}(\mathbf{x}_i) = \alpha\,\nabla_{\mathbf{x}_i}\Phi(\mathbf{x}_1,\ldots,\mathbf{x}_i,\ldots,\mathbf{x}_N) = \alpha\,\boldsymbol{\nabla}_{\mathbf{x}_i}\Big(\sum_{a<b}\phi_{ab}\Big) = \alpha\,\boldsymbol{\nabla}_{\mathbf{x}_i}\Big[\sum_{i<b}\phi_{ib} + \sum_{a<i}\phi_{ai}\Big] \tag{52}$$

$$= \alpha\,\boldsymbol{\nabla}_{\mathbf{x}_i}\Big[\sum_{i<b}\phi_{ib} + \sum_{a<i}\big(\pm\pi + \phi_{ia}\big)\Big] = \alpha\,\boldsymbol{\nabla}_{\mathbf{x}_i}\Big[\sum_{j\neq i}\phi_{ij} \pm \sum_{a<i}\pi\Big] \tag{53}$$

$$= \alpha\sum_{j\neq i}\boldsymbol{\nabla}_{\mathbf{x}_i}\arg(\mathbf{x}_i - \mathbf{x}_j\,;\hat{\mathrm{e}}_{\mathbf{x}}) = \alpha\sum_{j\neq i}\boldsymbol{\nabla}_{\mathbf{x}_i}\arg(\tilde{\mathbf{x}}_{ij})\,, \tag{54}$$

for which one can define a magnetic field and verify that there exist a relation with the charge density $n(\mathbf{x}_i)$ of the form

$$b(\mathbf{x}_i) = \boldsymbol{\nabla}_{\mathbf{x}_i} \times \boldsymbol{a}(\mathbf{x}_i) = \alpha \sum_{j \neq i} 2\pi\, \delta^{(2)}(\mathbf{x}_i - \mathbf{x}_j) \equiv 2\pi\alpha\, n(\mathbf{x}_i). \tag{55}$$

In other words, we naturally recover local flux attachment. The associated gauge transformation in this case is commonly known in literature as a statistical or singular transformation, and reduces to

$$\Psi(\mathbf{x}_1, \ldots, \mathbf{x}_N) = e^{i\alpha \sum_{m<l} \arg(\mathbf{x}_m - \mathbf{x}_l\,;\,\hat{\mathbf{e}}_{\mathbf{x}})}\, \Psi_c(\mathbf{x}_1, \ldots, \mathbf{x}_N), \tag{56}$$

where the sum in the exponent is over all particles. Hence, for a given pairwise exchange of particles $i \longleftrightarrow j$, where $1 \leq i < j \leq N$, a corresponding π phase from the argument function is collected by the composite wavefunction for every $e^{-i\alpha\phi_{ab}} e^{i\alpha\phi_{ba}}$ term $i \leq a < b \leq j$. This yields a statistical factor $\gamma_{ij} = \mp\alpha\pi\eta$, where $\eta \in \mathbb{Z}$ is the number of $a \longleftrightarrow b$ possible pairs. This is nothing but a many-particle Aharonov-Bohm phase for flux α.

2.3.1 Chern-Simons Gauge Theory

Alternatively, such a peculiar choice of gauge potential is provided by construction, if the correct term is incorporated at the level of a field theoretical Lagrangian. We make use of the quantised Abelian Chern-Simons term at level $1/\alpha$ with $\alpha \in \mathbb{Z}$, minimally coupled to matter via source term

$$S = \frac{1}{4\pi\alpha} \int dt\, d^2\mathbf{x}\, \epsilon^{\mu\nu\lambda} \hat{a}_\mu \partial_\nu \hat{a}_\lambda - \int dt\, d^2\mathbf{x}\, \hat{J}^\mu \hat{a}_\mu\,. \tag{57}$$

Computing the Euler-Lagrange equations for the gauge field in the presence of the matter source we are left with

$$\hat{J}^\mu = \frac{1}{2\pi\alpha} \epsilon^{\mu\nu\lambda} \partial_\nu \hat{a}_\lambda\,, \qquad \partial_\mu \hat{J}^\mu = 0\,, \tag{58}$$

where the current time component becomes nothing but a constraint equation or Gauss's law of the form

$$\boldsymbol{\nabla} \times \hat{\boldsymbol{a}}(t, \mathbf{x}) = 2\pi\alpha\, \hat{n}(t, \mathbf{x}) \tag{59}$$

which we can attempt to solve in the Coulomb gauge $\boldsymbol{\nabla} \cdot \boldsymbol{a} = 0$. This allows us to write the vector potential as $\boldsymbol{a} = \boldsymbol{\nabla} \times \varphi$, so that the Gauss's law becomes

$$n(t, \mathbf{x}) = \frac{1}{2\pi\alpha} \boldsymbol{\nabla}^2 \varphi(t, \mathbf{x}) \tag{60}$$

that can be solved using conventional Green's function methods to find that $\hat{\boldsymbol{a}}(t, \mathbf{x}) = \alpha\, \boldsymbol{\nabla}\, \hat{\Phi}(t, \mathbf{x})$ and

$$\hat{\Phi}(t, \mathbf{x}) = \int d^2\mathbf{x}'\, \varphi(t, \mathbf{x} - \mathbf{x}')\, \hat{n}(t, \mathbf{x}')\,, \tag{61}$$

where $\varphi(t, \mathbf{x}) = \tan^{-1}(y/x)$ is the conventional polar angle, and where we have assumed that the "charge" density is point-like so that the Chern-Simons gauge potential can be written as a pure gauge. This is nothing but the field theoretical version of flux attachment previously found in first-quantised language. The corresponding singular gauge transformation is

$$\hat{\Psi}(t, \mathbf{x}) = e^{i\alpha \hat{\Phi}(t, \mathbf{x})}\, \hat{\Psi}_C(t, \mathbf{x})\,. \tag{62}$$

Acknowledgements

I want to thank the lads from SciPost for creating and publicly sharing this beautiful template, which I borrowed and adapted for this set of notes. You can also find it and enjoy it at https://scipost.org/media/uploads/templates/SciPostPhysLectNotes/2023/SciPostPhysLectNotes_LaTeX_Template.tgz

A Connecting Geometry and Topology

The Wu-Yang dictionary [8] constituted the first explicit connection between the field of gauge theory in Physics and that of differential geometry in Mathematics.

Gauge Theory Language	Fibre Bundle Language
Gauge	Principal coordinate bundle $(A_x)_{x\in X}$
Gauge Transformation $(\mathcal{U})$	Transition function (Φ_x)
Gauge Field (A)	Connection on a principal fibre bundle (Γ)
Phase factor (Φ_{ab})	Parallel transporter or holonomy
Field strength (F)	Curvature (Ω)
Gauge group (G)	Structure group of the principal bundle
Minimal coupling	Covariant derivative

Table 1: Wu-Yang dictionary [8] of terminologies between Physics and Mathematics.

A relatively simple translation of terminology between concepts in electromagnetism and fibre bundle theory allowed devising a much deeper revelation between entire fields of apparently unrelated disciplines, i.e. a scientfic version of a *Rossetta Stone*. The recognition that a gauge field is nothing but a connection on a fibre bundle provided physicists with a passage to the true revolution taking place in differential geometry through the work of Shiing-Shen Chern [17, 18] — sometimes referred to as the “father of modern differential geometry” —, from which modern understanding of general relativity, condensed matter and quantum field theory have greatly benefited. Let me be a bit more concrete on why this is a big deal.

> *There exist ways to connect Geometry with Topology.*

The previous statement is incredibly non-trivial, since (local) Geometry and (global) Topology appear as disjoint subfields of Mathematics studying different properties of a given object . . . or so people thought. The direct connection between these fields was already hinted by the Gauss-Bonnet theorem, known since the mid 1800s. But it was not until the work of Shiing-Shen Chern more than 100 years later — through the introduction of the *Chern-Gauss-Bonnet Theorem* and *Characteristic Classes* —, that this connection was brought to its current full glory.

Gauss-Bonnet. The Gauss-Bonnet theorem establishes that, for a closed two-dimensional surface Σ, there is a relationship between the total curvature of surface and its Euler characteristic $\chi = 2-2g-b \in \mathbb{Z}$ such as

$$\int_\Sigma K = 2\pi\chi(\Sigma)\,, \tag{63}$$

where g is the genus, b is the number of boundaries, and $K = \kappa_1\kappa_2$ is the Gaussian curvature at a point $\mathbf{x}$ expressed in terms of the principal curvatures of the surface. Therefore, integrating a local property of a manifold we obtain knowledge about its global structure. This theorem is reminiscent of the definition of the first Chern number integrating the curvature of a gauge connection $\int_\Sigma F = 2\pi c_1$. This is not by chance, they are examples of elements of *Characteristic classes* classifying fibre bundles, whose existence is ensured by the *Chern-Weyl Theory* and the *Chern-Gauss-Bonnet theorem*. Characteristic classes are gauge-invariant cohomology classes associated to a base space X quantifying the *topological non-triviality* or the degree of *obstructions* of fibre bundles. They associate topologically-invariant numbers to the integral of the 2n-forms $\mathrm{Tr}(F^n)$, which are nothing but powers of the curvature 2-form $F = \mathrm{d}A + A \wedge A$ of a gauge connection. Amongst these classes we find the Euler class and the Chern class, to which the previous examples are particular instances of. The whole family of invariants is defined as follows.

Chern Classes. Given a 2n-dimensional manifold Σ the *n-th Chern number* is defined as the integral of the *n-th Chern class* $C_n(F) \in H^{2n}(\Sigma)$

$$c_n = \int_{\Sigma^{2n}} C_n(F)\,, \tag{64}$$

where the *total Chern class* is generated from

$$C(F) = \det\Big(\mathbb{I} + i\frac{F}{2\pi}\Big) = 1 + C_1(F) + C_2(F) + \dots \tag{65}$$

so

$$C_0(F) = 1 \tag{66}$$

$$C_1(F) = \frac{i}{2\pi}\mathrm{Tr}(F) \tag{67}$$

$$C_2(F) = \frac{1}{2}\Big(\frac{i}{2\pi}\Big)^2[\mathrm{Tr}(F)\wedge\mathrm{Tr}(F) - \mathrm{Tr}(F\wedge F)] \tag{68}$$

$$\dots \tag{69}$$

$$C_n(F) = \Big(\frac{i}{2\pi}\Big)^n \det(F)\,. \tag{70}$$

For compact manifolds Chern numbers are integers. Other important classes are those of Pontryagin and Euler.

B Topological Solitons

Topological quantum matter is no more than conventional quantum matter in which either the underlying fields find themselves in topologically non-trivial configurations, i.e. they are theories with defects, or the manifold they live in is itself topologically non-trivial. Topological defects come, mathematically, in the form of singularities, twists, discontinuities and, in general integer countable ill-definitenessess or oddities in functions. Such countable robust objects provide the origin of topological quantisation. They appear in Nature in the form of topological solitons or "lumps", examples of which are kinks, vortices, monopoles, domain walls, skyrmions, merons, hopfions, but also knots, dislocations and disclinations [4,6,19–22]. These defects are not necessarily point-like, but can be extended objects, with a characteristic dimension n. For instance, a vortex or a dislocation in 3d matter consists of a string of singularities. This implies that its effects can be felt non-locally, i.e. far away from where we would

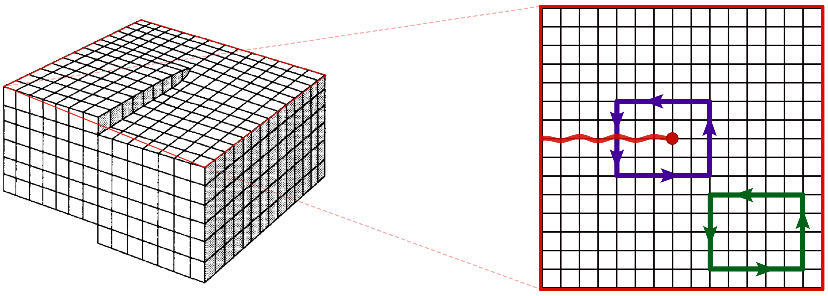

Figure 6: Screw dislocation in a lattice system. The topological defect is located in the bulk of the material, but its effects propagate to the boundary. (Left) Bulk of the material. The path around the disclination line is helical featuring a mismatch or slip. This part of the figure has been extracted and modified from Ref. [9]. (Right) Two topologically inequivalent Burgers circuits are shown. The purple circuit encloses a defect, which appears as a single line crossing in the contour, signaling a non-null holonomy.

locate the defect (see Figure 6). It is, thus, not surprising that their presence and properties are characterised by non-local observables such as the Burgers circuit for dislocations in solids or the Wilson and 't Hooft loops in parameter space of gauge theories. These defects stand out in an ordered configuration — or symmetry-broken phase — of the corresponding field and, in fact, can drive a transition, e.g. Berezinskii-Kosterlitz-Thouless (BKT) transition. This happens already in classical systems and it is relatively common in soft-matter physics [9]. An important quantity is the relative dimensionality between that of the defects and that of the system. This is known as *codimension* $\delta = d - n$.

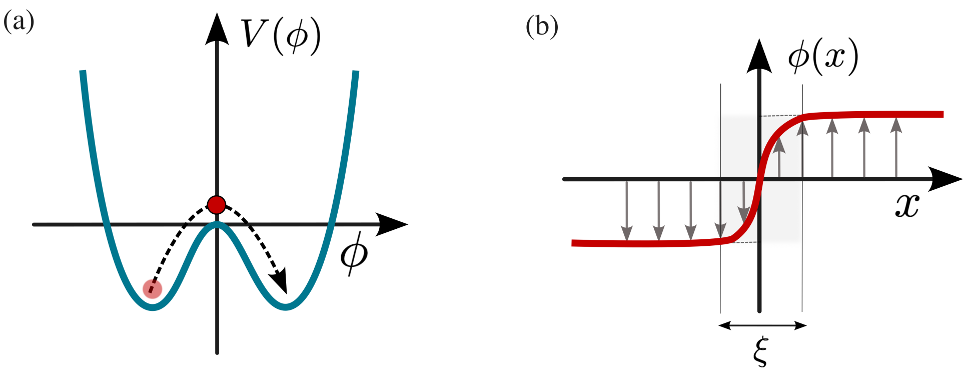

Figure 7: A topological soliton is a change of configuration in the (ordered) broken-symmetry phase. The groundstate has a $\mathbb{Z}_2$ degeneracy with equal weight. This situation describes the case for a Ginzburg-Landau theory of a real scalar field ϕ in a quartic potential.

It is particularly interesting to think, for instance, of spin models which live in a spatial d-dimensional lattice and have internal degrees of freedom with an associated symmetry group, for instance SO(3), corresponding to the group of three dimensional rotations in this internal space. Let us call the dimensionality of this internal space q. Now, in addition, defects can appear on either the physical or the internal degrees of freedom, each having an associated dimensionality n. For instance, a point defect has dimension $n = 0$. It is easy to realise that the interplay of d, q and n will be critical to determine the properties of the system.

A Topological Summary

Topological soliton configurations provide the bridge between conventional quantum matter and exotic topological phases. The interplay between relative dimensions of the order parameter q, the defect n and space d, determine the topological stability and classification of the solitons.

Sometimes topological solitons can have localised profiles in parameter space, can be stable against decay and cost a finite amount of energy to create. In these situations, they behave like particles and are effectively treated as such. In quantum matter, these topological solitons can appear as quantised excitations of the system.

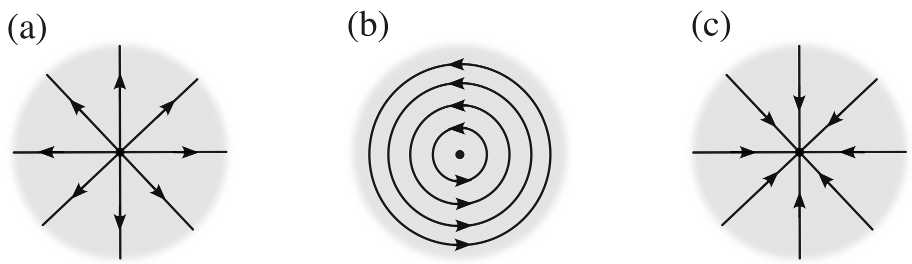

Figure 8: Topologically equivalent vortices for $k = +1$. (a) $\theta_0 = 0$, (b) $\theta_0 = \pi$ and (c) $\theta_0 = 2\pi$.

Equivalence of defects Consider a system living in two-dimensional space with an order parameter that has two components which can rotate in the plane. In other words $d = q = 2$. An example can be the magnetisation in an xy model in $\mathbf{x} = (r, \varphi)$ in polar coordinates

$$\langle \mathbf{m}(\mathbf{x}) \rangle = m_0 \big(\cos\theta(\mathbf{x}), \sin\theta(\mathbf{x}) \big) \tag{71}$$

so that $\boldsymbol{\nabla}\theta(\mathbf{x}) = \frac{1}{r}\hat{\mathbf{e}}_\varphi$ has a point singularity at $\mathbf{x} = \mathbf{0}$, a vortex. Provided $\varphi \in [0, 2\pi)$ the order parameter satisfies periodicity $\langle \mathbf{m}(\theta) \rangle = \langle \mathbf{m}(\theta + 2\pi k) \rangle$ with $k \in \mathbb{Z}$. We see that the vortex function is $\theta(\mathbf{x}) = k\varphi + \theta_0$. Different configurations of vortices are possible varying both k and θ_0. However, k labels topologically distinct vortices, while θ_0 labels vortices within the same equivalence class that can be smoothly deformed into one another (see Figure 8).

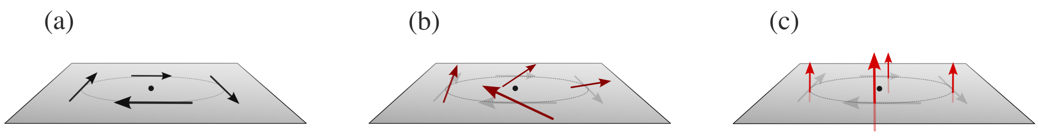

Figure 9: A vortex is smoothly lifted by ferromagnetic alignment perpendicular to the plane.

Pairs of opposite sign defects tend to distort the system close to them, but not globally. For instance, a pair of +1 vortex and −1 (anti-)vortex in a fluid distort the flow locally, but do not change the global topology of the system as $k_{\text{total}} = +1 - 1 = 0$. In fact, the pair vortex-antivortex can annihilate as they are a topologically equivalent configuration to no defects at all.

Stability of defects Homotopy theory constrains which configurations are topologically trivial or not. Non-trivial configurations are robust as there is an energy cost associated to a change in topological sector. According to this [9], we find that:

- When $q > \delta$ there are no stable defects. See Figure 9 as an illustration.
- When $q < \delta$ it is not possible to construct configurations in which the order parameter rotates continuously.

Fermionic Zero Modes

Let us consider a Dirac fermion in one spatial dimension in the presence of a kink or domain wall. This situation is a simplification of those treated by Jackiw and Rebbi [23], and Goldstone and Wilczek [24]. The system is governed by a modified Dirac equation of the form

$$\left[-i\hat{\sigma}^x \partial_x - m(x)\,\hat{\sigma}^z\right]\psi(x) = E\,\psi(x) \tag{72}$$

where we recall that in $d = 1$ a Dirac fermion is a two-component spinor variable. The mass term is a topological soliton of correlation length ξ and spatial profile

$$m(x) = m_0 \tanh\left(\frac{x}{\sqrt{2}\,\xi}\right). \tag{73}$$

An exact diagonalisation of this system reveals that the spectrum is complex. The real part of the energy is roughly *linear*, *gapped*, and presenting states at *zero energy*. A naïve linearisation of the spectrum gives $\mathcal{E}_k \approx \pm\sqrt{k^2 + m_0^2}$, which is a pseudo-relativistic spectrum with a symmetric gap $\Delta = \pm m_0$. This correctly captures the general features of the system but misses the unusual behaviour in the middle of the gap. Let us try to shed some light on this by considering the wavefunction ansatz

$$\psi(x) = \exp\left(\hat{\sigma}^y \int_0^x dx'\, \eta(x')\right)\psi_0\,, \tag{74}$$

where ψ_0 is a spinor constant and $\eta(x)$ is an arbitrary function. Upon taking the spatial derivative, we find that $\partial_x \psi(x) = \eta(x)\hat{\sigma}^y \psi(x)$. This identity can be rewritten as

$$\left[-i\hat{\sigma}^x \partial_x - \eta(x)\,\hat{\sigma}^z\right]\psi(x) = 0\,. \tag{75}$$

Comparing Eq. (75) with the initial Eq. (72), we realise that, for $\eta(x) = m(x)$,

$$\psi(x) = \exp\left(\hat{\sigma}^y \int_0^x dx'\, m(x')\right)\psi_0 \tag{76}$$

is a valid solution to the system and has energy $E = 0$. It is a *fermionic zero mode* or *zero-energy bound state*. We can show that it is a bound state as it is normalisable and it decays in the limits $x \to \pm\infty$. It is enough to choose a certain eigenstate, for instance $\hat{\sigma}^y \psi_0 = \sigma^y \psi_0$ for $\sigma^y = -1$ and $\psi_0 = \begin{pmatrix} u_0 & v_0 \end{pmatrix}^{\mathsf{T}} = \begin{pmatrix} 1 & i \end{pmatrix}^{\mathsf{T}}$, to find the explicit solutions (see Figure 10).

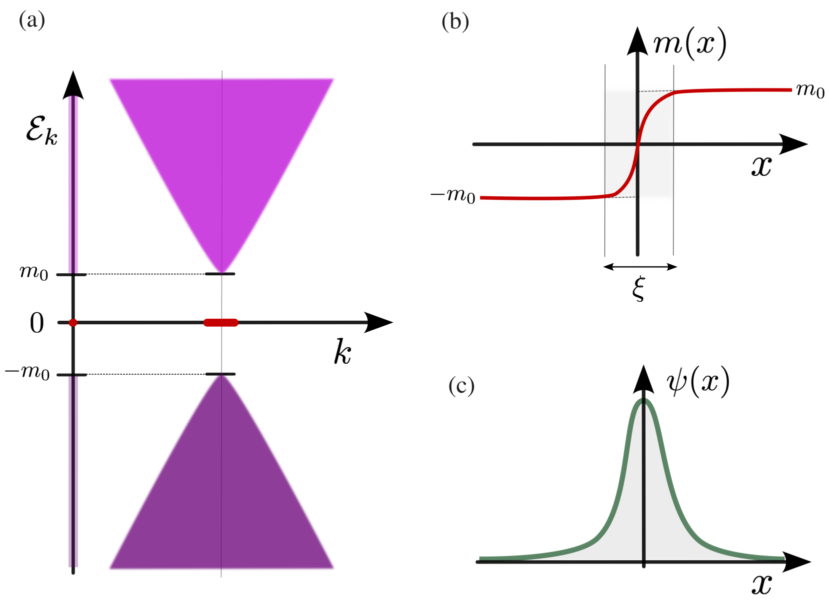

Figure 10: One-dimensional Dirac fermion in the presence of a kink. (a) Dirac dispersion relation featuring a zero energy state. (b) Kink mass profile. (c) Computed solution for the Dirac matter field.

References

[1] M. Nakahara, *Geometry, topology and physics*, CRC press (2018).

[2] C. Nash and S. Sen, *Topology and geometry for physicists*, Elsevier (1988).

[3] M. Stone and P. Goldbart, *Mathematics for physics: a guided tour for graduate students*, Cambridge University Press (2009).

[4] S. Coleman, *Aspects of symmetry: selected Erice lectures*, Cambridge University Press, doi:10.1017/CBO9780511565045 (1988).

[5] S. M. Girvin and K. Yang, *Modern condensed matter physics*, Cambridge University Press (2019).

[6] H. Kleinert, *Multivalued fields*, AIP Conference Proceedings **966**(1), 293 (2008), doi:10.1063/1.2837010.

[7] N. D. Mermin, *The topological theory of defects in ordered media*, Rev. Mod. Phys. **51**, 591 (1979), doi:10.1103/RevModPhys.51.591.

[8] T. T. Wu and C. N. Yang, *Concept of nonintegrable phase factors and global formulation of gauge fields*, Phys. Rev. D **12**, 3845 (1975), doi:10.1103/PhysRevD.12.3845.

[9] P. M. Chaikin, T. C. Lubensky and T. A. Witten, *Principles of condensed matter physics*, vol. 10, Cambridge university press Cambridge, doi:10.1017/CBO9780511813467 (1995).

[10] G. E. Volovik, *The universe in a helium droplet*, vol. 117, OUP Oxford (2003).

[11] L. P. Kadanoff and H. Ceva, *Determination of an operator algebra for the two-dimensional ising model*, Phys. Rev. B **3**, 3918 (1971), doi:10.1103/PhysRevB.3.3918.

[12] P. Hasenfratz, *A puzzling combination: Disorder × order*, Physics Letters B **85**(4), 338 (1979), doi:10.1016/0370-2693(79)91267-X.

[13] E. Fradkin, *Disorder operators and their descendants*, Journal of Statistical Physics **167**(3-4), 427 (2017), doi:10.1007/s10955-017-1737-7.

[14] E. C. Marino, *Quantum field theory approach to condensed matter physics*, Cambridge University Press, doi:10.1017/9781139696548 (2017).

[15] E. Fradkin, *Field Theories of Condensed Matter Physics*, Cambridge University Press, 2 edn., doi:10.1017/CBO9781139015509 (2013).

[16] M. V. Berry, *Quantal phase factors accompanying adiabatic changes*, P. Roy. Soc. Lond. A Mat. **392**(1802), 45 (1984), doi:10.1098/rspa.1984.0023.

[17] H.-h. Wu, *Shiing-shen chern: 1911–2004*, Bulletin of the American Mathematical Society **46**(2), 327 (2009).

[18] N. Hitchin, *Shiing-shen chern 1911–2004* (2006).

[19] N. Manton and P. Sutcliffe, *Topological solitons*, Cambridge University Press (2004).

[20] H. Kleinert, *Gauge fields in condensed matter*, vol. 2, World Scientific, doi:10.1142/0356 (1989).

[21] H. Nastase, *Classical field theory*, Cambridge University Press (2019).

[22] T. Vachaspati, *Kinks and domain walls: An introduction to classical and quantum solitons*, Cambridge University Press, doi:10.1017/CBO9780511535192 (2007).

[23] R. Jackiw and C. Rebbi, *Spin from isospin in a gauge theory*, Phys. Rev. Lett. **36**, 1116 (1976), doi:10.1103/PhysRevLett.36.1116.

[24] J. Goldstone and F. Wilczek, *Fractional quantum numbers on solitons*, Phys. Rev. Lett. **47**, 986 (1981), doi:10.1103/PhysRevLett.47.986.