IS THE VECTOR POTENTIAL REAL?

Dipankar Kaundilya*

Department of Physics, IISc Bangalore, Bangalore-560012, Karnataka, India

*Corresponding Author: kaundilyadipankar@gmail.com

Abstract: This article explores the fundamental relationship between symmetry principles and topological structures within a consistent theoretical framework. By demonstrating that gauge transformations are canonical, we establish that the vector potential is not merely a mathematical auxiliary, but rather it functions as a primary coordinate of the Hamiltonian and quantum mechanical descriptions. Through a mathematical analysis of solenoid and toroidal geometries, we derive the Aharonov-Bohm (AB) phase shift and energy spectra, illustrating how non-trivial topology induces measurable physical effects in strictly field-free regions. The discussion extends to the Dirac quantization condition via the Wu-Yang construction of magnetic monopoles and concludes by framing the potential-centric view as a necessary precursor to Lattice Gauge Theory and the geometric connections of General Relativity. This synthesis reinforces the shift in physical ontology from local fields to global, gauge-invariant holonomies as the foundational entities of modern physics.

Keywords: Aharonov-Bohm Effect, Gauge Symmetry, Topology, Vector Potential, Dirac Monopole, Holonomy.

1 Introduction

Symmetry principles are foundational to physics, where Abelian gauge symmetry gives rise to electromagnetism (EM). The unification of fundamental interactions, such as the electroweak theory, relies heavily on more complex, non-Abelian gauge fields and mechanisms for broken symmetry. In quantum mechanics (QM), local gauge transformations of wavefunctions necessitate the introduction of vector potentials, highlighting a deep link between seemingly abstract mathematical concepts and physical reality.

Historically, electromagnetism initially relegated potentials to the status of mathematical auxiliaries, favoring the directly measurable $\mathbf{E}$ and $\mathbf{B}$ fields. However, the advent of quantum mechanics revealed that the phase of the wavefunction couples directly to the potential $\mathbf{A}$, suggesting a deeper, non-local reality[1]. This article re-examines this relationship and proves that the vector potential is a real physical entity. This view is established in three steps, first by proving the canonical nature of gauge transformations and second, by demonstrating that local phase symmetry in quantum mechanics necessitates the existence of the gauge field; and third, by analyzing the topological influence of the Aharonov-Bohm effect in solenoid and toroidal geometries. We bridge these foundational concepts with modern developments in Dirac monopoles and Lattice Gauge Theory [2]. Ultimately, we argue for a potential-centric view underpinned by gauge theory and topology, moving toward the geometric descriptions found in General Relativity.

2 Gauge transformation is canonical

It is perhaps surprising that the Hamiltonian equations for a particle in an electromagnetic field involve potentials rather than fields [1]. Fields are defined in terms of forces, while potentials lack an obvious operational definition and are not uniquely defined [1]. In quantum mechanics, potentials govern both orbits and phase relations, affecting interference patterns even in field-free regions [1]. This clarifies the use of potentials in quantum mechanics, though it remains less clear in classical mechanics [1].

The potentials $\mathbf{A}$ and $\mathbf{\Phi}$ determine the fields according to the relations:

$$B = \nabla \times A \quad \text{and} \quad E = -\nabla\Phi - \frac{\partial A}{\partial t} \tag{2.1}$$

New potentials can be introduced via a gauge transformation:

$$A' = A + \nabla\Lambda \quad \text{and} \quad \Phi' = \Phi - \frac{\partial \Lambda}{\partial t} \tag{2.2}$$

where $\Lambda(r, t)$ is an arbitrary differentiable function [1].

This gauge transformation leaves the values of B and E unchanged, demonstrating that A and Φ are not uniquely determined by the forces they exert [1]. This non-uniqueness can simplify computations but may also introduce complications, as even simple fields can become time-dependent and complex under a gauge transformation with an arbitrary Λ [1].

Such transformations encompass a variety of physical symmetries. For instance, a rotation in phase space, defined by $Q = q\cos\alpha - p\sin\alpha$ and $P = q\sin\alpha + p\cos\alpha$, preserves the canonical relations while merely shifting the perspective of the observer. Similarly, hyperbolic transformations, such as the Bogoliubov transformation used in many-body theory, redistribute the identity of particles and excitations while maintaining the underlying commutation relations.

A transformation $(p_i, q_i) \to (P_i, Q_i)$ is canonical if it satisfies the following Poisson brackets:

$$[P_i, P_k]_{q,p} = 0 = [Q_i, Q_k]_{q,p} \quad \text{and} \quad [Q_i, P_k]_{q,p} = \delta_{ik} \tag{2.3}$$

Considering the Hamiltonians $H(p_i, q_i, t)$ and $H'(P_i, Q_i, t)$ derived from Lagrangians L and L', we have:

$$H(p_i, q_i, t) = \sum_j p_j \dot{q}_j - L(q_i, \dot{q}_i, t) \tag{2.4}$$

and

$$H'(P_i, Q_i, t) = \sum_j P_j \dot{Q}_i - L'(q_i, \dot{q}_i, t) \tag{2.5}$$

where P_i is related to p_i by a derivative of a function $f(q_i, t)$:

$$P_i = \frac{\partial L'}{\partial \dot{q}_i} = p_i + \frac{\partial}{\partial \dot{q}_i}\frac{df(q_i, t)}{dt} = p_i + \frac{\partial}{\partial q_i} f(q_i, t) \tag{2.6}$$

and $Q_i = q_i$.

To show that the transformation $(p_i, q_i) \to (P_i, Q_i)$ is canonical, we calculate the Poisson brackets for (P_i, Q_i) :

$$[Q_i, Q_k]_{q,p} = \sum_j \left[\frac{\partial Q_i}{\partial q_j}\frac{\partial Q_k}{\partial p_j} - \frac{\partial Q_i}{\partial p_j}\frac{\partial Q_k}{\partial q_j} \right] = 0 \tag{2.7}$$

as $\frac{\partial Q_i}{\partial p_j} = 0$.

$$[P_i, P_k]_{q,p} = \sum_j \left[\frac{\partial P_i}{\partial q_j}\frac{\partial P_k}{\partial p_j} - \frac{\partial P_i}{\partial p_j}\frac{\partial P_k}{\partial q_j} \right] = \frac{d}{dt}\left[\frac{\partial^2 f}{\partial q_k \partial q_i} - \frac{\partial^2 f}{\partial q_i \partial q_k} \right] = 0 \tag{2.8}$$

and finally

$$[Q_i, P_k]_{q,p} = \sum_j \left[\frac{\partial Q_i}{\partial q_j} \frac{\partial P_k}{\partial p_j} - \frac{\partial Q_i}{\partial p_j} \frac{\partial P_k}{\partial q_j} \right] = \delta_{ik} \tag{2.9}$$

Thus, gauge transformations that connect Lagrangians differing by a total time derivative of coordinate and time are canonical.

By proving that gauge transformations are canonical, we establish that the vector potential is not merely a mathematical convenience but a consistent coordinate of the Hamiltonian framework. This formal equivalence allows the potential to maintain its structural integrity as we transition from the classical equations of motion to the quantum mechanical wavefunction. That is the generalized coordinates (the potentials $\mathbf{A}, \mathbf{\Phi}$) we use to build the Schrödinger equation are as fundamental as the positions and momenta in classical description. Consequently, the mathematical freedom inherent in the choice of gauge provides the necessary degree of freedom to accommodate the local phase symmetries of quantum theory.

Once we grasp this idea, then we can appreciate that while the canonical momenta **p** is the fundamental coordinate of the Hamiltonian, it is the kinetic or mechanical momentum $\mathbf{\Pi} = \mathbf{p} - \mathbf{qA}$ that couples to the particle's velocity. We can easily verify the commutation relation, $[\Pi_i, \Pi_j] = q\epsilon_{ijk}B_k$, i.e. the kinetic momenta commute only if the magnetic field is zero. This reinforces the concept that the vector potential **A** functions like a coordinate in a field free region.

3 Local Phase Symmetry and the Gauge Field

This mathematical freedom within the gauge transformation finds its physical counterpart in the phase of the quantum mechanical wavefunction. To trace a deep link between Electromagnetism and Quantum Mechanics, we observe that the arbitrary nature of the vector potential **A** is not just a classical redundancy, but the very mechanism that allows the overall phase of Ψ to be changed locally without affecting physical observables. This suggests that the "background" gauge field is inextricably tied to the local symmetry of the particle's probability amplitude.

Let's analyze the vector potential.

$$B = \nabla \times A \tag{3.1}$$

$$A(x) \to A'(x) = A(x) + \nabla\Lambda(x) \tag{3.2}$$

Now we write the equation in the integral form:

$$\oint_\Gamma A \cdot dx = \iint_{S(\Gamma)} B \cdot ds = F(\Gamma). \tag{3.3}$$

That is, the line integral of the vector potential around a closed loop Γ is equal to the surface integral of the magnetic field over the area $S(\Gamma)$ enclosed by the loop. The magnetic flux $F(\Gamma)$ is then invariant under the Gauge transformation because it is directly related to B. Note also that the gradient term $\nabla\Lambda(x)$ does not contribute to the line integral over a closed loop.

The integral form in Equation (3.3) reveals that the vector potential is more than a computational aid; it functions as a **connection** on a $U(1)$ **fiber bundle**. In this geometric view, the phase of the wavefunction is not a global constant but a local property that "rotates" as the particle moves through space. The physical result of traversing a closed loop Γ is the acquisition of **holonomy**—the net group element $H(\Gamma) = \exp(\frac{iq}{\hbar} \oint_\Gamma \mathbf{A} \cdot d\mathbf{l})$. Even when the local magnetic field B (the curvature of the bundle) vanishes along the path, the holonomy can remain non-trivial ($H(\Gamma) \neq 1$) due to the **non-trivial topology** of the space. This identifies the Aharonov-Bohm effect as a global manifestation of gauge symmetry, providing a direct conceptual bridge to the **Wilson loops** of lattice gauge theory and the geometric connections of **General Relativity** discussed in the concluding sections of this article.

Now, let's analyze the phase of a wavefunction. We are free to change the phase under the following prescription [3]:

$$\Psi(x) \to \Psi'(x) = e^{i\theta}\Psi(x) \tag{3.4}$$

Note that the relevant physical quantities, e.g., the probability density $|\Psi|^2$, remain unchanged.

Something very interesting happens if we choose the phase angle θ independently in different places. That is, QM admits the local phase transformations:

$$\Psi(x) \to e^{i\theta(x)}\Psi(x) \tag{3.5}$$

where the angle θ is now allowed to be a function of the position x. Quantities such as probability density $\Psi^*\Psi$ and expectation values $\int \Psi^* V \Psi d^3x$ remain unchanged even under local phase transformations.

What about probability currents $\nabla\Psi$ and kinetic energy operators ∇^2 in the Schrödinger equation?

$$\nabla\Psi(x) \to \nabla\Psi'(x) = \nabla[e^{i\theta(x)}\Psi(x)] = e^{i\theta(x)}\nabla\Psi(x) + i\nabla\theta(x)e^{i\theta(x)}\Psi(x) \tag{3.6}$$

Note that the second term does not comply with the fact that everything simply changes by a phase factor. However, for a charged particle, the conjugate momentum p is replaced by[4]:

$$p \to p - qA(x) \tag{3.7}$$

This substitution is known as the **Minimal Coupling**. For a while it may appear as a post-hoc rule to link the particle's motion to the EM field, but it is necessary because it is the only coupling that allows the extra gradient term produced by a local phase shift $(\nabla\theta(x))$, to be exactly canceled by the gauge transformation of the vector potential $(\nabla\Lambda(x))$. Without this specific term, the Schrödinger equation would not be gauge invariant and the choice of gauge would indeed change the probability density.

Using canonical quantization, p is replaced by the operator $-i\nabla$. Hence, we define a covariant derivative D as:

$$-i\nabla \to -i[\nabla - iqA(x)] \equiv -iD \tag{3.8}$$

Now we perform the momentum operation on Ψ:

$$(\nabla - iqA)\Psi \to (\nabla - iqA')\Psi' = e^{i\theta(x)}(\nabla - iqA)\Psi + i\nabla(\theta(x) - q\Lambda(x))e^{i\theta(x)}\Psi \tag{3.9}$$

Note that the extraneous term would cancel out if we set:

$$\Lambda(x) = \frac{\theta(x)}{q} \tag{3.10}$$

Then everything simply acquires the same phase, i.e.,

$$D\Psi \to D'\Psi' = e^{i\theta(x)}[D\Psi] \tag{3.11}$$

where D' indicates that the covariant derivative is to be constructed using the transformed vector potential A'.

Thus, if the original wavefunction Ψ satisfies the Schrödinger equation:

$$\left[-\frac{1}{2m}(\nabla - iqA)^2 + V\right]\Psi(x) = E\Psi(x) \tag{3.12}$$

then the transformed wavefunction Ψ' also satisfies the same equation with A' and the same eigenvalue E, and hence both would describe the same physics.

A transformation that has no effect on the physics—that is to say, one that maps a solution of the equation into another solution—is called a **symmetry**; if the transformation can be implemented at different points, the symmetry is said to be a **local symmetry** or a **Gauge symmetry**.

This invariance demonstrates that the vector potential $\mathbf{A}$ is more than a mathematical auxiliary; it is a fundamental requirement for the local gauge symmetry of the wavefunction. While classical intuition suggests that the physical influence of electromagnetism is restricted to regions where the fields $\mathbf{E}$ and $\mathbf{B}$ are non-zero, this quantum mechanical coupling to the potential implies a "non-local" interaction. To move from this theoretical framework to an empirical proof, we must examine an experimental scenario where the potential acts in a truly field-free region: the Aharonov-Bohm effect.

4 Aharonov-Bohm Effect

The preceding derivation establishes the vector potential $\mathbf{A}$ as a mathematical necessity for the local gauge invariance of the Schrödinger equation. However, to confirm that this "background swirl" is a physical reality rather than a gauge-dependent artifact, we must identify a scenario where $\mathbf{A}$ exerts a measurable influence in the absence of a Lorentz force. This empirical test is provided by the **Aharonov-Bohm (AB) effect**, which occurs when a charged particle travels through a multiply connected space. By analyzing the specific geometries of the **solenoid** and the **toroid**, we can demonstrate how the global topology of the potential induces an observable phase shift, even when the particle is strictly confined to field-free regions.

4.1 Solenoid

We model an infinitely long solenoid as a cylinder of radius R centered on the z-axis, carrying a uniform internal magnetic field $B_0\hat{k}$ while the field outside is zero. Let's calculate the vector potential for the interior region $r \leq R$ and for the exterior region $r \geq R$.

4.1.1 Vector Potential for Solenoid

The vector potential must respect the cylindrical symmetry, because the source (the solenoid) is invariant under rotations about the z-axis and translations along it, the vector potential must reflect this cylindrical symmetry. We assume that it is purely azimuthal, such that $A = A(r)\hat{\phi}$. Then we utilize Stokes's theorem to relate the line integral of A around a circular loop of radius r to the magnetic flux F through the loop:

$$\oint_{\Gamma} A \cdot dl = \oint_{S(\Gamma)} B \cdot ds = \oint_{S(\Gamma)} (\nabla \times A) \cdot ds = F(\Gamma) \tag{4.1}$$

This choice of a purely azimuthal potential $\mathbf{A} = A(r)\hat{\phi}$ is dictated by the cylindrical symmetry of the infinite solenoid. While the magnetic field $\mathbf{B}$ is zero for $r > R$, the continuity of the vector potential across the boundary ensures that the information of the enclosed flux is "carried" into the exterior region as illustrated in Figure 1.

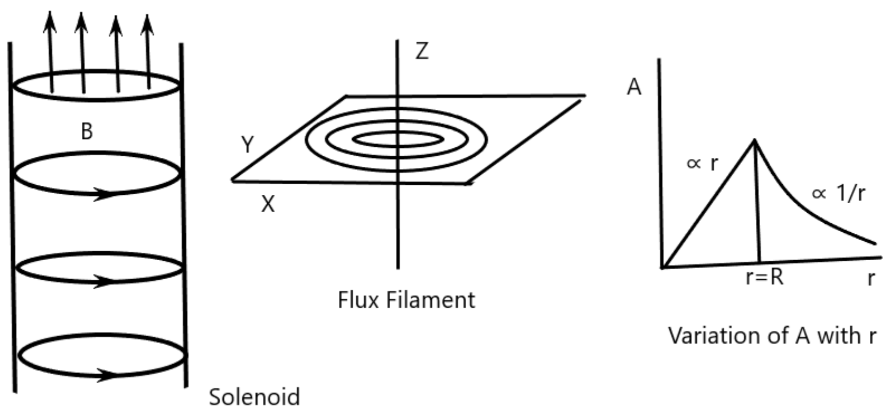

Figure 1: Vector potential of an infinite solenoid. The magnetic field $\mathbf{B}$ is uniform inside ($r < R$) and zero outside ($r > R$). However, the azimuthal vector potential $\mathbf{A}$ is non-zero everywhere, falling off as $1/r$ in the exterior region, which is the physical basis for the Aharonov-Bohm effect.

For the interior region $r \leq R$, the magnetic field inside the solenoid is uniform. The enclosed magnetic flux through a loop of radius r is:

$$\int A \cdot dl = A(2\pi r) = \int_0^r B_0(2\pi r')dr' = B_0\pi r^2 \tag{4.2}$$

$$A = \frac{B_0 r}{2}\hat{\phi} \tag{4.3}$$

For the exterior region $r \geq R$, the magnetic field outside the solenoid is zero. However, the loop of radius r still encloses the total flux contained within the cylinder of radius R:

$$\int A \cdot dl = A(2\pi r) = \int_0^R B_0(2\pi r')dr' = B_0\pi R^2 \tag{4.4}$$

$$A = \frac{B_0 R^2}{2r}\hat{\phi} \tag{4.5}$$

The vector potential for the given solenoid is:

$$A = \begin{cases} \frac{B_0 r}{2}\hat{\phi} & \text{for } r \leq R \\ \frac{B_0 R^2}{2r}\hat{\phi} & \text{for } r \geq R \end{cases} \tag{4.6}$$

Note that at the value at the boundary $r = R$, $A = \frac{B_0 R}{2}$ matches for the continuity of both inside and outside the solenoid. This proves that the vector potential is a ”well-behaved” coordinate that transitions smoothly between the hidden flux and the reachable space.

To check the differentiability of the vector potential at the boundary $r = R$, we calculate the radial derivative dA/dr from both sides using Equation (4.6):

$$\left.\frac{dA}{dr}\right|_{r\to R^-} = \frac{d}{dr}\left(\frac{B_0 r}{2}\right) = \frac{B_0}{2} \quad \text{vs.} \quad \left.\frac{dA}{dr}\right|_{r\to R^+} = \frac{d}{dr}\left(\frac{B_0 R^2}{2r}\right) = -\frac{B_0}{2} \tag{4.7}$$

Since the derivatives $\frac{B_0}{2}$ and $-\frac{B_0}{2}$ are not equal, the vector potential is **not differentiable** at $r = R$, resulting in a kink at the boundary.

4.1.2 Flux Filament

Now, if we shrink the radius of the solenoid to zero, we create a flux filament (an ideal line of flux) along the z axis. When the radius R shrinks to zero while the magnetic field B_0 increases such that the total magnetic flux F_0 remains constant, we create a flux filament. The total flux is defined by:

$$F_0 = \lim_{R\to 0}(B_0\pi R^2) \tag{4.8}$$

In this limit, the magnetic field B becomes a Dirac Delta function along the z axis:

$$B = F_0\delta(x)\delta(y)\hat{k} \tag{4.9}$$

Since the radius $R \to 0$, we only need to consider the region $r > R$. Then $\oint A \cdot dl = F_0$. Assuming azimuthal symmetry $A = A(r)\hat{\phi}$, the path integral over a circle of radius r is $A(r)2\pi r = F_0$, and hence:

$$A(r) = \frac{F_0}{2\pi r} \tag{4.10}$$

This expression is valid for all $r > 0$. Note that the vector potential is singular at $r = 0$, reflecting the infinite density of the flux line at the origin. Consider a vector potential defined by:

$$A_x = -\frac{F_0}{2\pi}\frac{y}{x^2+y^2}, \quad A_y = \frac{F_0}{2\pi}\frac{x}{x^2+y^2}, \quad A_z = 0 \tag{4.11}$$

Let's evaluate its curl in Cartesian coordinates:

$$\nabla \times A = \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}\right)\hat{i} + \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right)\hat{j} + \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right)\hat{k} \tag{4.12}$$

Since $A_z = 0$ and A_x, A_y do not depend on z, all derivatives involving z or A_z are zero. Thus, we only have to calculate the z component:

$$\frac{\partial A_y}{\partial x} = \frac{y^2 - x^2}{(x^2 + y^2)^2}, \quad \frac{\partial A_x}{\partial y} = \frac{y^2 - x^2}{(x^2 + y^2)^2} \tag{4.13}$$

Hence, the z component of the curl is also zero. Since all components are zero, $\nabla \times A = 0$ everywhere except at the origin $(0, 0)$ where the vector potential becomes singular. As $B = \nabla \times A$, this represents a region with zero magnetic field but non-zero vector potential.

We can see that A is ill-defined (singular) along the entire z axis. For any point on the z axis, the coordinates are $(0, 0, z)$. The denominator $x^2 + y^2$ becomes zero at $x = 0, y = 0$, meaning the vector potential is singular and the field blows up at the origin for all values of z.

This singularity represents an infinitely thin solenoid located exactly on the z axis. While the curl of A is zero everywhere else, it is a Dirac delta function at the origin:

$$B = \nabla \times A = F_0 \delta(x)\delta(y)\hat{k} \tag{4.14}$$

Because the z axis is ill-defined, we must "puncture" it from our coordinate system. This makes the space multiply connected, a topological requirement for the Aharonov-Bohm phase shift to exist.

Even though the curl is zero in the reachable region, the line integral around any loop Γ enclosing the z axis is non-zero, i.e., $\oint_\Gamma A \cdot dl = F_0$. To use Stokes' theorem, we first evaluate the circulation of A along a circular path Γ of radius R in the x-y plane, centered on the z axis.

We use polar coordinates where $x = R\cos\phi$, $y = R\sin\phi$, and $dl = (-R\sin\phi\hat{i} + R\cos\phi\hat{j})d\phi$. Then the vector potential is:

$$A = -\frac{\sin\phi}{R}\hat{i} + \frac{\cos\phi}{R}\hat{j} \tag{4.15}$$

Thus, the circulation is:

$$\int_0^{2\pi} \left(-\frac{F_0}{2\pi}\frac{\sin\phi}{R}\hat{i} + \frac{F_0}{2\pi}\frac{\cos\phi}{R}\hat{j} \right) \cdot (-R\sin\phi\hat{i} + R\cos\phi\hat{j})d\phi = \frac{F_0}{2\pi}\int_0^{2\pi} (\sin^2\phi + \cos^2\phi)d\phi = F_0 \tag{4.16}$$

According to Stokes' theorem, the line integral around a closed loop is equal to the flux of the curl through the surface enclosed by that loop:

$$\oint_\Gamma A \cdot dl = \iint_{S(\Gamma)} (\nabla \times A) \cdot ds = F_0 \tag{4.17}$$

Since we previously proved that $\nabla \times A = 0$ everywhere except at the origin, the entire value of the integral (F_0) must be concentrated at the singularity on the z axis.

To satisfy the condition that the integral over any area containing the origin is F_0 while the function remains zero elsewhere, we use the two-dimensional Dirac delta function. Thus, the curl of the vector potential on the z axis is:

$$B = \nabla \times A = F_0 \delta(x)\delta(y)\hat{k} \tag{4.18}$$

This represents an infinitely thin solenoid (a magnetic flux filament) carrying a total flux of F_0.

To understand why the z-axis must be 'punctured', consider the difference between a ball and a torus. In a ball (simply connected), any loop of string can be shrunk to a point. In a torus (multiply connected), a loop passing through the center hole cannot be shrunk to a point without leaving the surface. In the AB effect, the magnetic flux inside the solenoid acts as the 'hole'. Even if the particle never touches the flux, the fact that its path encircles the hole makes the interaction 'global' and topologically distinct from a path in a field-free, simply connected space.

4.1.3 Gauge Transformation and Phase Shift

Let's see how this result changes if we apply a gauge transformation. Transform the vector potential A into A' by adding the gradient of a scalar field $\Lambda(x, y, z)$:

$$A' = A + \nabla\Lambda \tag{4.19}$$

Applying the curl operator to both sides:

$$\nabla \times A' = \nabla \times A + \nabla \times (\nabla\Lambda) \tag{4.20}$$

Using the vector identity $\nabla \times (\nabla\Lambda) = 0$, we find:

$$\nabla \times A' = \nabla \times A = F_0\delta(x)\delta(y)\hat{k} \tag{4.21}$$

This demonstrates that the magnetic field is a gauge-invariant quantity. While A changes depending on the gauge, the physical observable—the magnetic flux concentrated on the z-axis—remains exactly the same.

Under a gauge transformation, the wavefunction of a particle undergoes a local phase shift to ensure that the physical predictions of the Schrödinger equation remain unchanged. When we change the potentials:

$$A' = A + \nabla\Lambda, \quad \Phi' = \Phi - \frac{\partial\Lambda}{\partial t} \tag{4.22}$$

The Hamiltonian for a particle with charge q becomes:

$$H' = \frac{1}{2m}(-i\hbar\nabla - qA')^2 + q\Phi' \tag{4.23}$$

To keep the physics invariant, we define the new wavefunction Ψ' as:

$$\Psi' = \Psi e^{i\frac{q}{\hbar}\Lambda} \tag{4.24}$$

This is a local phase shift because Λ depends on space-time. In the Aharonov-Bohm effect, this implies that even in a field-free region ($B = 0$), the non-zero potential A shifts the phase of the wavefunction. This shift is physically observable as an interference pattern displacement because the phase depends on the integral of A along the particle's path. While originally demonstrated with electrons, this effect has been precisely verified in modern contexts using ultracold atoms [5].

Let's calculate the total phase difference $\Delta\phi$ for the two paths around the solenoid. As a particle moves through a region in the presence of a vector potential, its wavefunction picks up a phase factor:

$$\phi = \frac{q}{\hbar}\int \mathbf{A} \cdot d\mathbf{l} \tag{4.25}$$

Consider a beam of charged particles split into two paths, path_1 and path_2, both starting at point S and ending at point P, encircling the z-axis as illustrated in Figure 2.

The phase accumulated along each path is:

$$\phi_1 = \frac{q}{\hbar}\int_{\text{path}_1} \mathbf{A} \cdot d\mathbf{l}, \quad \phi_2 = \frac{q}{\hbar}\int_{\text{path}_2} \mathbf{A} \cdot d\mathbf{l} \tag{4.26}$$

The observable phase difference is:

$$\Delta\phi = \phi_1 - \phi_2 = \frac{q}{\hbar}\left(\int_{\text{path}_1} \mathbf{A} \cdot d\mathbf{l} - \int_{\text{path}_2} \mathbf{A} \cdot d\mathbf{l}\right) \tag{4.27}$$

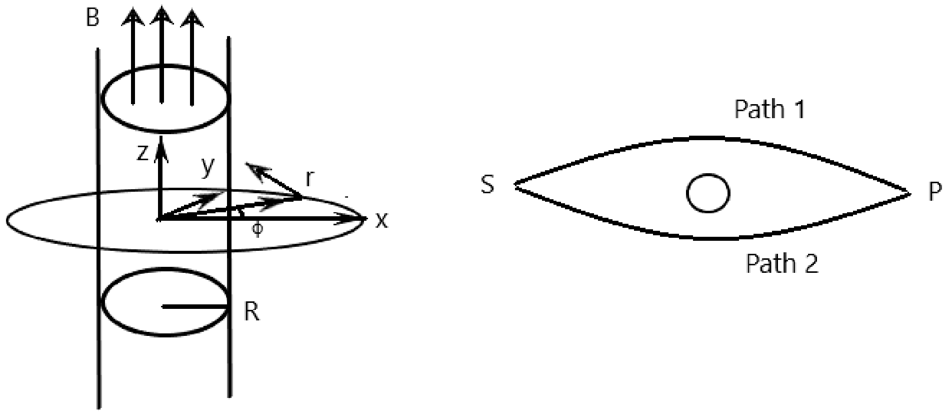

Figure 2: Interference paths encircling a magnetic flux filament. This diagram shows a particle beam splitting at source S into two paths that encircle a central flux filament before recombining at point P. It illustrates how the vector potential $\mathbf{A}$ induces a phase shift even in regions where the magnetic field $\mathbf{B}$ is zero.

Reversing the direction of path$_2$ turns the expression into a closed loop integral around the solenoid:

$$\Delta\phi = \frac{q}{\hbar} \oint_{\Gamma} \mathbf{A} \cdot d\mathbf{l} \tag{4.28}$$

Using Stokes' theorem, we know that the loop integral is the total magnetic flux enclosed by the paths. Thus, the total phase shift is:

$$\Delta\phi = \frac{q}{\hbar} F_0 \tag{4.29}$$

This result is extraordinary because it shows that the phase shift depends only on the total flux enclosed, even though the charged particle never passes through the region where the magnetic field is non-zero.

4.1.4 Centrifugal Barrier and Energy Spectrum

Now, suppose there is a repulsive barrier around the z-axis, so we do not have to worry about the singular nature of the vector potential there. The centrifugal repulsive barrier is an effective potential energy term that arises when we transform the kinetic energy from Cartesian into curvilinear cylindrical coordinates. Classically, for a particle with mass m and angular momentum L, the kinetic energy associated with its rotation is:

$$K_{\text{rot}} = \frac{L^2}{2mr^2} \tag{4.30}$$

As the particle approaches the center ($r \to 0$), the energy required to maintain that angular momentum increases quadratically, pushing the particle away from the axis.

In QM, we replace the classical L with the operator $\hat{L}_z = -i\hbar\frac{\partial}{\partial\phi}$. When acting on a wavefunction with angular dependence $e^{in\phi}$, the radial Schrödinger equation sees a term:

$$V_{\text{centrifugal}}(r) = \frac{n^2\hbar^2}{2mr^2} \tag{4.31}$$

This term prevents particles with non-zero angular momentum ($n \neq 0$) from reaching the origin, as the cost in kinetic energy becomes infinite.

However, in the presence of the vector potential $\mathbf{A}$, the kinetic momentum is shifted. The barrier is no longer governed by the canonical angular momentum $n\hbar$, but by the shifted mechanical angular momentum $(n\hbar - \alpha)$; hence, we obtain the effective potential as:

$$V_{\text{eff}} = \frac{(n\hbar - \alpha)^2}{2mr^2} + V(r) \tag{4.32}$$

where $\alpha = \frac{qF_0}{2\pi}$ is the flux coupling constant.

This is a sophisticated way to view the Aharonov-Bohm (AB) effect by observing how the hidden flux modifies the centrifugal barrier. We can write the vector potential in cylindrical coordinates as $\mathbf{A} = \frac{F_0}{2\pi r}\hat{\phi}$. The general vector potential for a solenoid with any flux F_0 in Cartesian components is:

$$\mathbf{A} = \frac{F_0}{2\pi}\left(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2}, 0\right) \tag{4.33}$$

In cylindrical coordinates (r, ϕ, z), assuming the motion is confined to the x-y plane ($p_z = 0$), the operator $(-i\hbar\nabla - q\mathbf{A})$ in the $\hat{\phi}$ direction is:

$$\hat{p}_\phi = -\frac{i\hbar}{r}\frac{\partial}{\partial\phi} - \frac{qF_0}{2\pi r} \tag{4.34}$$

We apply the Hamiltonian to a wavefunction of the form $\Psi(r,\phi) = R(r)e^{in\phi}$, where n is an integer to ensure the wavefunction is single-valued, $\Psi(\phi) = \Psi(\phi + 2\pi)$. The kinetic energy term $\frac{\hat{p}_\phi^2}{2m}$ acting on the angular part yields:

$$\frac{1}{2mr^2}\left(-i\hbar\frac{\partial}{\partial\phi} - \frac{qF_0}{2\pi}\right)^2 e^{in\phi} = \frac{1}{2mr^2}\left(n\hbar - \frac{qF_0}{2\pi}\right)^2 e^{in\phi} \tag{4.35}$$

Thus, we see that $V_{\text{centrifugal}} \propto \frac{(n\hbar-\alpha)^2}{r^2}$. This treatment reveals rich physics: while the canonical angular momentum is still $n\hbar$, the mechanical angular momentum is $(n\hbar - \alpha)$. Since α can be any real number, the particle behaves as if it has fractional angular momentum.

A profound question arises: how is the axially symmetric potential $\mathbf{A} = \frac{F_0}{2\pi r}\hat{\phi}$ consistent with perturbing the angular momentum? This is resolved by the distinction between canonical and kinetic angular momentum. Because the Hamiltonian is rotationally invariant ($[\hat{H}, \hat{L}_z] = 0$), the canonical angular momentum $\hat{L}_z = -i\hbar\frac{\partial}{\partial\phi}$ remains a conserved quantity with integer eigenvalues $n\hbar$. However, the kinetic angular momentum—the quantity associated with the particle's actual velocity—is shifted:

$$\hat{\Pi}_z = \mathbf{r} \times (\mathbf{p} - q\mathbf{A}) = -i\hbar\frac{\partial}{\partial\phi} - \frac{qF_0}{2\pi} \tag{4.36}$$

The vector potential adds a "background swirl" to the momentum operator. The axial symmetry ensures the shift is uniform for all ϕ, allowing n to remain a good quantum number while shifting the physical angular momentum from its vacuum value.

To find the energy spectrum, we evaluate the Hamiltonian for a particle constrained to a circular path of radius R. Assuming the repulsive barrier $V(r)$ restricts the particle to $r = R$, the Hamiltonian reduces to:

$$\hat{H} = \frac{1}{2mR^2}\left(-i\hbar\frac{\partial}{\partial\phi} - \frac{qF_0}{2\pi}\right)^2 \tag{4.37}$$

Using $\Psi_n(\phi) = \frac{1}{\sqrt{2\pi}}e^{in\phi}$, and defining the magnetic flux quantum $F_q = \frac{h}{q} = \frac{2\pi\hbar}{q}$, we obtain the energy spectrum:

$$E_n = \frac{\hbar^2}{2mR^2}\left(n - \frac{F_0}{F_q}\right)^2 \tag{4.38}$$

This spectrum is periodic in F_0 with period F_q. Increasing the flux by F_q shifts the state n into state $n-1$. At zero flux, $E_n = E_{-n}$, preserving the degeneracy between clockwise and counter-clockwise rotations.

4.2 Toroid

For completeness, let us discuss the same features for a toroid. If we merge the two open ends of a solenoid, we obtain a toroid. We know that the field inside an ideal solenoid is uniform and given by $B = \mu_0 nI$, where n is the number of turns per unit length and I is the current. Similarly, for a toroid with current I, turn density n, and mean radius R, we can calculate the uniform magnetic field inside the core.

4.2.1 Sigmoid Feature of Vector Potential

We consider a circular Amperian loop of radius r inside the toroid, concentric with the center. Due to axial symmetry, the magnetic field $\mathbf{B}$ must be tangential to this loop with a constant magnitude at a fixed radius. The total enclosed current I_{enc} depends on the total number of turns N. Given the mean circumference $L = 2\pi R$, the total number of turns is $N = nL = 2\pi Rn$. Since each turn carries current I, the total enclosed current is $I_{\text{enc}} = NI = 2\pi RnI$. Applying Ampère's law, $\oint \mathbf{B} \cdot d\mathbf{l} = \mu_0 I_{\text{enc}}$, for $r = R$ we obtain:

$$B(2\pi R) = \mu_0(2\pi RnI) \implies B = \mu_0 nI \tag{4.39}$$

$$B_0 = \mu_0 nI \tag{4.40}$$

For a given turn density n and current I, we define this fixed interior field as B_0. Consequently, the total trapped flux is F_0, allowing for a unified description consistent with the solenoid case.

Thus, the toroid contains the same magnetic field magnitude as the solenoid. However, while the solenoid field is axial, the toroid field is azimuthal (in the $\hat{\phi}$ direction) and entirely trapped within the device. This makes the toroid the "gold standard" for testing the AB effect; unlike the solenoid, which has fringe fields at the ends, the toroid's closed geometry ensures the magnetic field is essentially zero everywhere outside. The non-zero vector potential outside provides the cleanest proof that the phase shift is a result of the potential itself.

For a toroid centered at the origin in the x-y plane, symmetry dictates that the vector potential $\mathbf{A}$ outside must also be azimuthal to ensure its curl is zero while maintaining non-zero circulation. Assuming $\mathbf{A} = A(r)\hat{\phi}$, we use Stokes' theorem for a circular path of radius r:

$$\oint \mathbf{A} \cdot d\mathbf{l} = \int_0^{2\pi} A(r) r d\phi = 2\pi r A(r) = F_0 \tag{4.41}$$

Hence, $\mathbf{A} = \frac{F_0}{2\pi r}\hat{\phi}$. In Cartesian components, using $\hat{\phi} = -\sin\phi\hat{i} + \cos\phi\hat{j}$ and $r^2 = x^2 + y^2$, the vector potential for a toroid with internal flux F_0 is:

$$A_x = -\frac{F_0}{2\pi}\frac{y}{x^2+y^2}, \quad A_y = \frac{F_0}{2\pi}\frac{x}{x^2+y^2}, \quad A_z = 0 \tag{4.42}$$

Let's describe an ideal toroid with mean radius R, cross-sectional radius a, current flowing through the windings I with a number density n. Using Ampere's circuital law, the magnetic field $\mathbf{B}$ inside the toroid at a distance r from the center is directed azimuthally $(\hat{\phi})$ and is given by:

$$\oint \mathbf{B} \cdot d\mathbf{l} = \mu_0 I_{\text{enc}} \tag{4.43}$$

$$B(2\pi r) = \mu_0 n(2\pi R)I \implies B = \frac{\mu_0 nIR}{r} \tag{4.44}$$

Magnetic field is zero everywhere outside the core. A toroid divides the space in three distinct regions, the hole, the core and the outside. Magnetic flux $F = \int \mathbf{B} \cdot d\mathbf{s}$ is connected to the vector potential via Stokes' theorem so we can calculate the vector potential $\mathbf{A}$ once we have the magnetic flux, $\oint \mathbf{A} \cdot d\mathbf{l} = F$.

Inside the hole $(r < R - a)$ The flux is zero, because there is no magnetic field present.

$$F = \int \mathbf{B} \cdot d\mathbf{s} = 0 \implies \oint \mathbf{A} \cdot d\mathbf{l} = A(2\pi r) = 0 \implies A = 0 \tag{4.45}$$

Outside the core $(r \geq R + a)$ The flux is the integral of B over the enclosed area. The element of area at a distance r from the center is $ds = 2\sqrt{a^2 - (r-R)^2}dr$, and the field is $B(r) = \frac{\mu_0 nIR}{r}$. Thus the total flux enclosed is:

$$F_0 = \int_{R-a}^{R+a} \frac{\mu_0 nIR}{r} \left(2\sqrt{a^2 - (r-R)^2}\right) dr = 2\mu_0 nIR \int_{R-a}^{R+a} \frac{\sqrt{a^2 - (r-R)^2}}{r} dr \tag{4.46}$$

On evaluating the integral we get $F_0 = 2\pi\mu_0 nIR(R - \sqrt{R^2 - a^2})$. As total flux enclosed by the toroid core is F_0, the vector potential A relates to flux via $\oint \mathbf{A} \cdot d\mathbf{l} = F$. Hence:

$$A = \frac{F_0}{2\pi r} = \frac{\mu_0 nIR}{r}(R - \sqrt{R^2 - a^2})\hat{\phi} \tag{4.47}$$

For the thin shell approximation $(R \gg a)$:

$$A = \frac{\mu_0 nIR^2}{r} \left(1 - \left(1 - \frac{a^2}{R^2}\right)^{1/2}\right) \tag{4.48}$$

Using Binomial approximation:

$$A = \frac{\mu_0 nIa^2}{2r} = \frac{(\mu_0 nI)\pi a^2}{2\pi r} \tag{4.49}$$

Set total flux enclosed $F_0 = (\mu_0 nI)(\pi a^2)$. Thus we obtain $A = \frac{F_0}{2\pi r}\hat{\phi}$.

Inside the core $(R - a \leq r \leq R + a)$ In this region, a circular path of radius r encloses only a partial section of the flux. We integrate from the inner edge $(R - a)$ upto the present position r:

$$F(r) = \int_{R-a}^{r} \frac{\mu_0 nIR}{r'} \left(2\sqrt{a^2 - (r'-R)^2}\right) dr' \tag{4.50}$$

Again by Stokes's law we get $A(2\pi r) = F(r)$:

$$A(r) = \frac{\mu_0 nIR}{\pi r} \left[\int_{R-a}^{r} \frac{\sqrt{a^2 - (r'-R)^2}}{r'} dr'\right] \hat{\phi} \tag{4.51}$$

Let's see how the vector potential within the core is transiting smoothly from 0 value of the hole region to $\frac{F_0}{2\pi r}$ value of the outside region.

This transition occurs because the enclosed flux $F(r)$ is a continuous function of the radial distance r. As we move through the core, our path of radius r captures an increasing slice of the toroid's cross-sectional area as depicted in Figure 3.

To check the continuity of $A(r)$, we have $r = R - a$ at the inner edge. Thus the integration interval becomes $[R-a, R-a]$. The area enclosed is zero, so $F(R-a) = 0$. Thus $A(R-a) = 0$, matching the value with the hole region. At the outer edge we have $r = R + a$. The integration covers the entire cross-section $[R-a, R+a]$. Thus the flux enclosed is the total flux F_0. Thus $A = \frac{F_0}{2\pi(R+a)}$, which matches for the outside region for $r = R + a$.

To check the differentiability of the vector potential we ensure that there are no kinks on the boundaries. The transition is smooth because the integrand contains the term $\sqrt{a^2 - (r-R)^2}$. We know by the definition of flux that $\frac{dF}{dr} = B(r) \times \text{Height}(r)$. At the boundaries $r = R \mp a$, thus the height of the circular cross-section is exactly 0. This means that $\frac{dA}{dr}$ is continuous at the boundaries, ensuring there are no kinks in the vector potential as we enter or leave the core region.

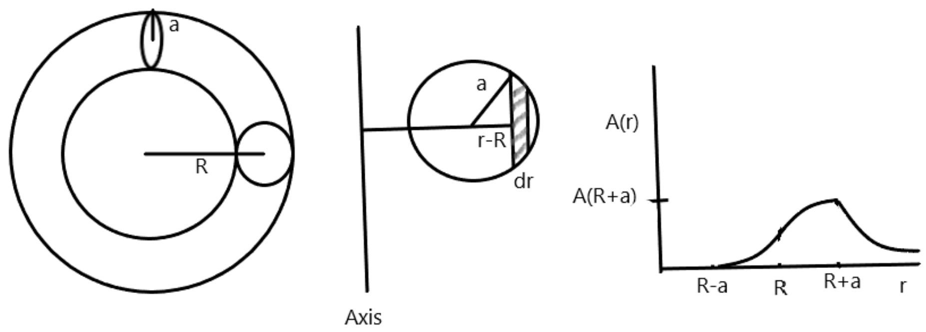

Figure 3: Radial transition of the vector potential $A(r)$. The plot demonstrates the continuous and differentiable sigmoid growth of the potential as it transitions from zero in the hole to its $1/r$ decay in the outside region, following the arcsin behavior in the core.

In the thin-shell limit ($R \gg a$), the growth inside the core is proportional to the area of a circular segment. If we define a normalized coordinate $u = \frac{r-R}{a}$, then limits $[R-a, R+a]$ will change to $[-1, +1]$. And then the vector potential inside the core will scale as:

$$A(u) \propto \frac{1}{2} + \frac{1}{\pi}(u\sqrt{1-u^2} + \arcsin u) \tag{4.52}$$

We can see that at $u = -1$ (inner edge), the term equals to 0, and at $u = +1$ (outer edge), the term equals 1. This $\arcsin u$ behavior is what creates that smooth and elegant sigmoid transition from the hole to outer region for the vector potential.

4.2.2 Phase Shift

In this setup, space is divided into three regions: the hole ($B = 0, A = 0$), the core ($B \neq 0, A \neq 0$), and the outside ($B = 0, A \neq 0$). To verify the AB effect, we utilize two paths—one passing through the hole and the other passing outside the toroid—ensuring the particle never overlaps with the $\mathbf{B}$ field as illustrated in Figure 4.

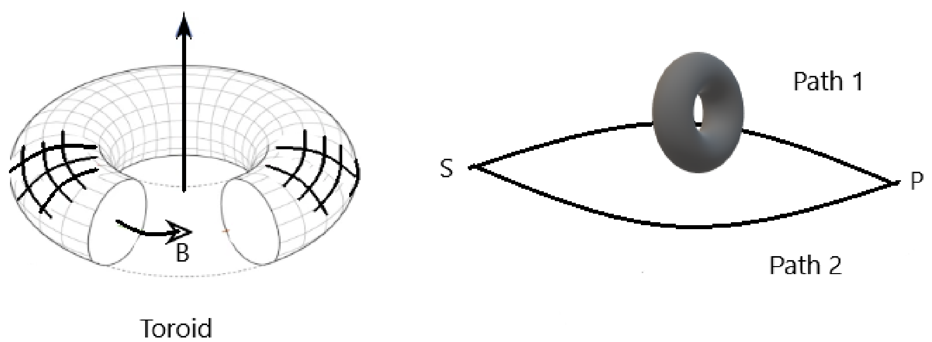

Figure 4: Interference paths around a toroidal magnetic flux. This setup shows paths passing through the center and outside of a toroid. Since the magnetic field is entirely trapped within the core, it proves the Aharonov-Bohm effect by showing the phase shift depends solely on the enclosed flux F_0.

The total wavefunction at the detection point P is the superposition $\Psi = \Psi_1 + \Psi_2$. The relative phase difference $\Delta\phi = \phi_1 - \phi_2$ is the observable that shifts the interference pattern. The phase accumulated along a path is:

$$\phi = \frac{q}{\hbar} \int_{\text{path}} \mathbf{A} \cdot d\mathbf{l} \tag{4.53}$$

The total phase shift is then represented as a closed loop integral Γ enclosing the toroid:

$$\Delta\phi = \frac{q}{\hbar}\oint_{\Gamma} \mathbf{A}\cdot d\mathbf{l} = \frac{q}{\hbar}\iint_{S(\Gamma)} \mathbf{B}\cdot d\mathbf{s} = \frac{qF_0}{\hbar} \tag{4.54}$$

This confirms that the phase shift depends solely on the total enclosed flux F_0, occurring even when the particle never enters the region where $\mathbf{B} \neq 0$. This serves as fundamental evidence that the vector potential $\mathbf{A}$ is a physical entity in quantum mechanics.

4.2.3 Angular Momentum and Energy Spectrum

When we make transition from the classical vector potential of a toroid to the quantum mechanical description of a charged particle interacting with it, we enter into the realm of Aharonov-Bohm effect. Because the magnetic field $\mathbf{B}$ is confined strictly within the core, a charged particle either in the hole or outside regions moves in a field free space ($\mathbf{B} = 0$) but feels the influence of the non-zero vector potential $\mathbf{A}$.

The Hamiltonian for a particle of mass m and charge q in the presence of a vector potential is $H = \frac{1}{2m}(\mathbf{p} - q\mathbf{A})^2$. For a particle constrained to a circular path of radius r in the outside region, where $\mathbf{A} = \frac{F_0}{2\pi r}\hat{\phi}$, the Hamiltonian in cylindrical coordinates becomes:

$$H = -\frac{\hbar^2}{2mr^2}\left(\frac{\partial}{\partial\phi} - i\frac{qF_0}{2\pi\hbar}\right)^2 \tag{4.55}$$

The angular momentum operator in the z direction is $\hat{L}_z = -i\hbar\frac{\partial}{\partial\phi}$. The wavefunction must satisfy the periodicity condition $\Psi(\phi) = \Psi(\phi + 2\pi)$ which leads to the standard integer quantization of the orbital angular momentum $\hat{L}_z\Psi = n\hbar\Psi$, with $n = 0, \pm1, \pm2, \ldots$. The eigenvalues of the mechanical angular momentum (which corresponds to actual velocity) are shifted by an amount $\Delta l_n = \hbar\left(n - \frac{qF_0}{2\pi\hbar}\right)$. Defining the magnetic flux quantum as $F_q = \frac{h}{q}$, the shift is written as $\hbar\left(n - \frac{F_0}{F_q}\right)$.

By substituting the angular momentum eigenvalues into the Hamiltonian, we obtain the energy levels for a particle at a fixed radius r:

$$E_n = \frac{\hbar^2}{2mr^2}\left(n - \frac{F_0}{F_q}\right)^2 \tag{4.56}$$

We infer several features from these results. The energy levels are periodic functions of the total enclosed flux F_0. If the flux is an integer multiple of the flux quantum F_q, the spectrum is identical to the free particle scenario. In a zero-field case n and $-n$ have the same energy. The toroid's vector potential breaks this symmetry unless F_0 is a half integer multiple of F_q. Since the ground state energy depends on the enclosed flux F_0, a particle in the ring will have a non-zero velocity even in the lowest energy state, creating a persistent current.

The AB effect and the use of toroid as an experimental setup have profound philosophical implications concerning the ontology (what is real) and epistemology (how we know it) of fundamental physics. We will touch upon several points hereunder.

4.2.4 Topological Invariance and the Field-Free Interaction

Having analyzed the toroid, we can now resolve the ontological conflict between fields $(\mathbf{E}, \mathbf{B})$ and potentials $(\Phi, \mathbf{A})$. The observable effect (a shift in the interference pattern) due to change in quantum phase even in the field free regions suggests that the vector potential has a signature of reality that transcends its classical role as a mathematical artifact.

Let's define two fundamental ideas in physics. Locality suggests that there is no action at a distance possible and thus physical effects should be caused by local interactions. Gauge Invariance requires that the physical predictions of a theory should not depend on the arbitrary choice of a mathematical gauge. The AB effect presents a conflict between these two ideas. Firstly, if the magnetic field is the only real entity, then its influence must be propagated non-locally to the charged particle passing outside the toroid. And secondly, if the vector potential is the real and local influence, its specific value at a point is not gauge-invariant as it can be changed arbitrarily, which violates the gauge principle for local quantities.

The resolution comes through the interpretation that the AB effect highlights the importance of the gauge potential as a more complete description of EM than the fields alone. The observable outcome of the experiment - the overall phase difference around a closed loop - is a gauge invariant quantity directly related to the total magnetic flux, $\Delta\phi = \frac{qF_0}{\hbar}$. While a specific value of $\mathbf{A}$ at any single point is not physical, the mathematical structure it defines (especially, its holonomy or circulation around a non-contractible loop) is a fundamental observable physical property linked to the topology of the space around the toroid. Attempts to explain the AB effect purely through non-local forces or interaction energies without potentials have been largely unsuccessful in fully accounting for experimental observations, reinforcing the potential-centric view.

The use of a toroid is crucial because, unlike an infinitely long solenoid, it completely traps the magnetic field inside, creating a truly field free external region for the particles. This setup provides the cleanest empirical evidence that the potentials, not just the fields, are physically significant in QM.

Gauge theory and topology provides the mathematical framework that underpins the AB effect and the physical reality of the vector potential. Gauge theory is built on the concept of symmetry. It suggests that the laws of physics should remain unchanged (invariant) even when certain mathematical descriptions (like potentials or wavefunctions) are transformed locally. In QM, we are free to change the overall phase of a wavefunction arbitrarily without affecting physical quantities like probability density $|\Psi|^2$. This is a global symmetry. But a problem arises if we make the phase angle θ a function of position and time, $\Psi(x) \rightarrow e^{i\theta(x)}\Psi(x)$, the standard Schrödinger equation changes in a way that is not simply a phase factor. The solution to this problem is given by the gauge potential. To restore this local symmetry, we must introduce a new field - the vector potential. This means that a local phase transformation in QM necessitates the idea of a vector potential in EM. The vector potential must simultaneously transform as $\mathbf{A} \rightarrow \mathbf{A}' = \mathbf{A} + \nabla\Lambda$, where Λ is the gauge function related to the phase angle by $\Lambda(x) = \frac{\theta(x)}{q}$. This ensures the Schrödinger equation remains invariant (a symmetry), mapping one valid physical solution to another. This deep connection between local gauge symmetry and the existence of force-carrying potentials is one of the most elegant links between EM and QM.

Topology is the mathematical study of shapes and spaces that remain the same under continuous deformations. In the context of the AB effect, it is essential for understanding why a zero-field region can still produce an effect. The setup for the AB effect (solenoid or toroid) requires the physical space where the particle travels to be punctured or multiply connected. The region of space is not simply connected because it has a hole (the core containing the magnetic field) that a closed loop cannot be shrunk to a single point without crossing a singularity. The vector potential itself is well behaved and curl-free in the reachable region, but it is singular (ill-behaved) precisely on the central axis where the flux filament is located (solenoid). Even if the magnetic field ($\mathbf{B} = \nabla \times \mathbf{A}$) is zero everywhere along the particle's path, the integral of the vector potential around a closed loop enclosing the flux ($\oint \mathbf{A} \cdot d\mathbf{l}$) is non-zero. Stokes' theorem links this circulation to the total flux (F_0) enclosed by the loop. This loop integral is a topological invariant - it doesn't depend on the specific path taken around the core, only on the fact that it encircles the core once. This invariant quantity is what determines the observable phase shift $\Delta\phi = \frac{qF_0}{\hbar}$. Thus, the topology of the experimental setup and the singular nature of the potential at the core are the mathematical requirements for the AB phase shift to exist, confirming that non-local topology plays a critical role in local quantum mechanics.

5 Magnetic Monopole

To model a magnetic monopole of charge g, we consider a semi-infinite flux filament (the string) starting at origin and extending upto $(-\infty)$ along the negative z axis. The end of this string at the origin acts as a source of magnetic flux, appearing as a magnetic monopole to the rest of the space.

5.1 Vector Potentials and Wu-Yang Construction

We have already seen the making of a flux filament from a long solenoid of radius R with internal field $B_0\hat{k}$ and the vector potential in the exterior region $(r \geq R)$ is $A = \frac{B_0 R^2}{2r}\hat{\phi}$. We define the total magnetic flux as

$F_0 = B_0 \pi R^2$. We then take the limit $R \to 0$ while keeping the total flux F_0 constant. Substituting $B_0 R^2 = \frac{F_0}{\pi}$ into the exterior vector potential we get $A = \frac{F_0}{2\pi r}\hat{\phi}$. In cartesian coordinates $A_x = -\frac{F_0}{2\pi}\frac{y}{x^2+y^2}$, $A_y = \frac{F_0}{2\pi}\frac{x}{x^2+y^2}$, $A_z = 0$. The magnetic field is $\mathbf{B} = \nabla \times \mathbf{A}$. Outside the z axis, $\nabla \times \mathbf{A} = 0$. However, at the origin, the curl is a Dirac delta function $\mathbf{B} = F_0 \delta(x)\delta(y)\hat{k}$. If this filament were semi-infinite (ending at the origin), the leakage of flux F_0 at the endpoint represents a magnetic monopole of strength g related to the flux via Gauss's Law as shown in Figure 5 following the global formulation of gauge fields proposed by Wu and Yang [6].

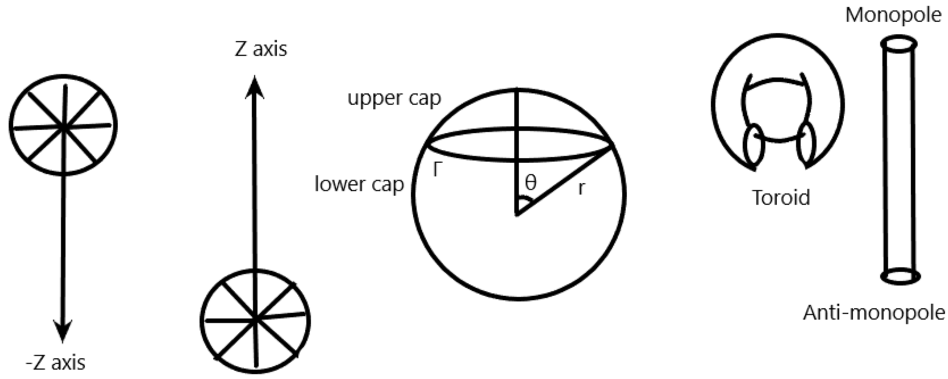

Figure 5: The Dirac string and magnetic monopole. A semi-infinite solenoid (the string) feeds magnetic flux $F_0 = 4\pi g$ into the origin. The endpoint acts as a point source for a radial magnetic field, effectively creating a magnetic monopole g while maintaining a singular vector potential along the negative z-axis.

The total flux F_0 emanating from the end of the string must satisfy Gauss' law for magnetism $\oint \mathbf{B} \cdot d\mathbf{s} = 4\pi g$. Hence we would get $F_0 = 4\pi g$. Note that in the SI units we write $\oint \mathbf{B} \cdot d\mathbf{s} = \mu_0 g$. We will use Gaussian units for our discussion here. This flux is fed into the origin by the semi-infinite filament. For a semi-infinite string on the negative z axis, we require a potential $\mathbf{A}$ whose curl gives the field of a point monopole $\mathbf{B} = \frac{g}{r^2}\hat{r}$, everywhere except the string. Thus a magnetic monopole of charge g at the origin produces a radial magnetic field in the space.

For finding the vector potential we choose a circular path Γ on a sphere of radius r, centered on the z axis at a polar angle θ. Assuming azimuthal symmetry $\mathbf{A} = A_\phi \hat{\phi}$, the integral around the loop of radius $r \sin\theta$ is $\oint_\Gamma \mathbf{A} \cdot d\mathbf{l} = A_\phi (2\pi r \sin\theta)$. By Stokes' theorem, this must be equal to the magnetic flux F passing through the spherical cap S bounded by Γ:

$$F = \int_0^\theta \int_0^{2\pi} (B_r \hat{r}) \cdot (r^2 \sin\theta' d\theta' d\phi')\hat{r} \tag{5.1}$$

Substituting $B_r = \frac{g}{r^2}$:

$$F = \int_0^\theta \int_0^{2\pi} \left(\frac{g}{r^2}\right) r^2 \sin\theta' d\theta' d\phi' = 2\pi g \int_0^\theta \sin\theta' d\theta' = 2\pi g(1 - \cos\theta) \tag{5.2}$$

Equating the path integral to the flux $A_\phi (2\pi r \sin\theta) = 2\pi g(1 - \cos\theta)$, we get:

$$A_\phi = \frac{g(1 - \cos\theta)}{r \sin\theta} \tag{5.3}$$

At $\theta = 0$ (North pole), we have a form of $0/0$. Using L'Hôpital's rule, $\lim_{\theta \to 0} \frac{g \sin\theta}{r \cos\theta} = 0$. Hence $A_\phi \to 0$. Thus the potential is well behaved at the north pole.

At $\theta = \pi$ (South pole), we have denominator $\sin\pi = 0$, while the numerator $1 - \cos\pi = 2$. Thus $A_\phi \to \infty$. This singularity represents the semi-infinite flux filament located on the negative z axis, which punctures the space. We can infer that the potential is singular along the path that carries the flux to the monopole. This property makes the space multiply connected allowing the AB phase shift to exist even where the magnetic field is zero.

Let's see how this result changes if we place the string on the positive z axis instead. To move the string to the positive z axis, we change our integration surface to the lower spherical cap (from θ to π). This feeds the flux of monopole from the north instead of south. We apply Stokes' theorem to the same circular path Γ at an angle θ, but this time we integrate the magnetic flux F' through the surface S' that includes the south pole ($\theta' = \pi$). The flux through the lower cap is:

$$F' = \int_{\theta}^{\pi} \int_{0}^{2\pi} \left(\frac{g}{r^2}\right) r^2 \sin\theta' d\theta' d\phi' = 2\pi g[-\cos\theta']_{\theta}^{\pi} = -2\pi g(1+\cos\theta) \tag{5.4}$$

The negative sign arises because the boundary Γ is traversed in the opposite direction relative to the outward normal of the lower cap. Equating the path integral to this flux:

$$A'_{\phi}(2\pi r\sin\theta) = -2\pi g(1+\cos\theta) \implies A'_{\phi} = -\frac{g(1+\cos\theta)}{r\sin\theta} \tag{5.5}$$

At $\theta = \pi$ (south pole), we have a form of $0/0$. Using L'Hôpital's rule, $\lim_{\theta\to\pi} \frac{g\sin\theta}{r\cos\theta} = 0$. Hence $A_{\phi} \to 0$. Thus the potential is well behaved at the south pole. At $\theta = 0$ (north pole), we have numerator $\cos 0 = 1$, while the denominator is $\sin 0 = 0$. Thus $A_{\phi} \to \infty$. So we get a singularity at the north pole. This singularity represents the semi-infinite flux filament located on the positive z axis.

Let's compare the two vector potentials:

- $A_N = \frac{g(1-\cos\theta)}{r\sin\theta}$: Well behaved at $\theta = 0$, Singular at $\theta = \pi$.
- $A_S = -\frac{g(1+\cos\theta)}{r\sin\theta}$: Well behaved at $\theta = \pi$, Singular at $\theta = 0$.

In the overlap region, they should differ only by a gauge transformation:

$$A_N - A_S = \frac{g(1-\cos\theta)}{r\sin\theta}\hat{\phi} + \frac{g(1+\cos\theta)}{r\sin\theta}\hat{\phi} = \frac{2g}{r\sin\theta}\hat{\phi} \tag{5.6}$$

5.2 Dirac Quantization

In spherical coordinates, the gradient of the azimuthal angle ϕ is $\nabla\phi = \frac{1}{r\sin\theta}\hat{\phi}$. Therefore, $A_N = A_S + \nabla(2g\phi)$. This shows that shifting the string is equivalent to a gauge transformation with gauge function $\Lambda = 2g\phi$.

To derive the Dirac quantization condition, we require that the shift between the two potentials be a physically invisible gauge transformation. When we transform the vector potential by $\mathbf{A} \to \mathbf{A} + \nabla\Lambda$, the wavefunction must transform by a local phase shift $\Psi_N = \Psi_S e^{i\frac{q}{\hbar}\Lambda}$. Substituting $\Lambda = 2g\phi$:

$$\Psi_N = \Psi_S e^{i\frac{2qg}{\hbar}\phi} \tag{5.7}$$

For the physics to be consistent, the wavefunction Ψ must be single-valued: $\Psi(\phi + 2\pi) = \Psi(\phi)$. Applying this condition to the phase factor:

$$e^{i\frac{2qg}{\hbar}(\phi+2\pi)} = e^{i\frac{2qg}{\hbar}\phi} \tag{5.8}$$

This holds only if $\frac{2qg}{\hbar}(2\pi) = 2\pi n$, where $n = 0, \pm1, \pm2 \ldots$. We arrive at the fundamental result:

$$\frac{2qg}{\hbar} = n \implies qg = \frac{n\hbar}{2} \tag{5.9}$$

In SI units, this is written as $qg = \frac{nh}{4\pi\mu_0}$. The existence of a single magnetic monopole g forces all electric charges q in the universe to be quantized in units of $\hbar/2g$.

For a magnetic monopole to truly appear as a point source, the semi-infinite flux filament (the string) that feeds it must be undetectable by any charged particle. If a particle could 'feel' the string via a phase shift as it passed by, the monopole would fail to be spherically symmetric. Requiring the phase shift around the string to be a multiple of (2π) making the string physically invisible—is precisely what forces the quantization of all electric charges in the universe.

The existence of a single magnetic monopole g forces all electric charges q in the universe to be quantized in units of $\frac{\hbar}{2g}$.

This elegant link mirrors the fact that the vector potential and the local phase of the wavefunction reside on an equal footing, where the topology of space (the string) dictates the allowable values of physical constants. This is what is meant when we say that symmetry dictates the laws of physics.

5.3 Monopole Pairs from Toroid

For the sake of completion, let's develop the idea of a monopole from the toroid. Now we make a transition from the open string of a solenoid to a closed topological defect of a toroid. While a solenoid's filament is semi-infinite, a toroid traps flux in a finite, closed volume, providing a model for quantized fluxoids and monopole-antimonopole pairs.

Outside a toroid with internal flux F_0, the vector potential is $\mathbf{A} = \frac{F_0}{2\pi r}\hat{\phi}$. A particle encircling the toroid picks up a phase shift:

$$\Delta\phi = \frac{qF_0}{\hbar} \tag{5.10}$$

If we imagine a toroid whose cross-sectional radius $a \to 0$ while the flux F_0 remains finite, we create a closed flux filament (a loop of flux). This loop creates a "hole" in space. Any path that encircles this loop is topologically non-trivial. From a distance, a small toroidal loop of flux looks like a magnetic dipole. However, unlike a standard dipole made of two poles, this is a current-free topological structure where the magnetic field is strictly zero everywhere except inside the singular loop.

For the vacuum surrounding the toroid to be physically consistent (i.e., for the vector potential to be "removable" by a gauge transformation everywhere except across the loop surface), the phase shift must be an integer multiple of 2π:

$$\frac{qF_0}{\hbar} = 2\pi n \implies F_0 = n\left(\frac{h}{q}\right) \implies F_0 = nF_q \tag{5.11}$$

This $F_q = \frac{h}{q}$ is the magnetic flux quantum. This implies that in a universe with discrete electric charges, trapped magnetic flux in closed loops (like those in superconductors) must be quantized. The connection to the monopole arises if we cut the toroid and pull the ends apart. By stretching the toroid into a long, thin tube, the two ends act as a monopole-antimonopole pair.

The body of the toroid becomes the Dirac string connecting them. If we pull one end to infinity, we recover the semi-infinite solenoid (the single monopole) as we derived earlier.

While the solenoid models the existence of a monopole, the toroid models the quantization of its flux. The toroid proves that the vector potential is not just a mathematical device but a topological swirl. In a toroidal configuration, the monopole is not a point, but the end point of a quantized flux tube, ensuring that magnetic charge and electric charge are fundamentally linked through the topology of the vector potential.

6 Lattice Gauge Theory

Consider electromagnetic fields in space-time, where the coordinates are denoted by $x = (x^0, \vec{x})$, with $x^0 = t$. The vector potential is part of the four-potential $A = (A^0, \vec{A}) = (\Phi, \vec{A})$, where Φ is the scalar potential.

6.1 Link Variables and Parallel Transport

The electromagnetic field tensor is given by:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{6.1}$$

which is the higher-dimensional generalization of a curl. The components of $F_{\mu\nu}$ are related to the electric and magnetic fields, e.g., $F_{10} = E_x$ and $F_{12} = B_z$.

As discussed earlier, the important physical quantity is the line integral of the potential around any closed loop Γ:

$$F(\Gamma) = \oint_\Gamma A_\mu(x) dx^\mu \tag{6.2}$$

which may be regarded as the generalization of the electromagnetic flux. Note that the flux can also be expressed in terms of the fields and their surface integrals in four dimensions.

If the wavefunction is subjected to a phase transformation:

$$\psi(x) \to \psi'(x) = e^{i\theta(x)}\psi(x) \tag{6.3}$$

then it is clear that one should also transform the potential:

$$A_\mu \to A'_\mu = A_\mu + \frac{1}{q}\partial_\mu \theta(x) \tag{6.4}$$

By introducing the covariant derivative $D_\mu = \partial_\mu - iqA_\mu$, we obtain:

$$D_\mu \psi(x) \to D'_\mu \psi'(x) = e^{i\theta(x)}[D_\mu \psi(x)] \tag{6.5}$$

where D'_μ indicates the derivative is constructed from A'_μ. This local transformation keeps the physics intact. This theory concerns the $U(1)$ symmetry group.

Consider a dynamical field ψ with N internal degrees of freedom, expressed as a column vector $(\psi_1, \psi_2, \ldots, \psi_N)^T$. For indistinguishable states, we transform the field via:

$$\psi(x) \to \psi'(x) = T\psi(x) \tag{6.6}$$

where T is an $N \times N$ matrix element of a Lie group G. If T varies with x, and matrices do not commute ($T_1T_2 \neq T_2T_1$), the theory is non-Abelian. This provides the framework for strong interactions (color charges) [7] and electro-weak unification [8].

While the AB effect is an Abelian phenomenon, where the phase shift is simply a number, nature also employs Non-Abelian symmetries. In these theories (like the Strong Force), the 'phase' is replaced by an internal orientation (like color) that can be rotated. Instead of a single number, the transformation requires $N \times N$ matrices that do not commute. This demonstrates that the 'background swirl' of the potential is a universal structural feature shared by all fundamental forces, not just electromagnetism.

A lattice gauge theory discretizes space-time into points $\{x_i\}$ with links (i, j). In this model, $\psi_i = \psi(x_i)$ and transformations are $\psi_i \to \psi'_i = e^{i\theta_i}\psi_i$. The derivative corresponds to:

$$\psi_j - \psi_i = \int_{x_i}^{x_j} \partial_\mu \psi(x) dx^\mu \tag{6.7}$$

We define the lattice link variable:

$$A_{ij} = \int_{x_i}^{x_j} A_\mu(x) dx^\mu \tag{6.8}$$

and the Wilson variables [9]:

$$U_{ij} = U(x_i, x_j) \equiv e^{iqA_{ij}} \tag{6.9}$$

where $U_{ji} = (U_{ij})^{-1}$. The accompanying transformation for link variables is:

$$A_{ij} \to A_{ij} + \frac{1}{q}(\theta_j - \theta_i) \tag{6.10}$$

or, in terms of Wilson variables, $U_{ij} \to e^{i(\theta_j - \theta_i)}U_{ij}$.

6.2 Wilson Loops and Gauge-Invariant Flux

For a closed loop (a K-gon Γ), the discrete analog of the flux is:

$$F(\Gamma) = \sum_{\Gamma} A_{i_n i_{n+1}} \tag{6.11}$$

This gauge-invariant quantity is expressed through Wilson variables as:

$$\prod_{\Gamma} U_{i_n i_{n+1}} = \exp\left(iq \sum_{\Gamma} A_{i_n i_{n+1}} \right) = \exp(iqF(\Gamma)) \tag{6.12}$$

Note that in the specific case where the lattice loop Γ encircles the solenoid or toroid core, the general flux $F(\Gamma)$ reduces to the quantized total flux F_0 derived in Sections 4 and 5.

Nontrivial physics occurs when $F(\Gamma) \neq 0$. In spatial planes, $F(\Gamma)$ relates to magnetic flux; in planes containing the time direction, it relates to electric flux. This Wilson loop is the discrete counterpart to the holonomy $H(\Gamma)$ discussed in Section 3. It confirms that even in a discretized spacetime, the gauge-invariant flux $F(\Gamma)$ remains the fundamental physical observable, reinforcing the potential-centric view established in the solenoid and toroid cases.

7 Geometric Synthesis between EM and GR

While the preceding analysis established the vector potential A_μ as a geometric connection within a $U(1)$ gauge theory, this interpretation gains its most profound justification through its structural equivalence to General Relativity (GR). By moving from the internal phase space of a particle to the manifold of spacetime itself, we can demonstrate that "force" is a universal manifestation of underlying geometric curvature.

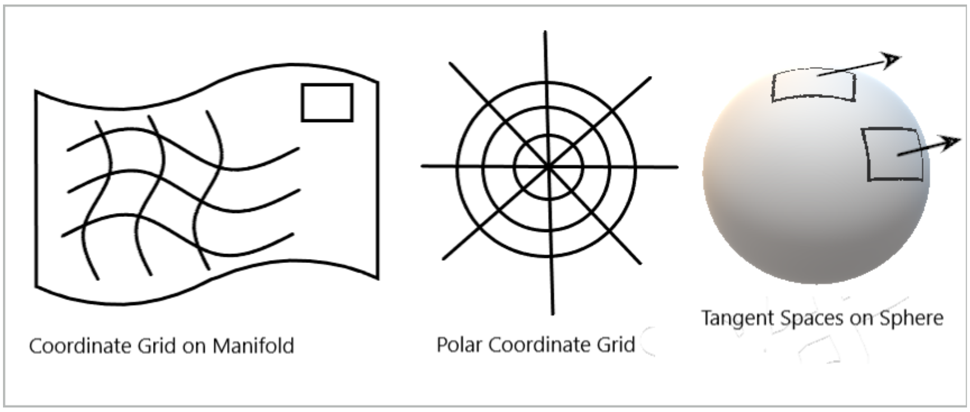

Figure 6: **Geometric Foundations of Curvature.** (Left) A coordinate grid overlaid on a general manifold, illustrating the principle of general covariance. (Center) The equivalence of Cartesian and polar grids on a flat manifold, where the underlying physics remains invariant under coordinate transformation. (Right) Representation of tangent spaces as "private" vector spaces at specific points on a curved manifold, where infinitesimal displacement vectors dx^μ are defined.

7.1 The Manifold and Coordinate Invariance

We define a **manifold** as a topological space that is locally homeomorphic to Euclidean or Minkowski space. On a global scale, the manifold may possess a complex, non-Euclidean topology, yet in a sufficiently small neighborhood, it remains "flat." We can think of it like the Earth: from a satellite, it is a sphere (curved), but to a person standing on a street corner, the ground looks like a flat plane (Euclidean). A Circle locally looks like a straight line. A Sphere locally looks like a flat sheet.

The principle of **General Covariance** dictates that physical laws must remain invariant under coordinate transformations. As seen in Figure [6], a manifold can be described by a variety of grids. For example, a flat 2D plane can be mapped using a Cartesian grid (x, y) or a polar grid (r, θ). While the numerical labels for points change, the physical relationship between them—the geometry—remains invariant.

7.2 The Riemannian Metric and the Concept of Distance

A Riemannian manifold is distinguished by the introduction of a **metric tensor** $g_{\mu\nu}$, which provides a rigorous way to measure distances. For two infinitesimally close points $P(x^\mu)$ and $Q(x^\mu + dx^\mu)$, the square of the interval ds is defined by a quadratic form:

$$ds^2 = g_{11}dx^1dx^1 + g_{12}dx^1dx^2 + g_{21}dx^2dx^1 + \cdots = g_{\mu\nu}dx^\mu dx^\nu \tag{7.1}$$

The metric $g_{\mu\nu}$ is an $N \times N$ symmetric matrix. In a curved manifold, these coefficients are not constant but vary from point to point. This metric is used to lower indices ($a_\mu = g_{\mu\nu}a^\nu$), while its inverse $g^{\mu\nu}$ is used to raise them ($a^\mu = g^{\mu\nu}a_\nu$).

In Cartesian coordinates (x, y), the metric tensor is the identity matrix:

$$g_{\mu\nu}^{\text{cart}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \implies ds^2 = g_{\mu\nu}dx^\mu dx^\nu = dx^2 + dy^2 \tag{7.2}$$

The transformation to Polar coordinates

$$(r, \theta)$$

uses the Jacobian matrix $\mathbf{J}$:

$$\mathbf{J} = \frac{\partial x^i}{\partial x'^j} = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{pmatrix} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} \tag{7.3}$$

The metric transforms as a rank-2 covariant tensor:

$$g_{\mu\nu}^{\text{polar}} = \mathbf{J}^T g_{\mu\nu}^{\text{cart}} \mathbf{J} \tag{7.4}$$

Performing the matrix multiplication:

$$g_{\mu\nu}^{\text{polar}} = \begin{pmatrix} \cos\theta & \sin\theta \\ -r\sin\theta & r\cos\theta \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} \tag{7.5}$$

$$g_{\mu\nu}^{\text{polar}} = \begin{pmatrix} \cos^2\theta + \sin^2\theta & -r\sin\theta\cos\theta + r\sin\theta\cos\theta \\ -r\sin\theta\cos\theta + r\sin\theta\cos\theta & r^2\sin^2\theta + r^2\cos^2\theta \end{pmatrix} \tag{7.6}$$

Simplifying via trigonometric identities yields the Polar metric tensor:

$$g_{\mu\nu}^{\text{polar}} = \begin{pmatrix} g_{rr} & g_{r\theta} \\ g_{\theta r} & g_{\theta\theta} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix} \implies ds^2 = dr^2 + r^2 d\theta^2 \tag{7.7}$$

$g_{rr} = 1$: This tells you that a change in the radial coordinate (dr) translates 1-to-1 into physical distance.

$g_{\theta\theta} = r^2$: This ”scales” the angular change. Since a degree of arc is physically longer the further you are from the center $(s = r\theta)$, the metric needs that r^2 factor to get the distance right.

$g_{r\theta} = 0$: This tells you the r and θ grid lines are perpendicular (orthogonal). If they weren't, these ”off-diagonal” terms would be non-zero.

Even though $dx \neq dr$ and $dy \neq d\theta$, the final physical distance is identical. This is the mathematical proof of Einstein's point: the coordinates are just ”scaffolding”. The metric tensor automatically adjusts to ensure the physical reality remains constant.

7.3 Tangent Spaces and Basis Vectors

On a curved surface, a finite displacement $x(x^1, x^2, \ldots, x^N)$ is not a vector. However, an infinitesimal displacement $dx(dx^1, dx^2, \ldots, dx^N)$ constitutes a true vector. These vectors do not live on the manifold itself but in a "private" vector space attached to each point, known as the **Tangent Space** (Figure [6], Right).

We define the basis vectors $\vec{e}_\mu$ of the tangent space as the partial derivative operators along the coordinate axes:

$$\vec{e}_\mu = \frac{\partial}{\partial x^\mu} \tag{7.8}$$

To account for the change in both the vector components and the basis vectors, we define the covariant derivative D_ν:

$$D_\nu A^\mu = \partial_\nu A^\mu + \Gamma^\mu_{\rho\nu} A^\rho \tag{7.9}$$

The connection coefficients $\Gamma^\mu_{\rho\nu}$ define how the basis vectors $\mathbf{e}_\rho$ vary across the manifold:

$$\partial_\nu \vec{e}_\rho = \Gamma^\mu_{\rho\nu} \vec{e}_\mu \tag{7.10}$$

The requirement of metric compatibility, $D_\lambda g_{\mu\nu} = 0$ ensures the physics (lengths and angles) remains invariant under transport:

$$\Gamma^\sigma_{\mu\nu} = \frac{1}{2} g^{\sigma\rho} \left(\partial_\nu g_{\rho\mu} + \partial_\mu g_{\rho\nu} - \partial_\rho g_{\mu\nu} \right) \tag{7.11}$$

7.4 Parallel Transport and the Connection

To connect different private vector spaces, we use the **connection 1-form** Γ_μ, which is an $N \times N$ matrix with components $\Gamma^\nu_{\mu\rho}$ known as the **Christoffel symbols**. The index μ describes the direction in which the basis vectors are rotated to align with the next tangent space.

The total parallel transport of a vector over a displacement Δx^μ is then represented by the operator $\exp(\Delta x^\mu D_\mu)$.

To relate the disjoint tangent spaces $T_x M$ and $T_{x+dx} M$, we introduce the connection 1-form Γ_μ, which is an $N \times N$ matrix representing the "rotation" of the basis:

$$(D_\mu \psi)^i = \partial_\mu \psi^i + (\Gamma_\mu)^i{}_k \, \psi^k \tag{7.12}$$

In component form, the connection carries three indices, where μ denotes the direction of displacement:

$$(\Gamma_\mu)^i{}_k = \Gamma^i_{k\mu} \tag{7.13}$$

The transport of a vector ψ along an infinitesimal displacement dx^μ involves the matrix multiplication:

$$\psi^i(x + dx) \approx \psi^i(x) - \Gamma^i_{k\mu} \psi^k dx^\mu \tag{7.14}$$

The matrix Γ_μ describes how the basis vectors $\mathbf{e}_k$ are connected across space:

$$\partial_\mu \vec{e}_k = \Gamma^i_{k\mu} \vec{e}_i \tag{7.15}$$

To connect the vector space $V(x)$ to $V(x + \Delta x)$, we define the parallel transport operator.This operator consists of an ordinary translation and a corrective rotation:

$$\exp(\Delta x^\mu D_\mu) = \exp(\Delta x^\mu \Gamma_\mu) \exp(\Delta x^\mu \partial_\mu) \tag{7.16}$$

Following the "Minimal Coupling Prescription," we define the **Covariant Derivative** D_μ as the sum of the ordinary derivative and the geometric rotation. To first order in Δx, this links the covariant derivative directly to the partial derivative and the connection matrix:

$$D_\mu = \partial_\mu + \Gamma_\mu \tag{7.17}$$

The connection matrix Γ_μ acts on the basis vectors $\vec{e}_j$ such that its components describe the "leakage" into other basis directions:

$$\partial_\mu \vec{e}_j = \Gamma^i_{j\mu} \vec{e}_i \tag{7.18}$$

In this framework, Γ_μ is an $N \times N$ matrix with entries $(\Gamma_\mu)^i{}_j = \Gamma^i_{j\mu}$. The covariant derivative of a vector field ψ is then:

$$D_\mu \psi = (\partial_\mu + \Gamma_\mu)\psi \tag{7.19}$$

7.5 The Non-Commutative Nature of Curvature

Curvature is the result of the non-commutative nature of the covariant derivatives.

$$[D_\mu, D_\nu]\psi = (D_\mu D_\nu - D_\nu D_\mu)\psi \tag{7.20}$$

Expanding $D_\mu D_\nu \psi$:

$$D_\mu(D_\nu \psi) = (\partial_\mu + \Gamma_\mu)(\partial_\nu \psi + \Gamma_\nu \psi) \tag{7.21}$$

$$= \partial_\mu \partial_\nu \psi + (\partial_\mu \Gamma_\nu)\psi + \Gamma_\nu \partial_\mu \psi + \Gamma_\mu \partial_\nu \psi + \Gamma_\mu \Gamma_\nu \psi \tag{7.22}$$

When we subtract the reverse order ($D_\nu D_\mu \psi$), the symmetric terms $\partial_\mu \partial_\nu \psi$ and the mixed terms like $\Gamma_\nu \partial_\mu \psi$ cancel out, leaving:

$$R_{\mu\nu} = \partial_\mu \Gamma_\nu - \partial_\nu \Gamma_\mu + [\Gamma_\mu, \Gamma_\nu] \tag{7.23}$$

In EM, because the $U(1)$ group is abelian, the commutator $[\Gamma_\mu, \Gamma_\nu]$ vanishes, reducing to the Faraday Tensor, simplifying to the standard field strength. In GR, the non-abelian nature of the connection accounts for the self-interaction of the gravitational field.

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{7.24}$$

7.6 Dynamics and the Unified Force Equation

For a free particle in GR, the linear momentum does not change ($d\vec{p} = 0$). Starting from the conservation of the momentum vector $\vec{p} = p^\mu \vec{e}_\mu$:

$$d(p^\mu \vec{e}_\mu) = (dp^\mu)\vec{e}_\mu + p^\mu (d\vec{e}_\mu) = 0 \tag{7.25}$$

Using the definition of the connection $\partial_\nu \vec{e}_\rho = \Gamma^\mu_{\rho\nu} \vec{e}_\mu$, we expand the second term:

$$(dp^\mu)\vec{e}_\mu + p^\rho \Gamma^\mu_{\rho\nu} dx^\nu \vec{e}_\mu = 0 \tag{7.26}$$

Factoring out the basis $\vec{e}_\mu$ and dividing by the proper time $d\tau$, we use $p^\mu = m\frac{dx^\mu}{d\tau}$:

$$m\left(\frac{d^2 x^\mu}{d\tau^2} + \Gamma^\mu_{\rho\nu} \frac{dx^\rho}{d\tau} \frac{dx^\nu}{d\tau}\right) = 0 \tag{7.27}$$

If the particle carries a charge q, it couples to the electromagnetic curvature $F^{\mu}{}_{\nu}$. The "zero" on the right-hand side is replaced by the Lorentz force, yielding the Unified Equation:

$$m\left(\frac{d^2x^\mu}{d\tau^2}+\Gamma^\mu_{\rho\nu}\frac{dx^\rho}{d\tau}\frac{dx^\nu}{d\tau}\right)=qF^{\mu}{}_{\nu}\frac{dx^\nu}{d\tau} \tag{7.28}$$

7.7 The Master Comparison: EM vs. GR

The following table summarizes the structural parallels and functional differences between the geometric descriptions of Electromagnetism and General Relativity.

Table 1: Geometric Synthesis: General Relativity (GR) vs. Electromagnetism (EM)

Feature	**General Relativity (GR)**	**Electromagnetism (EM)**
Mathematical Space	Tangent Space (Spacetime Manifold)	$U(1)$ Fiber Bundle (Internal Phase)
Geometric Connection	Christoffel Symbol ($\Gamma^\nu_{\mu\rho}$)	Gauge Potential (iqA_μ)
Covariant Derivative	$D_\mu=\partial_\mu+\Gamma_\mu$	$D_\mu=\partial_\mu+iqA_\mu$
Curvature / Field	Riemann Tensor ($R^\rho_{\sigma\mu\nu}$)	Faraday Tensor ($F_{\mu\nu}$)
Independent Unknowns	**10 components** (Symmetric $g_{\mu\nu}$)	**4 components** (Four-vector A_μ)
Field Dynamics	**Non-linear, 2nd Order** (Self-sourcing)	**Linear, 2nd Order** (Wave-like)
Commutator	$[D_\mu,D_\nu]=\mathcal{R}_{\mu\nu}$	$[D_\mu,D_\nu]=iqF_{\mu\nu}$
Non-Trivial Transport	Parallel Transport (Vector Rotation)	Holonomy (Phase Shift)
Physical Manifestation	Geodesic Deviation / Gravity	Aharonov-Bohm / Lorentz Force
The "Structural Wall"	**Curvature $\sim$ Source** ($G\sim T$)	**∂ of Curvature $\sim$ Source** ($\partial F\sim J$)
Quantization Link	Mass-Inertia Equivalence	Dirac String / Charge Quantization

7.8 Field Dynamics and the Structural Wall

Despite these kinematic parallels, the field dynamics reveal a fundamental divergence. Maxwell's equations relate the *derivative* of the curvature to the source ($\partial F\sim J$):

At this level, the theories differ significantly. While EM treats the connection as a field inhabiting spacetime, GR treats the connection as the definition of spacetime itself. This "wall" suggests that while the geometric language is shared, the dynamical origin of gravity is zero-order in curvature, whereas electromagnetism is first-order.

The dynamics of the Electromagnetic field are governed by the derivative of the Faraday tensor:

$$D_\mu F^{\mu\nu}=\partial_\mu F^{\mu\nu}+\cdots=J^\nu \tag{7.29}$$

Note: Since $F\sim\partial A$, this is a second-order equation in terms of the potential A_μ.

In contrast, General Relativity relates the curvature tensor directly to the energy-momentum source:

$$R_{\mu\nu}-\frac{1}{2}Rg_{\mu\nu}+\Lambda g_{\mu\nu}=8\pi GT_{\mu\nu} \tag{7.30}$$

Note: Since $R\sim\partial\Gamma+\Gamma^2$ and $\Gamma\sim\partial g$, this is also a second-order equation, but it links the *geometry itself* to the source.

While the Einstein equation (7.30) appears compact, it represents a **coupled set of 10 non-linear second-order partial differential equations**. This complexity arises because the Einstein tensor $G_{\mu\nu}=R_{\mu\nu}-\frac{1}{2}Rg_{\mu\nu}$ is a 4×4 symmetric matrix; since the indices μ and ν each range from 0 to 3, the 16 total components reduce to **10 independent entries** (4 diagonal and 6 off-diagonal). Furthermore, because the Christoffel symbols involve first derivatives of the metric and the Riemann curvature tensor involves derivatives of those symbols, the equations are inherently second-order. The non-linearity is a reflection of the gravitational field's self-interaction—a stark contrast to the Abelian, linear framework of Maxwellian electromagnetism.

In contrast to the 10 non-linear second-order equations of General Relativity, Maxwell's equations (7.29) represent a **linear system of 4 coupled partial differential equations** in terms of the fields, or 2 second-order equations in terms of the four-potential A^μ.

In the Lorenz gauge, the dynamics reduce to the wave equation $\Box A^\mu = \mu_0 J^\mu$, which is **linear** because the electromagnetic field does not act as its own source. This is a fundamental ontological divergence: while the metric $g_{\mu\nu}$ in GR is both the "scaffolding" and the "source" (leading to non-linearity), the vector potential A^μ in EM is a field that inhabits the scaffolding without deforming it. Consequently, while GR requires 10 independent components to describe the curvature of spacetime, EM requires only **4 components** of the four-potential to describe the "swirl" of the gauge field.

The "Structural Wall" is defined by this mapping:

$$\text{EM:} \quad \text{Derivative of Curvature} \sim \text{Source} \tag{7.31}$$

$$\text{GR:} \quad \text{Curvature} \sim \text{Source} \tag{7.32}$$

The AB effect provides the empirical proof that the potential A_μ is real. The phase shift $\Delta\phi$ acquired by a particle around a closed loop Γ is:

$$\Delta\phi = \frac{q}{\hbar}\oint_\Gamma A_\mu dx^\mu = \frac{q}{\hbar}\iint_{S(\Gamma)} F_{\mu\nu} d\sigma^{\mu\nu} \tag{7.33}$$

This phase shift is an example of **holonomy**—the failure of a state to return to its original value after transport around a loop. In the geometric language of both EM and GR:

$$\text{EM Holonomy:} \quad \exp\left(i\frac{q}{\hbar}\oint A_\mu dx^\mu\right) \tag{7.34}$$

$$\text{GR Holonomy:} \quad \text{Parallel Transport of a Vector } \vec{V} \tag{7.35}$$

The "Structural Wall" of field dynamics tells us *how* the swirl is created, but the Aharonov-Bohm effect tells us that the *swirl itself* is what the particle actually feels.

To address the dynamical mismatch between EM and GR, one may postulate a differential extension of the gravitational field equations. While Maxwell's equations relate the derivative of the curvature to the source ($dF = J$), the Einstein Field Equations relate the curvature tensor directly to the energy-momentum source ($G = T$). By shifting to a differential form ($\partial R = \partial T$):

$$\nabla_\lambda(G_{\mu\nu} - 8\pi G T_{\mu\nu}) = 0 \tag{7.36}$$

we align the two theories under a shared differential mapping. Integration of this form implies $G_{\mu\nu} - 8\pi G T_{\mu\nu} = \Lambda g_{\mu\nu}$, where the integration constant naturally emerges as the Cosmological Constant. While this "differential gravity" points toward a deeper geometric unity, it introduces extraneous solutions and potential instabilities common in higher-derivative theories. Thus, the question of whether the field dynamics should be unified at the level of curvature or its derivative remains an open challenge.

8 Outlook

Any theory is governed by the raw power of imagination, which is termed metaphysics. Metaphysics is the speculative playground where we explore possibilities and inquire about the nature of reality. Theory then comes to terms with the *ontology*—the inventory of entities we commit to in order for a theory to work, such as fields, potentials, flux, and locality.

The structural parallels summarized in Table 1 invite a deeper metaphysical question: if both gravity and electromagnetism are manifestations of geometric curvature, can they be unified within a single manifold? This inquiry leads naturally to the **Kaluza-Klein** framework, which posits that the $U(1)$ fiber bundle of electromagnetism is not merely an "internal" space, but a compactified fifth dimension of spacetime itself[10]. In this ontology, the vector potential A_μ emerges as a component of a higher-dimensional metric tensor, suggesting that the "background swirl" of the AB effect is a localized vibration of an extra-dimensional geometry. While the "structural wall" of field dynamics ($D_\mu F^{\mu\nu} = J^\nu$ vs. $G_{\mu\nu} = 8\pi G T_{\mu\nu}$) remains a significant hurdle, the transition from local fields to global holonomies—reinforced by the recent observation of the gravitational Aharonov-Bohm effect [11]—suggests that the future of physical description lies in a fully integrated, potential-centric geometry where the distinction between "force" and "space" finally dissolves.

Ontology is often constrained by the choices we make of what to select or reject. For example, General Relativity (GR) assumes the equivalence of inertial and gravitational mass, while Quantum Field Theory (QFT) posits the creation and annihilation of particles. The claims of a theory are verified by *epistemology*, which determines the conditions under which an ontology becomes justified knowledge through experimental validation.

A theory is a logically consistent framework that builds from actuality (ontology) and metaphysical principles (symmetry and causality) to create a logical synthesis with predictive power. There exists a feedback mechanism between ontological commitments and epistemological frameworks; experiments force us to rethink our metaphysics and expand our ontology.

For instance, the Aharonov-Bohm effect [1] was a theoretical prediction reached by preferring the ontological commitment of the vector potential $\mathbf{A}$ over magnetic fields. This is mathematically evidenced by the fact that the energy spectrum (Eq. 4.38 and 4.56) is determined solely by the ratio of the total flux to the flux quantum, regardless of the local field strength along the particle's path. The epistemological breakthrough was achieved by Tonomura [12], providing the experimental verification that shifted the vector potential from a mathematical construct to a physical reality.

To extend this toward General Relativity, we approach the concepts of fiber bundles and 1-form connections. The non-uniqueness of the vector potential $\mathbf{A}$ suggests the definition of a private vector space for each point in spacetime. Similar to how a tangent space is defined at every point on a spherical surface, an infinitesimal vector element resides within its own local space. To perform operations like addition or subtraction between vectors at different points, one must parallel transport them to a common space through a projection mechanism defined as a fiber bundle. In this geometric view, the potential acts as a connection that dictates how the phase of a wavefunction rotates during parallel transport. The failure of these transport operations to commute reveals a gauge-theoretic curvature mathematically analogous to **Riemann curvature** in GR. This link is further solidified by the recent experimental observation of a gravitational Aharonov-Bohm effect [11], suggesting that the "background swirl" of the vector potential is a local manifestation of the deep geometric principles governing the curvature of spacetime.

Acknowledgements

The author would like to thank the Physics Academy of the North East (PANE) for providing a platform to discuss these foundational concepts in physics education.

References

[1] Y. Aharonov and D. Bohm, "Significance of electromagnetic potentials in the quantum theory," *Physical review*, vol. 115, no. 3, p. 485, 1959.

[2] Y. Aharonov, I. L. Paiva, Z. Schwartzman-Nowik, A. C. Elitzur, and E. Cohen, "Time-symmetry and topology of the aharonov–bohm effect," *Journal of Physics A: Mathematical and Theoretical*, vol. 56, no. 47, p. 475302, 2023.

[3] P. A. M. Dirac, "Quantised singularities in the electromagnetic field," *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, vol. 133, no. 821, pp. 60–72, 1931.

[4] J. D. Jackson, *Classical Electrodynamics*. Wiley, New York, 2 ed., 1962.

[5] K. Shinohara, T. Aoki, and A. Morinaga, "Scalar aharonov-bohm effect for ultracold atoms," *Physical Review A*, vol. 66, no. 4, p. 042106, 2002.

[6] T. T. Wu and C. N. Yang, "Concept of nonintegrable phase factors and global formulation of gauge fields," *Physical Review D*, vol. 12, no. 12, p. 3845, 1975.

[7] D. J. Gross and F. Wilczek, "Ultraviolet behavior of non-abelian gauge theories," *Physical Review Letters*, vol. 30, no. 26, p. 1343, 1973.

[8] C.-N. Yang and R. L. Mills, "Conservation of isotopic spin and isotopic gauge invariance," *Physical review*, vol. 96, no. 1, p. 191, 1954.

[9] K. G. Wilson, "Confinement of quarks," *Physical review D*, vol. 10, no. 8, p. 2445, 1974.

[10] T. Kaluza, "Zum unitätsproblem der physik," *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)*, vol. 1921, no. arXiv: 1803.08616, pp. 966–972, 1921.

[11] C. Overstreet, P. Asenbaum, J. Curti, M. Kim, and M. A. Kasevich, "Observation of a gravitational aharonov-bohm effect," *Science*, vol. 375, no. 6577, pp. 226–229, 2022.

[12] A. Tonomura, N. Osakabe, T. Matsuda, T. Kawasaki, J. Endo, S. Yano, and H. Yamada, "Evidence for aharonov-bohm effect with magnetic field completely shielded from electron wave," *Physical review letters*, vol. 56, no. 8, p. 792, 1986.