

archived as [http://www.stealthskater.com/Documents/UNITEL\\_14.doc](http://www.stealthskater.com/Documents/UNITEL_14.doc)

(also ...[UNITEL\\_14.pdf](#)) => [doc](#) [pdf](#) [URL-doc](#) [URL-pdf](#)

more *UNITEL* is on the [/UNITEL.htm](#) page at [doc](#) [pdf](#) [URL](#)

*note: The following was archived from microfilm on November 27, 2009. This is NOT an attempt to divert readers. Indeed, the reader should only read this back-up copy if the original cannot be found at the original site.*

# Fiber Bundles and Quantum Theory

*A branch of mathematics that extends the notion of curvature to topological analogues of a Mobius strip can help to explain prevailing theories of the interactions of elementary particles.*

by Herbert J. Bernstein and Anthony V. Phillips  
*Scientific American* / July, 1981

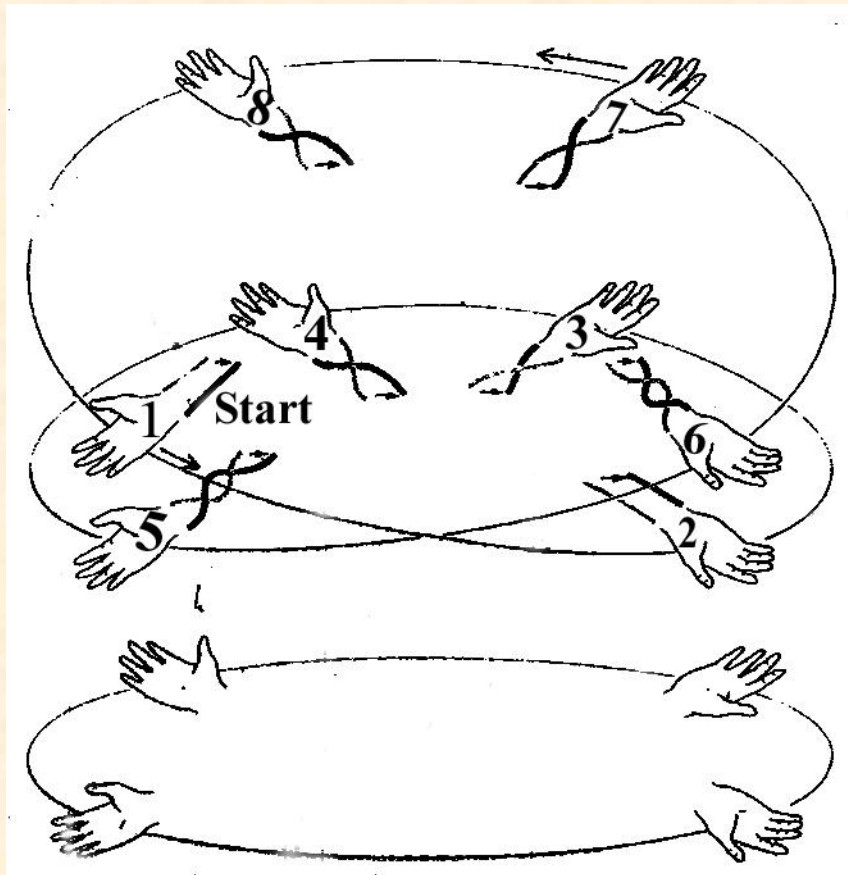
The intimate relation between Mathematics and Physics may seem surprising to the layman. But to someone well acquainted with either field, it is a natural evolutionary development. Physical problems have stimulated mathematical thinking at least since the Egyptians introduced geometry as a means of accurately measuring land. Newton's invention of the integral calculus was in part his practical-minded response to a difficulty in formulating the Law of Universal Gravitation. Infinite trigonometric series were devised to study the flow of heat. The abstract patterns investigated by present-day mathematicians are still based (albeit sometimes remotely) on the real patterns exhibited by the physicist's Universe.

Mathematics has not failed to repay the debt in kind. A theory invented by mathematicians to settle mathematical questions often turns out to be exactly what physicists need to advance their analyses and predictions of the ways of Nature. Tensor calculus -- the product of almost 100 years of development by such mathematicians as Karl Friedrich Gauss, Bernhard Riemann, and Tullio Levi-Civita -- was essential to Einstein's formulation of his General Theory of Relativity. The most recent beneficiaries of such mathematical research are the physicists who study the forces and Quantum Mechanical fields that mediate the interactions of elementary particles.

The fields that are most promising for this purpose are called **gauge fields**. Their utility lies chiefly in their ability to express underlying relations among forces that appear superficially to be quite distinct. Gauge fields have played a major role in recent attempts to devise a unified theory of 3 of the 4 basic forces known in Nature (namely the strong, the weak, and the electromagnetic forces). For the 4<sup>th</sup> basic force (gravity), there is as yet no Quantum Mechanical theory. But the General Theory of Relativity suggests that gravitation too may ultimately be described by a gauge field theory.

The idea of a gauge field was introduced by Hermann Weyl in the 1920s. But the present development of gauge field theory began in 1954. In that year, C. N. Yang and Robert L. Mills (who were then working at the Brookhaven National Laboratory) applied the gauge field concept to nuclear forces.

After almost 20 years of further refinement, physicists were able to express the concept of a gauge field in such a way that it could be recognized as an instance of more abstract structures known to mathematicians as connections in **fiber bundles**. The discovery of this equivalence has made it possible to apply mature and exceedingly powerful mathematical concepts to the description of physical reality.

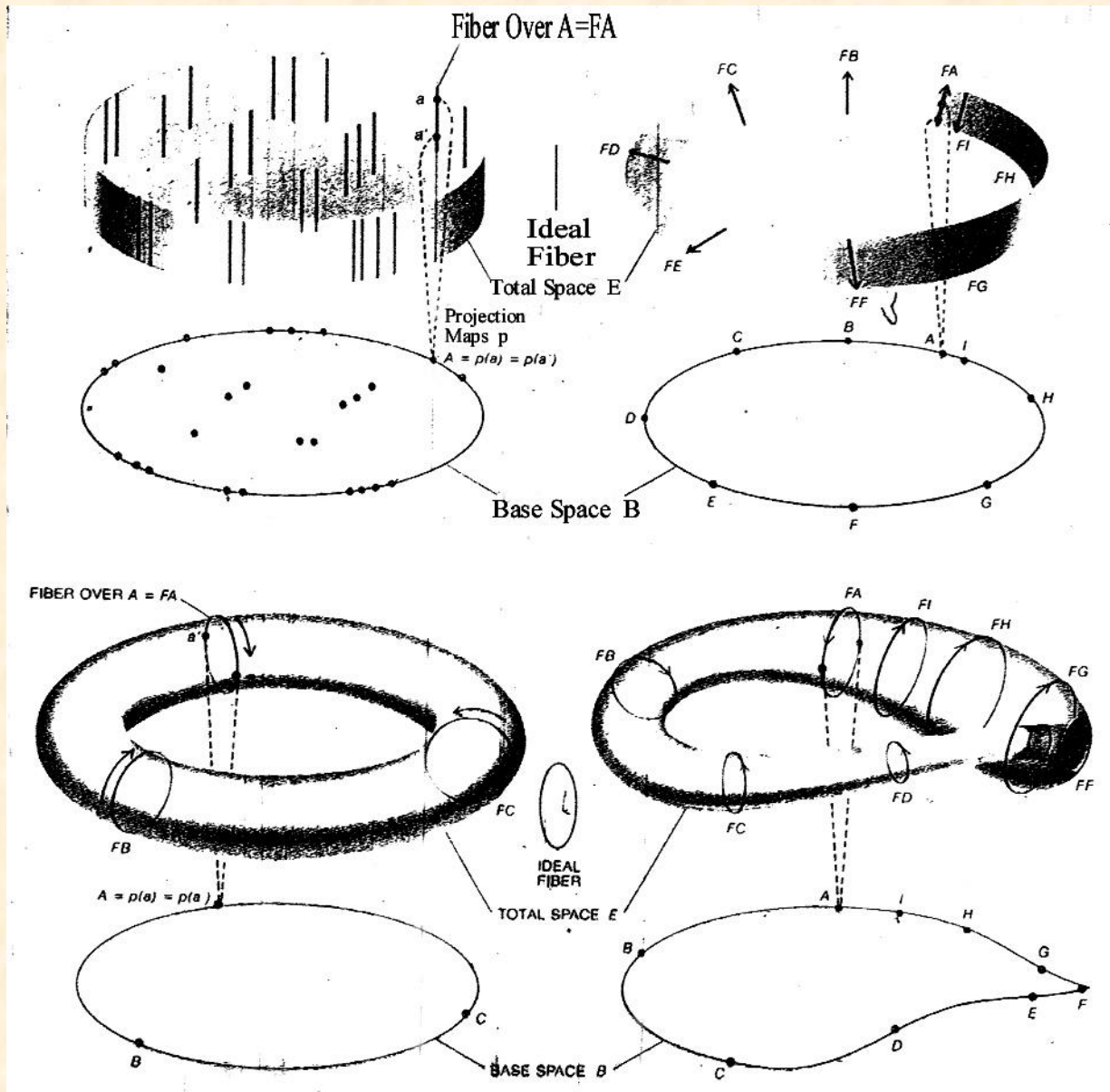


**Double-covering fiber bundle represents the set of relative rotations between the dancer's hand and the rest of her body (see illustration on opposite page). Arrows represent the fixed orientation of the dancer's feet. The circle at the bottom represents the orientations of the hand. Motion along the circle induces motion along one of the segments of the twisted curve directly above the circle. The topology of the latter curve shows that a counterclockwise relative rotation of 90 degrees is equivalent to a clockwise relative rotation of 630 degrees.**

What is a "gauge field"? What are "fiber bundles" and how do they enter into Physics? What does it mean for a fiber bundle to have a "connection" and how are the conceptions of a connection and of a gauge field related? We shall attempt to answer these questions by analyzing 2 physical experiments. One experiment shows what happens when neutrons are rotated 360 degrees by a magnetic field. And the other experiment shows the effect on 2 partial electron beams of a magnetic field in the region between them. Each experiment demonstrates in a different way how a fiber bundle can arise in Quantum Theory.

In the neutron-rotation experiment, the global structure of the fiber bundle is significant. Whereas in the electron-beam experiment, the central feature is a connection -- i.e., an intrinsic local structure that can be imposed on the bundle. The connection gives an elementary-but-fundamental example of a gauge field.

The study of fiber bundles is part of the branch of mathematics called Topology. But bundles have also been investigated in differential geometry because of their relation to the geometric concept of curvature. The idea of a connection in a fiber bundle grew out of attempts to generalize the notion of the curvature of a 2-dimensional surface (such as the surface of the Earth) to the curvature of a space with 3-or-more dimensions. Hence, another way of expressing the mathematical difference between the 2 experiments which we shall describe is to note that the neutron-rotation experiment concerns the topology of a fiber bundle whereas the electron-beam experiment concerns the geometry of a fiber bundle.



Fiber bundle consists of a base space, a total space, and a map that projects each point in the total space onto a point in the base space. The set of all the points in the total space that are mapped onto the same point in the base is called a fiber (colored lines and circle). The total space can resemble a sheaf of bundle of fibers. Every fiber in a fiber bundle must have the same topological structure. And so all the fibers can be represented by a single ideal fiber. For the solid cylinder (upper left) or the Möbius strip (upper right), the fibers are straight lines above the points in the bases. Whereas for the torus (lower left) or the Klein bottle (lower right), the fibers corresponding to each point are circles. Motion in the base may induce a change in the orientation of the fibers depending on the topological structure of the total space. A single circuit along a path in the base changes the orientation of the fibers of the Möbius strip and the Klein bottle. But the orientation of the fibers of the torus and the cylinder is not altered by motion along any such path. In moving from fiber-to-fiber on the Klein bottle, the bottle appears to pass through itself. The self-intersection cannot be avoided in a 3-dimensional drawing.

The 360-degree neutron-rotation experiment was proposed by one of us (Bernstein) in 1967. A similar thought experiment was described almost simultaneously by Yakir Aharonov and Leonard Susskind of Yeshiva University. The experiment demonstrates a highly counterintuitive effect whose mathematical equivalent is the one-sidedness of a Möbius strip.

At issue is the **spin** (or intrinsic angular momentum) of a subatomic particle. According to Quantum Theory, a neutron or another particle with spin does not return to initial state when its orientation is

rotated through 360 degrees. Instead, it takes 2 full turns (a 720-degree rotation) to restore the state of the particle to its initial condition.

To understand the experiment, it is necessary to know something of the quantum theory of "spin". Physicists have extrapolated the notion of intrinsic angular momentum from the spin of a top or gyroscope to the realm of elementary particles. In each case, spin is a vector quality. Which means that it has both a magnitude and a direction. By convention, the vector points along the spin axis in a direction determined by what is called the "right-hand rule".

If the fingers of the right hand are curled as if to grasp the axis of a spinning object with the fingers wrapped around the axis in the same sense as the spin, the direction of the thumb along the spin axis gives the direction of the spin vector.

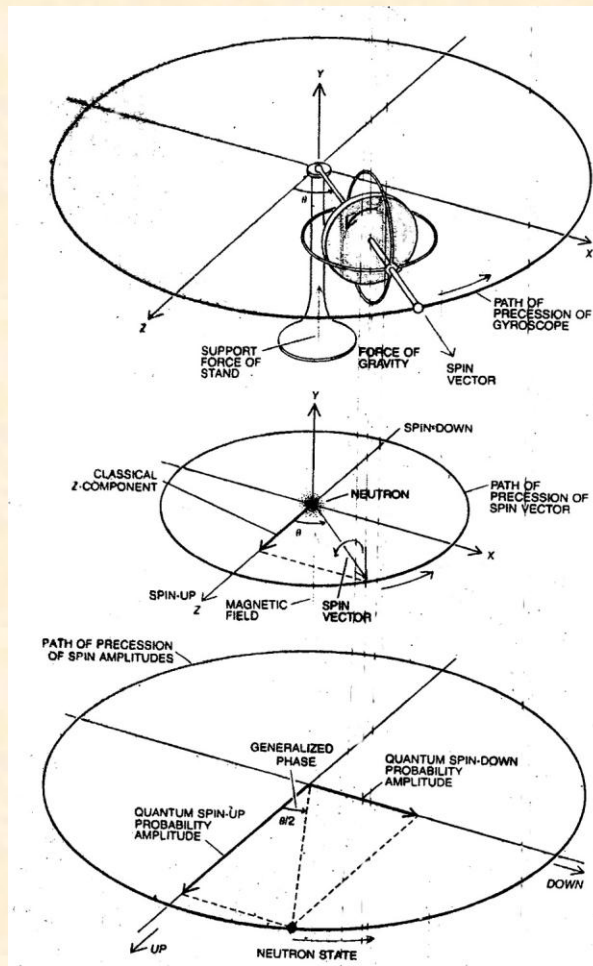
Unlike the spin of a top or a gyroscope, the spin of an elementary particle is quantized. Its magnitude can have only certain discrete values which are multiples of the smallest possible quantity of spin. Moreover, for any given particle, the magnitude of the spin can never change. It is one of the intrinsic properties that determine the identity of the particle. The electron, proton, and the neutron (and a number of other particles) have the smallest allowed nonzero quantity of spin. This minimum spin magnitude is equal to  $\hbar/2$  where  $\hbar$  is a form of Planck's constant with a value of about  $10^{-27}$  erg-second.

The constraint on the magnitude of the spin vector in Quantum Theory reflects a still more curious experimental finding. Like all vectors, the spin vector has components along the axes of any chosen coordinate system. An experimenter must select a particular axis along which to measure a spin component. No matter what direction is chosen as a reference axis, however, the only values ever found are  $+\hbar/2$  and  $-\hbar/2$ . No intermediate values are observed. In spite of this nonintuitive property of particle spins, the picture of a "spin vector" remains useful for describing neutrons.

It is possible to spin-polarize a series of neutrons so that ideally all the spin vectors point in the same direction. The experimenter can determine the direction by turning the axis of measurement to maximize the probability of detecting the value  $+\hbar/2$ . If all the neutrons are in the same state, all of them will give this result. For convenience, the reference along which the spins are polarized can be labeled the **z-axis** and the 2 possible states of the spin can be designated "spin-up" ( $+\hbar/2$ ) and "spin-down" ( $-\hbar/2$ ).

Suppose the spin vector of each neutron in the series rotates away from the fixed reference axis. The component along the z-axis cannot change continuously since measurements yield only the 2 discrete values  $+\hbar/2$  and  $-\hbar/2$ . Instead, the probabilities of finding spin-up and spin-down begin to change.

In the initial state of the spin vector, the probability of finding a spin-up neutron is 1 and that of finding a spin-down neutron is 0. After the spin vector has rotated a quarter-of-a-turn (or 90 degrees) from the z-axis, the Classical (i.e., non-Quantum Mechanical) model predicts that the z-axis component will vanish. The spin vector would then be oriented so that it points neither up nor down along the z-axis.



Precession of the spin vector of a neutron in a magnetic field resembles the precession of a gyroscope in a gravitational field. The magnetic torque on the spinning neutron causes it to precess at a rate proportional to the strength of the field and independent of the orientation of the neutron. If the initial direction of the neutron's spin vector is called "up" and the field is perpendicular to it, precession through an angle of 180 degrees will make the spin vector point down. At intermediate angles, Classical physics predicts that the component of the spin vector measured along the z-axis is equal to the perpendicular projection of the vector onto the axis. In Quantum Mechanics, the component of the spin measured along any axis can have one of only 2 values --  $+\frac{1}{2}$  or  $-\frac{1}{2}$  times Planck's constant  $\hbar$ . What changes during precession is the probability of detecting a neutron in the spin-up state ( $+\hbar/2$ ) or the spin-down state ( $-\hbar/2$ ). Each probability is determined by squaring a probability amplitude whose value can be either positive or negative. Hence, the quantum-mechanical precession of a neutron can be graphed as a rotation of the neutron state in an abstract space whose coordinate axes are the probability amplitudes of the spin-up and spin-down states of the spin vector.

For neutrons, however, a spin of zero is never found. According to the quantum mechanical model, a 90-degree rotation changes the state of the neutron to one in which both the spin-up and the spin-down probabilities are 0.5 .

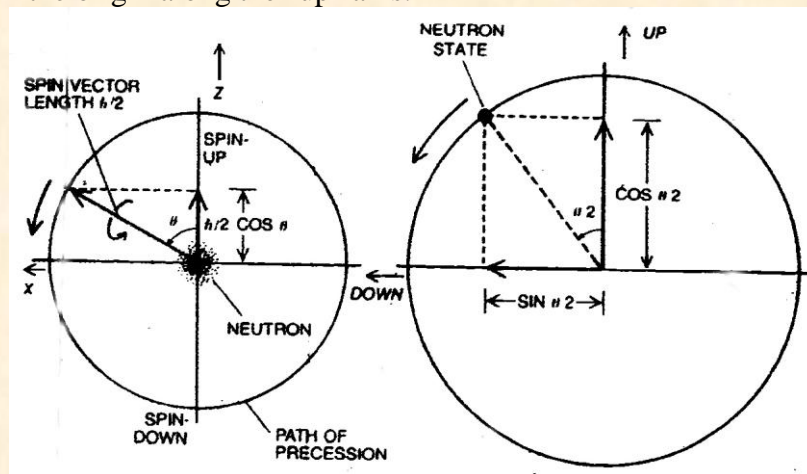
This outcome reconciles the quantization of spin with the classical description. If the z-axis components of a series of neutrons are measured while their spin vectors are oriented perpendicular to the z-axis, half have spin  $+\hbar/2$  and half have spin  $-\hbar/2$ . Thus the sum of all the measure components is zero and so is the average value (in agreement with the Classical result).

In a similar way, a rotation of 180 degrees orients the neutrons with their spin vectors down. This makes the spin-up probability 0 and the spin-down probability 1. After the neutron has swept out a full circle, the spin-up probability (i.e., the probability of finding the z component of the spin equal to  $+\hbar/2$ ) returns to 1 and the spin-down probability again becomes 0.

Physicists have taken the distinctness of spin states and the need for probabilities as being basic postulates of Quantum Theory. Probabilities are calculated from the mathematical description of each state of a particle as a **wave function**. At any point in space, there is a number called a probability amplitude of the neutron wave function for each distinct state. The term "amplitude" refers to the waves with which quantum theory describes material particles.

A probability amplitude can be positive or negative. Which reflects the observed capacity of waves to add constructively or destructively. The probability of finding a particle in a given state is the square of the corresponding probability amplitude. Squaring the amplitude ensures that the probability itself is always a positive number.

Since the neutron has 2 distinct spin states, the complete description of a neutron at a point in space consists of 2 numbers -- namely the probability amplitude for spin-up and the probability amplitude for spin-down. One way to show the amplitudes that respects the distinctness of the 2 states is to plot them on perpendicular axes marked "up" and "down" in an abstract space called **neutron-state space**. The state of the neutron can be represented by a point on this graph. If the initial state is spin-up, the spin-up probability is +1 and the spin-down probability amplitude is 0. This combination of values corresponds to a point one unit from the origin along the "up" axis.



Spin Vector of a neutron can precess. But the geometry of the precession cannot be pictured consistently by the Classical model (left) since a spin measurement along any given axis yields only the values  $+h/2$  or  $-h/2$ . In Quantum Mechanics, precession is manifested as a change in the probability of finding a neutron with spin  $+h/2$  (spin-up) or with spin  $-h/2$  (spin-down). The 2 amplitudes that determine the probability can be considered coordinates in an abstract space with axes labeled "up" and "down" (right). After a precession of 90 degrees from the z-axis, the spin vector points neither up nor down. If one measures the z-axis component of the vector, one find spin-up half of the time and spin-down half of the time. The average values of the spin is therefore zero in agreement with the classical result. Because the probabilities are equal, the probability amplitudes can be chosen to be equal. The corresponding point in neutron-state space is rotated 45 degrees from the "up" axis. Physical precession through any angle  $\theta$  causes a generalized phase shift  $\theta/2$  (represented as a rotation in neutron-state space).

As the orientation of the neutron spin is changed, the 2 probability amplitudes each vary continuously. The sum of the 2 probabilities must always add to 2, however, because spin-up and spin-down are the only possible states. Hence, the sum of the squares of the spin-up and spin-down probability amplitudes must be 1 and the point representing the state of the neutron must lie on a circle of radius 1. Every possible state of the neutron can then be labeled by the angle from the "up" axis to the corresponding point. The angle is called the **generalized phase** of the neutron state.

After a 90-degree of the physical spin vector away from the z-axis, the spin-up and spin-down probability amplitudes must be equal in absolute value since the corresponding probabilities must each be 0.5 . The point in the neutron-state space therefore lies halfway between the "up" and the "down"

axes. It follows that the generalized phase of the neutron state has changed by 45 degrees as a result of a 90-degree rotation of the neutron spin vector.

The half-angle relation continues. When the physical spin vector has rotated 180 degrees, it points down. The spin-down probability amplitude is +1 and the spin-up probability is 0. The corresponding point in the neutron-state space is 90 degrees from the initial direction. After a 360-degree physical rotation, both probability amplitudes change their sign. It takes 2 full turns (i.e., 720 degrees of rotation) to restore the probability amplitudes to their initial values.

This feature of Quantum Theory may at first seem paradoxical. When an ordinary object makes a complete rotation in space, it returns to the same state from which it started. A person's body or a spinning gyroscope is unchanged by a 360-degree rotation about any axis. This fact so deeply ingrained by common experience that although the theory of neutron spin is now some 50 years old, before 1967 even most physicists thought a 360-degree rotation could not have directly measurable consequences. Quantum Theory implies that one cannot measure a probability amplitude in any direct way and the change in the sign of the amplitude caused by such a rotation disappears when the amplitude is squared to compute the probability.

On the other hand, there are some circumstances in the realm of Macroscopic objects in which a 360-degree rotation has an observable effect. For example, if 2 objects are attached by a flexible ribbon, it is obvious that a full turn of one object does not restore the system to its original state. The ribbon ends up with a twist in it. What is not so obvious is that a second full turn in the same direction can bring such a system back to its initial state. The ribbon can be untwisted even though the relative rotation of the 2 objects undergoes no further change.

The effect can also be demonstrated by holding a wineglass in the palm of the hand and rotating the glass about its vertical axis (without moving the body as a whole). After a 360-degree rotation, the glass returns to its original orientation. But the arm is twisted. At further 360 degrees restores the glass and the arm to their initial position. All these phenomena (both Macroscopic and Quantum-mechanical) can be represented by the properties of a single fiber bundle.

A **fiber bundle** is a mathematical structure that consists of 2 distinct sets of points (called the base space  $B$  and the total space  $E$ ) and a rule  $p$  called the projection map that associates a point in  $B$  with every point in  $E$ . In the fiber-bundle model of the rotation of the wineglass, the points in the base space represent the possible orientations of the glass and the hand. The points in the total space represent the rotation the hand has undergone with respect to the rest of the body. The projection map defines an association between each rotation and the relative orientation determined by the rotation.

In the fiber-bundle model of neutron spin rotation, the points in the base space represent the orientation of the spin vector. In the part of the effect we are describing in detail, the vector rotates in a fixed plane and so any orientation can be described by the angle it makes with the z-axis. (Other possibilities can be treated analogously.) The points in the total space represent generalized phases of the neutron state and so they correspond to points on the unit circle in the neutron-state space with "up" and "down" coordinates. Each of these points can thus be described by its angular distance from the "up" axis.

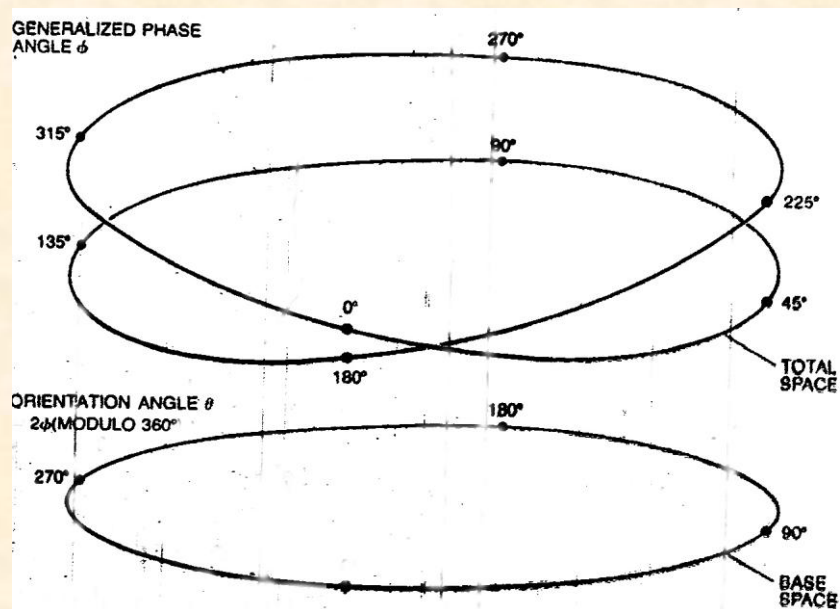
The projection map in the model assigns to each point in the total space a point in the base space according to the rule  $p(\Phi)=2\Phi(\text{modulo } 360)$ . The application of the rule is equivalent to wrapping the circle of generalized phases twice around the circle of orientations so that two phases  $\Phi$  and  $\Phi+180$  degrees are possible for each orientation.

The correspondence between points in the total space and points in the base is generally expressed by regarding the total space as being "over" the base. In this representation, the point or points in the total space that the projection map associates with a point in the base lie vertically above the base point. The set of points in the total space over a base point is called a **fiber**.

Another part of the definition of a fiber bundle requires that the fibers over any 2 points be topologically equivalent so that the topological structure of the fiber does not vary from one point in the base to another. In some instances, the fiber over each point in the base is a line. And it is this appearance that has given rise to the name "fiber bundle". When each fiber is a line, the total space looks like a bundle of fibers. In general, because the fibers are all topologically equivalent, they can all be described as copies of single fiber  $F$  (the ideal fiber of the fiber bundle).

The **ideal fiber** of the bundle representing neutron spin rotation is a space that consists of 2 distinct points. For example, over the pointed labeled 0 degrees in the base are the 2 points in the total space that correspond to the generalized phases 0 degrees and 180 degrees.

Hence the fiber over 0 degrees is the set that consists of 0 degrees and 180 degrees in the total space. Similarly, the fiber over 90 degrees is the set that consists of 45 degrees and 225 degrees in tot total space. In his bundle, both the total space and the base space are topologically equivalent to a circle. The projection map corresponds to the way the edge of a Mobius strip would project onto a circle at the center of the strip.



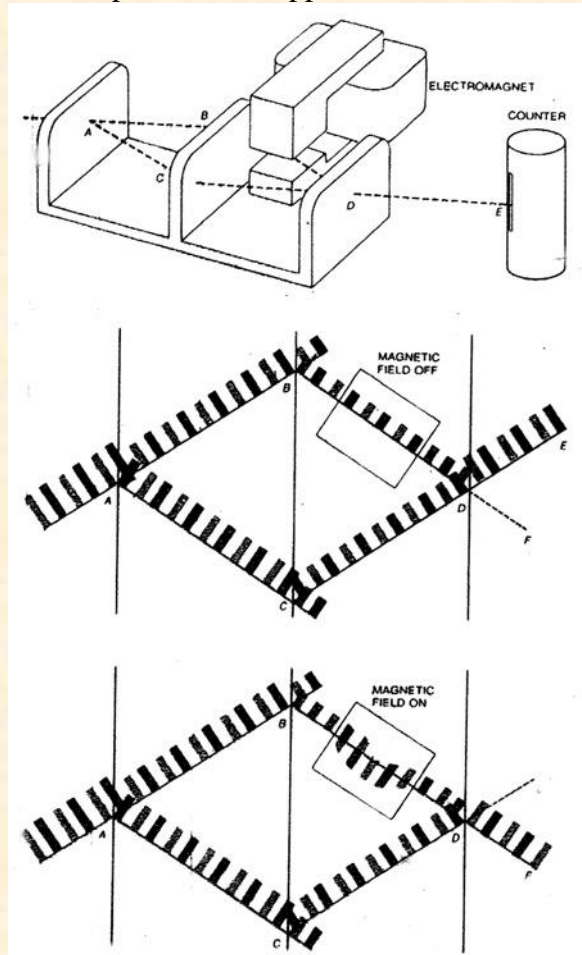
Fiber bundle of phase shifts shows the relation between the angular precession of a neutron and the shift in the generalized phase of the neutron spin state. Points in the base space of the bundle represent the orientation of the spin vector of a neutron. Points in the total space represent the relative phase shifts in neutron-state space that correspond to a given orientation. For instance, the projection map of the bundle assigns the point 45 degrees and 225 degrees in the total space to the point 90 degrees in the base. This means that generalized phase angles of 45 degrees and 225 degrees both correspond to an orientation of the spin vector 90 degrees from the z-axis. The topology of the total space, however, shows that attaining a phase shift of 225 degrees requires a precession of 450 degrees in the base (a 1¼ turn).

How does the fiber-bundle model of neutron spin rotation represent the relation between a rotation and a generalized phase shift? Suppose a neutron starts with its spin vector pointing along the positive z-axis ( $\theta=0$ ) so that its spin-up probability amplitude is +1 and its spin-down probability amplitude is 0. (This describes the point  $\Phi=0$  on the unit circle in the abstract space of neutron spin states.)



Now if the spin vector is physically rotated 90 degrees from the z-axis ending up at  $\theta=90$  degrees, out preceding discussion shows that the point on the generalized phase circle moves to  $\Phi=45$  degrees. If the spin vector rotates another 90 degrees to  $\theta=180$  degrees, the point on the phase circle moves to  $\Phi=90$  degrees.

The correspondence between rotations and generalized phase shifts can be described by saying that the point on the phase circle moves continuously in such a way that it always remains above the point on the orientation circle. This geometric principle -- together with the topological structure of the bundle -- accounts for the spin change of the neutron state as an effect of a phase reversal. One complete rotation in the base must shift the generalized phase to the opposite of what it was.



Neutron interferometer cut from a single perfect crystal of silicon has 3 projecting "ears". Each ear divides a beam of neutrons into a diffracted partial beam and a transmitted one. In this experiment, one partial beam passes through a magnetic field where the spin vectors of the neutrons are rotated and the generalized phase of the neutron state is shifted. When the 2 partial beams are recombined, they interfere according to the phase differences between them. If the magnetic field is zero, there is not phase difference and the beams interfere constructively. If the magnetic field rotates one of the beams 360 degrees, the phase difference between the 2 paths is 180 degrees and the beams interfere destructively at counter E. The crests of the neutron waves are represented by colored stripes and the troughs by gray stripes. The probability amplitudes of the spin states are indicated by the relative lengths of the stripes. Because the effect is illustrated for an incident spin-up neutron, a nonzero spin-down probability amplitude is shown below the black line only in the region of the magnetic field. Negative probability amplitudes are shown by converting crests into troughs and vice versa. The scattering of the neutrons at each ear is represented as 90-degree phase shift (skip one stripe) each time a neutron beam is diffracted and as no phase shift when it is transmitted.

How might one rotate the spin vector of a neutron 360 degrees? Current experimental designs take advantage of the magnetic properties of electrically neutral particles. A neutron has not only spin angular momentum but also a magnetic moment which makes it resemble a bar magnet spinning about its North-South axis.

Suppose the spin vector of the neutron is initially aligned with the z-axis and a magnetic field is introduced at right angles to that axis. The torque that aligns a bar magnet with an external field makes a spinning magnet precess about the direction of the field. The spin vector of the neutron will precess in the plane at right angles to the magnetic field just as a spinning gyroscope precesses in response to the pull of gravity. Hence to rotate the spin vector of a neutron away from z-axis, one can take advantage of its magnetic moment.

Actually, even if the magnetic field is not perpendicular to the initial spin, the neutron precesses at a rate that is proportional to the strength of the magnetic field. And that does not depend on the orientation of the neutron. Thus all the neutrons in an unpolarized beam passing through a magnetic field precess at the same rate. This rate is called the **Lamor frequency**.

Having a way to rotate neutrons 360 degrees is not enough, however. One must be able to compare the probability amplitudes for a rotated neutron with the amplitudes for the original state. The amplitudes for both the rotated and the unrotated states have the same magnitude but opposite signs.

The difference in sign can be detected because in Quantum Mechanics, it is possible to make a particle arrive at the same point in space along 2 different paths in the sense that there is a nonzero probability of detecting the particle along either one of the paths. As always, the probabilities are given by the squares of the probability amplitudes at any point along the paths.

At a point where 2 paths contribute to the probability of detecting the particle, the probability amplitude is the sum of the amplitudes for each of the paths. The addition is done before the amplitude is squared to get the probability. This rule - which embodies the phenomenon of **quantum interference** -- provides a method for demonstrating the sign change associated with a 360-degree rotation. The recent development of a neutron interferometer has made it possible to split a beam of neutrons so that the particles follow 2 paths and to recombine the partial beams.

Moreover, it is possible to rotate the spin vectors of the neutrons in one beam but not in the other. When the relative rotation is 360 degrees, the resulting sign change manifests itself through destructive interference of the 2 partial beams.

Experiments to detect 360-degree neutron rotations have been done by several groups. Helmut Rauch, Ulrich Bonse, and their colleagues first demonstrated the effect in 1975 at the Institut Laue-Langevin in Grenoble. An American team headed by Samuel A. Werner of the University of Missouri carried out a similar demonstration at about the same time. In 1976, Anthony Klein and G. I. Opat of the University of Melbourne employed a novel Fresnel diffraction technique to show the effect of neutron rotation (again at the Institut Laue-Langevin).

The heart of a typical neutron interferometer is a perfect cylindrical crystal silicon cut away so that 3 polished "ears" (or projections) stand up from a single intact base. When a beam of neutrons strikes the first ear at a suitable angle, it divides into a transmitted partial beam and a diffracted partial beam. At the second ear, each of these beams divides again. At the third ear, 2 partial beams recombine.

The 2 beams interfere constructively or destructively depending on their generalized phase. The recombined beam then splits again and counters (or detectors) placed beyond the 3<sup>rd</sup> ear record the number of neutrons in each of the 2 partial beams.

The probability of a neutron's arriving at a counter in the path of the recombined beam changes as the generalized phase of the rotated partial beam changes. The probability is measured by counting the

number of neutrons arriving per second. If the 2 partial beams are exactly in phase when there is no rotation, the initial counting rate is high. This situation results from constructive interference. The 2 amplitudes have the same sign and the square of their sum is at a maximum. Indeed, whenever the generalized phase difference is zero or a multiple of 360 degrees, there is constructive interference. After one full rotation of the spin vector, the phase difference is 180 degrees and the amplitudes have opposite signs. Ideally, the sum would be zero. In practice, the counting rate is at a minimum. This is destructive interference.

Rotation of the neutron spin vector is accomplished by placing an electromagnet in the region between the 2<sup>nd</sup> and 3<sup>rd</sup> ear. One of the partial beams passes through the field of the magnet but the other beam does not. Hence the field rotates spin along one partial beam, leaving the other beam unchanged.

The angle of rotation is proportional to the strength of the magnetic field. As a result, the generalized phase of the first partial beam increases continuously as the experimenter increases the current in the electromagnet from zero to its maximum value. As the phase shift increases, the counting rate at one counter at first declines. The interference changes smoothly from constructive to destructive.

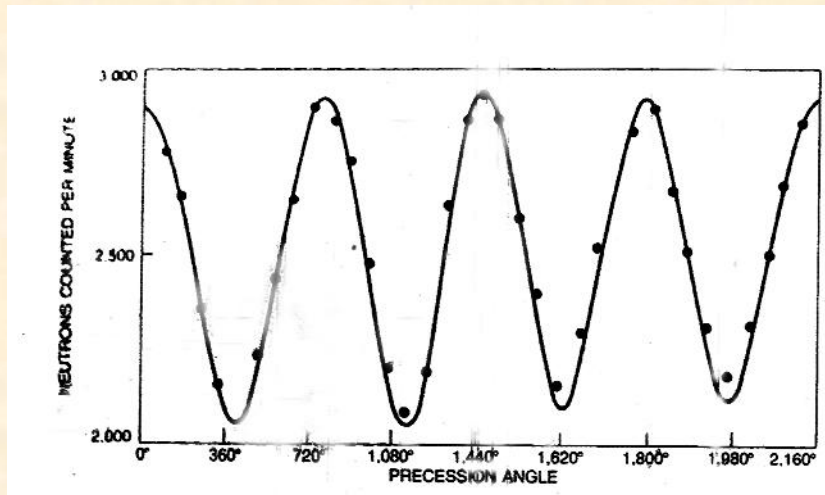
After reaching a minimum count, the interference tends back toward the maximum as the generalized phase approaches one full cycle. The resulting cycle of variation in the counting rate repeats as long as the current continues to rise. Since the rotation angle does not depend on initial spin orientation, the experiment does not require a polarized neutron beam.

The total rotation angle induced along the path through the magnetic field is equal to the Larmor Precession frequency multiplied by the time the neutrons spend in the field. The angle can therefore be calculated from measurements of the velocity of the beam; the intensity of the field; and the distance across the field.

In the version of the experiment done by Rauch, Bonse, and their colleagues, the neutrons travel through a magnetic field 1.5 cm wide at a speed of 2,170 meters/second so that each neutron spends a little less than 7 microseconds in the field.

When the electromagnet is operating at maximum current, the strength of field is 400 Gauss which corresponds to a Larmor frequency of 433 million degrees per second. At this rate, in 7 microseconds the spin vector of each neutron rotates about 8 full turns. If each 360-degree rotation of the spin vector restored a neutron to its original state, one would expect to observe 8 cycles of maximum and minimum counts. The actual result is significantly different. As the magnetic field increases from zero to its maximum, the number of neutrons detected at the counter passes through only 4 cycles.

**The outcome of the neutron-rotation experiment shows in a sense that fiber bundles exist in Quantum Mechanics and can be observed.** The fiber bundle associated with neutron rotation, however, is an extremely simple one because both its base space and its total space are one-dimension. (To repeat, the base space and tot total space are both circles with the total space wrapped around twice like the edge of a Mobius strip.)



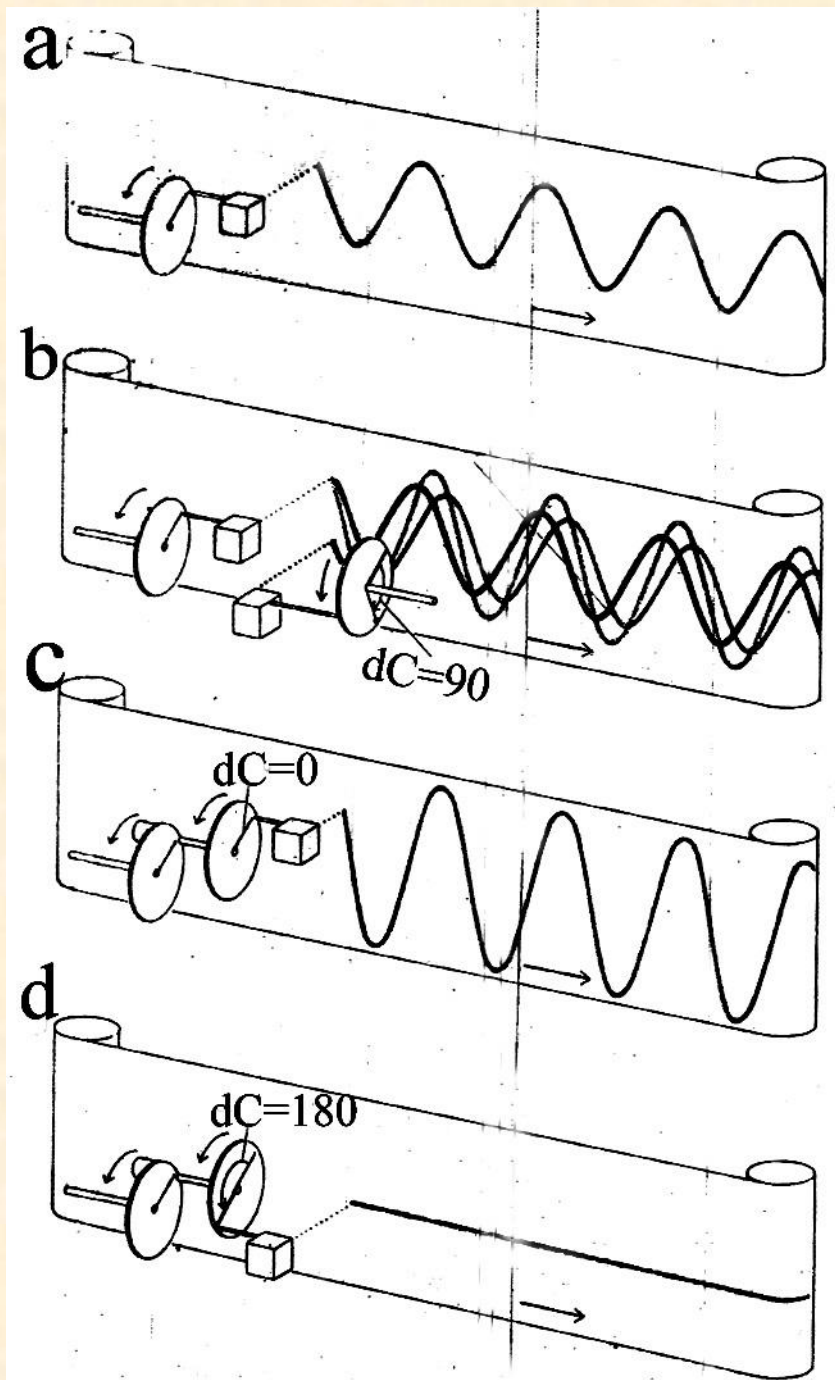
Several cycles of interference can be generated even with a moderate magnetic field. The total precession angle  $\theta$  of the rotated beam can be calculated from the strength of the magnetic field. When the intensity of the recombined beam is plotted against  $\theta$ , the experimental curve closely resembles the graph of the cosine of  $\theta/2$ . Background neutrons raise the minimum count above zero. 8 full rotations of the beam give only 4 intensity peaks.

The correspondence between rotations and phase shifts is embodied in the rule that the point in the total space should move continuously so that it always lies above the position of the point in the base. Because there is only one degree-of-freedom in the total space, the rule unambiguously specifies a phase shift for each rotation.

In more general fiber bundles, things are not so simple. For example, if each fiber is a line rather than a pair of points, motion in the base implies motion from fiber-to-fiber. But it does not specify which point in each fiber is to be traversed. To determine an unambiguous path through the total space in such a fiber bundle, additional structure is needed. A procedure that gives a path through the total space lying directly above a path in the base (once a starting point in the total space has been specified) is called a **path-lifting rule**.

The study of fiber bundles grew out of attempts to make more accessible to analysis the complexities of curvature in a "manifold" (i.e., an abstract topological space with an arbitrary number of dimensions). The idea of a fiber bundle was implicit in the work of the French mathematician Elie Joseph Cartan. But it was first explicitly stated in about 1935 by Hassler Whitney who now works at the Institute for Advanced Study. The concept of path lifting was developed as a systematic way of comparing the effects of curvature at different points of a manifold and was extended to fiber bundles by the French mathematicians Charles Ehresmann and Henri Cartan and by others in about 1950.

Consider what is perhaps the simplest curved manifold -- the 2-dimensional surface of a sphere. An observer standing on the surface has a complete circle of directions in which to look along the surface. (Assume that the observer looks only along the horizon -- never up or down.) There is such a circle of directions for each point on the surface. The circles taken together form a fiber bundle in a natural way called the bundle of directions on the surface of the sphere. The base space of the bundle is the surface itself. The fiber over a point in the base represents the set of all the directions on the surface of the sphere that can be surveyed from that point. Hence each fiber is a circle.



Phase of a wave (which is usually expressed as an angle) can be detected only as a difference between the phases of 2 waves. For sine and cosine waves, the waveform can be traced by projecting a point on a uniformly-rotating circle onto a screen moving uniformly at right angles to the line of projection. Any position on the circle can be chosen as zero degrees. The angle-of-rotation away from the arbitrarily-chosen position then labels the phase (d). The relative phase between the 2 waves is well-defined in that each crest and trough on one wave is the same number of degrees ahead of the corresponding feature of a second wave. When waves interfere, their amplitudes add at each instant. The resulting maximum height depends on the relative phase (b). The addition can be carried out automatically by centering the second circular motion on the perimeter of the first. If the interference is constructive (c), crests coincide with crests and troughs with troughs and the maximum height of the resulting pattern is equal to the sum of the heights of the original waves. If the interference is destructive -- corresponding to a phase difference of 180 degrees (d) -- crests coincide with troughs and waves cancel exactly.

It is not possible to draw a faithful picture that encompasses all the total space of the bundle of directions on the sphere. One reason such a drawing cannot be made is that it is topologically impossible to assign a reference direction to every point on the surface of a sphere in a continuous way.

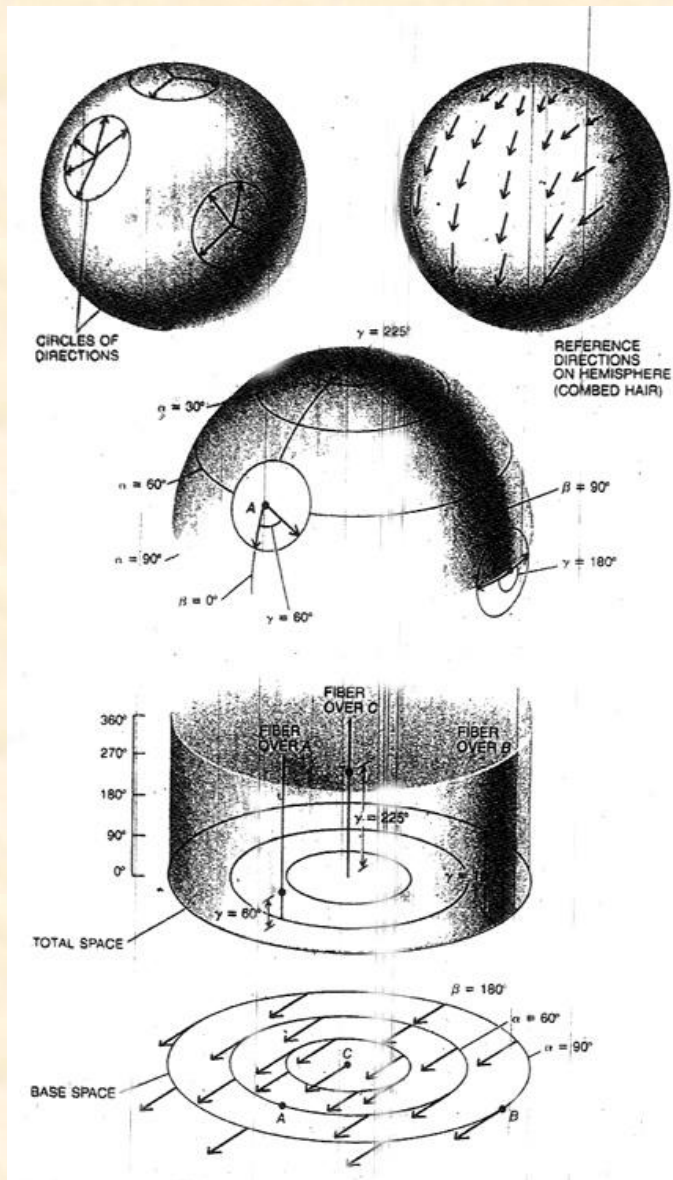
This fact is expressed whimsically the statement "You can't comb the hair on a sphere". At least one point on the sphere will always have a "cowlick".

On the other hand, it is possible to comb any patch of hair on the sphere even if the patch covers the entire surface with the exception of one point. For such a patch, it is possible to assign a continuous set of directions and so to draw a topologically-accurate picture of the total space of the bundle of direction over the patch.]

For example, on a flat map of the Northern Hemisphere, one can draw a continuous set of directions (say all pointing down and to the left). The set of directions can then be transferred back to the Northern Hemisphere of the sphere in order to comb its hair.

Combing the hair on the hemisphere gives a reference direction at each point. The total space of the bundle of directions on the hemisphere can then be illustrated by drawing each fiber vertically. If the bottom of each fiber corresponds to the reference direction, the angle any other direction makes with the reference direction can be represented by a height along the fiber.

The point halfway up the fiber represents 180 degrees from the reference direction and the point at the top of the fiber represent 360 degrees from the reference direction. Since the 360-degree direction coincides with the reference direction, the top and the bottom of each fiber represent the same point of the total space.



Bundle of directions on the surface of a sphere is an important example of a fiber bundle. At each point the sphere, there is a circle of directions along which one can look on the surface. To label these directions with angles, one must assign a reference direction to each point. If the reference directions could be assigned everywhere in a continuous manner, one could "comb the hair" on the sphere. But that is not possible. There must always be a "cowlick". Hair can be combed, however, over any region smaller than the entire surface. For example, on a flat map of the Northern Hemisphere, the description "downward and to the left" specifies a direction at each point and so defines a continuous set of reference directions on the hemisphere. A picture of the bundle of directions on the hemisphere can be made by adopting the flat map as a base space. Every direction at a point on the hemisphere appears on the vertical coordinate line above the corresponding point on the map and at a height that corresponds to the angle the direction makes with the reference direction. Heights 0 and 360 correspond to the same direction. The total space of the bundle is a cylinder where points at the top and the bottom of each vertical fiber are identical. The arrows that represent the reference directions on the map are parallel. But their counterparts on the sphere do not represent parallel transport.

How can one define a path-lifting rule for the bundle of directions on the sphere? A path in the base space of the bundle of directions is simply a path on the surface of the sphere. Lifting such a path into the total space requires that a direction chosen from the circle of directions be assigned to every point in the path.

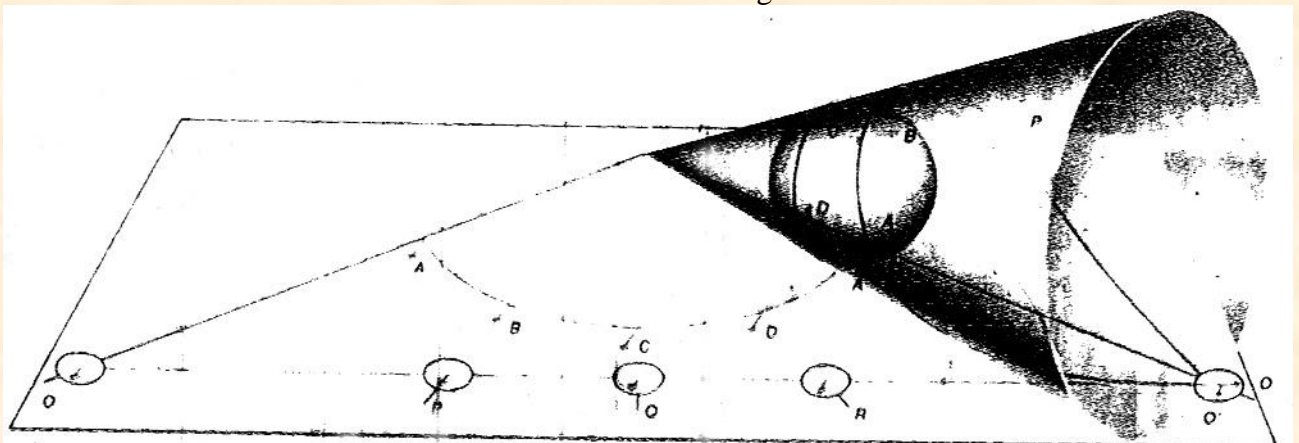
Think of a watch with one hand transported so that the center of the watch moves along wht path on the sphere while the hand moves freely around the dial. A path-lifting rule must determine the position of the hand at each point in the path once the hands' initial position is given.

If one considers only the topology of the sphere so that the surface can be stretched and deformed (but not torn) as if it were made of a sheet of rubber, there is no preferred way to give such a rule. The fundamental reason is that there is no topological relation between directions at one point on the surface of the sphere and directions at another point. If one considers the geometry of the sphere, however, there is a natural principle that will determine the movement of the watch hand. The principle is called **parallel transport**.

Parallel transport on a sphere can best be understood by imagining the sphere to be rolling on a flat surface. Suppose there are straight lines and curved lines drawn on the flat surface in wet ink. And suppose there are arrows spaced frequently along the line, all pointing, say, to the lower left. The term "parallel transport" is based on this picture. Since the arrows in the plane are all parallel, the arrows printed onto the sphere exemplify parallel transport along the printed curve on the sphere.

Parallel transport determines a path-lifting rule in the bundle of directions on the sphere. Given a curve on the sphere and an arrow representing a direction at the start of the ruve, one can place the sphere on the plane so that the starting point is the point of tangency.

Imagine there is an arrow at each point in the plane drawn in ink and parallel to the original arrow. If the sphere is rolled along the plane so that it prints the curve on the sphere onto the plane, the arrows from the plane will also be printed onto the sphere. The latter arrows give a path in the bundle of directions that starts with the initial arrow and lies above the given curve.



Parallel transport carries a direction along any curve in a plane so that an arrow points in the same direction everywhere on the curve. TO extend the idea to parallel transport along a curve on a surface that is not planar, one can imagine that parallel directions in the plane are printed onto the surface at the surface rolls on the plane without slipping or twisting about the vertical so that the point of tangency always remains on the curve. To roll a sphere along one of its circles of latitude, it is convenient to draw a cone tangent to the sphere along the circle. As the cone rolls on the plane, the sphere rolls along the circle of latitude. When a surface is rolled along a straight line in the plane, the curve printed onto the surface is a geodesic. If a geodesic is printed onto a cone of sufficiently small vertex angle, it can circumscribe the cone and intersect itself at an angle called the 'angular excess' (a measure of the curvature enclosed by the path).

When a straight line is printed onto the rolling sphere (or onto any other curved surface), the curve that it forms on the surface is called a **geodesic**. On the sphere, a geodesic is a "great circle". The shortest path on the surface between 2 points is a geodesic. The arrows that all point in the same direction on the plane retain a vestige of this property when they are printed onto a geodesic. At each



point along the geodesic curve, the angle between the arrow and the line tangent to the geodesic is the same.

Parallel transport can be described for a curved path made up of geodesic segments without reference to rolling. It can be carried out by maintaining a constant angle between the transported arrows and the tangents along successive geodesics. From a perspective above the surface of the sphere, however, parallel transport along a geodesic may seem anything but parallel. The arrows may appear to rotate. When an arbitrary curved line is printed onto the sphere, the rotation of the arrows as viewed from above the surface may appear to be even more chaotic.

Parallel transport provides a way of making quantitative and explicit the intuitive difference between a curved surface and a flat one. When an arrow is carried by parallel transport around a closed path in the plane, the directions of the arrow at the start and at the finish coincide.

Parallel transport around a closed path on a curved surface, however, may not lead to such coincidence. If there is a change in the direction of the arrow when it completes a single circuit of a closed path, the angle between the final direction and the initial one is called the **angular excess** of the path. It follows from the way parallel transport is defined that the angular excess does not depend on the initial direction of the arrow.

Mathematicians commonly express angular measure in radians instead of degrees. (Conversion from degrees-to-radians is made by multiplying the degree measure of an angle by the constant  $2\pi/360$ . One radian is therefore about 57 degrees.) When the angular excess is measured in radians, the results is a number called the "total curvature" of the region enclosed by the path. One can then define the average curvature of a region as its total curvature divided by its area.

By convention, the sign of the average curvature is given correctly when the arrow is transported along the path counterclockwise so that the region is on the left of the path. The curvature of a surface at a point can be defined as the limiting value of the average curvature of progressively smaller regions containing the point.

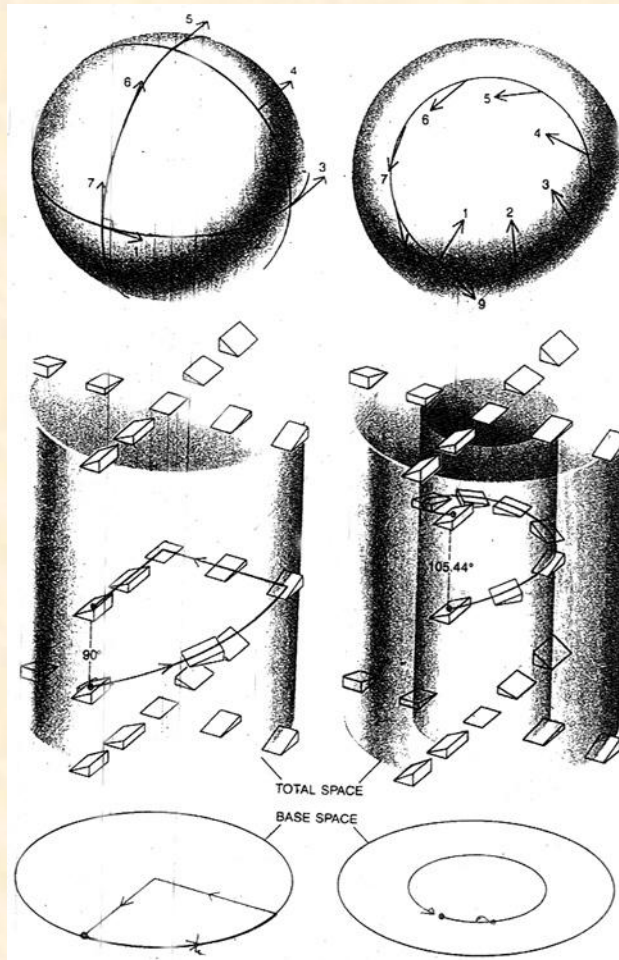
Parallel transport makes it possible to define a path-lifting rule from the surface of the sphere to the total space of all directions. In the total space, the angular excess is represented as the angular distance along the fiber that corresponds to the point on the closed path where the circuit of parallel transport begins and ends. Hence the total curvature of the region enclosed by a path in the base is represented by a distance along one of the fibers in the lifted path.

It turns out that replacing parallel transport by an arbitrary path-lifting rule generalizes the notion of curvature to other bundles on which the operation of parallel transport does not make sense. Such path-lifting rules have to be formulated without any reference to geodesics or angles.

Instead of focusing on the geometry of the base space (as one does in parallel transport), it is possible to lift a path by imposing structure on the total space. One way of doing this is to associate a set of parallel sloping planes with each fiber. The slopes of the planes determine how fast a lifted path rises or falls as it moves from fiber-to-fiber in the total space.

The planes must never be parallel to the fibers. Their slopes must vary continuously from point-to-point. They must have the same slope at every point along a given fiber. The latter condition is an analogue of the guarantee provided by parallel transport that the angular excess (and thus the curvature) is independent of the initial direction of the arrow being transported. Such a set of continuous planes in the total space is called a **connection** in the fiber bundle.

At the point on a fiber where a lifted path crosses the fiber, the path must be tangent to the sloping plane associated with the point. This is how the plane defines the slope of the lifted path at the point. The illustration below shows the connection that lifts paths on the sphere in the same way that parallel transport does.



Lifting a path in a fiber bundle is a means of finding a path in the total space starting at a given point and lying directly above the path in the base. For the bundle of directions on the surface of the sphere, parallel transport of directions on the surface gives a unique lifting for every path. The bundle of directions over the Northern Hemisphere can be represented as a cylindrical total space. Each direction is at a height corresponding to the angle it makes with the reference direction (here chosen to be pointing to the lower left in the flat map of the Northern Hemisphere). For the path around the spherical geodesic triangle, the reference directions point South along the meridian at the start of the path. The angle between the transported direction [colored arrow] and the tangent to a geodesic is constant. On the first leg of the triangle, the geodesic curves with respect to the reference direction so that the angle between the transported direction and the reference direction increases at a constant rate. Along the 2<sup>nd</sup> and 3<sup>rd</sup> legs of the triangle, the transported direction maintains a 180-degree angle with the reference direction. When the arrow returns to its starting point, its direction has changed by 90 degrees (the angular excess of the closed triangular path). The changes in transported directions are plotted as a lifted path in the bundle of directions on the hemisphere. For the path around the 45-degree latitude, the transported direction begins 180 degrees away from the reference direction and increases its angle with the reference direction at a constant rate. A connection can define a path-lifting rule without references to parallel transport by assigning planes to every point in the total sphere. The lifted path must be tangent to the planes. The slopes of the planes are the same for all points along a single fiber. But they vary continuously from fiber to fiber, the planes are never vertical. Such a collection of planes is called a "connection" in the fiber bundle.

The curvature of a connection can be defined by a procedure similar to the one employed for measuring the curvature of a surface. The aim of the procedure is to assign to each point in the base a number that represents the curvature at that point. (For high-dimensional spaces, the curvature is specified not by a single number but by a collection of numbers that are the components of a

mathematical object called a **tensor**.) The number that measures the curvature of the connection is obtained by finding the analogue to angular excess for progressively smaller paths enclosing the point. A connection has zero curvature over a region in the base if every sufficiently small closed curve in the region lifts to a closed curve in the total space. When this happens, the connection is called "flat" by analogy with parallel transport in plane geometry. If the connection has nonzero curvature, small closed paths in the base lift to yield curves in the total space that do not close.

Motion along a lifted path in a region where the connection is flat is analogous to motion on a hillside. Following a closed path may take one to various elevations above sea level. No matter how circuitous the route, however, when one returns to the surface coordinates (longitude and latitude) of the starting point, one has also returned to the same elevation. Motion in a curved region of the connection is analogous to motion in a cave. Some routes through the tunnels of the cave may lead back to the same longitude and latitude coordinates. But one's elevation may be quite different from where it was at the start of the path.

Mathematically, a connection above a region in the base is flat only if the planes of directions defining the connection are tangent to a family of surfaces. Each surface corresponds to the hillside in the analogy. The surfaces must conform to one another and pack together like spoons to fill up the part of the total space that is above the region of the base. Each surface must have the same number of dimensions as the base has.

It is now possible to give some account of how fiber bundles can represent the distinctive features of a gauge field. A more descriptive name for a gauge field would be a **phase-shift field**. In current gauge-field theories of nuclear forces, the phase shifts act on quantum mechanical waves to change the identity of the particle that the wave describes. For example, to turn the probability amplitudes for a proton field into those for a neutron field and back again (which is in effect to alter continuously the probability that a particle is a neutron or a proton), it is sufficient to shift generalized phases.

The Quantum theory of magnetism provides a much simpler example of the phase-shifting properties of a gauge field. The gauge field in question is called the "magnetic vector potential". It determines how electrons interact with a magnetic field. The clearest way to demonstrate the effect of the magnetic vector potential experimentally exploits the interference patterns of electron waves.

An electron in a beam can be represented by waves whose length is inversely proportional to the electron's momentum and whose frequency is proportional to its energy. At any point in space, such a wave has a definite height at each instant just as a water wave has an instantaneous height above-or-below the average surface level of the water. The sequence of heights at a point varies periodically from a maximum to a minimum and back again. A graph of the heights resembles a cosine curve.

Because the cosine is a function whose argument is an angle, the instantaneous height of a wave can be stated by giving the wave's maximum height and an angle that corresponds to the instantaneous height of the cosine curve. The curve attains its maximum height, for example, at 0 degrees and its minimum (or greatest negative value) at 180 degrees. It has a height of zero at 90 and 270 degrees.

The angle corresponding to the instantaneous height of an electron wave is called the **phase angle** of the wave. The phase angle does not affect the probability of finding an electron at a point because the probability amplitude is proportional only to the maximum height of the electron wave. (As usual, the probability is found by squaring the probability amplitude.) Thus if one shifts the phase of the electron wave arbitrarily at each point in space, the probability of finding an electron at any point does not change. A function that assigns such a local phase shift to each point in space is called a **gauge transformation**.

Although the overall phase of a single electron beam has no effect on observed quantities, the relative phase with which 2 partial beams arrive at the same point does have physical consequences. A phase difference between 2 interfering beams can alter the maximum height of the electron wave and so can change the probability amplitude.

An interference pattern shows up as a variation from one point to another in the probability of finding a particle. Hence wherever 2 partial beams overlap, they create an interference pattern. Suppose the experiment is symmetrical and the beams are exactly in phase at the center of a detecting screen. The value of the phase at the center of the screen can be changed at will. But the same pattern will then be made in each of the partial beams. Therefore, the interference is constructive under any gauge transformation at the point.

The interference pattern is formed because for points on the detecting screen (say, to the left of the center), the waves from the left partial beam have traveled a shorter distance to reach the screen than the waves from the right partial beam. As one searches farther from the center of the pattern, the phase difference or relative phase takes on greater values, creating the periodic variation in intensity that constitutes the interference pattern. The pattern is detected as a variation in the counting rate in the region of overlapping partial beams. If some physical effect introduces a phase difference between the 2 partial beams arriving at the center, the same phase difference will also be found at each point to the right or left of the center. The "fringes" in the interference pattern will be uniformly shifted.

For electrically-charged particles, the magnetic vector potential is a field that acts by shifting phases at each point in space. The vector potential determines the magnetic field. Indeed, magnetic effects on charged particles can be completely explained in terms of the phase shifts given by the vector potential.

The underlying reason that the vector potential field is a gauge field is that the magnetic force acts on a charged particle to change its direction without changing its energy. When an electron enters a magnetic field, the frequency of the electron waves therefore remains constant. But the spatial pattern of the waves is changed. It is as if the wave length were to vary from point-to-point. The phase shifts of an electron caused by a magnetic field, therefore, depends on the path of the electron.

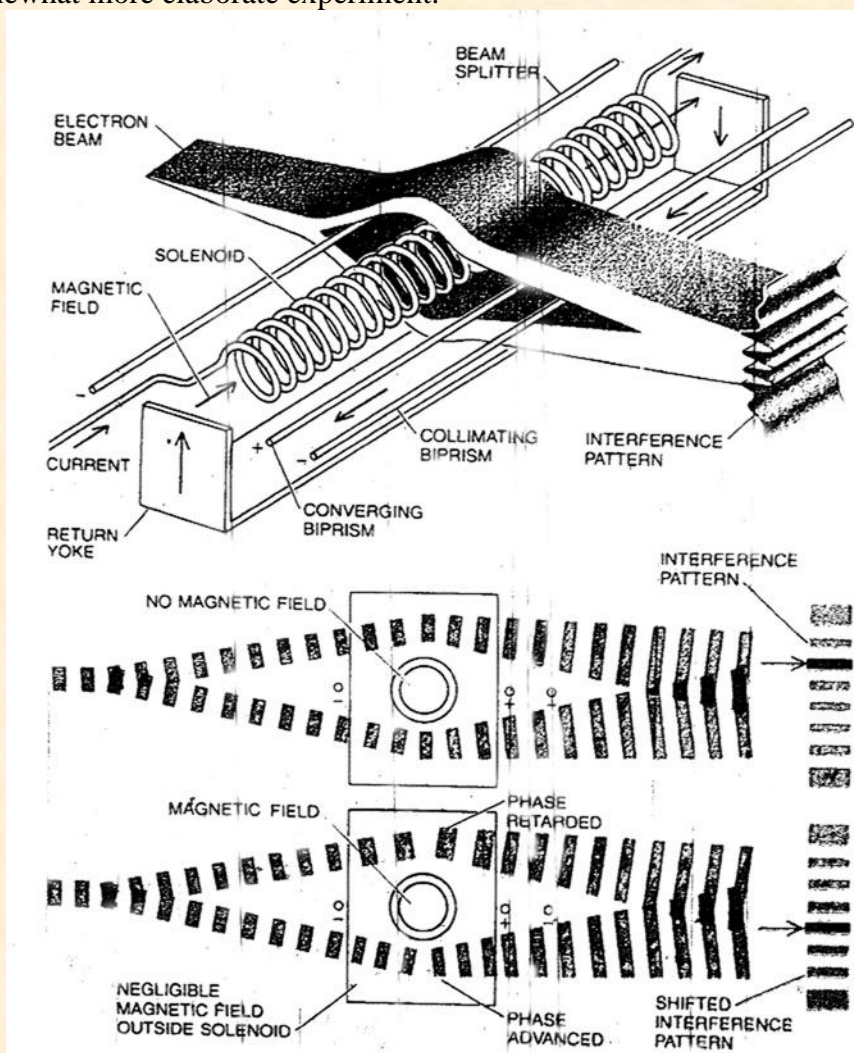
In the Classical theory of magnetism, the magnetic vector potential was conceived as an auxiliary device for calculating the magnetic field. The calculations showed that the magnetic vector potential could be nonzero in regions where there is no magnetic field. As a result, physicists thought the magnetic vector potential would not have observable consequences of its own.

In Quantum Mechanics, the magnetic vector potential field does have observable consequences. Its expected effect on phase was mentioned by W. Ehrenberg of the University of London and R. E. Siday of the University of Edinburgh in 1949. It was not until 1959, however, that Yakir Aharonov of Yeshiva University and David Bohm of the University of London proposed an experiment in which the effect could be observed directly.

Like the neutron-rotation experiment, the Aharonov-Bohm experiment calls for splitting a beam of subatomic particles; recombining the partial beams; and observing the resulting interference. The particles employed are electrons rather than neutrons. This presents a difficult technical problem because the maximum separation over which the 2 parts of a split electron beam remain coherent is only 60 micrometers.

Even to achieve such a small separation, the entire experiment must be set up inside an electron microscope. The experiment was first carried out by R. G. Chambers of the University of Bristol in

1960. The effect was confirmed in 1961 by Gottfried Mollenstedt and Werner Bayh of the University of Tübingen in a somewhat more elaborate experiment.

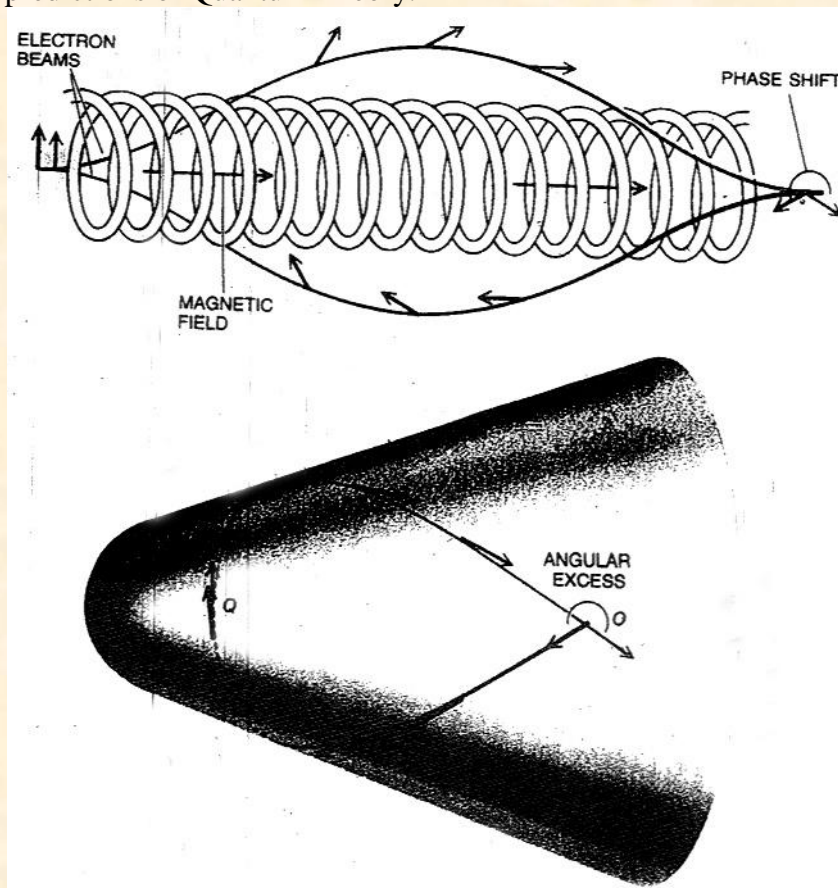


Electron phase shift can be induced not only by a magnetic field but also even by the passage of the electron through a region near a magnetic field. The effect can be detected by splitting an electron beam with a negatively-charged wire; passing the 2 partial beams around opposite sides of a solenoid within which a magnetic field has been confined; and then recombining the partial beams so that they interfere. As the current in the solenoid is increased, the magnetic field increases and the interference pattern of the overlapping beams is shifted. Because the electron beams remain coherent only if they are separated by no more than 60 micrometers, the experimental apparatus is exceedingly small. The diameter of the solenoid is less than  $1/7^{\text{th}}$  the thickness of a human hair. An experiment of this kind was first proposed in 1959 by Yakir Aharonov of Yeshiva University and David Bohm of the University of London.

The incoming electron beam travels toward a wire bearing a negative charge which splits the beam into 2 coherent partial beams. The diverging beams pass around opposite sides of a solenoidal electromagnet with an outside diameter of 14 micrometers (less than  $1/7^{\text{th}}$  the thickness of a human hair). The electromagnet is installed in such a way as to confine the magnetic field to the inside of the solenoid. Since the electrons remain outside the solenoid, they pass through a negligible magnetic field. Any phase changes they undergo must, therefore, be attributed to the magnetic vector potential field that surrounds the solenoid.

Beyond the solenoid in the electron microscope is a wire bearing a positive charge that brings the beams back together. Another negatively-charged wire deflects the beams so that they intersect at a small angle in order to increase the width of the fringes in the interference pattern. The resulting pattern has broad dark bands from diffraction at the beam-splitting wire and fine fringes from the interference.

As the current in the solenoid increases, the magnetic field becomes stronger and the fine lines of the interference pattern shift with respect to the broad diffraction bands. The magnetic flux can be calculated from the dimensions of the coil and the current passing through it. Hence, the experiment gives an exact numerical check on the relation between the magnetic flux and the phase shift. The results confirm the predictions of Quantum Theory.



Phase shift of the electron beams in the Ahronov-Bohm experiment can be modeled by the parallel transport of a direction on the surface of a truncated cone capped with a spherical dome. The shifting of phase along each beam is represented by the rotation of an arrow. The 2 partial beams begin in phase. The phase is then retarded along the colored path and advanced along the black path. The phases shift even though the magnetic field along both paths is zero. The shifts are directly proportional to the magnetic field between the paths. On the domed cone, the geometry of the conical region is like the geometry of a plane. The region can be slit longitudinally and unrolled onto a plane without stretching or compressing. Parallel transport along 2 paths around the cone, however, will not lead to the same direction when the paths meet again even though the curvature along both paths is zero. The paths shown form a curve that is a self-intersecting geodesic; the arrows represent the direction tangent to the curve. Because the tangent direction undergoes parallel transport along a geodesic, the angular excess of the curve is identical with the angle at which the 2 paths meet. The angular excess measured in radians is equal to the total curvature of the region between the paths. The curvature is concentrated on the dome just as the magnetic field is confined to the solenoid.

The fiber-bundle model of this experiment illuminates the general correspondence between gauge fields and connections in fiber bundles. The model describes parallel transport on the surface formed by truncating a cone and capping it with a dome.

Topologically, the surface formed in this way is equivalent to the surface of a hemisphere. But it has a quite different geometry. Note that if the conical region is slit longitudinally, the cone can be unrolled onto a planar surface without stretching or compressing. Lengths and angles measured on the unrolled cone must be the same as they are when they are measured on the cone itself. Thus a 2-dimensional observer measuring curvature on the surface of the cone by carrying out parallel transport over small

closed curves would always find the angular excess to be zero. The surface would appear to be flat. This part of the surface is analogous to the region outside the solenoid in the Aharonov-Bohm experiment where the magnetic field is negligible.

If sufficiently large paths on the cone are considered, however, the difference between the cone and the plane will become apparent. A straight line on the unrolled cone that intersects both sides of the slit at the same distance from the apex will roll up into a closed loop that is a geodesic on the cone.

Suppose a pair of 2-dimensional observers stand both facing forward at the point on the loop directly opposite the point where the loop crosses the slit. If the observers move along the loop (one walking forward and the other backward), they must think that they are moving away from each other along a straight line.

When they meet again at the point where the loop crosses the slit, however, they no longer face in the same direction. The angle  $\theta$  between them is equal to 360 degrees minus the angle at the top of the unrolled cone. The observers must conclude that the surface they live in is not flat after all because this angle is the angular excess of the closed curve formed by their paths. The region to their left must have a total curvature equal to the measure in radians of the angular excess.

What happens to the pair of 2-dimensional observers as they pass around opposite sides of the cone is similar to what happens to the pair of partial beams as they pass around opposite sides of the solenoid in the Aharonov-Bohm experiment. The 2 observers initially face in the same direction. Analogously, the 2 partial beams of electrons are in phase immediately after the beam is split. If the observers measure the curvature of the surface along either one of their paths, they will find it is zero because they are traveling on the flat part of the capped cone. Similarly, the magnetic vector potential -- which causes the beams to be out of phase when they are recombined -- cannot be detected by measurements made along either one of the separate paths.

The key to the similarity of the 2 effects is the interpretation of the magnetic vector potential as a connection in a fiber bundle. The bundle is called the "bundle of phases". Its base space is the 3-dimensional space in which the experiment takes place. The fiber over any point is the set of all possible phases of electrons at that point. And so the total space is made up of all possible phases of electrons at all points of 3-dimensional space. Since a phase can be described by an angle measured in degrees or radians, the fiber over each point is a circle just as it is in the bundle of directions on a surface. With a circular fiber for every point in 3 dimensional space, however, the total space is **4-dimensional**.

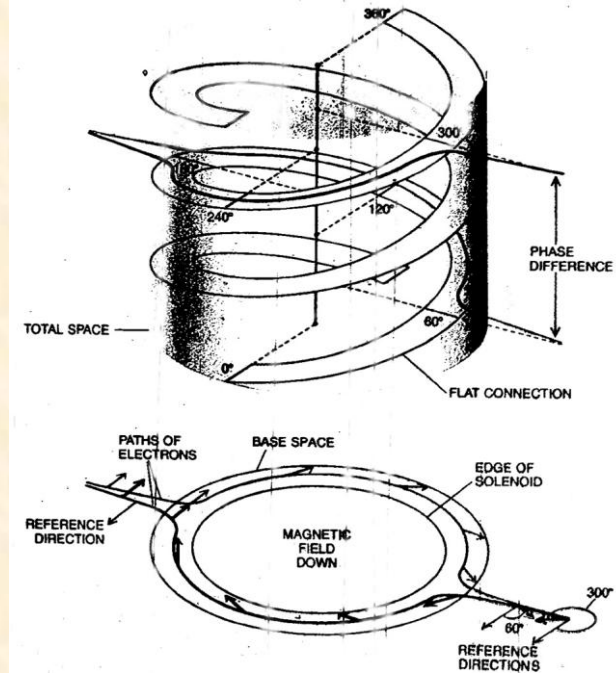
The 3-dimensional base space and 4-dimensional total space make it difficult to visualize the bundle of phases. A picture can be drawn, however, of the bundle of phases over a plane that passes through the experiment perpendicular to the solenoid. This bundle (which has a 3-dimensional total space) is completely adequate for describing the outcome of the Aharonov-Bohm experiment.

The bundle is now a cylindrical bundle over a surface. It can be drawn with vertical fibers like the bundle of directions on the hemisphere or on the capped cone. The connection in the bundle is a field of planes transverse to the fibers. The slope of one of the planes in a given direction (say, up or down the ramp) is proportional to the component of the magnetic vector potential field in that direction at the corresponding base point.

The curvature of the connection is proportional to the intensity of the magnetic field. Hence the curvature is different from zero only at points inside the solenoid. For an appropriate choice of the field strength inside the solenoid, the picture of the magnetic vector potential connection over the plane

perpendicular to the coil is identical with the picture of the parallel transport connection over the capped cone. The region of the base inside the solenoid corresponds to the spherical cap on the cone. There the connection is identical with the connection that is generated by parallel transport on a sphere.

In the region of base outside the solenoid, the curvature of the connection is zero since the magnetic field is zero there. The connection planes are tangent to a densely-packed family of ramps that spiral around the center. The phase difference between the 2 partial electron beams is exactly the height gained by following one of the ramps through one full rotation.



Fiber bundle of phase angles over a cross-section of the Aharonov-Bohm experiment is almost exactly like the fiber bundle of directions on the surface of a spherically-domed cone. Over each point in the base, the fiber in the total space represents the possible phase angles (from 0-to-360 degrees) of the electron at that point. For an appropriate choice of magnetic field intensity, the connection defined in the bundle of phases by the magnetic vector potential field is identical with the connection given by parallel transport on the domed cone. The curvature of the connection corresponds to the magnetic field in the solenoid. The region inside the solenoid corresponds to the spherical dome. There the rule for lifting a path is the same as the parallel-transport rule on the surface of the sphere. It can be described by same system of sloping planes (not shown). The region outside the solenoid corresponds to the truncated cone. There the connection is flat. The sloping planes are tangent to a family of spiral ramps that fill up the total space. The phase shift in the base is represented by the rotating arrows.

If the curvature is calculated for the entire bundle of phases, it is found to be a tensor since the base is 3-dimensional. The tensor has 3 components, each proportional to the corresponding components of the magnetic field. The construction can be extended to the bundle of phases over space-time where the base is 4-dimensional. The electric potential and the magnetic potential taken together give a connection that explain all electromagnetic interactions of charged particles. The curvature of the connection is a tensor with 6 components corresponding to the 3 components of the electric field and the 3 components of the magnetic field.

Each of the quantum gauge fields can be understood as a connection in a fiber bundle where the base is space-time. The fiber of the bundle is the set of internal symmetry transformations of particles that interact by means of the gauge field. Fiber bundles have now been constructed for Quantum ElectroDynamics; for the weak and the strong nuclear interactions; and for versions of theories that attempt to present a unified account of these forces. Although a complete scheme for the unification of all the fundamental forces has not come forth, it is clear that the mathematics of connections in fiber bundles is destined to play an important role in Theoretical Physics.



We believe the current usefulness and physical significance of such mathematical concepts is no accident. Neither mathematicians nor physicists are insulated from their cultural, political, and physical milieu. And the ideas and perceptions of workers in each discipline are influenced by the other discipline. Moreover, mathematicians and physicists unavoidably share unspoken assumptions about the everyday World and the logic by means of which the world is projected onto abstract science. Indeed, they share a passionate commitment to such rational work.

What seems most marvelous is not what has been called the "unreasonable effectiveness" of mathematical concepts in Physics or the fecundity of physical intuition as a source of new mathematics. Rather, one must admire the success of the common intellectual approach of mathematicians and physicists in creating a rich, coherent, and powerful image of the physical Universe.

**if on the Internet, press <BACK> on your browser to return to  
the previous page (or go to [www.stealthskater.com](http://www.stealthskater.com))**

**else if accessing these files from the CD in a MS-Word session, simply <CLOSE> this  
file's window-session; the previous window-session should still remain 'active'**