Annales de la Fondation Louis de Broglie, Volume 26, n° spécial, 2001

Topological Foundations of Electromagnetism

TERENCE W. BARRETT

BSEI
1453 Beulah Road
Vienna, VA 22182, USA

1 Introduction

Beginning with G.W. Leibniz in the 17th, L. Euler in the 18th, B. Reimann, J.B. Listing and A.F. Möbius in the 19th and H. Poincaré in the 20th centuries, "*analysis situs*" (Riemann) or "topology" (Listing) has been used to provide answers to questions concerning what is most fundamental in physical explanation. That question itself implies the question concerning what mathematical structures one uses with confidence to adequately "paint" or describe physical models built from empirical facts. For example, differential equations of motion cannot be fundamental, because they are dependent on boundary conditions which must be justified - usually by group theoretical considerations. Perhaps, then, group theory[1] is fundamental.

---

[1] Here we address the kind of groups addressed in Yang-Mills theory, which are *continuous* groups (as opposed to *discrete* groups). Unlike discrete groups, continuous groups contain an infinite number of elements and can be differentiable or analytic. Cf. Yang, C.N. & Mills, R.L., Conservation of isotopic spin and isotopic gauge invariance. *Phys. Rev.*, 96, 191-195, 1954.

Group theory certainly offers an austere shorthand for fundamental transformation rules. But it appears to the present writer that the final judge of whether a mathematical group structure can, or cannot, be applied to a physical situation is the topology of that physical situation. Topology dictates and justifies the group transformations.

So for the present writer, the answer to the question of what is the most fundamental physical description is that it is a description of the topology of the situation. With the topology known, the group theory description is justified and equations of motion can then be justified and defined in specific differential equation form. If there is a requirement for an understanding more basic than the topology of the situation, then all that is left is verbal description of visual images. So we commence an examination of electromagnetism under the assumption that topology defines group transformations and the group transformation rules justify the algebra underlying the differential equations of motion.

There are a variety of special methods used to solve ordinary differential equations. It was Sophus Lie (1842-99) in the 19th century who showed that all the methods are special cases of integration procedures which are based on the invariance of a differential equation under a continuous group of symmetries. These groups became known as Lie groups[2]. A symmetry group of a system of differential equations is a group which transforms solutions of the system to other solutions[3].

The relationship was made more explicit by Noether's theorems[4], which related symmetry groups of a variational integral to properties of its associated Euler-Lagrange equations. The most important consequences of this relationship are that (i) conservation of energy arises from invariance under a group of time translations; (ii) conservation of linear momentum arises from invariance under translational groups; and (iii) conservation of

---

[2] **Lie Group Algebras**

If a topological group is a group and also a topological space in which group operations are continuous, then Lie groups are topological groups which are also analytic manifolds on which the group operations are analytic.

In the case of Lie algebras, the parameters of a product are analytic functions of the parameters of each factor in the product. For example, $L(\gamma) = L(\alpha)L(\beta)$ where $\gamma = f(\alpha, \beta)$. This guarantees that the group is differentiable. The Lie groups used in Yang-Mills theory are *compact* groups, i.e., the parameters range over a closed interval.

[3] Cf. Olver, P.J., *Applications of Lie Groups to Differential Equations*, Springer Verlag, 1986.

[4] Noether, E., Invariante Variations Probleme. *Nachr. Ges. Wiss. Goettingen, Math.-Phys.* Kl. 171, 235-257, 1918.

angular momentum arises from invariance under rotational groups. Conservation and group symmetry laws have been vastly extended to other systems of equations, e.g., soliton equations. For example, the Korteweg de Vries "soliton" equation[5] yields a symmetry algebra spanned by the four vector fields of (i) space translation; (ii) time translation; (iii) Galilean translation; and (iv) scaling.

For some time, the present writer has been engaged in showing that the spacetime topology defines electromagnetic field equations[6] - whether the fields be of force or of phase. That is to say, the premise of this enterprise is that a set of field equations are only valid with respect to a set defined topological description of the physical situation. In particular, the writer has addressed demonstrating that the $\boldsymbol{A}_\mu$ potentials, $\mu = 0, 1, 2, 3$, are not just a mathematical convenience, but - *in certain well-defined situations* - are measurable, i.e., physical. Those situations in which the $\boldsymbol{A}_\mu$ potentials are measurable possess a topology, the transformation rules of which are describable by the *SU(2)* group[7] or higher-order groups; and those situations in which the $\boldsymbol{A}_\mu$ potentials are not measurable possess a topology, the transformation rules of which are describable by the *U(1)* group[8].

---

[5] Korteweg, D.J. & de Vries, G. On the change of form of long waves advancing in a rectangular canal, and on a new type of long stationary wave. *Philos. Mag.*, 39, 422-443, 1895.

[6] Barrett, T.W. Maxwell's theory extended. Part I. Empirical reasons for questioning the completeness of Maxwell's theory - effects demonstrating the physical significance of the *A* potentials. *Annales de la Fondation Louis de Broglie*, 15, 143-183, 1990;

______, Maxwell's theory extended. Part II. Theoretical and pragmatic reasons for questioning the completeness of Maxwell's theory. *Annales de la Fondation Louis de Broglie*, 12, 253-283, 1990;

______, The Ehrenhaft-Mikhailov effect described as the behavior of a low energy density magnetic monopole instanton. *Annales de la Fondation Louis de Broglie*, 19, 291-301, 1994;

______, Electromagnetic phenomena not explained by Maxwell's equations. pp. 6-86 in Lakhtakia, A. (Ed.) *Essays on the Formal Aspects of Maxwell's Theory*, World Scientific, Singapore, 1993;

______, Sagnac effect. pp. 278-313 in Barrett, T.W. & Grimes, D.M., (Ed.s) *Advanced Electromagnetism: Foundations, Theory, Applications*, World Scientific, Singapore, 1995;

______, The toroid antenna as a conditioner of electromagnetic fields into (low energy) gauge fields. In *Speculations in Science and Technology*, 21(4), 291-320, 1998.

[7] See appendix I on : ***SU(n)* Group Algebra**

[8] ***U(n) Group Algebra***

Table 1		
	U(1) Symmetry Form (Traditional Maxwell Equations)	**SU(2) Symmetry Form**
Gauss' Law	$\nabla \bullet \boldsymbol{E} = J_0$	$\nabla \bullet \boldsymbol{E} = J_0 - iq(\boldsymbol{A} \bullet \boldsymbol{E} - \boldsymbol{E} \bullet \boldsymbol{A})$
Ampère's Law	$\frac{\partial \boldsymbol{E}}{\partial t} - \nabla \times \boldsymbol{B} - \boldsymbol{J} = 0$	$\frac{\partial \boldsymbol{E}}{\partial t} - \nabla \times \boldsymbol{B} - \boldsymbol{J} + iq[A_0, \boldsymbol{E}]$ $- iq(\boldsymbol{A} \times \boldsymbol{B} - \boldsymbol{B} \times \boldsymbol{A}) = 0$
	$\nabla \bullet \boldsymbol{B} = 0$	$\nabla \bullet \boldsymbol{B} + iq(\boldsymbol{A} \bullet \boldsymbol{B} - \boldsymbol{B} \bullet \boldsymbol{A}) = 0$
Faraday's Law	$\nabla \times \boldsymbol{E} + \frac{\partial \boldsymbol{B}}{\partial t} = 0$	$\nabla \times \boldsymbol{E} + \frac{\partial \boldsymbol{B}}{\partial t} + iq[A_0, \boldsymbol{B}] =$ $iq(\boldsymbol{A} \times \boldsymbol{E} - \boldsymbol{E} \times \boldsymbol{A}) = 0$

Unitary matrices, U, have a determinant equal to ±1. The elements of $U(n)$ are represented by $n \times n$ unitary matrices.

U(1) Group Algebra

The one-dimensional unitary group, or *U(1),* is characterized by one continuous parameter. *U(1)* is also differentiable and the derivative is also an element of *U(1)*. A well-known example of a *U(1)* group is that of all the possible phases of a wave function, which are angular coordinates in a 2-dimensional space. When interpreted in this way - as the internal phase of the *U(1)* group of electromagnetism - the *U(1)* group is merely a circle ($0 - 2\pi$).

Table 2

U(1) Symmetry Form (Traditional Maxwell Theory)	SU(2) Symmetry Form
$\rho_e = J_0$	$\rho_e = J_0 - iq(\boldsymbol{A} \bullet \boldsymbol{E} - \boldsymbol{E} \bullet \boldsymbol{A}) = J_0 + q\boldsymbol{J}_z$
$\rho_m = 0$	$\rho_m = -iq(\boldsymbol{A} \bullet \boldsymbol{B} - \boldsymbol{B} \bullet \boldsymbol{A}) = -iq\boldsymbol{J}_y$
$g_e = \boldsymbol{J}$	$g_e = iq[A_0, \boldsymbol{E}] - iq(\boldsymbol{A} \times \boldsymbol{B} - \boldsymbol{B} \times \boldsymbol{A}) + \boldsymbol{J} = iq[A_0, \boldsymbol{E}] - iq\boldsymbol{J}_x + \boldsymbol{J}$
$g_m = 0$	$g_m = iq[A_0, \boldsymbol{B}] - iq(\boldsymbol{A} \times \boldsymbol{E} - \boldsymbol{E} \times \boldsymbol{A}) = iq[A_0, \boldsymbol{B}] - iq\boldsymbol{J}_z$
$\sigma = \boldsymbol{J} / \boldsymbol{E}$	$\sigma = \dfrac{\{iq[A_0, \boldsymbol{E}] - iq(\boldsymbol{A} \times \boldsymbol{B} - \boldsymbol{B} \times \boldsymbol{A}) + \boldsymbol{J}\}}{\boldsymbol{E}} = \dfrac{\{iq[A_0, \boldsymbol{E}] - iq\boldsymbol{J}_x + \boldsymbol{J}\}}{\boldsymbol{E}}$
$s = 0$	$s = \dfrac{\{iq[A_0, \boldsymbol{B}] - iq(\boldsymbol{A} \times \boldsymbol{E} - \boldsymbol{E} \times \boldsymbol{A})\}}{\boldsymbol{H}} = \dfrac{\{iq[A_0, \boldsymbol{B}] - iq\boldsymbol{J}_z\}}{\boldsymbol{H}}$

Historically, electromagnetic theory was developed for situations described by the *U(1)* group. The dynamic equations describing the transformations and interrelationships of the force field are the well known Maxwell equations, and the group algebra underlying these equations is *U(1)*. There was a need to extend these equations to describe *SU(2)* situations and to derive equations whose underlying algebra is *SU(2)*. These two formulations are shown in Table 1. Table 2 shows the electric charge density, ρ_e, the magnetic charge density, ρ_m, the electric current density, g_e, the magnetic current density, g_m, the electric conductivity, σ, and the magnetic conductivity, s.

2 Solitons[9]

Soliton solutions to differential equations require complete integrability and integrable systems conserve geometric features related to symmetry. Unlike the equations of motion for conventional Maxwell theory, which are solutions of *U(1)* symmetry systems, solitons are solutions of *SU(2)* symmetry systems. These notions of group symmetry are more fundamental than differential equation descriptions. Therefore, although a complete exposition is beyond the scope of the present chapter, we develop some basic concepts in order to place differential equation descriptions within the context of group theory.

Within this context, *ordinary differential equations are viewed as vector fields on manifolds or configuration spaces*[10]. For example, Newton's equations are second-order differential equations describing smooth curves on Riemannian manifolds. Noether's theorem[11] states that a diffeomorphism[12], ϕ, of a Riemannian manifold, C, indices a diffeomorphism, $D\phi$, of its tangent[13] bundle[14], TC. If ϕ is a symmetry of Newton's equations, then $D\phi$ preserves the Lagrangian, i.e.,

---

[9] A soliton is a solitary wave which preserves its shape and speed in a collision with another solitary wave. Cf. Barrett, T.W., 404-413 in Taylor, J.D. (ed.) *Introduction to Ultra-Wideband Radar Systems*, CRC Press, Boca Raton, 1995.

[10] Cf. Olver, P.J., *Applications of Lie Groups to Differential Equations*, Springer Verlag, 1986.

[11] Noether, E., Invariante Variations Probleme. *Nachr. Ges. Wiss. Goettingen, Math.-Phys. Kl.* 171, 235-257, 1918.

[12] A ***diffeomorphism*** is an elementary concept of topology and important to the understanding of differential equations. It can be defined in the following way:

If the sets U and V are open sets both defined over the space R^m, i.e., $U \subset R^m$ is open and $U \subset R^m$ is open, where open means nonoverlapping, then the mapping $\psi{:}U \rightarrow V$ is an infinitely differentiable map with an infinitely differential inverse, and objects defined in U will have equivalent counterparts in V. The mapping ψ is a diffeomorphism. It is a smooth and infinitely differentiable function. The important point is: conservation rules apply to diffeomorphisms, because of their infinite differentiability. Therefore diffeomorphisms constitute fundamental characterizations of differential equations.

[13] A vector field on a manifold, M, gives a ***tangent vector*** at each point of M.

$$\boldsymbol{L} \circ D\phi = \boldsymbol{L} \quad .$$

As opposed to equations of motion in conventional Maxwell theory, *soliton flows are Hamiltonian flows*. Such Hamiltonian functions define *symplectic structures*[15] for which there is an absence of local invariants but an infinite dimensional group of diffeomorphisms which preserve global properties. In the case of solitons, the global properties are those permitting the matching of the nonlinear and dispersive characteristics of the medium through which the wave moves.

In order to relate the three major soliton equations to group theory it is necessary to examine a Lax equation[16] or the *zero-curvature condition* (ZCC). The ZCC expresses the flatness of a connection by the commutation relations of the covariant derivative operators[17]:

$$A_t - B_x - [A, B] = 0 \quad .$$

Reformulated as a Lax equation the ZCC is[17]:

$$\left[\frac{\partial}{\partial x} - A, \frac{\partial}{\partial t} - B\right] = 0,$$

or

$$\left(\frac{\partial}{\partial x} - A\right)_t = \left[B, \frac{\partial}{\partial x} - A\right] .$$

Palais[17] shows that the generic cases of soliton - the *Korteweg de Vries Equation* (KDV), the *Nonlinear Schrödinger Equation* (NLS), the *Sine-Gordon Equation* (SGE) - can be given an *SU(2)* formulation. In each case,

---

[14] A ***bundle*** is a structure consisting of a manifold E, and manifold M, and an onto map: $\pi : E \to M$.

[15] ***Symplectic topology*** is the study of the global phenomena of symplectic symmetry. Symplectic symmetry structures have no local invariants. This is a subfield of topology: for example: McDuff, D. & Salamon, D., *Introduction to Symplectic Topology* Oxford: Clarendon Press, 1995.

[16] Lax, P.D., Integrals of nonlinear equations of evolution and solitary waves. *Comm. Pure Appl. Math.*, 21, 467-490, 1968;

Lax, P.D., Periodic solutions of the KdV equations, in *Nonlinear Wave Motion*, *Lectures in Applied Math*., 15, *American Mathematical Society*, 85-96, 1974.

[17] Palais, R.S., The symmetries of solitons. *Bull. Am. Math. Soc.* 34, 339-403, 1997.

below, V is a one-dimensional space that is embedded in the space of off-diagonal complex matrices, $\begin{pmatrix} 0 & b \\ c & 0 \end{pmatrix}$ and in each case $A(u) = \boldsymbol{a}\lambda + u$, where u is a potential, λ is a complex parameter, and $\boldsymbol{a}$ is the constant, diagonal, trace zero matrix $\boldsymbol{a} = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}$ - which links these equations to an *SU(2)* formulation.

From inverse scattering theory, a function is needed, defined:

$$B(\xi) = \sum_{n=1}^{N} c_n^2 \exp[-\kappa_n \xi] + \frac{1}{2\pi} \int_{-\infty}^{+\infty} b(k) \exp[ik\xi] dk, \ \textit{where}$$

$-\kappa_1^2,\ldots,-\kappa_N^2$ *are discrete eigenvalues of* u,

$c_1,\ldots c_N$ *are normalizing consts, and*

$b(k)$ *are reflection coefficients.*

In the ***first case*** (the KdV), if $u(x) = \begin{pmatrix} 0 & q(x) \\ -1 & 0 \end{pmatrix}$ and

$$B(u) = a\lambda^3 + u\lambda^2 + \begin{pmatrix} \frac{i}{2}q & \frac{i}{2}q_x \\ 0 & -\frac{i}{2}q \end{pmatrix}\lambda + \begin{pmatrix} \frac{q_x}{4} & \frac{-q^2}{2} \\ \frac{q}{2} & \frac{-q_x}{4} \end{pmatrix},$$

then ZCC is satisfied if and only if q satisfies the KdV in the form

$$q_t = -\frac{1}{4}\left(6qq_x + q_{xxx}\right).$$

In the ***second case*** (the NLS), if $u(x) = \begin{pmatrix} 0 & q(x) \\ -\bar{q}(x) & 0 \end{pmatrix}$ and

$$B(u) = a\lambda^3 + u\lambda^2 + \begin{pmatrix} \frac{i}{2}|q|^2 & \frac{i}{2}q_x \\ -\frac{i}{2}\bar{q}_x & -\frac{i}{2}|q|^2 \end{pmatrix},$$

then ZCC is satisfied if and only if *q(x,t)* satisfies the NLS in the form

$$q_t = \frac{i}{2}\left(q_{xx} + 2|q|^2 q\right).$$

In the ***third case*** (the SGE), if $u(x) = \begin{pmatrix} 0 & -\dfrac{q_x(x)}{2} \\ \dfrac{q_x(x)}{2} & 0 \end{pmatrix}$ and

$$B(u) = \frac{i}{4\lambda}\begin{pmatrix} \cos[q(x)] & \sin[q(x)] \\ \sin[q(x)] & -\cos[q(x)] \end{pmatrix},$$

then ZCC is satisfied if and only if q satisfies the SGE in the form $q_t = \sin[q]$.

Thus, if the Maxwell equation of motion with electric *and* magnetic conductivity is in soliton (SGE) form, the group symmetry is *SU(2)*. Solitons define Hamiltonian flows and their energy conservation is due to their *symplectic structure*.

In order to clarify the difference between conventional Maxwell theory which is of *U(1)* symmetry, and Maxwell theory extended to *SU(2)* symmetry, we can describe both in terms of mappings of a field $\psi(x)$. In the case of *U(1)* Maxwell theory, a mapping $\psi \to \psi'$ is:

$$\psi(x) \to \psi'(x) = \exp[ia(x)]\psi(x),$$

where $\boldsymbol{a}(x)$ is the conventional vector potential. However, in the case of *SU(2)* extended Maxwell theory, a mapping $\psi \to \psi'$ is:

$$\psi(x) \to \psi'(x) = \exp[iS(x)]\psi(x),$$

where $\boldsymbol{S}(x)$ is the action and an element of *SU(2)* defined:

$$\boldsymbol{S}(x) = \int A dx \ ,$$

and $\boldsymbol{A}$ is the matrix form of the vector potential. Therefore we see the necessity to adopt a matrix formulation of the vector potential when addressing *SU(2)* forms of Maxwell theory.

3 Instantons

Instantons[18] correspond to the minima of the Euclidean action and are pseudo-particle solutions[19] of *SU(2)* Yang-Mills equations in Euclidean 4 space[20]. A complete construction for any Yang-Mills group is also available[21]. In other words:

"It is reasonable... to ask for the determination of the classical field configurations in Euclidean space which minimize the action, subject to appropriate asymptotic conditions in 4-space. These classical solutions are the instantons of the Yang-Mills theory."[22]

In the light of the intention of the present writer to introduce topology into electromagnetic theory, I quote further:

"If one were to search ab initio for a non-linear generalization of Maxwell's equation to explain elementary particles, there are various symmetry group properties one would require. These are

(i) *external symmetries* under the Lorentz and Poincaré groups and under the conformal group if one is taking the rest-mass to be zero,

(ii) *internal symmetries* under groups like *SU(2)* or *SU(3)* to account for the known features of elementary particles,

(iii) *covariance* or the ability to be coupled to gravitation by working on curved space-time."[23]

The present writer applied the instanton concept in electromagnetism for the following two reasons: (1) in some sense, the instanton, or pseudo particle, is a compactification of degrees of freedom due to the particle's boundary conditions; and (2) the instanton, or pseudoparticle, might have the behavior of a real, high energy particle, but without the presence of high

---

[18] Cf. Jackiw, R, Nohl, C. & Rebbi, C. *Classical and semi-classical solutions to Yang-Mills theory*. Proceedings 1977 Banff School, Plenum Press.

[19] Belavin, A., Polyakov, A., Schwartz, A. & Tyupkin, Y., Pseudoparticle solutions of the Yang-Mills equations. *Phys. Lett.*, 59B, 85-87, 1975.

[20] Cf. Atiyah, M.F. & Ward, R.S., Instantons and Algebraic Geometry, *Commun. math. Phys.*, 55, 117-124, 1977.

[21] Atiyah, M.F., Hitchin, N.J., Drinfeld, V.G. & Manin, Yu.I., *Construction of instantons*. Phys. Lett., 65A, 23-25, 1978.

[22] Atiyah, M., p. 80 in *Michael Atiyah: Collected Works, Volume 5, Gauge Theories*, Clarendon Press, Oxford, 1988.

[23] Atiyah, M., p. 81 in *Michael Atiyah: Collected Works, Volume 5, Gauge Theories*, Clarendon Press, Oxford, 1988.

energy, i.e., the pseudoparticle would share certain behavioral characteristics in common with a real, high energy particle.

Therefore, I suggested[24] that the Mikhailov effect[25], and the Ehrenhaft effect (1879-1952) which demonstrate magnetic charge-like behavior, are actually descriptions or manifestations of instanton or pseudoparticle behavior.

4 Aharonov-Bohm Effect

As an example of the basic nature of topological explanation, we considered the Aharonov-Bohm effect. Beginning in 1959 Aharonov and Bohm[26] challenged the view that the classical vector potential produces no observable physical effects by proposing two experiments. The one which is most discussed is shown in Fig 1. A beam of monoenergetic electrons exists from a source at X and is diffracted into two beams by the slits in a wall at $Y1$ and $Y2$. The two beams produce an interference pattern at III which is measured. Behind the wall is a solenoid, the $\boldsymbol{B}$ field of which points out of the paper. The absence of a free local magnetic monopole postulate in conventional $U(1)$ electromagnetism ($\nabla \bullet \boldsymbol{B} = 0$) predicts that the magnetic field outside the solenoid is zero. Before the current is turned on in the solenoid, there should be the usual interference patterns observed at III, of course, due to the differences in the two path lengths.

Aharonov and Bohm made the interesting prediction that if the current is turned on, then due to the differently directed $\boldsymbol{A}$ fields along paths 1 and 2 indicated by the arrows in Fig. 1, additional phase shifts should be discernible at III. This prediction was confirmed experimentally[27] and the evidence for the effect has been extensively reviewed[28].

---

[24] Barrett, T.W., The Ehrenhaft-Mikhailov effect described as the behavior of a low energy density magnetic monopole instanton. *Annales de la Fondation Louis de Broglie*, 19, 291-301, 1994.

[25] A summary of the Mikhailov effect is: pp. 593-619 in Barrett, T.W. & Grimes, D.M. (Ed.s) *Advanced Electromagnetism: Foundations, Theory & Applications*, World Scientific, Singapore, 1995.

[26] Aharonov, Y. & Bohm, D., Significance of the electromagnetic potentials in quantum theory. *Phys. Rev.*, 115, 485-491,1959.

[27] Chambers, R.G., Shift of an electron interference pattern by enclosed magnetic flux. *Phys. Rev. Lett.*, 5, 3-5, 1960;

Boersch, H., Hamisch, H., Wohlleben, D. & Grohmann, K., Antiparallele Weissche Bereiche als Biprisma für Elektroneninterferenzen. *Zeitschrift für Physik* 159, 397-404, 1960;

It is the present writer's opinion that the topology of this situation is fundamental and dictates its explanation. Therefore we must clearly note the topology of the physical layout of the design of the situation which exhibits the effect. The physical situation is that of an *interferometer*. That is, there are two paths around a central location - occupied by the solenoid - and a measurement is taken at a location, III, in the Fig 1, where there is overlap of the wave functions of the test waves which have traversed, separately, the two different paths. (The test waves or test particles are complex wave functions with phase.) It is important to note that the overlap area, at III, is the only place where a measurement can take place of the effects of the ***A*** field (which occurred earlier and at other locations (I and II). The effects of the ***A*** field occur along the two different paths and at locations I and II, but they are *inferred*, and not measurable there. Of crucial importance in this special interferometer, is the fact that the solenoid presents a *topological obstruction*.

Möllenstedt, G & Bayh, W., Messung der kontinuierlichen Phasenschiebung von Elektronenwellen im kraftfeldfreien Raum durch das magnetische Vektorpotential einer Luftspule. *Die Naturwissenschaften* 49, 81-82, 1962;

Matteucci, G. & Pozzi, G., New diffraction experiment on the electrostatic Aharonov-Bohm effect. *Phys. Rev. Lett*., 54, 2469-2472, 1985;

Tonomura, A., et al, Observation of Aharonov-Bohm effect by electron microscopy. *Phys. Rev. Lett*., 48, 1443-1446, 1982;

_______________, Is magnetic flux quantized in a toroidal ferromagnet? *Phys. Rev. Lett*., 51, 331-334, 1983;

_______________, Evidence for Ahronov-Bohm effect with magnetic field completely shielded from electron wave. *Phys. Rev. Lett*., 56, 792-795, 1986;

_______________ & Callen, E., Phase, electron holography and conclusive demonstration of the Aharonov-Bohm effect. *ONRFE Sci. Bul*., 12, No 3, 93-108, 1987.

[28] Berry, M.V., Exact Aharonov-Bohm wavefunction obtained by applying Dirac's magnetic phase factor. *Eur. J. Phys*., 1, 240-244, 1980;

Peshkin, M., The Ahronov-Bohm effect: why it cannot be eliminated from quantum mechanics. *Physics Reports*, 80, 375-386, 1981;

Olariu, S. & Popescu, I.I., The quantum effects of electromagnetic fluxes. *Rev. Mod. Phys*., 157, 349-436, 1985;

Horvathy, P.A., The Wu-Yang factor and the non-Abelian Aharonov-Bohm experiment. *Phys. Rev*., D33, 407-414, 1986;

Peshkin, M & Tonomura, A., *The Aharonov-Bohm Effect*, Springer-Verlag, New York, 1989.

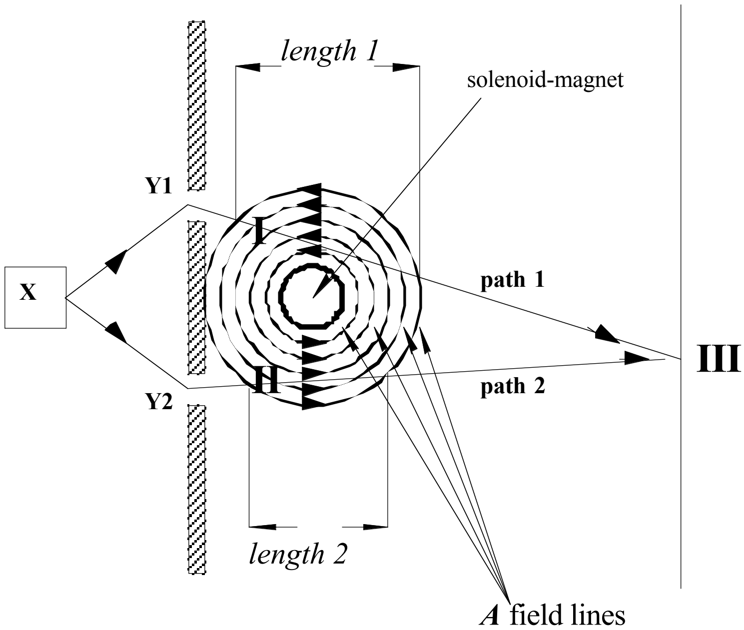

Fig 1. Two-slit diffraction experiment of the Aharonov-Bohm effect. Electrons are produced by a source at X, pass through the slits of a mask at Y1 and Y2, interact with the ***A*** field at locations I and II over lengths l_1 and l_2, respectively, and their diffraction pattern is detected at III. The solenoid-magnet is between the slits and is directed out of the page. The different orientations of the external ***A*** field at the places of interaction I and II of the two paths 1 and 2 are indicated by arrows following the right-hand rule.

That is, if one were to consider the two joined paths of the interferometer as a raceway or a loop and one squeezed the loop tighter and tighter, then nevertheless one cannot in this situation - unlike as in most situations - reduce the interferometer's raceway of paths to a single point. (Another way of saying this is: not all closed curves in a region need have a vanishing line integral, because one exception is a loop with an obstruction.) The reason one cannot reduce the interferometer to a single point is because of the existence in its middle of the solenoid, which is a positive quantity, and acts as an obstruction.

It is the present writer's opinion that the existence of the obstruction changes the situation entirely. *Without* the existence of the solenoid in the interferometer, the loop of the two paths *can be* reduced to a single point and the region occupied by the interferometer is then *simply-connected*. But *with* the existence of the solenoid, the loop of the two paths *cannot be* reduced to a single point and the region occupied by this special interferometer is *multiply-connected*. The Aharonov-Bohm effect only exists in the multiply-connected scenario. But we should note that the Aharonov-Bohm effect is a *physical* effect and simple and multiple connectedness are *mathematical descriptions* of physical situations.

The topology of the physical interferometric situation addressed by Aharonov and Bohm defines the physics of that situation and also the mathematical description of that physics. If that situation *were not* multiply-connected, but simply-connected, then there would be no interesting physical effects to describe. The situation would be described by $U(1)$ electromagnetics and the mapping from one region to another is conventionally one-to-one. However, as the Aharonov-Bohm situation is multiply-connected, there is a two-to-one mapping $(SU(2)/Z_2)$ of the two different regions of the two paths to the single region at III where a measurement is made. Essentially, at III a measurement is made of *the differential histories* of the *two* test waves which traversed the *two* different paths and experienced *two* different forces resulting in two different phase effects.

In conventional, i.e., normal $U(1)$ or simply-connected situations, the fact that a vector field, viewed axially, is pointing in one direction, if penetrated from one direction on one side, and is pointing in *the opposite direction*, if penetrated from the same direction, but *on the other side*, is of no consequence at all - because that field is of $U(1)$ symmetry and can be reduced to a single point. Therefore in most cases which are of $U(1)$ symmetry, we do not need to distinguish between the direction of the vectors of a field from one region to another of that field. However, the Aharonov-Bohm situation is not conventional or simply-connected, but special. (In other words, the physical situation associated with the Aharonov-Bohm effect has a non-trivial topology). It is a *multiply-connected situation* and of $SU(2)/Z_2$ symmetry. Therefore the direction of the $\boldsymbol{A}$ field on the separate paths is of crucial importance, because a test wave traveling along one path will experience an $\boldsymbol{A}$ vectorial component directed *against* its trajectory and thus be retarded, and another test wave traveling along another path will experience an $\boldsymbol{A}$ vectorial component directed *with* its trajectory and thus its speed is boosted. These "retardations" and "boostings" can be measured as phase changes, *but not at the time nor at the locations of, I and II, where their occurrence is separated along the two different paths*, but *later*, and at the *overlap location* of III. It is important to note that if measurements are attempted at locations I and II

in the Fig 1, these effects will not be seen because there is no two-to-one mapping at either I and II and therefore no referents. The locations I and II are both *simply-connected* with the source and therefore only the conventional *U(1)* electromagnetics applies at these locations (with respect to the source). It is only region III which is multiply-connected with the source and at which the histories of what happened to the test particles at I and II can be measured. In order to distinguish the "boosted" $\boldsymbol{A}$ field (because the test wave is traveling "with" its direction) from the "retarded" $\boldsymbol{A}$ field (because the test wave is traveling "against" its direction), we introduce the notation: $\boldsymbol{A}_+$ and $\boldsymbol{A}_-$.
Because of the distinction between the $\boldsymbol{A}$ oriented potential fields at positions I and II - which *are not* measurable and are *vectors or numbers* of *U(1)* symmetry - and the $\boldsymbol{A}$ potential fields at III - which *are* measurable and are *tensors or matrix-valued functions* of (in the present instance) $SU(2)/Z_2 = SO(3)$ symmetry (or higher symmetry) - for reasons of clarity we might introduce a distinguishing notation. In the case of the potentials of *U(1)* symmetry at I and II we might use the lower case, $\boldsymbol{a}_\mu$, $\mu = 0,1,2,3$ and for the potentials of $SU(2)/Z_2 = SO(3)$ at III we might use the upper case $\boldsymbol{A}_\mu$, $\mu = 0,1,2,3$. Similarly, for the electromagnetic field tensor at I and II, we might use the lower case, $\boldsymbol{f}_{\mu\nu}$, and for the electromagnetic field tensor at III, we might use the upper case, $\boldsymbol{F}_{\mu\nu}$. Then the following definitions for the electromagnetic field tensor are:

At locations I and II the Abelian relationship is:

$$f_{\mu\nu}(x) = \partial_\nu a_\mu(x) - \partial_\mu a_\nu(x), \tag{1}$$

where, as is well known, $f_{\mu\nu}$ is Abelian and gauge *invariant*; But at location III the non-Abelian relationship is:

$$F_{\mu\nu} = \partial_\nu A_\mu(x) - \partial_\mu A_\nu(x) - ig_m\left[A_\mu(x), A_\nu(x)\right], \tag{2}$$

where $F_{\mu\nu}$ is gauge *covariant*, g_m is the magnetic charge density and the brackets are commutation brackets. We remark that in the case of non-Abelian groups, such as *SU(2),* the potential field *can carry charge*. It is important to note that if the physical situation changes from *SU(2)* symmetry back to *U(1),* then $F_{\mu\nu} \to f_{\mu\nu}$

Despite the clarification offered by this notation, the notation can also cause confusion, because in the present literature, the electromagnetic field tensor is *always* referred to as $\boldsymbol{F}$, whether $\boldsymbol{F}$ is defined with respect to *U(1)* or

SU(2) or other symmetry situations. Therefore, although we prefer this notation, we shall not proceed with it. However, it is important to note that the ***A*** field in the *U(1)* situation is a *vector or a number*, but in the *SU(2)* or nonAbelian situation, it is a *tensor or a matrix-valued function*.

We referred to the physical situation of the Aharonov-Bohm effect as an interferometer around an obstruction and it is 2-dimensional. It is important to note that the situation is not provided by a toroid, although a toroid is also a physical situation with an obstruction and the fields existing on a toroid are also of *SU(2)* symmetry. However, the toroid provides a two-to-one mapping of fields in not only the *x* and *y* dimensions but also in the *z* dimension, and *without* the need of an electromagnetic field pointing in two directions + and -. The physical situation of the Aharonov-Bohm effect is defined only in the *x* and *y* dimensions (there is no *z* dimension) and in order to be of $SU(2)/Z_2$ symmetry *requires* a field to be oriented differentially on the separate paths. If the differential field is removed from the Aharonov-Bohm situation, then that situation reverts to a simple interferometric raceway which can be reduced to a single point and with no interesting physics.

How does the topology of the situation affect the explanation of an effect? A typical previous explanation[29] of the Aharonov-Bohm effect commences with the Lorentz force law:

$$\mathbf{F} = e\boldsymbol{E} + e\boldsymbol{v}\times\boldsymbol{B} \tag{3}$$

The electric field, $\boldsymbol{E}$, and the magnetic flux density, $\boldsymbol{B}$, are essentially confined to the inside of the solenoid and therefore cannot interact with the test electrons. An argument is developed by defining the $\boldsymbol{E}$ and $\boldsymbol{B}$ fields in terms of the $\boldsymbol{A}$ and ϕ potentials:

$$\boldsymbol{E} = -\frac{\partial\boldsymbol{A}}{\partial t} - \nabla\phi, \qquad \boldsymbol{B} = \nabla\times\boldsymbol{A} \; . \tag{4}$$

Now we can note that these conventional *U(1)* definitions of $\boldsymbol{E}$ and $\boldsymbol{B}$ can be expanded to *SU(2)* forms:

$$\boldsymbol{E} = -(\nabla\times\boldsymbol{A}) - \frac{\partial\boldsymbol{A}}{\partial t} - \nabla\phi, \qquad \boldsymbol{B} = (\nabla\times\boldsymbol{A}) - \frac{\partial\boldsymbol{A}}{\partial t} - \nabla\phi \; . \tag{5}$$

[29] Ryder, L.H., *Quantum Field Theory*, 2nd edition, Cambridge U. Press, 1996.

Furthermore, the *U(1)* Lorentz force law, Eq 3, can hardly apply in this situation because the solenoid is electrically neutral to the test electrons and therefore $\boldsymbol{E} = 0$ along the two paths. Using the definition of $\boldsymbol{B}$ in Eq 5, the force law in this *SU(2)* situation is:

$$\boldsymbol{F} = e\boldsymbol{E} + e\boldsymbol{v}\times\boldsymbol{B} = e\left(-(\nabla\times\boldsymbol{A}) - \frac{\partial \boldsymbol{A}}{\partial t} - \nabla\phi\right) + e\boldsymbol{v}\times\left((\nabla\times\boldsymbol{A}) - \frac{\partial \boldsymbol{A}}{\partial t} - \nabla\phi\right), \tag{6}$$

but we should note that Eqs 3 and 4 are *still valid* for the conventional theory of electromagnetism based on the *U(1)* symmetry Maxwell's equations provided in Table 1 and associated with the group *U(1)* algebra. They are *invalid* for the theory based on the modified *SU(2)* symmetry equations also provided in Table 1 and associated with the group *SU(2)* algebra.

The typical explanation of the Aharonov-Bohm effect continues with the observation that a phase difference, δ, between the two test electrons is caused by the presence of the solenoid:

$$\begin{aligned}\Delta\delta = \Delta\alpha_1 - \Delta\alpha_2 = \frac{e}{\hbar}\left(\int_{l_2} A \bullet dl_2 - \int_{l_1} A \bullet dl_1\right) = \\ \frac{e}{\hbar}\int_{l_2 - l_1} \nabla\times\boldsymbol{A} \bullet d\boldsymbol{S} = \frac{e}{\hbar}\int \boldsymbol{B} \bullet d\boldsymbol{S} = \frac{e}{\hbar}\varphi_M\end{aligned} \tag{7}$$

where $\Delta\alpha_1$ and $\Delta\alpha_2$ are the changes in the wave function for the electrons over paths 1 and 2, $\boldsymbol{S}$ is the surface area and φ_M is the *magnetic flux* defined:

$$\varphi_M = \iint \boldsymbol{A}_\mu(x)dx^\mu = \iint \boldsymbol{F}_{\mu\nu}d\sigma^{\mu\nu} \ . \tag{8}$$

Now, we can extend this explanation further, by observing that the local phase change at III of the wavefunction of a test wave or particle is given by:

$$\Phi = \exp\left[ig_m \iint A_\mu(x)dx^\mu\right] = \exp\left[ig_m\varphi_M\right]. \tag{9}$$

Φ, which is proportional to the magnetic flux, φ_M, is known as the *phase factor* and is *gauge covariant*. Furthermore, Φ, this phase factor measured at position III is the *holonomy* of the *connection*, $\boldsymbol{A}_\mu$; and g_m is the *SU(2) magnetic charge density*.

We next observe that φ_M is in units of volt-seconds *(V.s)* or *kg.m²/(A.s²)* = *J/A*. From Eq 7 it can be seen that $\Delta\delta$ and the phase factor, Φ, are dimensionless. Therefore we can make the prediction that if the magnetic flux, φ_M, is known and the phase factor, Φ, is measured (as in the Aharonov-Bohm situation), *the magnetic charge density, g_m, can be found by the relation*

$$g_m = \ln(\Phi)/(i\varphi_M). \tag{10}$$

Continuing the explanation: as was noted above, $\nabla \times \boldsymbol{A} = 0$ outside the solenoid and the situation must be redefined in the following way. An electron on path 1 will interact with the $\boldsymbol{A}$ field oriented in the positive direction. Conversely, an electron on path 2 will interact with the $\boldsymbol{A}$ field oriented in the negative direction. Furthermore, the $\boldsymbol{B}$ field can be defined with respect to a local stationary component $\boldsymbol{B}_1$ which is confined to the solenoid and a component $\boldsymbol{B}_2$ which is either a standing wave or propagates:

$$\begin{aligned} \boldsymbol{B} &= \boldsymbol{B}_1 + \boldsymbol{B}_2, \\ \boldsymbol{B}_1 &= \nabla \times \boldsymbol{A}, \\ \boldsymbol{B}_2 &= -\frac{\partial \boldsymbol{A}}{\partial t} - \nabla\phi. \end{aligned} \tag{11}$$

The magnetic flux density, $\boldsymbol{B}_1$, is the confined component associated with $U(1)\times SU(2)$ symmetry and $\boldsymbol{B}_2$ is the propagating or standing wave component associated *only* with *SU(2)* symmetry. In a *U(1)* symmetry situation, $\boldsymbol{B}_1$ = components of the field associated with *U(1)* symmetry, and $\boldsymbol{B}_2 = 0$.

The electrons traveling on paths 1 and 2 require different times to reach III from X, due to the different distances and the opposing directions of the potential $\boldsymbol{A}$ along the paths l_1 and l_2. Here we only address the effect of the opposing directions of the potential $\boldsymbol{A}$, i.e., the distances traveled are identical over the two paths. The change in the phase difference due to the presence of the $\boldsymbol{A}$ potential is then:

$$\Delta\delta = \Delta\alpha_1 - \Delta\alpha_2 =$$

$$\frac{e}{\hbar}\left[\int_{l_2}\left(-\frac{\partial \boldsymbol{A}_+}{\partial t} - \nabla\phi_+\right)dl_2 - \int_{l_1}\left(-\frac{\partial \boldsymbol{A}_-}{\partial t} - \nabla\phi_-\right)dl_1\right] \bullet d\boldsymbol{S} = \qquad (12)$$

$$\frac{e}{\hbar}\int \boldsymbol{B}_2 \bullet d\boldsymbol{S} = \frac{e}{\hbar}\varphi_M .$$

There is no flux density $\boldsymbol{B}_1$ in this equation since this equation describes events outside the solenoid, but only the flux density $\boldsymbol{B}_2$ associated with group *SU(2)* symmetry; and the "+" and "-" indicate the direction of the ***A*** field encountered by the test electrons - as discussed above.

We note that the phase effect is dependent on $\boldsymbol{B}_2$ and $\boldsymbol{B}_1$, but not on $\boldsymbol{B}_1$ alone. Previous treatments found no convincing argument around the fact that whereas the Aharonov-Bohm effect depends on an interaction with the ***A*** field outside the solenoid, ***B***, defined in *U(1)* electromagnetism as $\boldsymbol{B} = \nabla \times \boldsymbol{A}$, is zero at that point of interaction. However, when ***A*** is defined in terms associated with an *SU(2)* situation, that is not the case as we have seen.

We depart from former treatments in other ways. Commencing with a *correct* observation that the Aharonov-Bohm effect depends on the topology of the experimental situation and that the situation is not simply-connected, a former treatment then erroneously seeks an explanation of the effect in the connectedness of the *U(1)* gauge symmetry of conventional electromagnetism, but for which (1) the potentials are ambiguously defined, (the *U(1)* ***A*** field is gauge invariant) and (2) in *U(1)* symmetry $\nabla \times \boldsymbol{A} = 0$ outside the solenoid.

Furthermore, whereas a former treatment again makes a *correct* observation that the non-Abelian group, *SU(2),* is simply-connected and that the situation is governed by a multiply-connected topology, the author fails to observe that the non-Abelian group *SU(2)* defined over the integers modulo 2, $SU(2)/Z_2$, is, in fact, multiply-connected. Because of the two paths around the solenoid it is this group which describes the topology underlying the Aharonov-Bohm effect[30]. $SU(2)/Z_2 \cong SO(3)$ is obtained from the group *SU(2)* by identifying

---

[30] Barrett, T.W., Electromagnetic phenomena not explained by Maxwell's equations. pp. 6-86 in Lakhtakia, A. (Ed.) *Essays on the Formal Aspects of Maxwell's Theory*, World Scientific, Singapore, 1993;

______, Sagnac effect. pp. 278-313 in Barrett, T.W. & Grimes, D.M., (Ed.s) *Advanced Electromagnetism: Foundations, Theory, Applications*, World Scientific, Singapore, 1995;

pairs of elements with opposite signs. The $\Delta\delta$ measured at location III in Fig. 1 is derived from a *single* path in *SO(3)*[31] because the *two* paths through locations I and II in *SU(2)* are regarded as a *single* path in *SO(3)*. This path in $SU(2)/Z_2 \cong SO(3)$ cannot be shrunk to a single point by any continuous deformation and therefore adequately describes the multiple-connectedness of the Aharonov-Bohm situation. Because the former treatment failed to note the multiple connectedness of the $SU(2)/Z_2$ description of the Aharonov-Bohm situation, it *incorrectly* fell back on a *U(1)* symmetry description.

Now back to the main point of this excursion to the Aharonov-Bohm effect: the reader will note that the author appealed to topological arguments to support the main points of his argument. Underpinning the *U(1)* Maxwell theory is an Abelian algebra; underpinning the *SU(2)* theory is a non-Abelian algebra. The algebras specify the form of the equations of motion. However, whether one or the other algebra can be (validly) used can only be determined by topological considerations.

5 Summary

We have attempted to show the fundamental explanatory nature of the topological description of solitons, instantons and the Aharonov-Bohm effect – and hence electromagnetism. In the case of electromagnetism we have shown elsewhere that, given a Yang-Mills description, electromagnetism can, and should be extended, in accordance with the topology with which the electromagnetic fields are associated.

This approach has further implications. If the conventional theory of electromagnetism, i.e., "Maxwell's theory", which is of *U(1)* symmetry form, is but the simplest *local* theory of electromagnetism, then those pursuing a unified field theory may wish to consider as a candidate for that unification, not only the simple local theory, but other electromagnetic fields of group symmetry higher than *U(1)*. Other such forms include symplectic gauge fields of higher group symmetry, e.g., *SU(2)* and above.

______, The toroid antenna as a conditioner of electromagnetic fields into (low energy) gauge fields. *Speculations in Science and Technology*, 21(4), 291-320, 1998.

[31] See appendix II on : ***O(n)* Group Algebra**

Appendix I

***SU(n)* Group Algebra**

Unitary transformations, *U(n)*, leave the modulus squared of a complex wavefunction invariant. The elements of a *U(n)* group are represented by *n* x *n* unitary matrices with a determinant equal to ± 1 . Special unitary matrices are elements of unitary matrices which leave the determinant equal to +1. There are n^2 - 1 independent parameters. *SU(n)* is a subgroup of *U(n)* for which the determinant equals +1.

***SL(2,C)* Group Algebra**

The special linear group of 2×2 matrices of determinant 1 with complex entries is *SL(2,C)*.

***SU(2)* Group Algebra**

SU(2) is a subgroup of *SL(2,C)*. The are 2^2 - 1 = 3 independent parameters for the special unitary group *SU(2)* of 2×2 matrices. *SU(2)* is a Lie algebra such that for the angular momentum generators, J_i, the commutation relations are $[J_i, J_j] = i\varepsilon_{ijk}J_k, i, j, k = 1,2,3$. The *SU(2)* group describes rotation in 3-dimensional space with 2 parameters (see below). There is a well-known *SU(2)* matrix relating the Euler angles of *O(3)* and the complex parameters of *SU(2)* are:

$$\begin{matrix} \cos\left[\frac{\beta}{2}\right]\exp\left[\frac{i(\alpha+\gamma)}{2}\right] & \sin\left[\frac{\beta}{2}\right]\exp\left[\frac{-(\alpha-\gamma)}{2}\right] \\ -\sin\left[\frac{\beta}{2}\right]\exp\left[\frac{i(\alpha-\gamma)}{2}\right] & \cos\left[\frac{\beta}{2}\right]\exp\left[\frac{-i(\alpha+\gamma)}{2}\right] \end{matrix}$$

where α,β,γ are the Euler angles. It is also well known that a homomorphism exists between *O(3)* and *SU(2),* and the elements of *SU(2)* can be associated with rotations in *O(3);* and *SU(2)* is the *covering group* of *O(3)*. Therefore, it is easy to show that *SU(2)* can be obtained from *O(3)*. These *SU(2)* transformations define the relations between the Euler angles of group *O(3)* with the parameters of *SU(2)*. For comparison with the above, if the rotation matrix $R(\alpha,\beta,\gamma)$ in *O(3)* is represented as:

$$\begin{pmatrix} \cos[\alpha]\cos[\beta]\cos[\gamma]-\sin[\alpha]\sin[\gamma] & \sin[\alpha]\cos[\beta]\cos[\gamma]+\cos[\alpha]\sin[\gamma] & -\sin[\beta]\cos[\gamma] \\ -\cos[\alpha]\cos[\beta]\sin[\gamma]-\sin[\alpha]\cos[\gamma] & -\sin[\alpha]\cos[\beta]\sin[\gamma]+\cos[\alpha]\cos[\gamma] & \sin[\beta]\sin[\gamma] \\ \cos[\alpha]\sin[\beta] & \sin[\alpha]\sin[\beta] & \cos[\beta] \end{pmatrix}$$

then the orthogonal rotations about the coordinate axes are:

$$R_1(\alpha) = \begin{pmatrix} \cos[\alpha] & \sin[\alpha] & 0 \\ -\sin[\alpha] & \cos[\alpha] & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad R_2(\beta) = \begin{pmatrix} \cos[\beta] & 0 & -\sin[\beta] \\ 0 & 1 & 0 \\ \sin[\beta] & 0 & \cos[\beta] \end{pmatrix}$$

$$R_3(\gamma) = \begin{pmatrix} \cos[\gamma] & \sin[\gamma] & 0 \\ -\sin[\gamma] & \cos[\gamma] & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

An isotropic parameter, ϖ, can be defined:

$$\varpi = \frac{x - iy}{z},$$

where x, y, z are the spatial coordinates. If ϖ is written as the quotient of μ_1 and μ_2, or the homogeneous coordinates of the bilinear transformation, then:

$$\left|\mu_1^{'}\mu_2^{'}\right\rangle = \begin{bmatrix} \cos\left[\frac{\beta}{2}\right]\exp\left[\frac{i(\alpha+\gamma)}{2}\right] & \sin\left[\frac{\beta}{2}\right]\exp\left[\frac{-(\alpha-\gamma)}{2}\right] \\ -\sin\left[\frac{\beta}{2}\right]\exp\left[\frac{i(\alpha-\gamma)}{2}\right] & \cos\left[\frac{\beta}{2}\right]\exp\left[\frac{-i(\alpha+\gamma)}{2}\right] \end{bmatrix} \left|\mu_1\mu_2\right\rangle$$

which is the relation between the Euler angles of *O(3)* and the complex parameters of *SU(2)*. However, there is not a unique one-to-one relation, for 2 rotations in *O(3)* correspond to 1 direction in *SU(2)*. There is thus a many-to-one or homomorhpism between *O(3)* and *SU(2)*.

In the case of a complex 2-dimensional vector (u,v):

$$\begin{pmatrix} u' \\ \\ v' \end{pmatrix} = \begin{pmatrix} \cos\left[\frac{\beta}{2}\right]\exp\left[\frac{i(\alpha+\gamma)}{2}\right] & \sin\left[\frac{\beta}{2}\right]\exp\left[\frac{-(\alpha-\gamma)}{2}\right] \\ -\sin\left[\frac{\beta}{2}\right]\exp\left[\frac{i(\alpha-\gamma)}{2}\right] & \cos\left[\frac{\beta}{2}\right]\exp\left[\frac{-i(\alpha+\gamma)}{2}\right] \end{pmatrix} \begin{pmatrix} u \\ \\ v \end{pmatrix}$$

If we define:

$$a = \cos\left[\frac{\beta}{2}\right]\exp\left[\frac{i(\alpha+\gamma)}{2}\right],$$

$$b = \sin\left[\frac{\beta}{2}\right]\exp\left[\frac{-(\alpha-\gamma)}{2}\right]$$

then

$$\left|\mu_1^{'}\mu_2^{'}\right\rangle = \begin{bmatrix} a & b \\ -b^* & a^* \end{bmatrix} \left|\mu_1\mu_2\right\rangle,$$

where

$$\begin{bmatrix} a & b \\ -b^* & a^* \end{bmatrix}$$

are the well-known *SU(2)* transformation rules. Defining: $c = -b^*$ and $d = a^*$, we have the determinant:

$$ad - bc = 1 \text{ or } aa^* - b(-b^*) = 1 .$$

Defining the (x,y,z) coordinates with respect to a complex 2-dimensioanl vector (u,v) as:

$$x = \frac{1}{2}\left(u^2 - v^2\right), \quad y = \frac{1}{2i}\left(u^2 + v^2\right), \quad z = uv$$

then *SU(2)* transformations leave the squared distance $x^2 + y^2 + z^2$ invariant.

Every element of *SU(2)* can be written as:

$$\begin{bmatrix} a & b \\ -b^* & a^* \end{bmatrix} \quad |a|^2 + |b|^2 = 1.$$

Defining:

$$a = y_1 - iy_2, \qquad b = y_3 - iy_4,$$

the parameters y_1, y_2, y_3, y_4 indicate positions in *SU(2)* with the constraint:

$$y_1^2 + y_2^2 + y_3^2 + y_4^2 = 1,$$

which indicates that the group *SU(2)* is a 3-dimensional unit sphere in the 4-dimensional y-space. This means that any closed curve on that sphere can be shrunk to a point. In other words, *SU(2)* is *simply-connected*.

It is important to note that *SU(2)* is the quantum mechanical "rotation group".

Homomorphism of O(3) and SU(2)

There is an important relationship between *O(3)* and *SU(2)*. The elements of *SU(2)* are associated with rotations in 3-dimensional space. To make this relationship explicit, new coordinates are defined:

$$x = \frac{1}{2}\left(u^2 - v^2\right); \quad y = \frac{1}{2i}\left(u^2 + v^2\right); \quad z = uv.$$

Explicitly, the *SU(2)* transformations leave the squared 3-dimensional distance $x^2 + y^2 + z^2$ invariant, and invariance which relates 3-dimensional rotations to elements of *SU(2)*. If a,b of the elements of *SU(2)* are defined:

$$a = \cos\frac{\beta}{2}\exp\frac{i(\alpha+\gamma)}{2}, \quad b = \sin\frac{\beta}{2}\exp\frac{-i(\alpha-\gamma)}{2},$$

then the general rotation matrix $R(\alpha,\beta,\gamma)$ can be associated with the *SU(2)* matrix:

$$\begin{pmatrix} \cos\frac{\beta}{2}\exp\frac{i(\alpha+\gamma)}{2} & \sin\frac{\beta}{2}\exp\frac{-i(\alpha-\gamma)}{2} \\ -\sin\frac{\beta}{2}\exp\frac{i(\alpha-\gamma)}{2} & \cos\frac{\beta}{2}\exp\frac{-i(\alpha+\gamma)}{2} \end{pmatrix}$$

by means of the Euler angles.

It is important to indicate that this matrix does not give a unique one-to-one relationship between the general rotation matrix $R(\alpha,\beta,\gamma)$ and the *SU(2)* group. This can be seen if (i) we let $\alpha = 0,\ \beta = 0,\ \gamma = 0$, which gives the matrix:

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

and (ii) $\alpha = 0,\ \beta = 2\pi, \gamma = 0$, which gives the matrix:

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Both matrices define zero rotation in 3-dimensional space, so we see that this zero rotation in 3-dimensional space corresponds to 2 different *SU(2)* elements depending on the value of β. There is thus a *homomorphism*, or many-to-one mapping relationship between *O(3)* and *SU(2)* – where "many" is 2 in this case - but not a one-to-one mapping.

***SO(2)* Group Algebra**

The collection of matrices in Euclidean 2-dimensional space (the plane) which are orthogonal and moreover for which the determinant is +1 is a subgroup of *O(2)*. *SO(2)* is the special orthogonal group in two variables.

The rotations in the plane is represented by the *SO(2)* group:

$$R(\theta) = \begin{pmatrix} \cos[\theta] & -\sin[\theta] \\ \sin[\theta] & \cos[\theta] \end{pmatrix} \theta \in \Re$$

where $R(\theta)R(\gamma) = R(\theta+\gamma)$. S^1, or the unit circle in the complex plane with multiplication as the group operation is an *SO(2)* group.

Appendix II

***O(n)* Group Algebra**

The orthogonal group, *O(n),* is the group of transformation (including inversion) in an n-dimensional Euclidean space. The elements of *O(n)* are represented by $n \times n$ real orthogonal matrices with $n(n-1)/2$ real parameters satisfying $AA^t = 1$.

***O(3)* Group Algebra**

The orthogonal group, *O(3),* is the well-known and familiar group of transformations (including inversions) in 3-dimensional space with 3 parameters, those parameters being the rotation or Euler angles (α, β, γ). *O(3)* leaves the distance squared, $x^2 + y^2 + z^2$, invariant.

***SO(3)* Group Algebra**

The collection of matrices in Euclidean 3-dimensional space which are orthogonal and moreover for which the determinant is +1 is a subgroup of *O(3)*. *SO(3)* is the special orthogonal group in three variables and defines rotations in 3-dimensional space.

Rotation of the Riemann sphere is a rotation in $\Re^3$ or $\xi - \eta - \zeta$ space, for which

$$\xi^2 + \eta^2 + \zeta^2 = 1, \quad \xi = \frac{2x}{|z|^2 + 1}, \quad \eta = \frac{2y}{|z|^2 + 1}, \quad \zeta = \frac{|z|^2 - 1}{|z|^2 + 1}, \quad z = x + iy = \frac{\xi + i\eta}{1 - \zeta}.$$

$$U_\xi(\alpha) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} e^{i\alpha/2} & 0 \\ 0 & e^{-i\alpha/2} \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} \cos\alpha/2 & i\sin\alpha/2 \\ i\sin\alpha/2 & \cos\alpha/2 \end{pmatrix}$$

or $\pm U_\xi(\alpha) \to R_1(\alpha)$,

$$U_\eta(\beta) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix}\begin{pmatrix} e^{i\beta/2} & 0 \\ 0 & e^{-i\beta/2} \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} = \begin{pmatrix} \cos\beta/2 & -\sin\beta/2 \\ \sin\beta/2 & \cos\beta/2 \end{pmatrix}$$

or $\pm U_\eta(\beta) \to R_2(\beta)$,

$$U_\zeta(\gamma) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} e^{i\gamma/2} & 0 \\ 0 & e^{-i\gamma/2} \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \cos\gamma/2 & -\sin\gamma/2 \\ \sin\gamma/2 & \cos\gamma/2 \end{pmatrix}$$

or $\pm U_\zeta(\gamma) \to R_3(\gamma)$.

which are mappings from *SL(2,C)* to *SO(3)*. However, as the *SL(2,C)* are all unitary with determinant equal to +1, they are of the *SU(2)* group. Therefore *SU(2)* is the covering group of *SO(3)*. Furthermore, *SU(2)* is simply connected and *SO(3)* is multiply connected.

A simplification of the above is:

$$U_\xi(\alpha) = e^{i(\alpha/2)\sigma_1}, \quad U_\eta(\beta) = e^{-i(\beta/2)\sigma_2}, \quad U_\zeta(\gamma) = e^{i(\gamma/2)\sigma_3},$$

where $\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$, $\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

$\sigma_1, \sigma_2, \sigma_3$ are the Pauli matrices.

(Manuscrit reçu le 15 septembre 2000)