INTRODUCTION TO GAUGE THEORY

Tai Tsun Wu[+]
CERN, Geneva

1. Introduction

Professor Bleuler has kindly asked me to give an elementary overall view of gauge theory from a physicists's point of view. Gauge theory consists of two types:

a) Abelian gauge theory, which is essentially electromagnetism of Faraday and Maxwell, and

b) Non-Abelian gauge theory, or Yang-Mills theory (20).

Both types can be studied from three points of view:

i) Differential formelism,
ii) Integral formelism, and
iii) Global formalism.

It is in the global formalism where fiber bundles play a central role. This classification is summarized in Fig. 1.

2. Differential Formalism for Abelian Gauge Theory

The concept of a gauge transformation was first introduced by Weyl (15) in his attempt to unify gravitation and electromagnetism. By gauge transformation he means a scale change that depends on the space-time point x^μ. Consider two neighbouring space-time points x^μ and $x^\mu+dx^\mu$. In going from x^μ to $x^\mu+dx^\mu$ the gauge transformation gives a factor $1+s_\mu dx^\mu$ say. Since a function f becomes $f+\frac{\partial f}{\partial x^\mu}dx^\mu$, combination with the scale change gives

$$f + dx^\mu\left(\frac{\partial}{\partial x^\mu} + s_\mu\right) f \tag{2.1}$$

Weyl attempted to identify s_μ with the electromagnetic vector potential A_μ, but he was not successful. It is remarkable that this attempt of Weyl was made before the advent of quantum mechanics.

As we know now, there is a missing factor of i. As studied by Fock (8), London (10) and Weyl (16) himself, what is needed is a phase change instead of a scale change. Instead of identifying s_μ with A_μ the paper identification is

[+] On sabbatical leave from Harvard University. Work supported in part by U.S. Department of Energy Contract No. EY-76-S-02-3227.

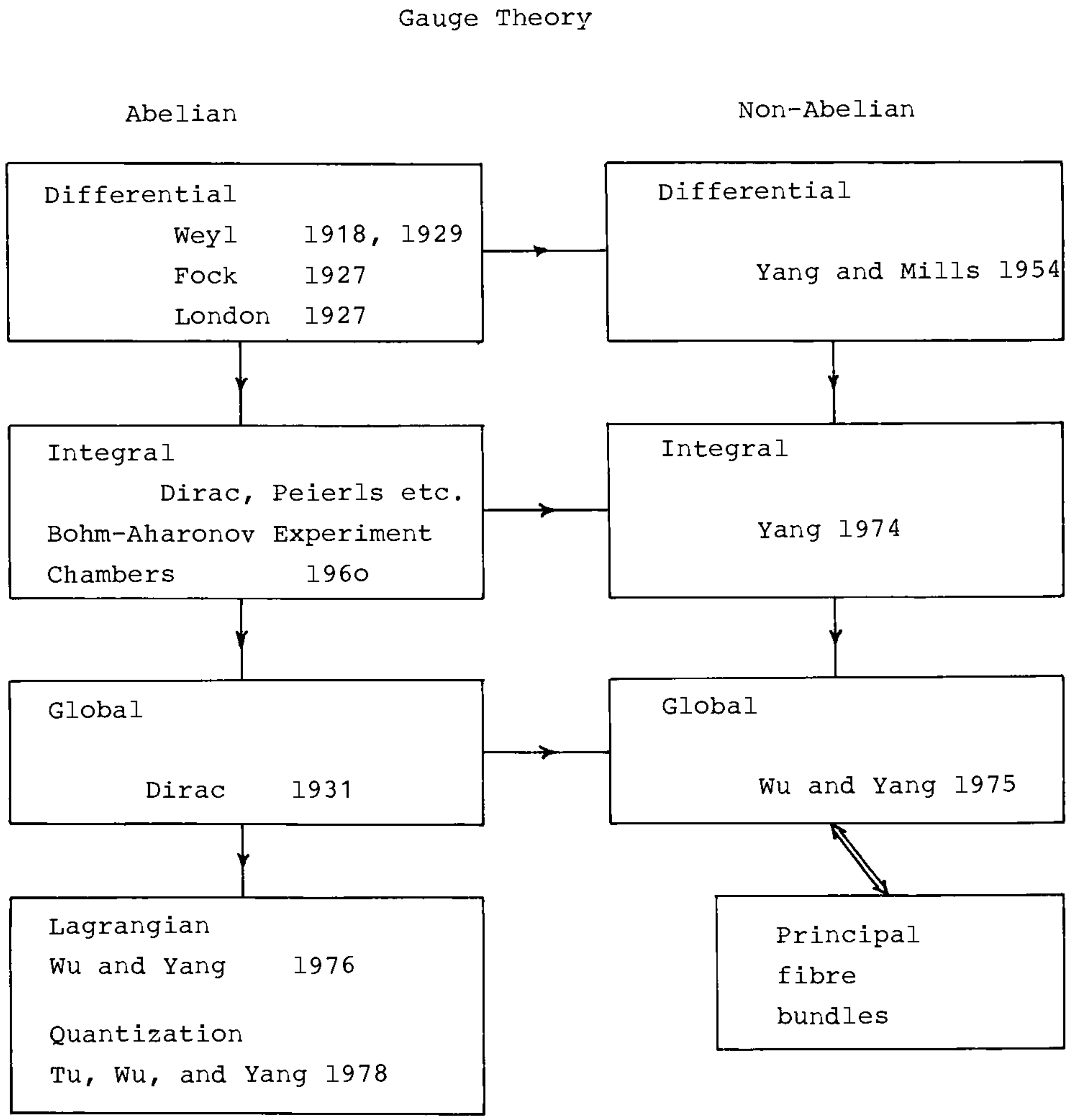

Fig. 1 Schematic outline of gauge theory

$$s_\mu = -i \frac{e}{\hbar c} A_\mu \quad . \tag{2.2}$$

Strictly speaking, the term gauge transformation is no longer appropriate but the term sticks.

3. Differential Formalism for Non-Abelian Gauge Theory

The next conceptual advance came quarter of a century later, when Yang and Mills (20) generalized space-time-dependent gauge transformation to the non-Abelian case SU(2) . It was known at that time that, in the absence of electromagnetic interaction, isotopic spin is conserved. In other words, strong interactions are invariant under isospin rotation. This means that there is freedom of convention to choose, for example, what we call a proton rather than a neutron or a linear combination. Yang and Mills pointed out that there is an independent freedom for each space-time point, and generalized

$$\frac{\partial}{\partial x^{\mu}} - \frac{ie}{\hbar c} A_{\mu} \tag{3.1}$$

of electromagnetism to

$$\frac{\partial}{\partial x^{\mu}} - \frac{i\varepsilon}{\hbar c} B_{\mu} \ , \tag{3.2}$$

which operators on, for example, the two-component wave function

$$\psi = \binom{\psi_p}{\psi_n} \tag{3.3}$$

of the proton and neutron. Here B_μ has isospin indices and is the Yang-Mills field.

4. Integral Formalism for Abelian Gauge Theory

Maxwell's equations are partial differential equations for $f_{\mu\nu}$, which consists of $\vec{E}$ and $\vec{H}$. In classical mechanics, the motion of an electron in an electromagnetic field is governed by the Lorentz force, which is expressable in terms of $f_{\mu\nu}$ and the electron world line. Therefore, classically the introduction of the vector potential A_μ is entirely for mathematical convenience. The situation is quite different in quantum mechanics, where the Dirac equation for the electron in an electromagnetic field involves A_μ explicitly. Therefore, here A_μ plays an essential role.

If $f_{\mu\nu}=0$ in the entire spac-time continuum, then the obvious choice for A_μ is $A_\mu=0$. More generally, if $f_{\mu\nu}=0$ in a simply connected space-time region, then we can choose $A_\mu=0$ in this region. The important point made by Aharonov and Bohm (1), also previously by Ehrenberg and Siday (7), is that if $f_{\mu\nu}=0$ in a doubly connected (or more generally a multiply connected) space-time region, then it may not be consistant to choose $A_\mu=0$.

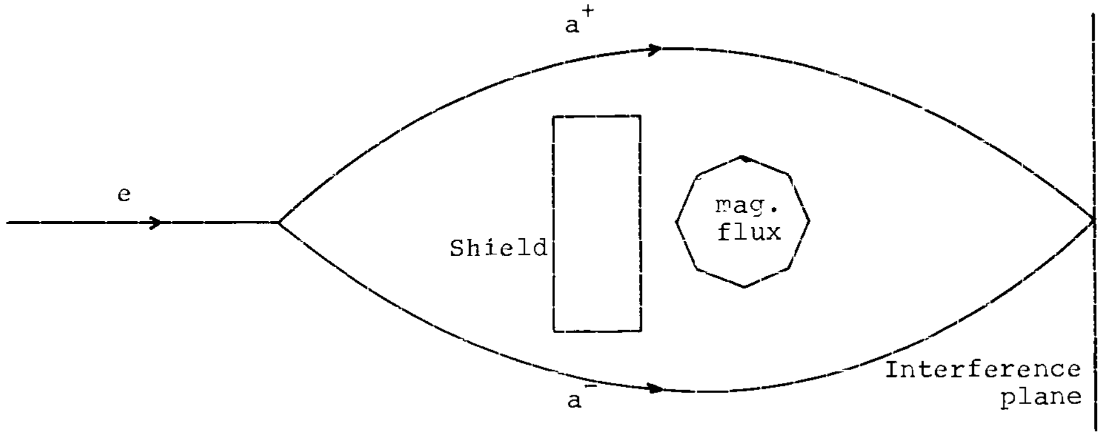

Fig.2 Bohm-Aharonov experiment

The Bohm-Aharonov experiment (1) first carried out by Chambers (3), is shown schematically in Fig. 2. Suppose the static magnetic field is confined in a cylindrical region in space, with the flux return path far away. Scattering experiment with an incident electron beam is then carried out such that the electron does not enter the region of magnetic flux. Interference pattern is then observed on the screen. In the absence of the magnetic flux, let $a^+(a^-)$ be the electron wave function at the screen when a shielding block is introduced to prevent the electron from reaching the screen via the path below (above) the region of magnetic flux. Then, when the magnetic flux is turned off, the amplitude at the screen is

$$a^+ + a^- \,. \tag{4.1}$$

If the magnetic flux is turned on, the interference pattern is determined by

$$a^+ + a^- \exp\left(\frac{ie}{\hbar c}\,\Omega\right) \tag{4.2}$$

where Ω is the total magnetic flux in the cylinder. Since

$$\Omega = \oint A_\mu dx^\mu \tag{4.3}$$

with the line integral taken outside the cylinder, interference pattern depends on the phase factor

$$\exp \quad \frac{ie}{\hbar c} \oint A_\mu dx^\mu \tag{4.4}$$

A more detailed consideration (17) shows that $f_{\mu\nu}$ underdescribes electromagnetism (ie. , different physical situations in a space--time region may have the same $f_{\mu\nu}$), and $\oint A_\mu dx^\mu$ overdescribes electromagnetism (i.e., different $\oint A_\mu dx^\mu$ in a space-time region may describe the same physical situation). The phase factor $e^{\frac{ie}{\hbar c} \oint A_\mu dx^\mu}$ is just right to describe electromagnetism. Note that this phase factor gives $f_{\mu\nu}$ in the limit where the line of integration is an infinitesimal loop.

For theoretical treatments, this phase factor (4.4), which is invariant under the gauge transformation

$$A_\mu \rightarrow A_\mu + \frac{\hbar c}{e} \frac{\partial \alpha}{\partial x^\mu} , \tag{4.5}$$

is less easy to use than the concept of a path-dependent phase factor

$$e^{\frac{ie}{\hbar c} \int_P^Q A_\mu dx^\mu} . \tag{4.6}$$

Such a path-dependent (or non-integrable) phase factor is not gauge invariant. Instead under (4.5),

$$e^{\frac{ie}{\hbar c} \int_P^Q A_\mu dx^\mu} \rightarrow e^{i\alpha(Q)} \, e^{\frac{ie}{\hbar c} \int_P^Q A_\mu dx^\mu} \, e^{-i\alpha(P)} \tag{4.7}$$

Electromagnetism is thus the gauge-invariant manifestation of a non-integrable phase factor (2, 11, 17).

5. Integral Formalism for Non-Abelian Gauge Theory

The generalization of integral formalism to Yang-Mills theory was carried out by Yang (19) in 1974. The starting point of this consideration is the non-integrable phase factor

$$\left(e^{\int_P^Q \frac{-\varepsilon}{\hbar c} b_\mu^k X_k dx^\mu} \right)_{\text{ordered}} \tag{5.1}$$

where b_μ^k is the Yang-Mills field (20) and X_k are the generators of the underlying Lie group. If the group is U_1, (5.1) reduces to (4.6). The Yang-Mills field is the gauge-invariant manifestation of this non-integrable phase factor.
For the properties of (5.1), see Yang (19).

6. Global Formalism for Abelian Gauge Theory

The early work of Dirac (6) in 1931 leads to the recognition that in general A_μ and the phase factor (4.4) can only be properly defined in each of several overlapping space-time regions. These overlapping space-time regions are like coordinate neighbourhoods and their union is the entire space-time region of interest. Analogous to the static electric field

$$\vec{E} = \frac{e}{r^2}\hat{r} \tag{6.1}$$

of a point charge, Dirac (6) considered the static magnetic field

$$\vec{H} = \frac{g}{r^2}\hat{r} \tag{6.2}$$

of a magnetic monopole. The important point is that this magnetic field (6.2) cannot be obtained from a vector potential A_μ which is regular for $r>0$. For example, the choices in spherical coordinate

$$\vec{A}_a = \frac{g}{r\sin\theta}(1-\cos\theta)\hat{\phi} \tag{6.3}$$

$$\vec{A}_b = -\frac{g}{r\sin\theta}(1+\cos\theta)\hat{\phi} \tag{6.4}$$

are respectively singular at $\theta=\pi$ and $\theta=0$. However, if we define two regions R_a and R_b by

$$R_a: \quad 0\leqslant\theta<\frac{\pi}{2}+\delta\ , \ r>0 \qquad 0\leqslant\phi<2\pi\ ,$$

and

$$R_b: \ \frac{\pi}{2} - \delta < \theta \leqslant 2\pi, \ r>0 \qquad 0 \leqslant \phi < 2\pi \tag{6.5}$$

with $0<\delta\leqslant\frac{\pi}{2}$, then we get a global description of the electromagnetic field of a magnetic monopole by using $\vec{A}_a$ in R_a and $\vec{A}_b$ in R_b (17). In the overlap $R_a \cap R_b$, the gauge transformation, according to (4.7), is given by

$$e^{i\alpha} = e^{\frac{2ieg}{\hbar c}\phi} \tag{6.6}$$

This $e^{i\alpha}$ is single-valued, and hence well-defined, if and only if

$$\frac{2eg}{\hbar c} = \text{integer} \ . \tag{6.7}$$

This is the Dirac quantization condition (6).

7. Global Formalism for Non-Abelian Gauge Theory

The necessity of dealing with several overlapping space-time regions has nothing to do with the nature of the underlying group for the gauge theory. For the non-Abelian as well as the Abelian case, the Yang-Mills field (20) b_μ^k and the non-integrable phase factor (5.1) are defined in each of the overlapping space-time regions. In the overlap of any two such regions, there is a gauge transformation relating the phase factors defined for the two regions. The phase factor takes on a slightly more complicated form if the points P and Q are in different space-time regions. With this setup, there is clearly a very close relation with fibre bundles (9,13). A short dictionary (17) between the physical and mathematical terminologies is given in Table 1.

8. Discussions

We have described briefly the development of the language appropriate for gauge theory. It is my belief that this basic language needs no further major revision.

The next problem is to develop the dynamics of gauge theory on the basis of the global formalism. Some steps have been taken in this direction in the context of electromagnetism. The Lagrangian with positrons, monopoles and electromagnetic field has been written down (18). To avoid the Rosenbaum paradox (12), the condition needs to be imposed that the positron world lines and the monopole world lines do not intersect. This condition causes no difficulty because it excludes a set of probability zero. The quantum field theory for this system has also been carried out recently (14) within the global formalism.

There is the following most chellenging and important problem that is as yet unsolved. Because of the Dirac (6) quantization condition (6.7) which plays a central role in the global Lagrangian formalism it is not possible for the two coupling constants $\frac{e^2}{\hbar c}$ and $\frac{g^2}{\hbar c}$ to be both small. How can we develop a systematic approximate scheme for calculating the interaction of positrons, magnetic monopoles, and photons in the context of quantum field theory?

I thank CERN and the University of Bonn for their kind hospitality. I am greatly indebted to Professor Chen Ning Yang from whom I learn gauge theory.

References

1. Y. Aharonov and D. Bohm, Phys. Rev. 115, 485 (1959).

2. I. Białynicki-Birula, Bull. Acad. Pol. Sci., Ser. Sci. Math. Astron. Phys. 11, 135 (1963).

3. R.G. Chambers, Phys. Rev. Letters 5, 3 (1960).

4. Shiing-shen Chern, Ann. Math. 45, 747 (1944).

5. Shiing-shen Chern, Geometry of Characteristic Classes, Proc. 13th Biennial Seminar, Canadian Math. Congress (1972).

6. P.A.M. Dirac, Proc. Roy. Soc. London A 133, 60 (1931).

7. W. Ehrenberg and R.E. Siday, Proc. Phys. Soc. London B 62, 8 (1949).

8. V. Fock, Zeit. f. Physik 39, 226 (1927).

9. S. Kobayashi and K. Nomizu, Foundations of Differential Geometry (Interscience, New York, Vol. I -1963, Vol. II-1969).

10. F. London, Zeit. f. Physik 42, 375 (1927).

11. S. Mandelstam, Ann. Phys. (N.Y.) 19, 1 and 25 (1962).

12. D. Rosenbaum, Phys. Rev. 147, 891 (1966)

13. N. Steenrod, The Topology of Fibre Bundles (Princeton University Press, Princeton, N.J., 1951).

14. Tung Sheng Tu, Tai Tsun Wu, and Chen Ning Yang. Acad. Sinica (to be published)

15. H. Weyl, Stzber. Preuss. Akad. Wiss. 465 (1918)

16. H. Weyl, Zeit. f. Physik 56, 330 (1929)

17. Tai Tsun Wu and Chen Ning Yang, Phys. Rev. D 12, 3845 (1975)

18. Tai Tsun Wu and Chen Ning Yang, Phys. Rev. D 14, 437 (1976)

19. Chen Ning Yang, Phys. Rev. Letters 33, 445 (1974)

20. Chen Ning Yang and Robert L. Mills, Phys. Rev. 96, 191 (1954).

Table I. Translation of Terminology

Gauge Theory	Bundle
gauge (or global gauge)	principal coordinate bundle
gauge type	principal fibre bundle
gauge potential b_μ^k	connection on principal fibre bundle
field strength $f_{\mu\nu}^k$	curvature
phase factor (5.1)	parallel displacement
source J_μ^k	?
electromagnetism	connection on U_1 bundle (i.e. principal fibre bundle with group U_1).
Yang-Mills field for isospin	connection on SU_2 bundle
Dirac's monopole quantization	classification of U_1 bundle according to first chern class (4,5)
electromagnetism without monopole	connection on trivial U_1 bundle
electromagnetism with monopole	connection on non-trivial U_1 bundle