MODELS AND MECHANISMS IN GAUGE THEORIES

A. M. Polyakov
Landau Institute of Theoretical Physics
Moscow, U.S.S.R.

INTRODUCTION

Our theoretical activity in particle physics can be roughly divided between invention of new physical mechanisms, and application of these mechanisms to the construction of different models. The most familiar example is the Higgs mechanism, and it is amazing in how many different models this simple phenomenon has been used. In my opinion at the present time we lack new mechanisms, and not models.

In this talk, I describe several pieces of information concerning the dynamics of gauge theories. Since a complete picture of this dynamics is absent up to now, the best I can do is to present several disjointed ideas which look beautiful and promising. Certainly the choice is highly subjective and I do not try to conceal this.

Gauge fields are used for the construction of QCD and QFD. In both cases the most important question is what phases are realized if the gauge group is given. Different possibilities are known: confinement, total spontaneous breakdown, partial spontaneous breakdown and their combinations. Some unknown options also are not excluded. At the moment we have some superficial understanding of the qualitative features of different phases, but we do not know under what circumstances this or that phase is realized. I want to begin the discussion by describing some of the features mentioned above.

GEORGI-GLASHOW TYPE PHASE AND SUICIDAL MONOPOLES

There is a strong appeal for grand unification of all interactions, which needs no explanation. One of the nice features of such unifications is that electric charge becomes a generator of some semi-simple group and is quantized in the same fashion as angular momentum in quantum mechanics. This certainly provides a natural and beautiful explanation of one of the most fundamental experimental facts. At the same time in any unified scheme we should have magnetic monopoles as stable elementary particles. This fact is a consequence of a simple topological consideration: each grand unification group should be spontaneously broken almost completely, so that only the unbroken U(1) subgroup remains. This U(1) has the topology of the circle since it is a part of some compact nonabelian group. This in turn implies the angular nature of the vector potential and the inevitable existence and stability of magnetic monopoles. A rough estimate of their mass can be obtained as the energy in the magnetic field:

$$M \sim \int H^2 \, d^3x, \qquad H \sim \frac{g}{x^2} = \frac{1}{ex^2}, \qquad x \gtrsim m_w, \tag{1}$$

so that $M \sim m_w/e^2$. (Here g is magnetic charge, e^2 the fine structure constant, and m_W is the mass of the heaviest vector boson in the theory.) Of course, there could be considerable numerical factors in Eq. 1 depending on the detailed strucutre of the theory. A somewhat disturbing observation has been made by Zeldovich and Khlopov.[1] They estimated the number of relic magnetic monopoles produced in the big bang. Their result is that this number is about ten orders of magnitude larger than the experimental limit, almost independently of the monopole mass. Here I would like to suggest some mechanism which may be or may not be relevant for the solution of this paradox. Namely, monopoles at finite temperature have some suicidal tendencies which make them unstable. I shall demonstrate this by an example in one-dimensional field theory. Let us consider the standard one-dimensional field theory, described by the Lagrangian

$$\mathcal{L} = \tfrac{1}{2} (\partial_\mu \phi)^2 + V(\phi^2) \quad , \tag{2}$$

where $V(\phi^2)$ is a Higgs potential, such that in the weak coupling limit the asymptotic value of ϕ^2 is fixed: $\langle\phi\rangle = \pm \eta$. It is well known that apart from the usual particles, this theory contains kinks as stable objects. The stability of kinks is of topological nature and can be explained as follows. Let us consider the transition amplitude for the possible decay of the kink. According to general rules it is given by the functional integral:

$$\text{Amplitude} \sim \int_{\phi_{in}}^{\phi_{out}} e^{iS(\phi)} \, \mathcal{D}\phi \quad . \tag{3}$$

(Here $S(\phi)$ is the classical action.) The in-state should present the kink, i.e. $\phi_{in}(-\infty) = -\eta$ and $\phi_{in}(+\infty) = +\eta$. The out-state is topologically trivial: $\phi_{out}(+\infty) = \phi_{out}(-\infty) = \pm\eta$. There exists no continuous interpolating field, $\phi(x,t)$, such that $\phi(x,-\infty) = \phi_{in}(x)$ and $\phi(x,+\infty) = \phi_{out}(x)$. Because of inevitable discontinuities in $\phi(x,t)$, the action $S[\phi]$ is infinite on all possible classical paths. This proves topological stability of a kink as a quantum particle. We see that the most crucial point was spontaneous symmetry breaking due to which states with and without kinks were readily distinguishable. Let us now proceed to the case of finite temperatures. At any nonzero temperature T, the system is dramatically different from T = 0. Namely, for any finite T we have:

$$\langle\phi\rangle_{T\neq 0} = 0 \quad . \tag{4}$$

This restoration of symmetry takes place because at $T\neq 0$ we have a finite density $\bar{n}$ of kinks in the system, given by the Boltzmann formula:

$$\bar{n} \sim \exp\left[-\left\{\frac{M(\text{kink})}{T}\right\}\right] \quad . \tag{5}$$

This density creates a finite correlation length in the system, $r_c \sim \bar{n}^{-1}$, thus destroying the long range order. One can say that at finite T, kinks play the role of instantons, while being solitons at T=0. We see that at $T\neq 0$ we have no more reasons for the topological stability, and hence one should expect that as elementary excitations they acquire finite lifetime. It seems reasonable to conjecture that the decay rate Γ is proportional to the background kink density:

$$\Gamma(\text{kink}) \sim \exp\left[-M(\text{kink})/T\right] \quad . \tag{6}$$

Of course a more detailed derivation is highly desirable and, I believe, possible. The resumé of the whole story is that kinks become suicidal: it is the finite background density of kinks which causes their decay.

In the real case of the four-dimentional Georgi-Glashow type model, my conclusions are less definite at the moment. It is usually believed that the restoration of symmetry takes place in this model at some finite temperature T_c. This conclusion is based on perturbation theory and does not take into account nonperturbative effects coming from monopoles. If we recall our previous two-dimensional model, we find that in this model perturbation theory predicts the phase transition at $T_c \neq 0$, but kinks shift it to $T_c=0$. Is this also true in four dimensions? The answer is unclear at present. The problem is whether monopoles (which become instantons at $T \neq 0$) disorder the system enough to restore the original symmetry. It is a quantitative question. One should compute the monopole contribution to the order parameter in WKB approximation. Also, a strong coupling expansion must be very helpful. This has not yet been done. If the answer is affirmative, it means that the symmetry is restored at any $T \neq 0$ as in our previous example and that monopoles, as well as kinks, have suicidal tendencies and their decay rate (i.e. decay rate of the magnetic charge) is given by Eq. 6. This instability may resolve the paradox of relic monopoles. But obviously a certain amount of work is still necessary to verify our conjectures, and the fate of monopoles is still unclear.

There are many other interesting questions connected with thermal effects. One of them is further investigation of the process of quark liberation - the effect found in References 2 and 3. An especially interesting problem is to understand the nature of the transition when quarks are taken into account. Presumably it becomes a first order transition because no symmetry is broken in it. But the question is still unexplored.

Another problem is the thermal influence on the θ-term, CP-nonconservation and baryon number nonconservation through the axial anomalies. It seems that since temperature alleviates tunnelling, these effects become stronger at high T. If so, they may play an important role in cosmology. I think that all these problems present a very tempting area for future investigations.

There are also some unconventional properties of the standard Weinberg-Salam model which I would like to mention. Namely, this model has a topologically stable particle in its spectrum. To understand how this happens, consider a very large sphere S^2. To have finite energy in the Higgs field, the isospinor $\psi(x)$ should satisfy the condition $\psi^\dagger(x)\psi(x) = 1;\ x\varepsilon S^2$, thus defining a point on the sphere S^3 in the isotopic space. All mappings of S^2 to S^3 are trivial. However, one should recall, that due to U(1) gauge symmetry all points of S^3 which are connected by the U(1) transformation have the same energy. So our field is, in fact, the quotient space:

$$S^2 = \frac{S^3}{U(1)}$$

which is obtained from S^3 by identification of all U(1) equivalent points. As a result we have a topological charge in the Weinberg-Salam model, associated with the mappings:

$$S^2 \to \frac{S^3}{U(1)} = S^2$$

The analytic expression for this charge is:

$$n = \oint_{S^2} d\sigma^{\mu\nu}\ (\psi^\dagger \vec{\tau}\psi)\ \vec{F}_{\mu\nu} \quad .$$

Here $\vec{\tau}$ denotes the Pauli matrices, $\vec{F}_{\mu\nu}$ the SU(2) field strength, and $d\sigma_{\mu\nu}$ is integration over a large sphere. One still has to check that these topological excitations have finite energy. Preliminary estimates confirm this point.

INSTANTONS

There are several important recent results and misconceptions concerning instantons which I would like to discuss. First is the computation by Frolov, Fateev and Shwartz[4] of the one loop corrections to the many instanton solution in the O(3) σ-model, and the generalization of these results to the CP^N-models obtained in References 5 and 6. I shall describe the result itself, a method of derivation alternative to the original one, and general implications of this calculation. Let us start from the O(3) σ-model, described by the lagrangian:

$$\mathcal{L} = \frac{1}{2g_o^2}\ (\partial_\mu \vec{n})^2\ ,\ \vec{n}^2 = 1 \quad . \tag{7}$$

The many-instanton solution is given by[7]:

$$w \equiv (\tan \theta/2)e^{i\phi} = \prod_{k=1}^{n} \left(\frac{z-a_k}{z-b_k}\right)\ , \tag{8}$$

$$z = x_1 + ix_2 \quad ,$$

with θ,ϕ polar and azimuthal angles. The result obtained in Reference 4 is that the contribution of this classical solution to the partition function Z is given by:

$$Z^{(n)} = \frac{1}{n!}\ e^{-4\pi n/g_o^2} \int \prod_{k=1}^{n} d^2a_k\ d^2b_k$$
$$\times \exp\left\{-\sum_{i<j} \log\left[|a_i-a_j|^2 + \log\ |b_i-b_j|^2\right] + \sum_{i,j} \log\ |a_i-b_j|^2\right\} \tag{9}$$

With great surprise, one recognizes in this expression the partition function of the two-dimensional Coulomb gas, taken at the special temperature at which it is equivalent to free massive fermions. This system does not have any infrared divergences. Instantons created enough disorder for symmetry restoration which is the confinement analog in the σ-model. This destroys the common misconception that instantons cannot be relevant for confinement. The latter statement is true only if one uses the dilute gas approximation and considers only small size instantons. The mysterious part of Eq. 9 is why free fermions appeared. I don't know any satisfactory answer to this question and the best I can do is to present a derivation in which the Dirac operator appears naturally. Let us introduce a moving frame consisting of the background classical field $\vec{n}_{c\ell}(x)$ and two orthogonal unit vectors $\vec{e}_a(x)$, a=1,2. We have the Cartan expansion:

$$\partial_\mu \vec{n}_{c\ell}(x) = \sum_a B_\mu^a(x)\ \vec{e}_a \quad ,$$

$$\partial_\mu \vec{e}_a(x) = A_\mu(x) \sum_b \varepsilon_{ab}\vec{e}_b - B_\mu^a(x)\vec{n}_{c\ell} \quad . \tag{10}$$

Here the connections $A_\mu(x)$ and $B_\mu^a(x)$ describe completely the original background field $\vec{n}_{c\ell}$. If we write

$$\vec{n}(x) = \sqrt{1-\phi^2}\ \vec{n}_{c\ell}(x) + \sum_{a=1}^{2} \phi_a(x)\ \vec{e}_a(x)\ , \tag{11}$$

and expand up to quadratic terms in ϕ we obtain:

$$S = \int (\partial_\mu \vec{n})^2 \, d^2x$$

$$= S_{c\ell} + \frac{1}{2} \int d^2x \left\{ (\partial_\mu \phi_a + A_\mu \varepsilon_{ab} \phi_b)^2 + (B^a_\mu B^b_\mu - (B^c_\mu)^2 \, \varepsilon_{ab}) \, \phi_a \phi_b \right\} . \qquad (12)$$

If we take into account the duality condition:

$$\partial_\mu \vec{n} = \varepsilon_{\mu\nu} \left[\vec{n} \times \partial_\mu \vec{n} \right] \quad , \qquad (13)$$

or:

$$B^a_\mu = \varepsilon_{\mu\nu} \, \varepsilon_{ab} \, B^b_\nu$$

and the zero curvature condition for A and B:

$$\partial_\mu A_\nu - \partial_\mu A_\mu = \varepsilon_{ab} B^a_\mu B^b_\nu \quad , \qquad (14)$$

we find that the differential operator in Eq. 12 is just the square of the Dirac operator:

$$(\gamma_\mu (\partial_\mu + iA_\mu))^2 = (\partial + iA)^2 + \gamma_5 \, F_{\mu\nu} \qquad (15)$$

and the calculation of the determinant of any self-dual configuration is reduced to the calculation of the determinant in the Schwinger model (massless QED in two dimensions), which in turn is equal to:

$$\log \det \, (\gamma_\mu (\partial_\mu + iA_\mu)) =$$

$$\frac{1}{2\pi} \int F_{\mu\nu}(x) \, \log |x - x'| \, F_{\mu\nu}(x') d^2x d^2x' \, . \qquad (16)$$

However, the precise connection between Schwinger fermions and the fermions in the σ-model is not quite clear. I feel that understanding this connection will give us some insight into the structure of the σ-model.

In the CP^{n-1}-model, one loop calculations (Ref. 5 and 6) have led to the picture in which the instanton contains n constituents, so to say instanton quarks. This model is not as completely investigated as the previous one. However, it is reasonable to believe that instantons dissociate in this case also, and we have a plasma phase as in the O(3) case. Calculations in the CP^{n-1}-model disprove the misconception that in the large N limit instantons are irrelevant, their contribution being exponentially small. This appears to be untrue, because in the partition function important distances are such that instantons contribute like powers of N and not like exponentials. It seems that instantons and the 1/N expansion are two complementary ways of describing the system.

CONFINEMENT AS AN EXACT SYMMETRY CENTER OF THE GAUGE GROUP

It is useful to remember that the question of confinement can be formulated as whether or not some exact symmetry of the gauge theory is spontaneously broken. It has been derived in Ref. 3 that the self-energy of a heavy quark in the gauge vacuum is given by the formula:

$$e^{-\beta \Delta M_I} = \langle \chi_I(\Omega(\vec{x})) \rangle \quad . \qquad (17)$$

Here ΔM_I is the self-energy, β is the inverse temperature, I is the representation of the gauge group, and $\chi_I(\Omega(\vec{x}))$ is the trace of the gauge matrix $\Omega(\vec{x})$ in this representation. Averaging in Eq. 17 means:

$$\langle \quad \rangle = Z^{-1} \int \mathcal{D}\Omega(\vec{x}) \; e^{-W[\Omega(\vec{x})]} \quad , \qquad (18)$$

with $W[\Omega(\vec{x})]$ given by the functional integral with the twisted boundary condition:

$$e^{-W[\Omega(\vec{x})]} = \int \mathcal{D}A_n(\vec{x},\tau) \exp\{-\int d\vec{x} d\tau (\dot{A}_n^2 + H^2)\} \quad ,$$

$$A_n(\vec{x},\beta) = A_n^\Omega(\vec{x},0) \quad ,$$

$$A_n^\Omega = \Omega A_n(x)\Omega^{-1} + \Omega \partial_n \Omega^{-1} \quad . \qquad (19)$$

From Eq. 19 we see that $W[\Omega(\vec{x})]$ is invariant under:

$$\Omega(\vec{x}) \rightarrow {}^{\pm 2\pi i/3} \, \Omega(\vec{x}) \quad . \qquad (20)$$

If this symmetry is unbroken, then for all representations with nonzero triality we find that $\langle \chi_I \rangle = 0$ and $\Delta M_I = \infty$. Quark liberation in this language is just spontaneous breakdown of the symmetry in Eq. 20.

Apart from the kinematical considerations given above, there was an attempt to use dynamical instantons, associated with the Z_3 subgroup of SU(3)[7]. These instantons form closed surfaces in four dimensions and closed rings in three. In my opinion there is a difficulty in using such instantons because while being relevant stable minima of the action in the Z(3) gauge theory, they become unstable extrema when Z(3) is embedded into SU(3). The only way to stabilize them is to consider the theory with the Higgs fields in the zero triality representation[8] or to modify the action in the lattice theory. It may happen, however, that some other subgroup instantons are important, as we shall see later.

LOOP SPACE FORMULATION OF GAUGE THEORIES

It has been shown in Ref. 9 that gauge theories can be considered as chiral theories in loop space. Physically, this means that bare gluons or glue partons have the geometry of infinitely thin closed rings. Mathematically, if we introduce the Wilson loop functional:

$$\psi(C) = P \, \exp \int_C A_\mu dx_\mu \qquad (21)$$

(here C is a closed path and P is the ordering operation) it is possible to introduce a trivial connection in the loop space:

$$F_\mu(s,C) = \frac{\delta\psi\,(C)}{\delta x_\mu(s)} \, \psi^{-1}(C) \quad , \qquad (22)$$

(here $x_\mu = x_\mu(s)$ in some parametrization of the contour C). This connection satisfied the zero curvature condition:

$$\frac{\delta F_\mu(s,C)}{\delta x_\nu(s')} - \frac{\delta F_\nu(s',C)}{\delta x_\mu\;(s)} + \left[F_\mu(s,C), \; F_\nu(s',C) \right] = 0 \qquad (23)$$

It can be shown that this zero curvature condition in the loop space corresponds to the Bianchi identity in the ordinary space. Next, the Yang-Mills equations in terms of $F_\mu(s,C)$ are written as:

$$\frac{\delta F_\mu(s,C)}{\delta x_\mu\;(s)} = 0 \quad . \qquad (24)$$

Equations 23 and 24 are equations of motion for the chiral theory in the loop space. Their quantum analog, Ward identities in the loop space, can also be given. Instead of going into these details, we discuss here some possible perspectives on this approach. First of all, it is possible that there are infinitely many hidden symmetries in the gauge theories, since such symmetries are present in the usual chiral models in two dimensions, thus allowing their exact solution. For the case of three dimensional space-time it is indeed possible to find an infinite sequence of functionally conserved currents:

$$\frac{\delta F_{\mu}^{(k)}(s,C)}{\delta x_{\mu}(s)} = 0 \quad , k = 1,2\ldots\infty \quad . \tag{25}$$

The form of Eq. 25 is precisely the conservation equation one expects in the theory of strings[10]. However, in the four-dimensional case the presence of the hidden symmetry is still questionable. If it is there, we may be able to solve Yang-Mills theory exactly.

If we are not so ambitious, the loop approach may still be useful. Using this approach we are able to reformulate the usual perturbation theory in a manifestly gauge invariant way by considering propagation of rings of glue instead of usual particles. This new form of the theory may give us some new approximations. Another use of the loop space is to consider SU(N) theory with large N. In this case, as was shown by Migdal, we have the remarkable decoupling property of the loop functional

$$\langle \mathrm{Tr}\ \psi(C_1)\ \mathrm{Tr}\ \psi(C_2)\rangle \underset{N\to\infty}{\approx} \langle \mathrm{Tr}\ \psi(C_1)\rangle\langle \mathrm{Tr}\ \psi(C_2)\rangle \quad . \tag{26}$$

When substituted into the quantum equations of motion mentioned above, Migdal's decoupling transforms the chain of equations for different Green's functions in the loop space into a closed equation for the $\langle \mathrm{Tr}\ \psi(C)\rangle$, thus reducing considerably the complexity of the problem[11].

Another useful feature of the loop approach is that it extends our understanding of possible field theories. One can consider different field theories in the loop space, for example with non-zero curvature (Eq. 23) or with fields being elements of homogeneous spaces. Who knows what king of field theory will be necessary at the next fundamental level?

REFERENCES

1. Ya. Zeldovich and M. Khlopov, Phys. Lett. B79, 362, (1979).
2. L. Susskind, SLAC preprint (1978).
3. A. Polyakov, Phys. Lett. B72, 477 (1978).
4. I. Frolov and A. Shvarts, ZhETF Pisma 28, 274 (1978) [JETP Lett. 28, 249 (1979)].
5. V. Fateev, I. Frolov and A. Shvarts, Nucl. Phys. B154, 1 (1979).
6. B. Berg and M. Luscher, DESY preprint (1979).
7. G. t'Hooft, Nucl. Phys. B138, 1 (1978).
8. A. Polyakov, ZhETF 68, 1990 (1975) [Sov. Phys. JETP 41, 988 (1975)].
9. A. Polyakov, Phys. Lett. B82, 247 (1979), and Nuclear Physics (to be published).
10. C. Marshall and P. Ramond, Nucl. Phys. B138, 1 (1978).
11. Y. Makeenko and A. Migdal, ITEP preprint (1979).

DISCUSSION

H. Thacker (Fermilab) It seems that there is sort of a two-pronged attack on QCD. One is led by you and Mandelstam and other people who want to solve the theory exactly. Another attack is perturbation theory. Do you think that any of the properties that you could discover in perturbation theory, for example factorization or something like that, have anything to do with or could give you any insight into the possible exact integrability of QCD?

A. Polyakov Well, one example was this Midgal decoupling property which was checked in perturbation theory and it was very useful. Also I think that the development might go in the opposite direction. You can consider perturbation theory in the loop space - you can consider small loops and represent propagation of small loops and by expanding, just as in ordinary perturbation theory, you can do perturbation theory for the chiral field in the loop space as well. I believe this perturbation theory will be manifestly gauge invariant. So my belief is that maybe working in the loop space will give you a better way of computation than conventional Feynman diagrams with gauge-fixing and ghosts, etc. This resembles a little the situation with QED. Old QED was not manifestly Lorentz covariant. Then people invent manifestly Lorentz covariant QED which was equivalent to the old one, but much more convenient. I believe that in the loop space we would have a manifestly gauge invariant perturbation theory, which might be much more convenient for understanding and computation than the usual one. Certainly that's our duty, I think, to check everything in perturbation theory before it is admitted.

N. Christ (Columbia) I would like to raise a question about your remarks on the U(1) problem. It would seem to me that if you consider a very large sphere for space-time, a 4-dimensional sphere whose radius was quite a bit bigger than the Compton wavelength of the massive meson that was causing the U(1) problem, you would have a context in which the U(1) problem could be addressed. I think that in that case the analysis that was done on the CP^1 model involves just the superposition of instanton or anti-instanton configurations with definite topological charge, so the θ dependence would always have the normal period. I believe that the difficulties Crewther raised refer to an unexpected θ dependence of the quantum numbers of the Goldstone particles rather than the need to go beyond the normal collection of instantons and anti-instantons.

A. Polyakov Well, let me say that for any finite volume you would be right. A problem precisely analogous to what happens when you introduce fermions with chirality into a gas of instantons is just the introduction of external fractional charge into a plasma. And what happens there? Certainly we have the Gauss law. If you take a finite sphere, as you like, then if you are outside the sphere the total charge will not be integer, there is no miracle here - the charge will be fractional, but all the fractional part of the charge will be situated at the surface of the sphere. So if you are inside, and presumably we are inside, then what you would observe is unusual θ-dependence. In a plasma you have unusual θ-dependence, but it appears only in the infinite volume limit. It is a fact of infinite volumes. So you cannot really consider finite volume as an example. In this case you would have a kind of spontaneous symmetry breaking which appears only in the infinite volume limit.