THE CONFINEMENT PHENOMENON IN QUANTUM FIELD THEORY

Gerard 't Hooft

Instituut voor Theoretische Fysica der Rijksuniversiteit
De Uithof
Utrecht, The Netherlands

0. ABSTRACT

In these written notes of four lectures it is explained how the phenomenon of permanent confinement of certain types of particles inside bound structures can be understood as a consequence of local gauge invariance and the topological properties of gauge field theories.

I. INTRODUCTION

The lagrangian of a quantum field theory describes the evolution of a certain number of degrees of freedom of the system, called *fields* as a function of space and time. In many cases this evolution is only straightforward in a small region of space-time and therefore these fields should be interpreted only as the *microscopic* variables of the theory. Only in the simplest cases these microscopic variables also correspond to actual physical particles (the *macroscopic* objects) but very often the connection is less straightforward for two reasons. One reason is that there may be a local gauge invariance. This is a class of transformations that transform a set of fields into another set of fields with the postulate that the new set describes the same physical situation as the old one. Therefore the physical fields only constitute some orthogonal subset of the original set of fields and one of the problems that we will study in these lectures is how to associate this subset to observable objects.

However even these observable objects ("transient particles") may in some cases not yet be the macroscopic physical particles. That is one second point : various kinds of Bose-condensation may take place after which the spectrum of physical particles may look entirely different once again. We will study these condensation phenomena as we go along.

The most challenging application of our theoretical considerations is "quantum chromodynamics" or generalizations thereof. The "microscopic" lagrangian there contains only vector particles ("gluons") and spinors ("quarks") but neither of these are really physical. The spectrum of physical particles always consists of bound states of certain numbers of quarks and/or antiquarks with unspecified numbers of gluons. We will obtain a qualitative understanding of this transition from the microscopic to the macroscopic dynamical variables.

The first models that we will consider may seem to be a far way off from that desired goal but studying them will turn out to be crucial for obtaining a suitable frame and language in order to put the more advanced systems in a proper perspective.

II. SCALAR FIELD THEORY

Let us consider the lagrangian of a complex scalar field theory :

$$\mathcal{L}(\phi,\phi^*) = -\partial_\mu\phi^*\partial_\mu\phi - m^2\phi^*\phi - \frac{\lambda}{2}(\phi^*\phi)^2 \qquad (2.1)$$

Equivalently we can use real variables :

$$\phi = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2) \; ; \qquad \phi^* = \frac{1}{\sqrt{2}}(\phi_1 - i\phi_2). \qquad (2.2)$$

We write ϕ^* rather than $\phi^\dagger$ because here the fields are c-numbers, not operators, and the lagrangian must be seen in the context of a functional integral as described by B. de Wit in his lectures at this School.

The lagrangian (2.1) describes a system with two species of Bose-particles both with the same mass m but distinguishable in the quantum mechanical sense. One can either consider ϕ_1 and ϕ_2 as the two species, both equal to their own antiparticles, or consider ϕ as a particle and ϕ^* as its distinct antiparticle. At this level these descriptions are equivalent.

Obviously there is a *symmetry* under exchange of ϕ_1 and ϕ_2, or more generally

$$\phi \to e^{i\Lambda}\phi$$
$$\phi^* \to e^{-i\Lambda}\phi^* \qquad (2.3)$$

This is a group of rotations in the complex plane called U(1) and is only global, that is, Λ must be independent of space-time.

By Noether's theorem the model contains a conserved current,

$$j_\mu = \phi^* \partial_\mu \phi - (\partial_\mu \phi^*)\phi \tag{2.4}$$

with $$\partial_\mu j_\mu = 0 \tag{2.5}$$

Classically (2.5) is only true if ϕ, ϕ^* are required to obey the Euler-Lagrange equations generated by (2.1). Quantum mechanically (2.5) follows if we substitute for ϕ and ϕ^* the corresponding operators ϕ and $\phi^\dagger$.

Our model is trivial in the sense that, if $m^2 > 0$, the microscopic fields ϕ_1 and ϕ_2 directly correspond to the expected macroscopic physical particles (apart from possible stable bound states) and the symmetry (2.3) is also a symmetry between these particles.

III. BOSE CONDENSATION

Bose condensation is a well-known phenomenon in quantum statistical physics. Just in order to make the connection with our case of interest let us first consider an ideal non-relativistic Bose gas. The states are

$$|n(\vec{k}_1), n(\vec{k}_2), \ldots > \tag{3.1}$$

where

$$\vec{k}_i = \frac{\pi}{L}(l_1, l_2, l_3) \tag{3.2}$$

and l_1, l_2, l_3 are positive integers. L is the length of a side of the box in which the particles are contained. The energy of the states is

$$E = \sum_i n(\vec{k}_i) k^2/2M \tag{3.3}$$

where M is the mass of the (non-relativistic) particles. We take for the thermodynamic free energy F,

$$e^{-\beta F} = \sum_{\{n(\vec{k})\}} e^{-\beta[E(\{n(\vec{k})\}) - \mu N]} \tag{3.4}$$

where $N = \sum_i n(\vec{k}_i)$, $\beta = 1/T$

T is the temperature in natural units, and μ is the chemical potential.

We can easily solve (3.4) :

$$e^{-\beta F} = \prod_k \sum_{n=0}^{\infty} e^{\beta(\mu - k^2/2M)n} = \prod_k \frac{1}{1 - \exp \beta(\mu - k^2/2M)} \qquad (3.5)$$

and in the limit of infinite volume :

$$F = (2\pi)^{-3} TV \int d^3k \log(1 - \exp\beta(\mu - k^2/2M)) \qquad (3.6)$$

$$V = L^3$$

The particle number density is

$$\frac{N}{V} = -\frac{1}{V}\frac{\partial F}{\partial \mu} = \int d^3k \, (\exp \beta(-\mu + k^2/2M) - 1)^{-1} \qquad (3.7)$$

One easily notes that the formulas (3.5) - (3.7) explode if the chemical potential μ becomes larger than or equal to zero, a typical property of Bose gases.

Since this is a non-relativistic system it is convenient to introduce a field $\phi(\vec{x})$ in the following way :

$$\phi(\vec{x}) = 2\sqrt{\frac{2}{V}} \sum_{\vec{k}} a(\vec{k}) \prod_{i=1,2,3} \sin \frac{\pi(k_i x_i)}{L} \qquad (3.8)$$

where $a(\vec{k})$ is an operator that annihilates one particle with momentum $\vec{k}$ in the usual way. The hamiltonian is then

$$H = \sum_{\vec{k}} \left(\frac{k^2}{2M} - \mu\right) a^\dagger(\vec{k})\, a(\vec{k}) \qquad (3.9)$$

$$= \int d^3\vec{x} \left(\frac{1}{2M} \vec{\partial}\phi^\dagger(\vec{x})\vec{\partial}\phi(\vec{x}) - \mu\phi^\dagger(\vec{x})\phi(\vec{x})\right) \qquad (3.10)$$

where for convenience we included the chemical potential term so that this hamiltonian describes the complete system :

$$e^{-\beta F} = \mathrm{Tr}\, e^{-\beta H}$$

It is obvious here that μ should not be allowed to be positive.

Now however we can take into account the *repulsive* forces between the particles. When many particles are close together we expect an extra, positive contribution to H. A simple model for that is

$$H = \int d^3x \left(\frac{1}{2M}\, \partial\phi^\dagger \partial\phi \;-\; \mu\phi^\dagger\phi + \frac{\lambda}{2}\, \phi^{\dagger 2}\phi^2\right) \tag{3.11}$$

As long as μ is negative the λ term is just a small perturbation. But if μ is positive then the λ term is only one that can stabilize the system.

Of course it is difficult to find the free energy of this revised system exactly, but an easy approximation, valid for λ not too large, is to substitute φ by a c-number (as defined in (3.8) it was an operator), and subsequently minimize H. We then find approximately the energy of the lowest eigenstate :

$$\phi^\dagger\phi \equiv \rho \tag{3.12}$$

$$\frac{\partial}{\partial\rho}\left(-\mu\rho + \frac{\lambda}{2}\,\rho^2\right) \simeq 0 \tag{3.13}$$

$$\rho \simeq \frac{\mu}{\lambda} \tag{3.14}$$

$$E_o \simeq -\frac{\mu^2}{2\lambda}\, V \tag{3.15}$$

$$e^{-\beta F} = \mathrm{Tr}\; e^{-\beta H} \simeq e^{-\beta E_o} \tag{3.16}$$

$$F \simeq E_o \tag{3.17}$$

$$\frac{N}{V} = \frac{1}{V}\int \phi^\dagger\phi\, d^3x = \rho \simeq \mu/\lambda \;, \quad (\mu > 0) \tag{3.18}$$

to be contrasted with (3.7), still approximately valid for $\mu < 0$. As μ turns from negative to positive a phase transition is said to take place; we suddenly get large values for the fields φ already in the lowest eigenstate of H. This is called Bose condensation. It takes place whenever the pressure is so high that the chemical potential becomes negative. Also the λ term must be sufficiently small.

The model described by our hamiltonian (3.11) resembles somewhat the model of the previous section. Note however that the mass

M comes in the derivative term and never vanishes. In the relativisstic field theory M is replaced by 1/2, and $-\mu$ by the mass-squared, m^2. In that theory Bose-condensation can take place also : we must extend the allowed values for m^2 to negative values, clearly a more profound change in the particle properties. But what we get in return is that in the relativistic model, the vacuum itself, without any external pressure, can become a Bose-condensate.

IV. GOLDSTONE PARTICLES

In the previous section we discussed Bose-condensation in a statistical system only in a sketchy way because we will not need the details (for instance, ϕ and $\phi^\dagger$ were not canonical variables in the usual sense). We will be more precise for the case that is more relevant to us : the relativistic complex scalar case. We return to the lagrangian (2.1) and now assume m^2 to be negative. It is convenient to rewrite it as

$$\mathcal{L}(\phi,\phi^*) = -\partial_\mu\phi^*\partial_\mu\phi - \frac{\lambda}{2}(\phi^*\phi - F^2)^2 \tag{4.1}$$

where $m^2 = -\lambda F^2 < 0$ automatically, and an irrelevant constant, $-\lambda/2\ F^4$, has been added to the lagrangian.

The hamiltonian density of the system is

$$\mathcal{H} = \pi^\dagger\pi + \partial_i\phi^\dagger\partial_i\phi + \frac{\lambda}{2}(\phi^\dagger\phi - F^2)^2 \tag{4.2}$$

where $\pi, \pi^\dagger$ are the canonical momenta associated with $\phi, \phi^\dagger$ which are now operators rather than fields. The $\phi, \phi^\dagger$ dependent part has an extremum for

$$\phi^\dagger\phi = F^2 \tag{4.3}$$

from which

$$\phi = Fe^{i\omega}$$

$$\phi^\dagger = Fe^{-i\omega} \tag{4.4}$$

where ω is arbitrary but fixed to the same value everywhere in space.

Now if there were no $\pi^\dagger\pi$ term in $\mathcal{H}$ then (4.4) would be the exact solution to the Schrödinger equation for the lowest energy state. Every ω between 0 and 2π would describe a lowest-energy eigenstate of $H = \int \mathcal{H} d^3x$, each with eigenvalue $E_o = 0$. The question is whether the $\pi^\dagger\pi$ term causes sufficient fluctuations to lift this degeneracy. The answer is not so simple : in 1 space - 1 time dimension, yes; in 2 or more space dimensions, no. So let us limit ourselves to 3 space + 1 time dimensions. Then the vacuum (= lowest energy state) is degenerate and characterized by a phase angle ω. But (4.4) is not exactly valid due to the $\pi^\dagger\pi$ term. We replace it by

$$< 0|\phi|0 > = F' e^{i\omega} \tag{4.5}$$

where $|0>$ is the vacuum state and ω is the vacuum angle. From now on we will consider only the world surrounded by a vacuum with $\omega = 0$. Further, F' is close to F. In fact, because of ultraviolet divergences, subtractions must be made in F and ϕ, and we could choose these such that $F' = F$.

Actually our theory only makes sense if either λ is chosen to be rather small and F of order $1/\sqrt{\lambda}$, or if an adequate ultraviolet cutoff has been introduced. The reasons for this are deeper field theoretic arguments connected with the renormalization group that I will not go into. Let us assume that λ is rather small. Then the fluctuations of ϕ around F are also relatively small and it makes sense to split

$$\phi = F + \eta \tag{4.6}$$

and the lagrangian becomes

$$\mathcal{L} = - \partial_\mu \eta^* \partial_\mu \eta - \frac{\lambda}{2}(F(\eta+\eta^*) + \eta^*\eta)^2 \tag{4.7}$$

Writing

$$\eta = \frac{1}{\sqrt{2}}(\eta_1 + i\eta_2)$$

This becomes

$$\mathcal{L} = - \frac{1}{2}(\partial_\mu \eta_1^2 + \partial_\mu \eta_2^2) - \lambda F^2 \eta_1^2 + \text{int} \tag{4.8}$$

where "int" stands for higher order terms in η_1, η_2. Notice now that one particle, η_1, obtained a mass

$$M_\eta = F\sqrt{2\lambda} \tag{4.9}$$

But its companion η_2 became massless.

The occurrence of a massless particle as soon as the vacuum expectation value of a field ϕ is not invariant under the continuous symmetry (2.3) has been first observed by J. Goldstone[1], and it is an exact property of the system, not related to our perturbative approximation (no higher order mass corrections). We conclude that after the phase-transition caused by Bose-condensation, the symmetry (2.3) is *spontaneously broken* (the degeneracy of ϕ_1 and ϕ_2 is not reproduced in η_1, η_2) and at the same time a massless particle appears : the Goldstone particle. Here we see the first example where the microscopic fields ϕ,ϕ^* in the lagrangian do not reflect accurately the physical spectrum, but the transition towards the η fields was still very simple. It is correct to characterize the vacuum by

$$< 0|\phi(\vec{x})|0 > = F \neq 0$$

and the vacuum is infinitely degenerate. Characterization of the *Higgs* mode, next section (V) will be very different!

V. THE HIGGS MECHANISM

We now switch on electromagnetic interactions*) simply by adding the Maxwell term to the lagrangian and replacing derivatives by covariant derivatives :

$$\mathcal{L}(\phi,\phi^*,A_\mu) = -\tfrac{1}{4}F_{\mu\nu}F_{\mu\nu} - (D_\mu\phi)^* D_\mu\phi - \tfrac{\lambda}{2}(\phi^*\phi - F^2)^2 \qquad (5.1)$$

where

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

$$D_\mu\phi = (\partial_\mu + iqA_\mu)\phi \qquad (5.2)$$

q is the electric charge of the particle ϕ.

Indeed the previously introduced current j_μ, eq (2.4), is now the conserved electromagnetic current. But the invariance (2.3) can now be replaced by

$$\phi \to e^{i\Lambda(x)}\phi \ , \quad \phi^* \to e^{-i\Lambda(x)}\phi^*$$

$$A_\mu \to A_\mu - 1/q\ \partial_\mu\Lambda(x) \qquad (5.3)$$

---

*) It may seen to a superficial reader that these notes are just repeating the story[2] of the early 70's. However we are now not primarily interested in pertubative quantization but rather *non-perturbative characterization* of what happens.

which is a *local* invariance

As already mentioned in the Introduction, the consequence of this local invariance is that only a subspace of all (ϕ, ϕ^*, A_μ), namely the gauge-nonequivalent values, correspond to physical observables.

Traditionally, one now proceeds by choosing a gauge fixing procedure so that most, if not all, degeneracy is removed. One then performs perturbation expansions in λ and q^2. At zero λ and q^2 one can again ask whether or not

$$< \phi >_o = F \neq 0 \ ? \tag{5.4}$$

and since the higher order corrections to $< \phi >_o$ are of higher order in λ and q^2 the qualitative distinction whether or not $< \phi_o > = 0$ remains valid at every order. And so we get the local variant on the Goldstone mechanism : after Bose-condensation of charged particles we get the Higgs mechanism [3]. However, there is a difficulty with (5.4), because ϕ is not gauge invariant. Smearing (5.4) over all of space-time may yield zero or not, depending on the gauge chosen. In a trivial gauge

$$\begin{aligned} \mathrm{Re}(\phi) &> 0 \\ \mathrm{Im}(\phi) &= 0 \end{aligned} \tag{5.5}$$

we have

$$< \phi >_o > 0$$

always. We therefore propose to use criterion (5.4) *only* in perturbative considerations, where it is correct (as good as it can be) but *not* as an absolute non-perturbative criterion for the Higgs mode. Another criterion that cannot be used is whether or not the vacuum is degenerate. The problem there is that transformation (5.3) yields physically equivalent states, contrary to its global equivalent (2.3)!! Therefore all those vacuum states corresponding to different ω angles are now one and the same state. The vacuum is never degenerate if the symmetry is local. Local symmetries are *never* "spontaneously broken". Then why is this phrase so often used in connection with gauge theories ? Because, as I will show now, there certainly is such a thing as a Higgs mode and it usually can be described in some or other reasonable perturbation expansion around a Goldstone (= global) field theory.

Let us return to perturbation theory momentarily. We then write as usual

$$\phi(x) = F + \eta(x) \tag{5.6}$$

$$\mathcal{L} = -\tfrac{1}{4}F_{\mu\nu}F_{\mu\nu} - q^2F^2A_\mu^2 - \sqrt{2}\, qFA_\mu D_\mu \eta_2 - (D_\mu\eta)^* D_\mu\eta - \lambda F^2\eta_1^2 + \text{int} \tag{5.7}$$

A convenient renormalizable gauge is obtained by adding the gauge fixing term

$$\mathcal{L}^c = -\tfrac{1}{2}(\partial_\mu A_\mu + \sqrt{2}\, qF\eta_2)^2 \tag{5.8}$$

so that

$$\mathcal{L} + \mathcal{L}^c = -\tfrac{1}{2}(\partial_\mu A_\nu)^2 - \tfrac{1}{2}M_A^2A^2 - D_\mu\eta^* D_\mu\eta - \tfrac{1}{2}M_\eta^2\eta_1^2 - \tfrac{1}{2}M_A^2\eta_2^2 + \text{int} \tag{5.9}$$

with

$$M_A = \sqrt{2}\, qF \qquad\qquad M_\eta = \sqrt{2\lambda}\, F \tag{5.10}$$

It is easily read off from this lagrangian that the vector particle A has a mass M_A and η_1 a mass M_η. The longitudinal component of the vector field A_μ and η_2 are ghosts, which both cancel against the Faddeev-Popov-De Wit ghost [4,5] all having the same mass M_A.

So perturbation theory suggests that the Higgs theory behaves in a way very different from the symmetric or Coulomb theory : one of the two scalar fields ϕ disappears and the vector field obtains a mass so that the photon field is short-range only. This *is* a distinction that should survive beyond perturbation theory. Thus the criterion that electromagnetic forces become short-range is much more fundamental than either the vacuum value of the scalar field (5.4) or the "degeneracy of the vacuum". But, there is yet another new phenomenon in the Higgs mode contrary to the "unbroken" or Coulomb mode. This is important because the above does not yet distinguish a Higgs theory from just any non-gauge theory with massive vector particles.

VI. VORTEX TUBES

The non-relativistic version of the theory of the previous section is the superconductor : if electrically charged bosons (Cooper's bound state of an electron pair) Bose-condense then there the electric fields become short-range. Also magnetic fields are repelled completely (Meissner effect). Except when they become too strong. Then, because of magnetic flux conservation, they have to be allowed in. What happens is that a penetrating magnetic field forms narrow flux tubes. These flux tubes carry a multiple of a precisely defined quantum of magnetic flux. A little amendement to the perturbative Higgs theory can explain this.

By gauge transformations the "vacuum value" of the Higgs field ϕ can be changed into

$$< \phi(x) >_o = Fe^{i\omega(x)} \tag{6.1}$$

if $\omega(x)$ is a continuous, differentiable function of space and time. However we can also consider perturbation theory around field configurations

$$\phi(x) = \rho(x)\ e^{i\omega(x)} \tag{6.2}$$

where $\rho(x) \simeq F$ nearly everywhere, but at some points ρ may be zero. At such points ω needs not be well defined and therefore in all the rest of space ω could be multivalued. For instance, if we take a closed contour C around a zero of $\rho(x)$ then following ω around C could give values that run from 0 to 2π instead of back to zero. The energy of such a field configuration is only finite provided that $D_i\phi(x)$ goes to zero sufficiently rapidly at ∞. Since $\partial_i\phi$ does not go to zero fast enough there must be a supplementary vector potential $A_i(x)$.

The easiest way to find that is by taking a gauge transformation that is regular at ∞ but singular at the origin.

$$e^{-i\Lambda} = \phi(x)/\rho(x) \tag{6.3}$$

In the new gauge $\partial_i\phi$ may vanish rapidly at ∞ and therefore $A_i(x)$ also. So in the old gauge

$$\oint A_i dx^i = \oint \frac{d\lambda}{q} = \frac{2\pi}{q} \tag{6.4}$$

which is the magnetic flux. One can compute the energy of the flux tube by assuming cylindrical symmetry and substituting (6.2) into (5.1). One then varies A_μ and ρ with the boundary condition (6.4) fixed, and minimizes the hamiltonian derived from (5.1). One typically finds that the energy of a flux with length ℓ is [6)]

$$E = \alpha\ell \tag{6.4}$$

$$\text{with } \alpha = O(F^2) = O(M_A^2/q^2) \tag{6.5}$$

$$\text{if } q^2 = O(\lambda) \tag{6.6}$$

If finally a magnetic field is admitted inside a superconductor it can only come in some multiple of these vortices, never spread out because of the Meissner effect. Classically one may consider this as an aspect of the infinite conductivity of the material that is only broken down in sufficiently strong magnetic fields. The stable vortex configurations that we discussed here were first derived in the relativistic theory by Nielsen, Olesen and Zumino[6]. The existence of these macroscopic stable objects can be used as another characterization of the Higgs mechanism. They should also survive beyond perturbation expansion.

VII. DIRAC'S MAGNETIC MONOPOLES

At this stage it is useful to introduce the notion of a single magnetic charge à la Dirac[7]. It is not (yet) a dynamic particle but just a source or sink of magnetic flux, a spectator particle not dynamically involved in the lagrangian of the theory. A Dirac monopole can be visualized as the end point of an infinitely thin coil carrying a large electric current. The vector potential $\vec{A}$ is very large close to the coil, because of this electric current :

$$\oint \vec{A}d\vec{x} = \Phi \qquad (7.1)$$

where Φ is the magnetic flux of the coil and the integral is over any contour going closely around it. Close to the coil dx is small, therefore $\vec{A}$ becomes large.

Nevertheless the effect of the coil on its surroundings comes only through the end points, if a gauge transformation exists that removes this large vector potential :

$$\oint \frac{d\Lambda}{q} = \Phi \qquad (7.2)$$

Such gauge transformations Λ would be multivalued, but we require that $e^{i\Lambda}$ in (5.3) remains single-valued. So the jumps that Λ is allowed to make are multiples of 2π. Therefore the gauge transformation (7.2) turns single-valued field configurations into single-valued field configurations if

$$\Phi = 2\pi n/q \qquad (7.3)$$

This is how Dirac found that the total amount of magnetic flux carried by a magnetic monopole must be quantized in units $2\pi/q$ where q is the smallest possible electric charge in the universe. This condition must be satisfied whenever we want a rotationally invariant quantized theory with magnetic monopoles and single valued fields. It is illustrative now to see what would happen with such a spectator particle inside a Higgs theory (or superconductor).

It is not accidental that the monopole quantum $2\pi/q$ coincides with the Nielsen-Olesen-Zumino vortex quantum. This implies that the monopole will be sitting at the end of an integer number of such vortices. Antimonopoles may be sitting at the other ends. Now the energy of such configurations is approximated by eq. (6.4). It is proportional to their separation distance. And so we notice that the monopoles inside a superconductor are kept together by an infinite potential well, the potential being simply linearly proportional to their separation. This is the first observation of a confinement feature in quantum field theory, although the confined objects were as yet spectators, not any of the participants of the field equations. That will come later (sect. XIII).

VIII. THE UNITARY GAUGE

So far we limited ourselves strictly to Abelian gauge theories. We knew what the microscopic field variables are, and now we know what particles and vortices survive at macroscopic distance scales. The latter depend critically on what kind, if any, of Bose condensation took place. Now in our introduction we also mentioned the *microscopic physical* variables. Formally the space H of these variables is given by

$$H = R/G \qquad (8.1)$$

where R is the space of field variables and G is the (local) gauge group. How do we enumerate the variables in H ?

Traditionally one imposes a gauge condition on the fields in R, thus obtaining a subspace in R which could be representative for H. In section V we used the gauge fixing term (5.8). This is good enough if one intends to do perturbation expansion[2,5]. The ghost particles one obtains cancel each other and can be dealt with. However we claim that if a non-perturbative characterization of the physical variables is required then this is not good enough. Imagine that one tries to solve the Dyson-Schwinger equations of the theory in some nonperturbative way. Whenever a computed S-matrix element shows a pole one can never be sure whether or not this is due to a ghost or whether it is physical. Furthermore as we will see the ghosts will produce their own topological features called "phantom solitons" which are entirely non-physical. Therefore if we want some understanding of the physical variables we must go to a "unitary gauge" (a gauge with no ghosts)

Often the axial gauge

$$A_0 = 0 \qquad (8.2)$$

is used in order to understand the physical Hilbert space. However, this leaves invariance with respect to time-independent gauge transformations :

$$\Lambda(\vec{x},t) = \Lambda(\vec{x}) \tag{8.3}$$

and so there is still a redundancy in our set of variables. It is not suitable for our purposes.

A completely ghost-free gauge can be formulated if we have a charged scalar field ϕ (if no such field is present one may consider building such a field by composing, say, two fermion fields). We do *not* require the Higgs phenomenon to take place. Regardless what condensation takes place at large distance scales one can look at the gauge

$$\begin{aligned} \mathrm{Re}(\phi) &= \rho > 0 \\ \mathrm{Im}(\phi) &= 0 \end{aligned} \tag{8.4}$$

This fixes the gauge function Λ locally, point by point in space-time, contrary to gauges such as eq. (5.8) where the condition on Λ requires solving a second order partial differential equation (the cause of the ghosts).

Within the unitary gauge (8.4) all components of the vector field A_μ are entirely observable. The complex scalar field ϕ is reduced to a real field ρ that can only take positive values. This would be a convenient description of the space of microscopic physical field variables were it not for one deficiency in the condition (8.4) : the original space of variables R certainly allows the scalar field ϕ to vanish at certain points in space-time. These points, defined by (for any $\phi \in R$)

$$\begin{aligned} \mathrm{Re}(\phi) &= 0 \\ \mathrm{Im}(\phi) &= 0 \end{aligned} \tag{8.5}$$

have the topological structure of a set of closed curves in 3-space, or closed surfaces in 3 + 1 dimensional space-time. At these points the condition (8.4) becomes singular : if $\phi = \rho e^{i\theta}$ then we must choose

$$\Lambda = -\theta \tag{8.6}$$

but the gradient of θ is easily seen to explode close to a zero of ϕ and therefore the vector potential $\vec{A}$, transforming as the gradient of Λ, will grow as the inverse power of the distance to this zero. Thus we find that the string-like structures, defined by (8.5) are separate degrees of freedom, giving a boundary condition on ρ ($\rho = 0$) and a prescribed singular boundary behavior of A_μ. This completes our discussion of the microscopic physical degrees of freedom for any Abelian gauge theory : we have observable vector fields A_μ, a truncated scalar field ρ ($\rho > 0$) and all possible closed strings *), on which there is a boundary condition for both ρ and A_μ

Note that only in the Higgs theory these physical variables are in the same time the *macroscopic* physical variables, although of course the macroscopic variables will be "dressed with a bound of virtual particles" (the string becomes a vortex with finite thickness). In the "unbroken" Abelian theory of electromagnetism the macroscopic variables are harder to discuss. It appears that our vortices "Bose-condense" to form long range, non-energetic magnetic field lines : the ordinary magnetic field $\vec{B}$.

IX. PHANTOM SOLITONS

The gauge (8.5) is called "unitary gauge" because in that gauge all surviving fields will be physically observable. Their quanta will all contribute in the unitarity relation

$$\sum_n \langle a|S^\dagger|n\rangle\langle n|S|b\rangle = \langle a|b\rangle \qquad (9.1)$$

However as soon as practical calculations are considered smoother gauge conditions are required. (8.5) is hard to implement if ϕ oscillates wildly at small distances. We will now argue that after a transition towards "smoother" gauge conditions not only ghost particles arise but also what we call "phantom solitons" : extended structures which are stable for topological reasons but nevertheless unphysical gauge artefacts [8].

Intermediate between the "renormalizable gauge" (5.8) and the unitary gauge we could choose

$$\arg(\phi) + \kappa\partial_\mu A_\mu = 0 \qquad (9.2)$$

where $\arg(\phi) = \mathrm{Im}(\log\phi)$ and κ is an arbitrary gauge parameter. The gauge condition (9.2)is smoother than the unitary gauge because at small distances, by power counting, the second term dominates and we come close to the renormalizable Lorentz gauge. One finds ghosts in this gauge which propagate with a mass

$$m_{gh} = (q/\kappa)^{1/2} \qquad (9.3)$$

so for small κ they become unimportant.

---

*) That is, strings without ends; they could run from ∞ to ∞.

Now imagine a Nielsen-Olesen-Zumino vortex tube in the form of a closed curve. What do the field configurations in the gauge (9.2) look like ? The gauge(9.2) is that particular gauge for which

$$W = \int d^4x(q^{-1}(\arg(\phi))^2 + \kappa A_\mu^2) \qquad (9.4)$$

has an extremum. Let us assume this is a minimum. The system then likes to arrange $\arg(\phi)$ to be zero as much as possible but not with too large vector potentials A_μ.

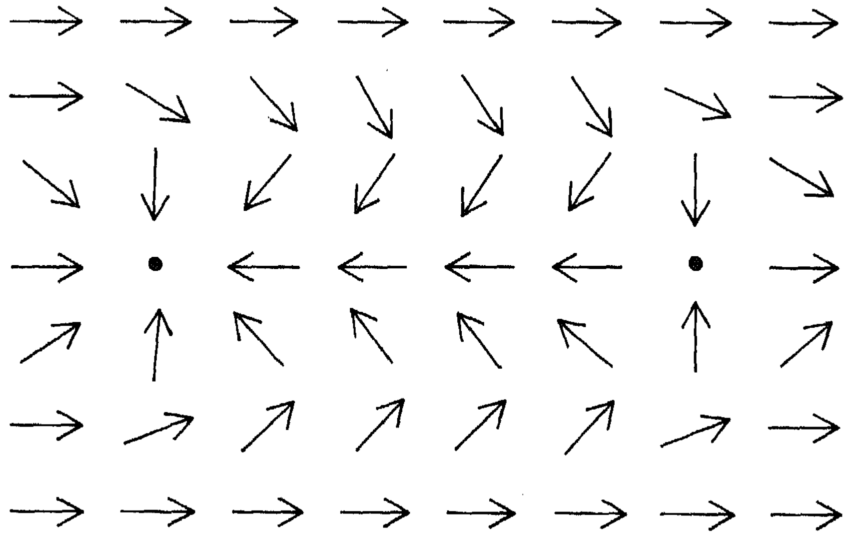


Fig 1

In fig 1 we pictured a cross section of the vortex. The plane is intersected twice, in opposite directions. At the intersection points the scalar field ϕ makes one complete rotation, again in opposite directions. The configuration in the figure is close to the optimal gauge (9.2), keeping ϕ as much as possible oriented towards the positive real axis. We see that the topology of the complete "twist" from top to bottom had to be preserved. This twist will cover an entire sheet spanned by the vortex. Certainly (9.2) will be obtained (i.e. (9.4)will be minimal) if that sheet has minimal surface. The equations for the field configurations inside the sheet are easy to solve if the sheet is considered to be locally sufficiently flat. Clearly the sheet is a gauge artifact. We think that structures of this sort will further obscure the physical interpretation of whatever solutions will be found to the Dyson-Schwinger equations in the gauge (9.2) or completely renormalizable gauges. For instance, bubbles made out of these sheets will form a whole Regge-like family of phantom particles.

X. NON-ABELIAN GAUGE THEORY

We now intend to perform the same procedure in non-Abelian gauge theories. For simplicity we restrict ourselves to the case that the gauge group is SU(N), with arbitrary N. The *microscopic* field variables are : a matrix vector field $A_\mu^{ij}(x)$ where μ is a Lorentz index and i, j run from 1 to N; and there may be Dirac spinor fields ψ_{iF}, $\bar{\psi}_{iF}$ where F is a flavor index. Scalar fields are not assumed to play a significant role but may be present too.

An element of the gauge group G (here SU(N)) is a space-time dependent unitary matrix $\Omega(x)$. Any transformation of the form

$$A'_\mu = \Omega(x)(A_\mu + \frac{i}{g}\partial_\mu)\Omega^{-1}(x)$$

$$\psi'_F = \Omega(x)\psi_F$$

$$\bar{\psi}'_F = \bar{\psi}_F\Omega^{-1}(x) \qquad (10.1)$$

is postulated to describe the same physical situation as before, but of course gives a different set of values to the microscopic variables. If R is the space of microscopic variables, then R/G is the space of microscopic *physical* variables. Again we ask the question how to categorize or enumerate these physical variables.

In perturbation theory it is customary to impose a gauge condition, which implies that we find a subspace of R (the set of fields in R that satisfy the gauge condition) representative for R/G. A renormalizable gauge condition is

$$\partial_\mu A_\mu = 0 \qquad (10.2)$$

but it is easy to see that this subspace of R does not accurately describe the physical degrees of freedom in R/G (even though it is accurate enough for perturbation theory). To see this, consider an infinitesimal perturbation in R

$$A_\mu \to A_\mu + \delta A_\mu \qquad (10.3)$$

where δA_μ is entirely localized in a particular region δV of space-time. We may however have

$$\partial_\mu \delta A_\mu = f(x) \neq 0 \qquad (10.4)$$

In order to impose the gauge condition (10.2) we now find an infinitesimal gauge transformation $\Omega = e^{ig\Lambda}$ that restores (10.2). It must satisfy

$$\partial_\mu D_\mu \Lambda = f(x) \tag{10.5}$$

but the inverse of the operator $\partial_\mu D_\mu$ is non-local. So in our subspace of R satisfying the gauge condition we find a perturbation which is spread out all over space-time, well outside δV. Clearly this perturbation outside δV is unphysical and in perturbation theory we learnt how to deal with this : the theory has "ghosts".

We know claim that beyond perturbation theory these ghosts obscure the physical contents of our theory. Therefore we shall look for a "unitary gauge". Our strategy is to determine this gauge in two steps. Let L be one of the largest Abelian subgroups of G, in our case $L = U(1)^{N-1}$.

$$L = U(1)^{N-1} \tag{10.6}$$

We call G/L the "non-Abelian part" of the gauge group and our first step will be to fix the "non-Abelian part" of the gauge redundancy. We choose a gauge condition C that reduces the space R into a subspace that could be called

$$H_1 = R/(G/L) \tag{10.7}$$

If the gauge group G has N^2-1 generators and L has N-1 generators, then we choose N^2-N real components for the gauge condition C, all invariant under $L \subset G$ but not G itself. The second step is the choice of an N-1 component gauge condition that fixes the remaining invariance under L :

$$H = H_1/L = R/G \tag{10.8}$$

But this second step is precisely the same as fixing the gauge in an ordinary Abelian gauge theory such as electromagnetism and is therefore much more trivial. To understand the physical contents of the theory one could just as well stop after obtaining (10.7) which is expected to describe Abelian charged particles and photons.

A variant on the Lorentz condition that reduces R to R/(G/L) is easy to find :

$$D^o_\mu A^{ch}_\mu = 0 \tag{10.9}$$

there D^o_μ is the L-covariant derivative, containing the diagonal part of the vector field A_μ only. A^{ch}_μ is the set of off-diagonal elements of A_μ only. Eq. (10.9) has indeed N^2-N components.

This gauge suffers from the ghost problem as much as the ordinary Lorentz gauge, and is therefore not suitable for understanding all physical degrees of freedom.

XI. UNITARY GAUGE

A unitary gauge must be picked in a way similar to the Abelian case. We need a field that transforms without derivatives under gauge transformations. We will limit ourselves to the case that this field, call it X, transforms as the adjoint representation under G.

$$X \to \Omega X \Omega^{-1} \tag{11.1}$$

Such a field namely can always be found. The simplest choice would be

$$X^{ij} = G^{ij}_{12} \tag{11.2}$$

which is one of the components of the covariant curl $G_{\mu\nu}$. This choice has the disadvantage of not being Lorentz-invariant. One may choose a composite field :

$$X^{ij} = G^{ik}_{\mu\nu} G^{kj}_{\mu\nu} \tag{11.3}$$

This however does not work if G = SU(2) because then X would be proportional to the identity matrix. We need a non-vanishing isovector part. We could choose

$$X^{ij} = G^{ik}_{\mu\nu} D^2 G^{jk}_{\mu\nu} \tag{11.4}$$

but this choice looks rather complicated. Perhaps the most practical choice would be to take ones refuge to an extra scalar field in the theory, giving it a sufficiently high mass value so that the theory is not changed perceptibly at low energies.

Our gauge condition will be that X is diagonal :

$$X = \begin{pmatrix} \lambda_1 & & & & 0 \\ & \lambda_2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \\ 0 & & & & \lambda_N \end{pmatrix} \tag{11.5}$$

where the eigenvalues λ_i may be ordered :

$$\lambda_1 > \lambda_2 > \dots > \lambda_N \tag{11.6}$$

What is the subgroup of the gauge transformations Ω under which (11.5) is invariant ? If we require

$$X' = \Omega X \Omega^{-1} = X$$

then

$$[X,\Omega] = 0 \tag{11.7}$$

therefore, Ω is also diagonal :

$$\Omega = \begin{pmatrix} e^{i\omega_1} & & & 0 \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & e^{i\omega_N} \end{pmatrix} \tag{11.8}$$

and since $\det\Omega = 1$, we have

$$\sum_{i=1}^{N} \omega_i = 0 \tag{11.9}$$

Indeed, this is the largest Abelian subgroup L of G. If we write

$$\Omega = e^{i\omega} \tag{11.10}$$

then the diagonal part A^o_μ of A_μ transforms as

$$A^o_\mu \to A^o_\mu - \frac{1}{g}\partial_\mu \omega \tag{11.11}$$

and the off-diagonal part A^{ch}_μ as

$$A^{ch\ ij}_\mu \to e^{i(\omega_i - \omega_j)} A^{ch\ ij}_\mu \tag{11.12}$$

Apart from the gauge transformations (11.11) and (11.12) all our fields are physically observable. So our physical degrees of freedom are

- N-1 "massless" photons
- 1/2 N(N-1) "massive" charged vector fields
- N scalar fields λ_i with the restriction : $\lambda_1 > \lambda_2 > \dots > \lambda_N$.

There is of course another constraint : depending on our choice for

X, we have that

$$\begin{pmatrix} \lambda_1 & & & \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ & & & & \lambda_N \end{pmatrix}$$

satisfies (11.2), (11.3), or (11.4). Of course we still have the local $U(1)^{N-1}$ symmetry to be removed by either conventional gauge fixing, or the procedure described in the previous chapters.

XII. A TOPOLOGICAL OBJECT

So far we assumed that the eigenvalues $\lambda_1, \ldots, \lambda_N$ coincide nowhere. What if they do, at some set of points in space-time ? At those points, the invariance group is larger and a problem emerges with our enumeration procedure. We now argue that these exceptional points : a) are pointlike in 3-space, describing particle-like trajectories in space-time, and b) correspond to singularities in the fields A_μ and ψ. The argument for the singularity is similar to the Abelian case, but the fact that these things are particle-like differs from the Abelian case, where the singular points were string-like.

The reason why the singularities in the generic case are pointlike is that the dimensionality of the space of field variables with two coinciding eigenvalues λ is three less than the space with non-coinciding eigenvalues. Therefore the dimensionality of the points in space or space-time where two eigenvalues coincide is three less than that of space-time itself. This statement holds for any generic N x N hermitian matrix field X(x).

What is the physical nature of such a particle-like singularity ? Since the eigenvalues λ_i were ordered, we only need to consider the case that two successive λ's coincide :

$$\lambda_j = \lambda_{j+1} \tag{12.1}$$

for certain j. Let us consider a close neighbourhood of such a point. Prior to the gauge-fixing we may take X to be (12.2), where D_1 and D_2 may safely be considered to be diagonalized because the other eigenvalues did not coincide. The three fields $\varepsilon_a(x)$ are small because we are close to the point where they vanish. With respect to that SU(2) subgroup of SU(N) that corresponds to rotations among the j^{th} and $j + 1^{st}$ components, the fields ε_a form an isovector.

$$X = \begin{array}{c} \\ j \\ \\ j+1 \\ \\ \end{array} \left[\begin{array}{c|cc|c} D_1 & & 0 & 0 \\ \hline & \lambda + \varepsilon_3 & \varepsilon_1 - i\varepsilon_2 & \\ 0 & & & 0 \\ & \varepsilon_1 + i\varepsilon_2 & \lambda - \varepsilon_3 & \\ \hline 0 & & 0 & D_2 \end{array}\right] \qquad (12.2)$$

$$\qquad\qquad j \qquad j+1$$

One may write the center block as

$$X = \lambda I + \vec{\varepsilon}.\vec{\sigma} \qquad (12.3)$$

where σ_a are the Pauli spin matrices. Close to a zero point of this $\vec{\varepsilon}$ field, the field $\vec{\varepsilon}$ has a hedgehog configuration. But gauge fixing i.e. diagonalization of X, corresponds to rotating ε_1 and ε_2 away such that ε_3 is positive $(\lambda_j > \lambda_{j+1})$. Thus is our unitary gauge,

$$\vec{\varepsilon} = \begin{pmatrix} 0 \\ 0 \\ +|\varepsilon_3| \end{pmatrix} \qquad (12.4)$$

By now the reader may recognize this field configuration as the one for a magnetic monopole[9]. Indeed, fixing the L-gauge as well cannot be done without accepting a string-like singularity connecting zeros of opposite signature : the Dirac string.

The magnetic charges of the monopole can most easily be characterized with respect to the $U(1)^N$ subgroup of the extended gauge group U(N) :

$$\vec{m} = \left(0, \ldots, 0, \frac{2\pi}{g}, -\frac{2\pi}{g}, 0, \ldots, 0\right) \qquad (12.5)$$

where the $\pm\, 2\pi/g$ are at the j^{th} and $j+1^{st}$ position. g is here the fundamental electric charge of the elementary representation. We then see that $\vec{m}$ actually only acts in the subgroup $U(1)^N/U(1)$ of SU(N) because the sum of all its charges vanishes. It is constructive to notice a subtile difference between this magnetic charge spectrum and the spectrum of the electrically charged gauge particles A_μ^{ij} : from (11.12) we read off that these have electric charges

$$\vec{Q} = (0, \ldots, 0, +g, 0, \ldots, 0, -g, 0, \ldots, 0) \tag{12.6}$$

where ± g occur at arbitrary, not necessarily adjacent, positions i and j. Again however the sum of the charges vanishes, so that we are really working in the Cartan group $U(1)^{N-1}$, not $U(1)^N$. Magnetic monopoles with

$$\vec{m} = (0, \ldots, 0, \frac{2\pi}{g}, 0, \ldots, 0, -\frac{2\pi}{g}, 0, \ldots, 0) \tag{12.7}$$

are possible only if three or more eigenvalues coincide. The dimensionality of such points is at least 8 less than space-time so in general they do not occur at single points. Rather, they should be considered as bound states of "elementary" monopoles (12.5).

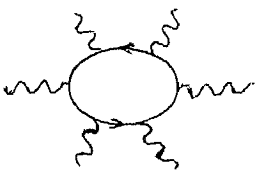
We conclude that we arrive at a picture where an Abelian gauge theory is enriched with magnetic monopoles, but because of the slightly different charge spectrum this picture is in general not "self dual". Contrary to the case discussed in section VII, the monopoles we have here are "dynamical", that is, they will inevitably take part in the dynamics of the system.

XIII. THE MACROSCOPIC VARIABLES

We have now arrived at a point where we could sketch a possible strategy for precise calculations for the dynamics of the system :

1) Consider the physical degrees of freedom in the space $H_1 = R/(G/L)$. We find N-1 sets of Maxwell fields, electrically charged fields (among which vector fields), and magnetically charged particles. The particular case of interest is now the possibility that magnetically charged particles "Bose-condense". If we ever are to understand such a mechanism in detail, the following step is probably necessary :

2) Eliminate the electric charges. With as much precision as possible we must compute all light-by-light scattering amplitudes and express them in term of an effective interaction lagrangian for the photon fields :

$\Gamma(A_\mu) = \sum$  + higher orders (13.1)

3) Now perform the "dual transformation". Since we have only Maxwell fields and magnetic charges interacting with them, we could replace $\vec{B}$ by $\vec{E}$ and $\vec{E}$ by $-\vec{B}$, then introduce operator fields in the usual way for the monopole particles, which now look like ordinary electrically charged objects.

4) Work out the self-interactions among these magnetic monopoles. Set up a perturbation theory now in terms of $2\pi/g$. Then the question is :

5) Does, in terms of this perturbation theory, Bose condensation occur among these monopoles ? Is it reasonable to start with

$$< 0|\phi_{mon}|0 > \neq 0 \quad ? \tag{13.2}$$

If so, then the vacuum is a magnetic superconductor. The monopoles formally have a negative mass-squared. In this magnetic superconductor electric charges are confined. The descriptions of section VII apply qualitatively, after the interchange electric $\leftrightarrow$ magnetic.

XIV. THE DIRAC CONDITION IN THE ELECTRIC-MAGNETIC CHARGE SPECTRUM

Fig 2 represents the spectrum of possible charges in the case that the gauge group G is SU(2).

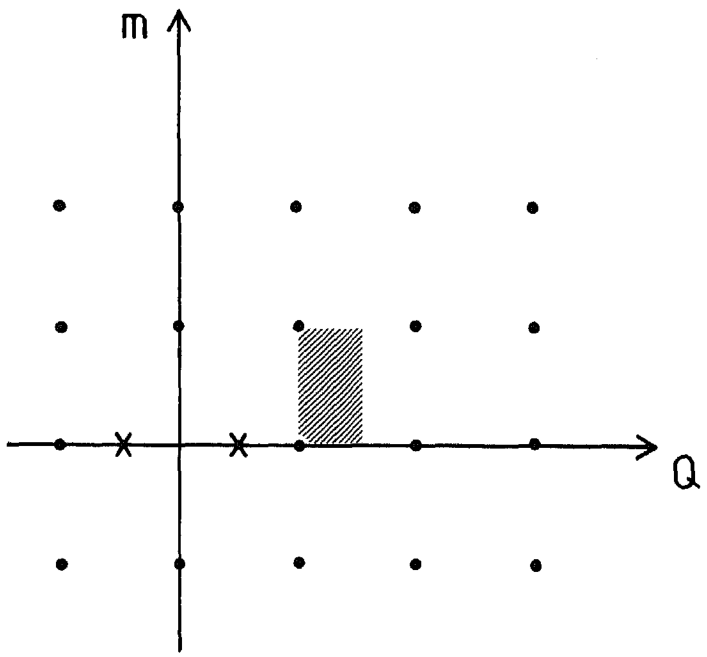


Fig 2

Horizontally is plotted the electric charge Q, vertically the magnetic charge $\vec{m}$. Elementary and bound state charges are indicated. The crosses represent a fundamental SU(2) doublet which may or may not have been added to the theory. The pure Maxwell field equations are now invariant under rotations of the figure about the origin. This is why one is always free to postulate that the fundamental fields carry no magnetic monopole charge. But does the lattice obtained in fig 2 have to be rectangular ? We will now argue that Dirac's quantization condition allows more kinds of lattices and "oblique" lattices indeed may result in a non-Abelian gauge theory.

The Dirac condition for a magnetic charge quantum m and the electric charge quantum q was

$$qm = 2\pi n \tag{14.1}$$

where n is integer. To be precise this corresponds to a quantization condition for the Lorentz force that the magnetically charged particle exerts on the electrically charged object. But now consider two particles 1 and 2 both with various kinds of magnetic and electric charges. Then the Lorentz force quantization corresponds to

$$\sum_i (q_i^{(1)} m_i^{(2)} - q_i^{(2)} m_i^{(1)}) = 2\pi n_{12} \tag{14.2}$$

where the index i refers to the label of the species of photons and n_{12} is an integer relevant for particles 1 and 2, to be referred to as the Dirac quantum of particle (1) with respect to particle (2).

In our specific case of SU(N) broken down to $U(1)^{N-1}$ we usually require

$$\sum_{i=1}^{N} q_i = 0 \qquad\qquad \sum_{i=1}^{N} m_i = 0 \tag{14.3}$$

Our charge lattice in this case will be spanned by 2(N-1) basic charges, to be labelled by an index A = 1, ..., 2N-2. Because of the invariance

$$\begin{pmatrix} m_i \\ \\ q_i \end{pmatrix} \rightarrow \begin{pmatrix} \cos\phi_i & \sin\phi_i \\ \\ -\sin\phi_i & \cos\phi_i \end{pmatrix} \begin{pmatrix} m_i \\ \\ q_i \end{pmatrix} \tag{14.4}$$

we may always take

$$m_i^{(A)} = 0 \qquad \text{for } A = 1, \ldots, N-1 \tag{14.5}$$

The gluons will provide us with a basis of electric charges:

$$q_i^{(A)} = g\delta_i^A - g\delta_i^{A+1} \quad \text{for } A = 1, \ldots, N-1 \tag{14.6}$$

(The fundamental representation, if it occurs, could have

$$q_i^{(k)} = g\delta_{ik} - g/N \text{).} \tag{14.7}$$

The magnetic monopoles give the remaining basic charges:

$$m_i^{(A)} = \frac{2\pi}{g}\,\delta_i^{A+1-N} - \frac{2\pi}{g}\,\delta_i^{A+2-N} \quad \text{for } A = N, \ldots, 2N-2 \tag{14.8}$$

It was Witten[10)] who observed that monopoles may also carry electric charges. He found

$$q_i^{(A)} = \frac{\theta g^2}{4\pi^2}\, m_i^{(A)}, \quad \text{for } A = N, \ldots, 2N-2 \tag{14.9}$$

where θ is the instanton angle of the theory, $0 \leqslant \theta < 2\pi$. Notice that for any value of θ the Dirac condition (14.2) is fulfilled.

It can be seen that this phenomenon, eq. (14.9), follows from the lagrangian

$$\mathcal{L} = -\frac{1}{4}G^a_{\mu\nu}G^a_{\mu\nu} + \frac{\theta i g^2}{32\pi^2}\, G^a_{\mu\nu}\tilde{G}^a_{\mu\nu} + A^a_\mu J^a_\mu \tag{14.10}$$

there

$$\tilde{G}^a_{\mu\nu} = \frac{1}{2}\,\varepsilon_{\mu\nu\alpha\beta}\, G^a_{\alpha\beta} \tag{14.11}$$

and $G^a_{\mu\nu}\tilde{G}^a_{\mu\nu}$ corresponds to

$$4\sum_a \vec{E}^a . \vec{B}^a \tag{14.12}$$

The canonical argument can be found in refs 10), 11), 8).

In the case of SU(2) the charge lattice indeed becomes tilted now (Fig 3). It is remarkable that if θ runs from 0 to 2π then the charge lattice indeed turns back into itself, but the "elementary" monopole labelled by (2) in fig 3 is replaced by the "monopole gluon bound state" labelled (3). It seems that no fundamental distinction

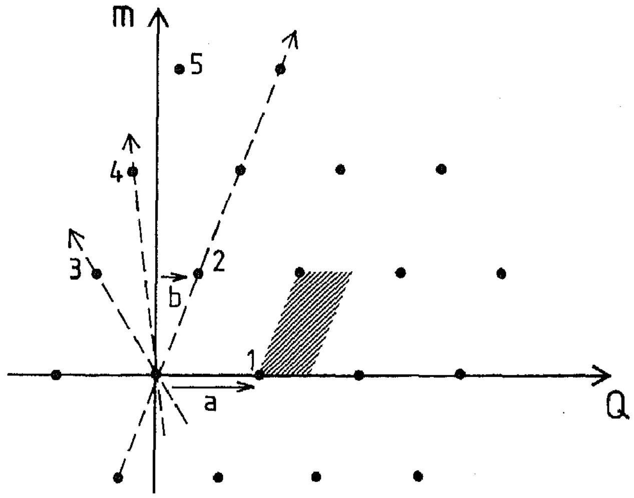


Fig 3

will be possible between monopoles and monopole-gluon bound states.

XV. *OBLIQUE CONFINEMENT*

We can now ask which of the objects in the lattice of fig 3 will form a Bose-condensate. If it is a purely electrically charged object, (1) in fig 3, then we get the familiar Higgs theory or electric superconductor. All charges that are not on the horizontal axis will then be confined by linear potentials, because of the arguments presented in section VII. By duality we now expect that also monopoles may conceivably undergo Bose condensation, for instance charge #(2) in fig 3. One cannot however have both electric and magnetic charges Bose-condense because if the electric ones condense then the magnetic ones have $m^2 \to +\infty$, not negative, and vice versa. In the case of larger N, Bose condensation can only take place among charges forming any linear sublattice of the original charge lattice, as long as its members all have vanishing Dirac quanta with respect to each other.

Now if θ is switched on, then point (3) gradually takes the place of (2). (2) and (3) cannot both Bose condense because they have a non-vanishing relative Dirac quantum. So it is either (2) or (3). It is likely that Bose condensation in the (2)-direction is replaced by condensation in the (3)-direction at $\pi < \theta < 3\pi$. This would then be a *phase transition* in θ, possibly of first order, just like the transition between Higgs and confinement.

Various attempts at dynamical calculations however indicate that at $\theta \simeq \pi$ the confinement mechanism is not strong. An explanation could simply be that the monopoles then carry large electric charges and therefore may have larger self-energies contributing positively to their mass-squared. Suggestions have been made that at $\theta \simeq \pi$ the Higgs mode reappears [12] or a Coulomb mode (no Bose condensation at all).

I suggest yet a different condensation mode that could possibly occur in theories with θ close to π. If neither (2) nor (3) condense because they carry large (but opposite) electric charges, then perhaps (4) which is a bound state of these two with much smaller electric charge condenses. This would only be possible if the lattice is oblique ($\theta \neq 0$) so this mode is referred to as "oblique confinement". A theory with oblique confinement shows some peculiar features. We stress that these will not occur in ordinary QCD because there we know that $\theta \simeq 0$. Our observations may be relevant for certain models with "technicolor" as we will show shortly.

Returning to the case that our gauge group was SU(2), we first argue that the "quarks" (or "preons") in this oblique confinement mode are not confined in the usual sense. That is, if we attach a flavor quantum number to every type of preon, then physical particles transforming as the fundamental representation of the flavor group do occur. The preons are the crosses in fig 4.

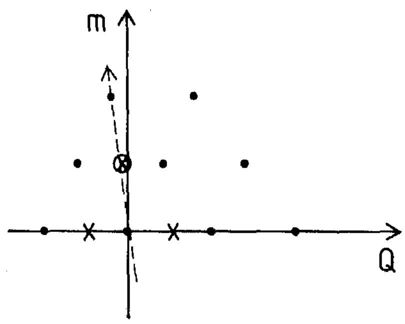


Fig 4.

Since they are not on the line connecting the origin to the condensed object (arrow) they are confined. However, a bound state of a monopole (without flavor quantum numbers) and one preon does occur on that line (⊗), and hence is physical, and has the same flavor as the preon itself. But a price had to be paid : spin and statistics-properties of the physical object are opposite to that of the original preon! If the preon is a fermion, the liberated object is a boson and vice-versa. This is a consequence of a general rule : if two particles have an odd relative Dirac quantum, then the *orbital* angular momentum in any bound state of the two is half-odd-integer, a well-known property of the Schrödinger equation of an electrically charged particle in the field of a magnetic point-source[13)]. That also the *statistics* of the bound state gets an extra fermionic contribution has been shown by Goldhaber[14)]. Let me give an outline of the argument.

If we wish to consider the statistics of two identical composite states A and B, both composed of a magnetic monopole M and an electric charge Q, (fig 5) then we would like to separate the center-of-mass motions of A and B from the orbital motions of M and Q inside A and B.

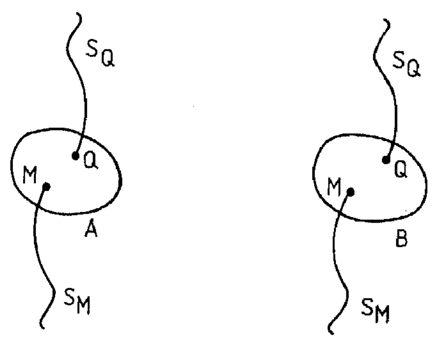


Fig 5

The particles Q see two Dirac strings, ending at both M's. The M's see two Dirac strings ending at both Q's. Now the center-of-mass motion of A can only be split from the orbital motion of M and Q if the magnetic strings run in a direction opposite to that of the electric strings (see fig 5), so that if M hits an electric string, in the same time Q hits a magnetic string. Now let us ignore the motion of M and Q inside A and B, but only consider A and B as a whole. Then A feels both strings at B (and vice versa), in fact, these two are connected in such a way that one string results, running from infinity to infinity. This could be expected because A and B have the same electromagnetic charge combination (they were identical) so *their* relative Dirac quantum vanishes. Their relative motion (A against B) is as if they only were electrically charged. Therefore we obtain the familiar Coulomb Schrödinger equation by removing this complete string by a single gauge transformation :

$$\psi_{AB} \to e^{i\phi_{AB}}\psi_{AB} \tag{15.1}$$

where ϕ_{AB} is the angle by which A rotates around the B-string.

Now notice that if we interchange A and B this angle is 180°, so that

$$e^{i\phi_{AB}} = -1$$

This is how one can see an extra minus sign appearing in the commutation properties of the particles A and B.

XVI. FERMIONS OUT OF BOSONS AND VICE-VERSA

Some exotic models can be constructed if we use oblique confinement as a starting point. Consider for instance an SU(3) gauge theory without fermions but with a scalar field ϕ in the fundamental representation. Let θ be close to π, and let us assume that the familiar Higgs mechanism takes place, described by

$$<\phi> = \begin{pmatrix} 0 \\ 0 \\ F \end{pmatrix} \tag{16.1}$$

(through formally incorrect, as explained in section V, the notation is useful and adequate for our purpose). SU(3) then "breaks spontaneously" into SU(2). One heavy neutral Higgs particle arises. The original octet of gauge fields splits into one neutral heavy vector particle and one SU(2) complex doublet of heavy vector particles.

The SU(2) triplet remains massless, and now we will assume that it produces oblique confinement in the SU(2) sector. The SU(2) doublet of heavy vector bosons will appear as physical particles but disguised as fermions! Here we have an example of fermionic gauge "bosons" without invoking anything resembling supersymmetry.

Another model worth considering is a simplistic weak interaction theory, based upon $SU(2)^{tc} \times (SU(2) \times U(1))^{ew}$. There "tc" stands for "technicolor" and ew for "electro-weak". Quarks and fermions are all in the usual representation of $(SU(2) \times U(1))^{ew}$ and singlets under $SU(2)^{tc}$. Now add one fermion multiplet transforming just like all other fermions under $(SU(2) \times U(1))^{ew}$ but also as a 2 under tc. Assume oblique confinement. Then this fermion will be liberated, but disguised as a boson. Probably it will have a spin-zero component. Since it is in line with the condensing monopole bound states it may Bose-condense itself. So we obtain a scalar field transforming just as the fermions under $(SU(2) \times U(1))^{ew}$ and a non-vanishing vacuum expectation value : a model for the Higgs particle. Indeed, its Bose condensation could well be responsible for the oblique confinement mode in the first place, so here it was not even necessary to consider the monopole-dyon bound state.

Unfortunately, elegant as it may be, this model seems to suffer from the same shortcomings as the more conventional technicolor ideas : it is hard to reproduce the required Yukawa couplings between this scalar field and the other fermions.

XVII. OTHER CONDENSATION MODES

It will be clear from the previous sections, by looking at the electric-magnetic charge lattice, that even more exotic forms of oblique confinement can be imagined. Just assume condensation of bound states with three or more monopoles. Such a condensation mode would be required for instance if we would wish to liberate the fundamental triplets in an SU(3) gauge theory. A fundamental exercise tells us then that these triplets do not switch their spin-statistics properties [8]. In any case, all these different confinement modes will be separated from each other by sharp phase transition boundaries, which should show up in the solutions of the theory when the parameter θ is varied.

In principle each point on the electric-magnetic charge lattice may correspond to a possible phase of the system. There may however be features which cannot easily be understood in terms of our intermediate physical degrees of freedom with Abelian electric and magnetic charges. We have in mind a condensation mode studied in more detail by Bais[15]. The simplest example is an SU(2) gauge theory with an isospin-two Higgs field, ϕ^{ab} (a, b = 1, 2, 3).

Let us assume

$$< \phi^{ab} > = F^{(a)}\delta^{ab} \qquad (17.1)$$

If we were just dealing with a global symmetry, we would say that SU(2) is spontaneously broken into a subgroup D (the invariance group of eq. (17.1)). Now D is the discrete subgroup of SU(2) corresponding to rotations of spinors over 90° :

$$D = (\pm I, \pm i\sigma^1, \pm i\sigma^2, \pm i\sigma^3) \qquad (17.2)$$

If the symmetry is local then D is not really a global invariance of the vacuum. What we do see is that magnetic vortex tubes can be constructed which are characterized by the following boundary condition at infinity :

$$A_\mu \to \frac{i}{g}\Omega\partial_\mu\Omega^{-1} \qquad (17.3)$$

where Ω is multivalued. If we go around the vortex once, then Ω turns into itself multiplied with an element D_1 of D. These vortices are non-commuting. Physically this means the following (fig 6)

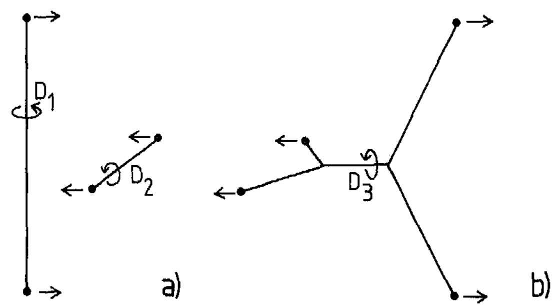


Fig 6

If two strings, characterized by different, non commuting elements D_1 and D_2, approach each other at right angles (fig 6a), then they cannot pass each other without leaving a connecting string D_3 (fig 6b). The element D_3 is given by

$$D_3 = D_1 D_2 D_1^{-1} D_2^{-1} \qquad (17.4)$$

Clearly the non-commuting properties of the original gauge group were crucial for understanding this phenomenon, so that our Abelian physical variables are not useful here. Indeed, we could ask the question whether the dual of this "Bais mode" exists, with electric strings having similar properties ? As yet, the answer to that question is unknown.

REFERENCES

1. J. Goldstone, Nuovo Cim. 19 154 (1961)
2. F.S. Abers and B.W. Lee, Phys Reports 9C, 1 (1973) and references therein.
3. P.W. Higgs, Phys. Lett. 12, 132 (1969), Phys. Lett. 13, 508 (1964), Phys. Rev. 145, 1156 (1966)
   F. Englert, R. Brout, Phys. Rev. Lett. 13, 321 (1964)
   G.S. Guralnik, C.R. Hagen and T.W.B. Kibble, Phys. Rev. Lett. 13 585 (1964)
   T.W.B. Kibble, Phys. Rev. 155, 627 (1967)
4. B.S. De Wit , Phys. Rev. 162, (1967), 1195, 1239
5. G. 't Hooft and M. Veltman, "DIAGRAMMAR", CERN report 73/9 (1973)
6. H.B. Nielsen, P. Olesen, Nucl. Phys. B61, 45 (1973)
   B. Zumino, in "Renormalization and Invariance in Quantum Field Theory ed. E.R. Caianiello Plenum Press New York (1974) p. 367
7. P.A.M. Dirac, Proc. Roy. Soc. A133, 60 (1931)
   Phys. Rev. 74, 817 (1948)
8. G. 't Hooft, Nucl. Phys. B, to be published
9. A.M. Polyakov JETP Lett. 20, 194 (1974)
   G. 't Hooft, Nucl. Phys. B79, 276 (1974)
10. E. Witten, Phys. Lett. 86B, 283 (1979)
11. A. Salam and J. Strathdee, Lett. in Mathematical Physics 4, 505 (1980)
12. C. Callan, private communication (1979)
13. R. Jackiw and C. Rebbi, Phys. Rev. Lett. 36, 1116 (1976)
    G. 't Hooft and P. Hasenfratz, Phys. Rev. Lett 36, 1119 (1976)
14. A.F. Goldhaber, Phys. Rev. Lett. 36, 1122 (1976)
15. F.A. Bais, private communication.