PHYSICAL REVIEW D VOLUME 9, NUMBER 12 15 JUNE 1974

Vacuum polarization and the absence of free quarks in four dimensions

J. Kogut*
Laboratory of Nuclear Studies, Cornell University, Ithaca, New York 14850

Leonard Susskind†‡
Belfer Graduate School of Science, Yeshiva University, New York, New York 10033 and Tel Aviv University, Tel Aviv, Israel
(Received 25 March 1974)

We speculate that the mechanism of one-dimensional quantum electrodynamics for prohibiting quark production can be generalized to three-dimensional space. The important feature of one-dimensional gauge theories which trap quarks is that electric flux lines cannot spread out and can only end on a charge. The generalization to three dimensions involves dynamical nonlinearities which trap electric flux lines into tubelike configurations. It is suggested that the required nonlinearities arise in Yang-Mills theories of "colored" quarks. The large-scale properties of hadrons would be described by a stringlike model similar to the conventional dual model.

I. INTRODUCTION

It has been suggested that vector gauge theories containing fundamental quark fields may prohibit the existence of free quarks.[1,2] The crucial idea underlying this suggestion is the possibility of the Schwinger phenomenon[3] occurring in gauge theories of "colored" quarks. Previously, we have not addressed in much detail the important question of whether the Schwinger phenomenon is possible outside of one-dimensional field theories. Unfortunately, very little physical insight into this problem exists. Here, however, we shall try to create some semiclassical intuitions which we hope will inspire a more complete answer to this important question.

We will say (by definition) that the Schwinger phenomenon[3] occurs when a gauge theory undergoes a transition to a new, "third" phase. This phase is different from the conventional and Higgs-Englert-Brout (HEB) phases.[4] In the conventional phase free charges can exist accompanied by their long-range Coulomb fields. The spectrum of states includes massless gauge bosons whose source is the generator of the gauge transformation (of the first kind).

In the HEB phase[4] gauge invariance of the first kind is spontaneously broken and the vacuum is not a group singlet. In this case the vector gauge particle gains a mass and the long-range Coulomb force is converted to a short-range force. The particles generally form nondegenerate multiplets under the symmetry group as if the symmetry were broken.

In the Schwinger phase the massless gauge boson is eliminated either by disappearing altogether or becoming massive. Accordingly, the long-range Coulomb force is also eliminated. However, in contrast with the HEB case, the vacuum remains a singlet. The gauge symmetry is *not* spontaneously broken, and instead, all states with nonvanishing charge (nonsinglets in the non-Abelian theories) have infinite energy.

Consider the possibility that the massless gauge bosons disappear. Then it follows by well-known arguments[1,3] that the long-range r^{-1} component of the Coulomb field must also be absent. However, it is a special feature of gauge theories that the strength of the long-range field is connected to the total value of the charge through Gauss's theorem. It follows that the total charge of any isolated system must vanish. If the relevant charge for strong interactions is color, then all colored states are forbidden and free quarks will not exist.

It is interesting to see why this logic can fail in the HEB phase. Consider an Abelian gauge field A_μ coupled to two fermion fields ψ and χ,

$$\mathcal{L} = -\tfrac{1}{4}F^{\mu\nu}F_{\mu\nu} + \bar{\psi}\, i\gamma^{\mu}\partial_{\mu}\psi + \bar{\chi}\, i\partial^{\mu}\gamma_{\mu}\chi + m\bar{\psi}\psi + m\bar{\chi}\chi - e(\bar{\psi}\gamma^{\mu}\psi - \bar{\chi}\gamma^{\mu}\chi)A_{\mu}\,. \tag{1}$$

By construction the fermions ψ and χ have opposite charges. Now suppose that the symmetry operation

$$\psi \to e^{i\alpha}\psi, \quad \chi \to e^{-i\alpha}\chi \tag{2}$$

is spontaneously broken by a vacuum expectation value of $\bar{\psi}\chi + \bar{\chi}\psi$. In this event the photon becomes massive and the previous argument would indicate that all states must have zero charge. But all zero-charge states must be constructed from an even number of fermion operators which would lead to the (erroneous) conclusion that this model has no fermions.

The trouble with the argument is that the vacuum expectation value $\langle \bar{\psi}\chi + \bar{\chi}\psi \rangle_0$ causes transitions from positively charged to negatively charged fermions. The physical fermions are more like superpositions,

$$|f_{\pm}\rangle = |\psi\rangle \pm |\chi\rangle \ . \tag{3}$$

The two states $|f_{\pm}\rangle$ are split in mass and have zero average charge. However, neither is an eigenvector of the charge with 0 eigenvalue.

In order to conclude that there are no fermions in models we require not only that the average charge of any state vanish, but also that the states be eigenvectors of the charge. In other words, we require that the symmetry not be spontaneously broken.

The differences between the Schwinger and HEB phenomena have not been sufficiently appreciated in the literature. The crucial mechanisms which cancel the long-range effects in the photon propagator are in fact completely different. In the HEB case the effect is due to a massless scalar bound state in the proper vacuum-polarization tensor, while in the Schwinger case a vector positronium-like state shifts the photon pole.[3]

A question which is closely related to quark confinement by the Schwinger mechanism is the infrared behavior in non-Abelian gauge theories.[5] It is known that the infrared divergences of these theories do not add up as innocently as in quantum electrodynamics. More likely the infrared singularities mount up and produce a dramatic change in the structure of the vacuum. One way for the infrared problem to resolve itself is through a spontaneous breakdown in the manner of Higgs, Englert, and Brout. Another way, which occurs in one-dimensional quantum electrodynamics (which also is violently infrared-divergent), is through the Schwinger mechanism.

II. WHY VECTOR FIELDS?

Consider a quark pair produced at the space-time origin with high relative energy. The quarks begin to recede from one another with almost the speed of light (define $c=1$). Assuming that the quarks are coupled to some gluon field, this field will become excited in the region between the outgoing quarks. Three possibilities exist:

(1) The field energy never becomes too large and the quarks escape.

(2) The field energy grows indefinitely between the two quarks, giving rise to a force which eventually prevents the escape.

(3) Possibility (2) is modified by the field finding a way to lower its energy by pair creation. The produced pairs combine with the outgoing quarks to produce hadrons.

In cases (2) and (3) it is essential that non-negligible fields be produced in the region between the outgoing quarks and that these fields not fall to zero as the energy of the process increases.

We shall represent the outgoing quarks as c-number sources of the gluon field. When the energy tends to infinity the sources become points on the light cone. First, consider the case of scalar gluons satisfying

$$(\Box + m^2)\,\phi = \rho\,, \tag{4}$$

where ρ describes a scalar source composed of a pair of points separating with velocity $v \approx 1$. Lorentz contraction of the source requires

$$\int \rho\, dz = \text{const} \times (1 - v^2)\,. \tag{5}$$

Thus,

$$(\Box + m^2)\,\phi = (1 - v^2)[\delta(z - vt) + \delta(z + vt)]\,\delta(x)\,\delta(y) \tag{6}$$

for $t > 0$. As $v \to 1$ in the high-energy limit, the resulting field between the quarks tends to zero as $(1 - v^2)$. For this reason the field energy in the scalar case does not inhibit quark production.[6]

In the vector-gluon case the source is a vector current. As the source approaches the light cone its dimensions again become Lorentz-contracted. However, this time the Lorentz-contraction factor is precisely offset (for the z and t components) by the transformation properties of vector currents. Thus, in the region between the quark pair the vector potential satisfies an equation similar to

$$\begin{aligned} \Box A_0 &= [\delta(z - t) - \delta(z + t)]\,\delta(x)\,\delta(y)\,, \\ \Box A_z &= [\,\delta(z - t) + \delta(z + t)]\,\delta(x)\,\delta(y)\,. \end{aligned} \tag{7}$$

Thus the field does not fall to zero as $v \to 1$.

Another way to see this in a gauge theory is to observe that each charge is a source or sink of g lines of electric flux. These lines must begin on one quark and end on the other. Furthermore, causality requires the lines to be contained within a sphere of radius t (see Fig. 1). Since the

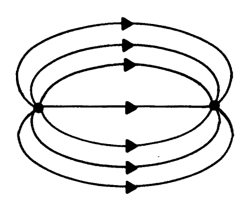


FIG. 1. Electric field of a high-energy pair a short time after their production.

electric field is the number of flux lines per unit area, the field midway between the quarks is at least

$$E \sim g/t^2 . \tag{8}$$

The important point is that this field does not fall to zero with increasing energy.

However, in three dimensions, if the field develops according to the standard linear Maxwell equations, the field energy (after ultraviolet self-energy effects are removed) will remain finite even as the quarks recede to infinity. It is this fact which distinguishes three-dimensional quantum electrodynamics from its one-dimensional counterpart: In one dimension the Coulomb energy is proportional to the distance between receding particles.

If quarks are not to be produced in three dimensions, it is essential that nonlinearities come into play and distort the field lines. In particular, we shall try to make plausible the idea that nonlinearities prevent the flux lines from radiating away from a charge and forming the customary Coulomb field. Instead, the lines of flux are squeezed into one-dimensional tubelike configurations as in Fig. 2. In this case the field configuration associated with a receding pair would consist of a tube of flux stretched between the quarks as in Fig. 3. When the flux tube is very long, we may calculate its energy by assuming an energy per unit length for the uniform portion of the tube far from its ends. The energy would then be proportional to the distance between the quarks as in one-dimensional gauge theories.

As in one-dimensional quantum electrodynamics (QED), we expect pair production in the field between the quarks.[1] The new pair provides new end points and allows the flux tube to break as in Fig. 4. This process screens the long-range force between the quarks, and, if it works as in the one-dimensional theory, only a short-range force remains.

III. FLUX TUBES AND STRINGS IN ONE-DIMENSIONAL QED

In this section we will give a qualitative description of the "quark-elimination" mechanism

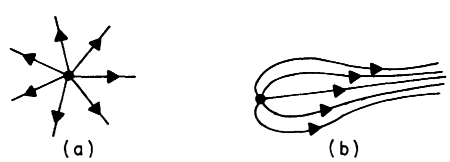


FIG. 2. (a) Flux lines associated with the spherically symmetric Coulomb field. (b) Distortion of the Coulomb field into a flux tube.

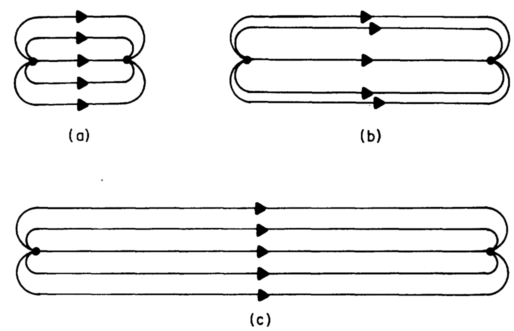


FIG. 3. Time development of the electric field of a high-energy pair in the flux-tube model. (a)–(c) represent successive stages of evolution of the tube.

in one-dimensional gauge models.[1] Consider a pair of oppositely charged quarks separated by distance Δz in one-dimensional quantum electrodynamics. The charge of each quark is $\pm g$. In the interval between the quarks there is a constant electric field of magnitude g. We will picture this situation by saying that each charge is a source or sink of a single flux line which must not end except on a charge. The energy stored in the flux line is proportional to its length. This follows either from the one-dimensional Coulomb law,

$$V(z_1 - z_2) = g^2 |z_1 - z_2| , \tag{9}$$

or from the fact that the energy density of the field is the square of the electric field. If we consider the flux tube (Fig. 5) to be a material system, then it is characterized by having a tension

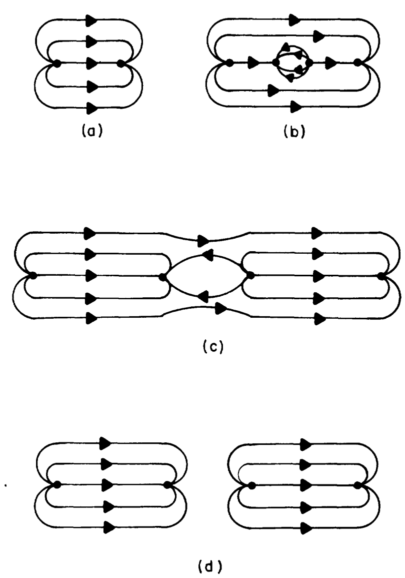


FIG. 4. The breaking of a flux tube and the associated production of a pair. (a)–(d) represent successive stages of evolution.

FIG. 5. A flux line in one-dimensional gauge theory.

which is independent of its length.[7]

As we have mentioned, stretching the flux line can result either in increasing its length until the available energy runs out or in breaking the flux tube into one or more segments. Since a flux line can only begin or end on a charge, breaking the flux line is equivalent to producing a pair. A detailed analysis in one-dimensional QED shows that at high energy the available energy is used up in breaking the flux line into a number of segments.[1] The segments each form a meson, and the resulting meson distribution uniformly populates the available rapidity interval. In space-time the process looks roughly like Fig. 6.

The shaded region of space-time in Fig. 6 is occupied by flux. The figure is very similar to diagrams which are drawn in the string model of hadrons. In this theory the shaded region would be called a *world sheet*. In fact, one-dimensional QED is a special, if trivial, example of the string model. To make this clear we can imagine computing an amplitude involving a single fermion loop. The amplitude in the one-loop approximation for a quantity like

$$\langle 0|\,\bar{\psi}(x)\,\psi(x)\,\bar{\psi}(0)\,\psi(0)|\,0\rangle \tag{10}$$

is a path integral over all closed paths[8] passing through the points 0 and x. The action consists of kinetic and potential terms, which can be identified with the boundary and interior of the loop, respectively. The kinetic term arises from the free Lagrangian of a propagating fermion and has the form of a line integral around the loop. More interestingly, the potential energy gives rise to an integral,

$$\int V\,dt = g^2 \int |z_1 - z_2|\,dt\;. \tag{11}$$

Here z_1 and z_2 are the coordinates of the intersections of the loop with a line of constant time (Fig. 7). Evidently, this piece of the action is given by the surface area of the interior of the loop. It may be identified with the action of a world sheet stretched between the fermion boundary. The result that the action of a world sheet is proportional to its area is well known in the string model and may be taken as a logical starting point for the derivation of the dual resonance model in higher dimensions.[9]

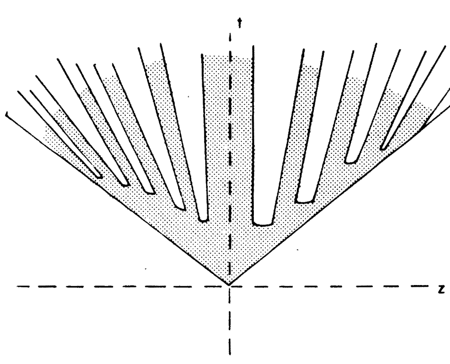


FIG. 6. World history of an annihilation event into hadrons in one-dimensional gauge theory. The shaded region is occupied by electric flux.

IV. ELECTROSTATIC PICTURE OF A FLUX TUBE

In Sec. V we will construct a phenomenological-covariant field theory which exhibits the flux tube effect. Here we introduce the basic ideas of the model in the simpler context of classical electrostatics of continuous media. (See the added note at the end of the paper.) Our discussion parallels that in Ref. 10. Recall that within a dielectric material polarization charges, ρ_P, are associated with a nonuniform polarization vector $\vec{P}$,

$$\rho_P = -\vec{\nabla}\cdot\vec{P}\,, \tag{12}$$

where $\vec{P}$ is the dipole moment per unit volume in the medium. Defining

$$\vec{D} = \vec{E} + 4\pi\vec{P}\,, \tag{13}$$

it follows that

$$\vec{\nabla}\cdot\vec{D} = 4\pi\rho_{\text{free}}\,, \tag{14}$$

where ρ_{free} is the external charge density implanted into the dielectric.[11] In a linear dielectric $\vec{D}$ and $\vec{E}$ are related by the dielectric permeability ϵ,

$$\vec{D} = \epsilon\vec{E}\,. \tag{15}$$

For our applications, ϵ may depend on spatial position. Similarly,

$$\vec{P} = \left(\frac{\epsilon - 1}{4\pi}\right)\vec{E} = \chi\vec{E}\,, \tag{16}$$

where χ is known as the dielectric susceptibility.

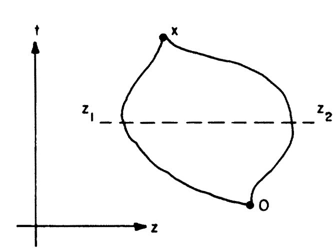


FIG. 7. A closed fermion loop in coordinate space.

The polarization charge ρ_P is then

$$\rho_P = -\vec{\nabla}\cdot\vec{P} = -\vec{\nabla}\cdot(\chi\vec{E})\,. \tag{17}$$

Finally, recall that the energy contained in the field is

$$W = \frac{1}{4\pi}\int \vec{E}\cdot\vec{D}\,d^3x\,. \tag{18}$$

The dielectric permeability ϵ can be proved to be greater than unity.[10] The idea behind the proof is that when an electric field is introduced into a dielectric the positive charges move in the direction of the field. This effect always reduces $\vec{E}$, the electric field in the medium. As a consequence the energy of a free charge is always less than its energy in the vacuum.

We will think of quarks as free charges embedded in a dielectric medium. However, in contrast with the conventional dielectric, we shall allow the susceptibility to be negative and will suppose that regions of space can exist in which ϵ may approach zero. Although this behavior is impossible in a conventional electrodynamical system, there are reasons to believe that similar effects occur in non-Abelian Yang-Mills theories. For simplicity allow ϵ to take on only the values zero or unity. In addition to the field energy, the dielectric medium has internal energy which depends on ϵ. We will take the internal energy density to be lowest when $\epsilon = 0$. Then the energy of the system can be written as

$$\int \frac{\epsilon E^2}{4\pi}\,d^3x + \text{const}\times\int\epsilon(x)\,d^3x = \int\frac{D^2}{4\pi\epsilon}\,d^3x + \text{const}\times\int\epsilon(x)\,d^3x\,. \tag{19}$$

In a region of vanishing ϵ it is prohibitively energetically unfavorable for D to be nonvanishing, while a finite $\vec{E}$ field costs nothing.

Consider the possibility of a free charge of magnitude g at the origin. Suppose that the electromagnetic field and $\epsilon(x)$ are radially symmetric. The $\vec{D}$ field is then

$$\vec{D} = g\hat{r}/r^2\,. \tag{20}$$

If $\epsilon(x)$ vanishes in any finite region, then the term in the energy $\int[D^2/(4\pi\epsilon)]\,d^3x$ diverges. On the other hand, if ϵ is everywhere 1, then the field energy is finite, but the internal energy of the dielectric is infinite and proportional to the volume of space. Now consider a configuration of the dielectric in which a narrow tube of radius a emanates from the charge. Within the tube $\epsilon = 1$ and outside it is 0. The electrostatic equations governing this system read

$$\begin{aligned}\vec{\nabla}\cdot\vec{D} &= 4\pi\rho_{\text{free}} = 4\pi g\,\delta(\vec{x})\,,\\ \vec{\nabla}\times\vec{E} &= 0\,.\end{aligned} \tag{21}$$

These equations may be solved as follows (let the tube lie along the z axis). The electric field E is constant over all of space including the tube and points in the z direction. It contributes nothing to the energy density at points lying outside the tube. Inside the tube $\epsilon = 1$. Therefore, $\vec{D} = \vec{E}$ and it is easy to see that $\vec{\nabla}\cdot\vec{D} = 0$ everywhere except at the origin. The magnitude of the E field can be adjusted until $\vec{\nabla}\cdot\vec{D} = 4\pi\rho_{\text{free}}$. This is done by setting the total D flux, $\pi a^2 D = \pi a^2 E$, equal to the charge $4\pi g$. The true charge $(\rho_P + \rho_{\text{free}})$ is the source of the electric field and vanishes since the electric field is a constant. Thus, the dielectric has provided a polarization charge which exactly compensates the free charge.

The energy density in this solution is zero outside the flux tube, while inside the flux tube it is $g^2/(8\pi^2 a^4) + \text{const}$. Therefore, the total energy is proportional to the length of the flux tube (infinite). This means that the energy of a free charge diverges linearly with the characteristic linear dimension of space. Therefore, this configuration is energetically preferable to the radial solution whose energy was proportional to the volume of space. Thus we see an example of the formation of a flux tube associated with the field D in a classical nonrelativistic model.[12]

The physical picture described here is very similar to the "bag" model of hadrons developed by Chodos, Jaffe, Johnson, Thorn, and Weisskopf.[13]

V. A RELATIVISTIC MODEL WITH FLUX TUBES

It has been conjectured that the nonlinearities of quantized Yang-Mills theory cause a long-range force to develop between quarks[12] which may not fall off at spatial infinity. In this section we will construct a phenomenological field theory which exhibits this effect. We show that the long-range force is due to the formation of flux tubes.[14]

Our approach to the problem of quark trapping is to study the problem in two stages. In the first stage we consider a pair of quarks and imagine computing their interaction by completely summing the effects of the pure Yang-Mills gauge field. It is at this stage that we hope to find a long-range force mediated by flux tubes. We also expect to find that the energy needed to remove a quark from a color-neutral system is proportional to the separation distance.

The next stage introduces the effects of pair production by the gauge field. It allows the flux tube to break and form hadrons and at the same time purges all long-range forces from the theory by screening.

In the next section we will discuss the calculations which support the possibility of a strong long-range force in Yang-Mills theory.

For now we just mention that renormalization effects typically require dimensionless interaction constants to depend on the distance scales being studied. For example, in QED the interaction between electrons is much more intense at close distances than at large distances.[15] This effect is above and beyond the fact that the Coulomb force varies as e^2/r^2. Roughly speaking, the interactions between charges should be written

$$F(r) = e^2(r)/r^2, \tag{22}$$

where $e(r)$ is a dimensionless electric charge which depends on the distance between electrons. The physical origin of the r dependence of $e(r)$ is vacuum polarization which partially screens the charge of an electron. When a second electron is far away, it sees only a portion of the bare electronic charge. However, when the second electron is brought up close so as to penetrate the screening cloud it sees a more intense interaction. The result is that in QED the effective charge $e(r)$ is always a decreasing function of r. (This is equivalent to the fact that $\chi > 0$.)

In pure Yang-Mills theory, however, the effects which renormalize the coupling constant go in the opposite direction.[16] If the force between charges is characterized by a dimensionless coupling $g(r)$, then it is found that $g(r)$ *increases* with r. This raises the important possibility that $g(r)\to\infty$ and thereby produces a long-range force which prevents quark production.

In order to mimic these conditions in an Abelian gauge field, we will have to introduce an element which reverses the direction of renormalization effects. A simple possibility is to make the electric charge imaginary. Then, like charges attract and opposites repel. The effect causes an antiscreening which builds up the force law at large distances. However, the theory with imaginary charges probably has no ground state since the energy is not bounded from below.[17] In the model we construct here, the positivity constraints are violated as if the charges were imaginary, but the Hamiltonian is manifestly positive.

Our model incorporates three fields: the Abelian gauge field A_μ which may be used to construct a gauge-invariant Maxwell field,

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \,, \tag{23}$$

a charged fermion field ψ to be identified as the quark, and a local dielectric susceptibility $\chi(x)$.[10] The dielectric susceptibility will be expressed as a function of a canonical scalar field ϕ. For illustrative purposes $\chi(x)$ will be a fourth-order polynomial in ϕ,

$$\chi(x) = \alpha\phi + \beta\phi^2 + \gamma\phi^3 + \delta\phi^4 \,. \tag{24}$$

The Lagrangian is

$$\mathcal{L} = -\tfrac{1}{4} F_{\mu\nu} F^{\mu\nu} + \tfrac{1}{2}\partial_\mu \phi \partial^\mu \phi - V(\phi) - j^\mu A_\mu + i\bar{\psi}\partial^\mu \gamma_\mu \psi \,, \tag{25}$$

where V is a function of ϕ.

The current j consists of two parts. The first is the standard fermion current

$$j_F^\mu = g\bar{\psi}\gamma^\mu\psi \,, \tag{26}$$

which will be thought of as free[11] charge. The second term is analogous to polarization charge[10] and is given by an expression analogous to Eq. (17):

$$j_P^\nu = \partial_\mu [\chi(x) F^{\mu\nu}(x)] \,. \tag{27}$$

The Lagrangian may be conveniently rewritten

$$\mathcal{L} = -\tfrac{1}{4}(1+\chi) F_{\mu\nu} F^{\mu\nu} + j_F^\mu A_\mu + \mathcal{L}_\phi + \mathcal{L}_\psi \,. \tag{28}$$

The equation of motion for the gauge field is

$$\partial^\mu [(1+\chi) F_{\mu\nu}] = j_F^\nu \,. \tag{29}$$

Define the electromagnetic induction tensor[18] $D_{\mu\nu}$ by

$$D_{\mu\nu} = (1+\chi) F_{\mu\nu} \,. \tag{30}$$

The fermion current is then the source of $D_{\mu\nu}$,

$$\partial_\mu D^{\mu\nu} = j_F^\nu \,. \tag{31}$$

The term $(1+\chi) F_{\mu\nu} F^{\mu\nu}$ in $\mathcal{L}$ can be written in terms of D:

$$(1+\chi) F_{\mu\nu} F^{\mu\nu} = F_{\mu\nu} D^{\mu\nu} = D_{\mu\nu} D^{\mu\nu} (1+\chi)^{-1} \,. \tag{32}$$

This shows that a nonzero $D_{\mu\nu}$ field is very costly in action wherever $(1+\chi)$ is close to zero. On the other hand a nonzero finite $F_{\mu\nu}$ makes very little action.

The first stage of the quark-trapping problem treats the fermions as unquantized c-number sources. We demonstrate that the fields A_μ and ϕ create a long-range force mediated by flux tubes. The second stage shows how the long-range force is screened by production of fermion pairs.

We begin stage one by considering the classical physics of the fields A and ϕ in the presence of classical free charges. The classical model will be studied for the special case in which

$$1+\chi = [(\phi - \phi_0)/\phi_0]^4 \tag{33}$$

and $V(\phi) \sim \mu(\phi - \phi_0)^2$ for $(\phi - \phi_0)$ small. As in the previous section, we allow χ to be negative. As we shall see later this causes the charge-renormalization effects to be opposite to conventional

QED.[19]

The Hamiltonian for our model is positive and is given by

$$W=\frac{1}{2}\int[\vec{E}\cdot\vec{D}+\vec{B}\cdot\vec{H}]\,d^3x+\frac{1}{2}\int[(\vec{\nabla}\phi)^2+\dot{\phi}^2+V(\phi)]\,d^3x\,, \tag{34}$$

where

$$\begin{aligned} &E_i=F_{i0}\,, \qquad D_i=D_{i0}\,,\\ &B_i=\epsilon_{ijk}F_{jk}\,, \quad H_i=\epsilon_{ijk}D_{jk}\,. \end{aligned} \tag{35}$$

Now consider the energy of an isolated charge g. Then for a static field

$$\vec{H}=0\,,\quad \vec{\nabla}\cdot\vec{D}=g\,\delta^3(x)\,. \tag{36}$$

The energy W is given by

$$W=\frac{1}{2}\int[D^2\phi_0^{\,4}(\phi-\phi_0)^{-4}+(\vec{\nabla}\phi)^2+V(\phi)]\,d^3x\,, \tag{37}$$

where for small $(\phi-\phi_0)$ we assume $V(\phi)=\mu(\phi-\phi_0)^2$.

Let us suppose that the D field is spherically symmetric. From Eq. (36) it follows that

$$\vec{D}=g\hat{r}/4\pi r^2\,. \tag{38}$$

Assume now that $(\phi-\phi_0)^2\to r^p$, where p is a constant. Using Eq. (37) gives

$$W=\frac{1}{2}\int dr\,\{Ar^{-2-2p}+Br^p+Cr^{p+2}\}\,. \tag{39}$$

This integral diverges at large r as $r^{5/3}$ when $p=-\frac{4}{3}$. For any other value of p the divergence is worse.

On the other hand, if a flux-tube solution is possible, the flux tube will have an energy proportional to its length. For a single isolated charge this means an energy which diverges only linearly with r and is therefore energetically favored.

Let us consider a section of a possible flux-tube solution far from any free charge. The flux tube and $\vec{D}$ are oriented along the z axis. Define ρ to be the distance from the axis of the tube (circular symmetry of the tube is assumed). Both D_z and ϕ are functions of ρ which tend to zero and ϕ_0 respectively as $\rho\to\infty$.

The energy per unit length is

$$\tilde{W}=\frac{1}{2}\int\rho\,d\rho\left[D^2(\rho)\,\phi_0^{\,4}(\phi-\phi_0)^{-4}+\mu(\phi-\phi_0)^2+\left(\frac{d\phi}{d\rho}\right)^2\right]. \tag{40}$$

To find a static solution of the equations of motion we must minimize W subject to the constraint that a given total flux passes through the tube. This constraint is included by means of a Lagrange multiplier λ. Hence

$$0=\delta\left[\tilde{W}+\lambda\int\rho\,d\rho\,D(\rho)\right]. \tag{41}$$

Variation with respect to D gives

$$D_z=\frac{1}{2}\lambda(\phi-\phi_0)^4\,\phi_0^{\,-4}=\frac{1}{2}\lambda(1+\chi)\,. \tag{42}$$

Thus from Eq. (42) it is evident that λ is twice the electric field strength. Furthermore, the electric field is constant as in the simplified model of the previous section.

Variation with respect to ϕ gives

$$\frac{d}{d\rho}\left(2\rho\,\frac{d\phi}{d\rho}\right)-2\mu(\phi-\phi_0)-\lambda^2(\phi-\phi_0)^3\,\phi_0^{\,-4}=0\,. \tag{43}$$

The differential equation must be supplemented with boundary conditions stating

$$\left.\frac{d\phi}{d\rho}\right|_{\rho=0}=0\,,\quad \lim_{\rho\to\infty}\phi(\rho)=\phi_0\,. \tag{44}$$

These boundary conditions uniquely determine ϕ, which behaves like

$$(\phi-\phi_0)\sim\exp(-\text{const}\times\rho) \tag{45}$$

for large ρ. This ensures that the flux tube is spatially well defined. Since $D(\rho)$ is given by Eq. (42), we see that it, as well as the energy density, also falls exponentially to zero.

We would like to study the field near the end of the flux tube assuming it terminates on a point charge. First consider a disk of charge given by

$$j_F^0=\delta(z)\,F(\rho)\,. \tag{46}$$

If $F(\rho)$ is chosen equal to $D_z(\rho)$ for the infinite tube, then the equations of motion for D are satisfied by merely terminating the tube at $z=0$ (see Fig. 8). The true charge (free plus polarization) is defined to be $\vec{\nabla}\cdot\vec{E}$. Using $\vec{E}=\vec{D}/(1+\chi)$ we get

$$j^0_{\text{true}}=\vec{\nabla}\cdot\vec{D}(1+\chi)^{-1}-\vec{\nabla}\chi\cdot\vec{D}(1+\chi)^{-2}\,. \tag{47}$$

Therefore, if $\phi=\phi_0$ outside the tube then the two terms in Eq. (47) cancel. Thus, as in the previous section the true charge is exactly neutralized by a polarization charge.

Point charges are expected to produce fields similar to those shown in Fig. 9. Very close to

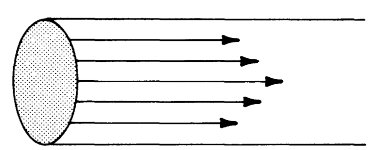


FIG. 8. A truncated flux tube capped by a disk of free charge. The free charge is neutralized by a similar disk of polarization charge.

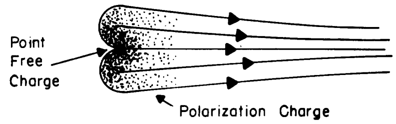


FIG. 9. A flux tube ending on a point free charge. A cloud of polarization charge neutralizes the free charge.

the charge, the field would be radial. The free charge is again compensated by polarization charge. However, the polarization charge does not form a point charge which sits right on top of the free charge. Instead, it is distributed over a volume near the end of the tube (Fig. 9).

Now consider a pair of charges at positions z_1 and z_2. The charges are connected by a long flux tube with an energy proportional to its length. This may be translated into a covariant statement concerning the action of a closed current loop in space-time. According to the same argument used in Sec. IV, the action of such a loop is proportional to its space-time area.

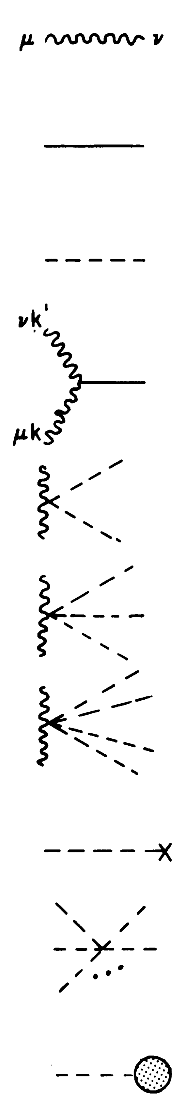
We now return to the quantum-mechanical treatment of the fields ϕ and A in the presence of classical free currents. We will temporarily allow $\chi(\phi)$ to be a general fourth-order polynomial in ϕ, as in Eq. (24). $V(\phi)$ will continue to have a minimum at ϕ_0. Thus, the vacuum expectation value of ϕ satisfies

$$\langle\phi\rangle_0 = \phi_0 + \text{quantum corrections}. \tag{48}$$

The model defined by Eq. (28) is severely nonrenormalizable, but will be regarded as a phenomenological effective Lagrangian to be used only in the tree-graph approximation. In fact, if the phenomenological dielectric field χ has any meaning it is as an approximate substitute for all the higher-order loop diagrams which renormalize the Yang-Mills theory. The Feynman rules and diagrams are enumerated in Fig. 10.

First we shall compute the gauge field propagator in the approximation of neglecting fermions. The relevant diagrams are shown in Fig. 11 for the vacuum polarization tensor $\Pi_{\mu\nu}$. There are four tree diagrams, which give

$$\begin{aligned}\Pi_{\mu\nu}(\text{trees}) &= (g_{\mu\nu}k^2 - k_\mu k_\nu) \\ &\quad\times(\alpha\phi_0 + \beta\phi_0^{\,2} + \gamma\phi_0^{\,3} + \delta\phi_0^{\,4}) \\ &= (g_{\mu\nu}k^2 - k_\mu k_\nu)\chi(\phi_0).\end{aligned} \tag{49}$$



$$= \left[g_{\mu\nu} - \frac{k_\mu k_\nu}{k^2}\right]\frac{1}{k^2}$$

$$= \left[\gamma_\mu p_\mu + m\right]\frac{1}{p^2 + m^2}$$

$$= \left[q^2 + \mu^2\right]^{-1}$$

$$= \alpha\left[g_{\mu\nu}k\cdot k' - k'_\mu k_\nu\right]$$

$$= \beta\left[g_{\mu\nu}k\cdot k' - k'_\mu k_\nu\right]$$

$$= \lambda\left[g_{\mu\nu}k\cdot k' - k'_\mu k_\nu\right]$$

$$= \delta\left[g_{\mu\nu}k\cdot k' - k'_\mu k_\nu\right]$$

$$= \left[\frac{\partial V}{\partial\phi}\right]_{\phi = 0}$$

$$= \left[\frac{\partial^n V}{\partial\phi^n}\right]_{\phi = 0}$$

$$= \left[q^2 + \mu^2\right]\phi_0$$

FIG. 10. Feynman rules for the model of Sec. V.

The gauge field propagator in the first stage is

$$(g_{\mu\nu} - k_\mu k_\nu/k^2)k^{-2}[1 + \chi(\phi_0)]^{-1}. \tag{50}$$

If our model satisfies conventional positivity constraints, then $\chi > 0$. However, if we relax these constraints and allow $1 + \chi(\phi_0) = 0$ as in the classical discussion, the effective propagator is infinite.

To understand the meaning of this result we

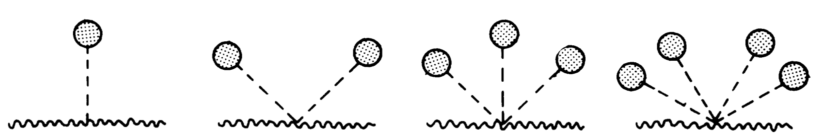


FIG. 11. Tree graphs contributing to $\Pi_{\mu\nu}$.

should recall that the theory constructed here is intended to describe only long-wavelength behavior. For example, if the effective vertices in the Lagrangian were smeared in coordinate space over a cutoff distance μ^{-1}, we would expect little or no change in long wavelengths. As an example, we can introduce momentum-dependent cutoffs $\mu^2(k^2+\mu^2)^{-1}$ for each external line of a vertex. The contribution to $\Pi_{\mu\nu}$ would now be modified to read

$$\Pi_{\mu\nu}=(g_{\mu\nu}-k_\mu k_\nu/k^2)\chi'(0)\mu^4(k^2+\mu^2)^{-2}, \tag{51}$$

and the propagator becomes

$$(g_{\mu\nu}-k_\mu k_\nu/k^2)k^{-2}[1+\chi(\phi_0)\mu^4(k^2+\mu^2)^{-2}]^{-1}. \tag{52}$$

For $k\ll\mu$, this may be written[20]

$$\sim(g_{\mu\nu}-k_\mu k_\nu/k^2)\mu^2 k^{-4}, \tag{53}$$

where we have used $\chi(\phi_0)=-1$. We could also obtain the same effect from loop corrections to $\Pi_{\mu\nu}$ evaluated with cutoffs. Therefore, for very large distances the propagator indeed diverges. One may conclude that at large distances the theory becomes superstrongly coupled.

Of course, the k^{-4} behavior of the propagator is not consistent with positivity requirements and would certainly lead to inconsistencies in the Abelian theory. We can take two possible views with regard to this situation. The first view is to allow the propagator to be infinite at this stage and to rely on fermion loops to render it finite. We shall see how this works later in this section. The second view is to say that in Yang-Mills theory the gauge field propagator does not satisfy positivity constraints and that there is nothing wrong with a k^{-4} behavior.

Although the propagator is infinite for all k in the tree approximation, the force law between charges is not. The single-photon exchange is not a good guide to the long-distance force because, as we have seen, the theory is superstrongly coupled in the infrared. To study the force law resulting from tree graphs, consider calculating the amplitude for a fermion loop with the loop treated as a c-number current,

$$F=\left\langle 0\left| T\exp\left(ig\oint A^\mu dx_\mu\right)\right|0\right\rangle. \tag{54}$$

Wilson[8] suggests that if the effective force is short-range, then the amplitude is proportional to e^{-P}, where P is the perimeter of the loop. A long-range quark-confining force would result if the amplitude were proportional to e^{-a}, where a is the area of the loop.[21]

To compute the amplitude in the tree approximation we make use of a theorem which states that the sum of all connected tree graphs is exactly equal to the classical least action.[22] Combining this theorem with the observation that the full loop is the exponential of connected graphs gives

$$F=\exp(-A_{\text{classical}}). \tag{55}$$

It is clear from our discussion of the energetics of the classical theory that the classical action is finite and proportional to the area of the loop. Moreover, the classical action is not sensitive to a cutoff μ for large μ. Therefore, the actual force law, unlike the gauge field propagator, is also cutoff-insensitive in the tree approximation.

At this stage we should attempt to eliminate the unphysical behavior of the gauge field propagator and the long-range force by accounting for pair production. As we have already remarked, pair production would allow the flux tube to break, thus eliminating the long-range force. We can also see how pairs would cure the peculiar properties of the gauge field propagator. Let us therefore consider the corrections due to simple fermion loops. The graphs in question are indicated in Fig. 12. The double wavy line represents the complete photon propagator in the absence of fermions and behaves like $k^{-2}[1+\chi(\phi_0)]^{-1}$. Define

$$\int e^{ik\cdot x}\langle 0|Tj_F^\mu(x)j_F^\nu(0)|0\rangle|_{1\text{PI}}\,dx=(g^{\mu\nu}k^2-k^\mu k^\nu)\Pi(k^2), \tag{56}$$

where the matrix element is computed in the one-fermion-loop approximation.

$\Pi(k^2)$ may be expressed in terms of a positive spectral function $R(m^2)$,

$$\Pi(k^2)=(a-k^2)\int_0^\infty R(m^2)(k^2+m^2)^{-1}(a+m^2)^{-1}dm^2, \tag{57}$$

where a is a convenient but arbitrary Euclidean renormalization point.[23,24] For our present application we may not use $a=0$ since we are expecting infrared singularities at $k^2=0$. Summing the graphs in Fig. 12 gives

$$(g_{\mu\nu}-k_\mu k_\nu/k^2)k^{-2}[1+\chi(\phi_0)+\Pi(k^2)]^{-1}. \tag{58}$$

Now we may allow $\chi(\phi_0)=-1$. Since $R(m^2)$ is positive we know $\Pi(0)\neq 0$, and therefore the gauge field propagator now has the form

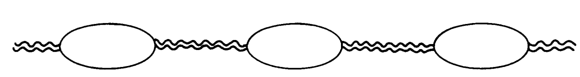


FIG. 12. Fermion-loop corrections to the gauge-field propagator. The double wavy line is the full propagator in tree-graph approximation.

$$(g_{\mu\nu}-k_\mu k_\nu/k^2)\,k^{-2}\,\Pi(0)^{-1}\,. \tag{59}$$

This approximation is not consistent, however. The very-low-k^2 behavior of the fermion loop is controlled by very large fermion paths in space-time. Therefore, we must not ignore the long-range force which acts between the pair. This force comes from diagrams like Fig. 13. That is to say, the fermion pair is not free but is instead bound by a flux tube.

In order to guess the results of such a modification in the fermion loop we will recall a result from the string model. Namely, in this model there exists a massless-vector bound state. If such a state actually occurs in the spectrum of the flux tube, it will make $\Pi(k^2)$ behave as

$$\Pi(k^2)\sim k^{-2}\,. \tag{60}$$

Then the propagator is completely finite at $k^2=0$.

The effect which eliminates the long-range force and renders the propagator finite at $k^2=0$ is the classic Schwinger phenomenon. The relevant bound state in this theory is a fermion pair held together by a flux tube. The same role is played in one-dimensional QED by a bound electron pair.

VI. YANG-MILLS THEORY AND INFRARED SLAVERY

In this section we will describe the reasons for believing that Yang-Mills theory might generate a long-range force between unscreened quarks.

The expression for a single quark loop propagating around a closed curve Γ in coordinate space space involves the expression

$$F\equiv \mathrm{tr}\left\langle 0\left| T\exp\!\left(ig\tau^\alpha \oint A^\mu_\alpha\, dx_\mu\right)\right|0\right\rangle , \tag{61}$$

where dx_μ is a line element along the curve Γ, μ is a space-time index, α is a color index, τ^α are the generators of the color group in the quark representation, and T denotes time ordering. The vector $|0\rangle$ is the vacuum of the pure Yang-Mills theory. By definition, the action is $A=\ln F$.

One important advantage of using F in characterizing the range of the interquark force is that it is gauge-invariant. A perturbation series in g may be derived, and it is found that F is a sum of all graphs of the type shown in Fig. 14. The gauge field propagators and vertices are standard and may be calculated in any gauge. Each vertex at which the fermion loop absorbs a gauge boson involves a factor $g\tau^\alpha$. The graphs may be computed in coordinate space and each external gauge boson may be attached to a point on the circumference of the loop in an obvious fashion. To each order in g the amplitude is found to be gauge-invariant.

Let us consider a loop Γ of a given shape and size. A scale factor λ will be used to vary the size of the loop. For example, if the reference loop is a unit circle, then λ is the radius of an arbitrary similar contour. Our aim is to determine whether the action can grow as the area λ^2.

Dimensional arguments assure us that in a given order of perturbation theory

$$\ln F\sim \sum_{m=0}^{d(n)} C_{mn}\, g^m (\ln\lambda)^m\,, \tag{62}$$

where, for given n, $\sum_m$ is a finite sum. Therefore, in order to build up the required λ^2 behavior, graphs of progressively higher order will have to be important as the loop increases in size. In view of the expected area rule, it is reasonable to hope that the most important graphs will have a two-dimensional or semiplanar structure as λ $\lambda\to\infty$. 't Hooft has made some interesting speculations along these lines[25] based on the unusual limit of Yang-Mills theory in which the internal-symmetry group is

$$\lim_{n\to\infty} \mathrm{SU}(n)\,.$$

The function F should be invariant under changes of the renormalization subtraction point and therefore should satisfy the renormalization equations

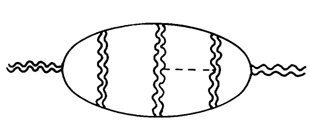


FIG. 13. Fermion-loop correction allowing fermions to interact.

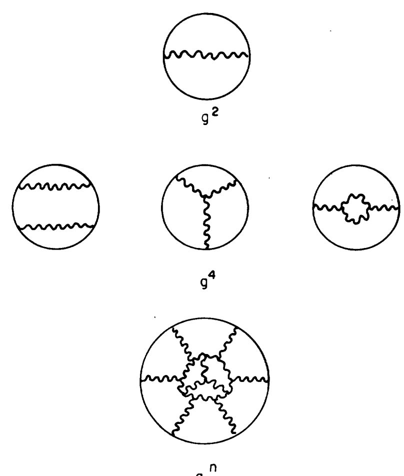


FIG. 14. Graphs contributing to F in the Yang-Mills theory.

$$\lambda \frac{\partial}{\partial \lambda} F(\lambda, g) = -\beta(g) \frac{\partial}{\partial g} F(\lambda, g) \tag{63}$$

or

$$\frac{\partial}{\partial t} A(t, g) = -\beta(g) \frac{\partial}{\partial g} A(t, g), \tag{64}$$

where $t = \ln\lambda$.

Equation (66) states that a change of scale can be exactly compensated for by an appropriate change of coupling constant. The change in coupling constant needed to compensate a given scale change is governed by the Gell-Mann–Low equation[23]

$$\frac{\partial}{\partial t} g(t) = \beta(g(t)) . \tag{65}$$

Equivalently,

$$t - t_0 = \int_{g_0}^{g} \frac{dx}{\beta(x)} , \tag{66}$$

or

$$\frac{\lambda_1}{\lambda_0} = \exp\left[\int_{g_0}^{g} \frac{dx}{\beta(x)} \right] . \tag{67}$$

Therefore, according to this last expression a change of coupling constant from g_0 to g_1 is equivalent to a rescaling of coordinates by a factor λ_1/λ_0. This means that if we knew $\beta(g)$, we could substitute the question: "How does A vary with coupling constant?" for the question: "How does A vary with λ ?"

We will assume that for fixed λ a finite coupling constant leads to a finite A. Thus, to make A approach infinity for fixed λ, it is necessary to allow $g \to \infty$. Accordingly, if the action A is to grow to infinity with λ, then so must $g(\lambda)$. From Eq. (69) it is seen that $\beta(g)$ must be negative for this to occur. It is well known that positivity requirements force $\beta(g)$ to be positive in Abelian gauge theories. However, recent calculations in pure Yang-Mills theories have shown that $\beta(g)$ is negative for g sufficiently small.[16]

Of course, in order to verify the flux-tube model or the area rule, it will be necessary to know a great deal more about β and $\partial A/\partial g$ for large g.

In Yang-Mills theories, as in one-dimensional QED, we expect fermion-antifermion production to screen long-range forces. However, in non-Abelian theories there is no clear connection between the screening of physical forces and the gauge field propagator. This is so because the Yang-Mills field itself is colored and does not correspond to the propagation of physical states if the Schwinger phenomenon occurs. The behavior of the Yang-Mills propagator would be similar to that of the charged-particle propagator in one-dimensional QED. Recall that for one-dimensional QED this propagator is gauge-dependent, and its infrared behavior is therefore not physically significant. The behavior can in fact be made arbitrary by gauge transformations.

The fact that $\beta(g) \geq 0$ in Abelian theories does not necessarily preclude the occurrence of the Schwinger mechanism. It is only that the two-stage method of solving the theory would not be sensible. In fact, Wilson[8] has given arguments that strongly coupled QED does undergo the Schwinger phenomenon.

ACKNOWLEDGMENTS

We would like to thank many people for discussions which led to the ideas contained in this paper. Our ideas about the importance of vector fields originated in discussions with J. D. Bjorken and A. Casher. The flux tube idea was greatly reinforced by discussions with G. 't Hooft, K. G. Wilson, and A. Chodos, who explained the idea of the "bag" model to us. We also thank M. Weinstein and J. D. Bjorken for helpful conversations concerning the differences between Higgs-Englert-Brout and Schwinger mechanisms. Finally, the need for an intermediate stage with a long-range force was emphasized to us by R. P. Feynman.

Note Added. A magnetic-flux-tube model has been suggested by H. B. Nielsen and P. Olesen [Nucl. Phys. B61, 45 (1973)]. The magnetic flux tube is analogous to vertex lines of superconductors and accompanies theories with a spontaneous breakdown of symmetry of the HEB variety. The electric flux tube discussed in the present paper should not be confused with magnetic vortices.

Note added in proof. A string model of hadrons based on one-dimensional QED has been developed by C. E. Carlson, L.-N. Chang, F. Mansouri, and J. F. Willemsen, Phys. Lett. B (to be published).

---

*Work supported in part by the National Science Foundation.

†Work supported in part by the National Science Foundation under Grant No. NSF-GP-38863.

‡Present address: Laboratory of Nuclear Studies, Cornell University, Ithaca, New York 14850.

[1] A. Casher, J. Kogut, and Leonard Susskind, Phys. Rev. Lett. 31, 792 (1973); Phys. Rev. D 9, 706 (1974).

[2]S. Weinberg, Phys. Rev. Lett. 31, 494 (1973).

[3]J. Schwinger, Phys. Rev. 128, 2425 (1962); *Theoretical Physics* (International Atomic Energy Agency, Vienna, 1963), p. 89.

[4]P. W. Higgs, Phys. Lett. 12, 132 (1964); Phys. Rev. Lett. 13, 508 (1964); Phys. Rev. 145, 1156 (1966); F. Englert and R. Brout, Phys. Rev. Lett. 13, 321 (1964); G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *ibid.* 13, 585 (1964); T. W. B. Kibble, Phys. Rev. 155, 1554 (1967).

[5]T. Appelquist, Harvard report, 1973 (unpublished) and references cited therein.

[6]For a discussion based upon the analysis of Feynman graphs see J. Kogut, D. K. Sinclair, and Leonard Susskind, Phys. Rev. D 7, 3637 (1973); 8, 2746(E) (1973).

[7]A very stimulating and similar physical picture has been proposed by E. P. Tryon, Phys. Rev. Lett. 28, 1605 (1972).

[8]K. G. Wilson (unpublished). We thank Professor Wilson for enlightening discussions.

[9]L.-N. Chang and F. Mansouri, Phys. Lett. 39B, 375 (1972); P. Goddard, J. Goldstone, C. Rebbi, and C. B. Thorn, Nucl. Phys. B56, 109 (1973); C. E. Carlson, L.-N. Chang, F. Mansouri, and J. F. Willemsen, Phys. Lett. B (to be published).

[10]L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media* (Addison-Wesley, Reading, Mass., 1960).

[11]The terms "free charge" and "true charge" are defined and discussed by W. K. H. Panofsky and M. Phillips, *Classical Electricity and Magnetism* (Addison-Wesley, Reading, Mass., 1955).

[12]One can view one-dimensional QED as the theory of electrodynamics in a narrow tube of high dielectric permeability surrounded by a region of vanishing permeability.

[13]A. Chodos, R. L. Jaffe, K. Johnson, C. B. Thorn and V. F. Weisskopf, Phys. Rev. D 9, 3471 (1974).

[14]The resulting physical picture is similar to that discussed in Ref. 13.

[15]The classic paper on screening in a field-theoretic, perturbation-series framework is by V. F. Weisskopf, Phys. Rev. 56, 72 (1939).

[16]G. 't Hooft, in proceedings of the Marseilles Conference on Gauge Theories, 1972 (unpublished); H. D. Politzer, Phys. Rev. Lett. 30, 1346 (1973); D. J. Gross and F. Wilczek, *ibid.* 30, 1343 (1973).

[17]The unstable character of QED with an imaginary electric charge has been discussed by F. J. Dyson, Phys. Rev. 85, 631 (1952).

[18]In general the dielectric permeability is a fourth-rank tensor field.

[19]As far as we can see the model is consistent at the classical or tree-graph level.

[20]K. Kaufman and R. P. Feynman have considered other theories with k^{-4} propagators. These are discussed by R. P. Feynman, in *Proceedings of the Fifth Hawaii Topical Conference on Particle Physics, 1973*, edited by Leo Pilachowski (Univ. of Hawaii Press, Honolulu, 1973).

[21]This means the minimal area of a surface bounded by the loop.

[22]Y. Nambu, Phys. Lett. 26B, 626 (1968).

[23]M. Gell-Mann and F. E. Low, Phys. Rev. 95, 1300 (1954).

[24]K. Symanzik, Commun. Math. Phys. 23, 49 (1971); C. G. Callan, Phys. Rev. D 5, 3202 (1972).

[25]G. 't Hooft, private communication and CERN report, 1974 (unpublished).