Nuclear Physics B72 (1974) 461–473. North-Holland Publishing Company

A PLANAR DIAGRAM THEORY FOR STRONG INTERACTIONS

G. 't HOOFT
CERN, Geneva

Received 21 December 1973

Abstract: A gauge theory with colour gauge group U(N) and quarks having a colour index running from one to N is considered in the limit $N \to \infty$, $g^2 N$ fixed. It is shown that only planar diagrams with the quarks at the edges dominate; the topological structure of the perturbation series in $1/N$ is identical to that of the dual models, such that the number $1/N$ corresponds to the dual coupling constant. For hadrons N is probably equal to three. A mathematical framework is proposed to link these concepts of planar diagrams with the functional integrals of Gervais, Sakita and Mandelstam for the dual string.

1. Introduction

The question we ask ourselves in this paper is how to construct a field theory of strong interactions in which quarks form inseparable bound states. We do not claim to have a satisfactory solution to that problem, but we do wish to point out some remarkable features of certain (gauge) field theories that make them an interesting candidate for such a theory.

First we have the singular infra-red behaviour of massless gauge theories [1] that makes it impossible to describe their spectra of physical particles by means of a perturbation expansion with respect to the coupling constant. It is not inconceivable that in an infra-red unstable theory long range forces will accumulate to form infinite potential wells for single quarks in hadrons.

The Han-Nambu quark theory [2] gives a qualitative picture of such forces between quarks: a very high, or infinite, energy might be required to create a physical state with non-zero "colour" quantum number. It is natural to take the symmetry corresponding to this quantum number to be a local gauge symmetry of some group SU(N). In that case, a formal argument in terms of functional integrals has been given by Amati and Testa [3] that supports the conjecture that "coloured" states have infinite energy.

In this paper we put the emphasis on an interesting coincidence. If we consider the parameter N of the colour gauge group SU(N) as a free parameter, then an expansion of the amplitudes at $N \to \infty$ arranges the Feynman diagrams into sets which have exactly the topology of the quantized dual string with quarks at its ends. The analogy with the string can be pursued one step further by writing the planar diagrams in the light cone reference frame. In sect. 6, we write down a Hamiltonian that generates all planar diagrams, in a Hilbert space of a fixed number of quarks. The quarks are inseparable if and only if the spectrum of this Hamiltonian becomes discrete in the presence of the interactions.

2. U (N) gauge theory

In order to show that the set of planar diagrams may play a leading rôle if certain physical parameters have certain values, we first formulate a possible gauge theory for strong interactions in which the parameters N and g have arbitrary values.

The quarks p_i, n_i and λ_i form three representations of the group U(N); $i = 1, \ldots, N$. Let us assume that an observer can distinguish between p, n and λ, but that he cannot distinguish the different colour components (see also sect. 3) *.

There is an anti-Hermitian gauge (vector) field

$$A_i{}^j{}_\mu (x) = -A^*{}_j{}^i{}_\mu (x), \tag{2.1}$$

and the Lagrangian is

$$\mathcal{L} = \tfrac{1}{4}\, G_{\mu\nu\, i}{}^j\, G_{\mu\nu\, j}{}^i - \overline{q}^{ai} (\gamma_\mu D_\mu + m_{(a)})\, q^a{}_i\,, \tag{2.2}$$

where

$$\begin{aligned} G_{\mu\nu\, i}{}^j &= \partial_\mu A_i{}^j{}_\nu - \partial_\nu A_i{}^j{}_\mu + g\, [A_\mu, A_\nu]_i{}^j\,; \\ D_\mu q^a{}_i &= \partial_\mu q^a{}_i + g\, A_i{}^j{}_\mu\, q^a{}_j\,. \end{aligned} \tag{2.3}$$

The index a runs from one to three

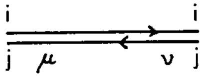
$$q^1 = p;\quad q^2 = n;\quad q^3 = \lambda. \tag{2.4}$$

For sake of simplicity we do not make the restriction that the trace of the gauge field, $A^i_{i\mu}$ should vanish, and so we will have a photon corresponding to the Abelian subgroup U(1) of U(N), and coupling to baryon number. Of course we could dispose of it, either by replacing U(N) by SU(N), or when we switch on weak and electromagnetic interactions through the Higgs mechanism. But for the time being it is there and we must keep it in mind when we finally interpret the results of our calculations.

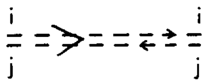
The Feynman rules [4, 5] may be formulated as usual in any suitable gauge. Let us take the Feynman gauge. We add to the Lagrangrian

$$\tfrac{1}{2}\, \partial_\mu A_i{}^j{}_\mu\, \partial_\nu A_j{}^i{}_\nu - \partial_\mu\, \phi_j^{*i}\, (\partial_\mu \phi_i{}^j + g\, [A_\mu, \phi]_i{}^j), \tag{2.5}$$

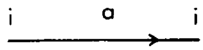
* We do not know whether this assumption is really essential for the theory, but it does simplify the arguments in sect. 3.



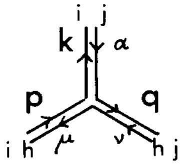
$-\dfrac{\delta_{\mu\nu}}{k^2-i\epsilon}$ (vector field)



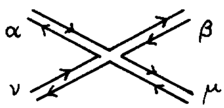
$\dfrac{1}{k^2-i\epsilon}$ (F.P. ghost)



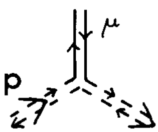
$\dfrac{1}{m_{(a)}+i\gamma k-i\epsilon}$ (quark q_i^a)



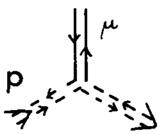
$ig\{\delta_{\alpha\nu}(k-q)_\mu+\delta_{\alpha\mu}(p-k)_\nu+\delta_{\mu\nu}(q-p)_\alpha\}$



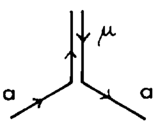
$g^2\{2\delta_{\alpha\mu}\delta_{\beta\nu}-\delta_{\alpha\beta}\delta_{\mu\nu}-\delta_{\beta\mu}\delta_{\alpha\nu}\}$



igp_μ



$-igp_\mu$



$-g\gamma_\mu$

-1 (Fermi statistics)

Fig. 1. Feynman rules for U(N) gauge theory in Feynman gauge.

where ϕ is the Feynman-DeWitt-Faddeev-Popov ghost field. Now the bilinear parts of the Lagrangian generate the propagators and the interaction parts the vertices.

In order to keep track of the indices, it is convenient to split the fields $A_{\mu i}{}^{j}$ into complex fields for $i > j$ and real fields for $i = j$. One can then denote an upper index by an incoming arrow, and a lower index by an outgoing arrow. The propagator is then denoted by a double line. In fig. 1, the vector propagator stands for an $A_{\mu i}{}^{j}$ propagator to the right if $i > j$; an $A_{\mu j}{}^{i}$ propagator to the left if $i < j$ and a real propagator if $i = j$. The extra minus sign in this propagator is a consequence of the anti-Hermiticity of the field A (eq. (2.1)). The ghost fields satisfy no Hermiticity condition and therefore their propagators have an additional arrow (fig. 1).

The vertices always consist of Kronecker delta functions connecting upper and lower indices, and thus connect ingoing with outgoing arrows. The quark propagators consist of a single line.

As usual, amplitudes and Green functions are obtained by adding all possible (planar and non-planar) diagrams with their appropriate combinatory factors. Note now that the number N does not enter in fig. 1 (this would not be the case if we would try to remove the photon).

But, of course, the number N will enter into expressions for the amplitudes, and that is when an index-line closes. Such an index loop gives rise to a factor

$$\sum_i \delta_i{}^i = N.$$

3. The $N \to \infty$ limit

In sect. 2 we assumed that the observer is colour-blind. This can be formulated more precisely: only gauge-invariant quantities can be measured. A measuring apparatus can formally be represented by a c number source function $J(x)$ which is coupled to a gauge invariant current, for instance

$$\sum_i \bar{p}^i n_i \, . \tag{3.1}$$

We observe from fig. 2 that index lines never stop at a gauge invariant external source, but they continue. "Index loops" going through an external source also obtain a factor N, because of the summation in (3.1).

We are now in the position that we can classify the diagrams with gauge invariant sources according to their power of g and their power of N. Let there be given a connected diagram. First we consider the two-dimensional structure obtained by attaching little surfaces to each index loop. We get a big surface, with edges formed by the quark lines, and which is in general multiply connected (contains "worm holes"). We close the surface by also attaching little surfaces to the quark loops separately.

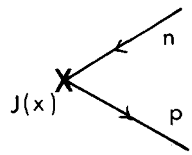


Fig. 2. Gauge invariant source function.

Let that surface have F faces, P internal lines or propagators, and V vertices. Here $F = L + I$, where L is the number of quark loops and I the number of index loops; and $V = \Sigma_n V_n$, where V_n is the number of n-point vertices. The diagram is associated with a factor

$$r = g^{V_3 + 2V_4} N^I . \tag{3.2}$$

By drawing a dot at each end of each internal line, we find that the number of dots is

$$2P = \sum_n n\, V_n , \tag{3.3}$$

and eq. (3.2) can be written as

$$r = g^{2P-2V} N^{F-L} . \tag{3.4}$$

Now we apply a well-known theorem of Euler:

$$F - P + V = 2 - 2H , \tag{3.5}$$

where H counts the number of "holes" in the surface and is therefore always positive (a sphere has $H = 0$, a torus $H = 1$, etc.). And so,

$$r = (g^2 N)^{\frac{1}{2} V_3 + V_4} N^{2-2H-L} . \tag{3.6}$$

Suppose we take the limit

$$N \to \infty , \quad g \to 0 , \quad g^2 N = g_0^2 \text{ (fixed)} . \tag{3.7}$$

If the sources are coupled to quarks, then there must be at least one quark loop: $L \geqslant 1$. The leading diagrams in this limit have $H = 0$ and $L = 1$, they are the planar diagrams with the quark line at the edges (fig. 3).

Note, however, that the above arguments not only apply to gauge fields but also to theories with a global U(N) symmetry containing fields with two U(N) indices, but from the introduction, it will be clear why we concentrate mainly on gauge fields.

It is interesting to compare our result with that of Wilson [6], who considers gauge fields on a dense lattice and also finds structures with the topology of a two-

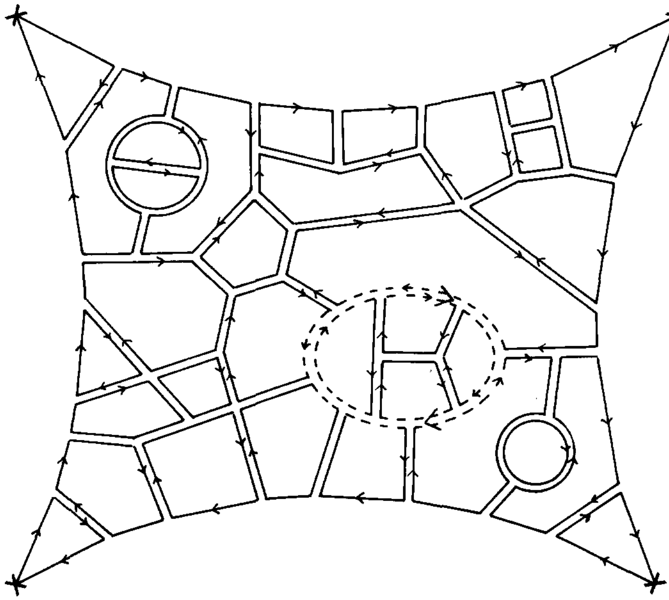


Fig. 3. One of the leading diagrams for the four-point function.

dimensional surface. It is not difficult to show that also Wilson's surfaces are associated with factors $1/N^2$ and $1/N$ for each worm hole or fermion loop, respectively.

The dual topology of the set of planar diagrams has been noted before [7]. Here we see that the analogy with dual models goes even further; the expansion in powers of $1/N$ corresponds to the expansion with respect to the dual coupling constant in

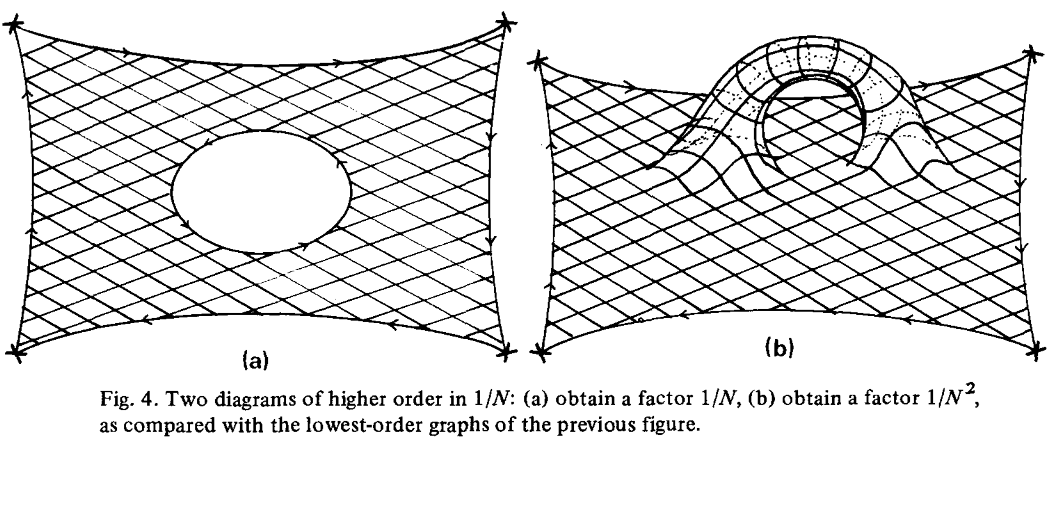


Fig. 4. Two diagrams of higher order in $1/N$: (a) obtain a factor $1/N$, (b) obtain a factor $1/N^2$, as compared with the lowest-order graphs of the previous figure.

dual models. If we adopt the Han-Nambu picture of hadrons [2] then N is very likely to be three. This seems to give a reasonable order of magnitude for the dual coupling constant.

Let us now formulate our theory more precisely. We assume that there is a local gauge group of the type U(3), (or SU(3)) for which no preferred reference frame in the form of a Higgs field exists. Such a theory is infra-red unstable [1] which implies that infra-red divergences accummulate instead of cancel, and the physical spectrum is governed by long range forces. A simple-minded perturbation expansion with respect to g_0 in eq. (3.7) does not describe the spectrum and the S-matrix. But the $1/N$ expansion may be a reasonable perturbation expansion, in spite of the fact that N is not very big.

4. Planar diagrams in the light-cone frame

The theory implies that we have to sum all planar diagrams in order to get the leading contributions to the amplitudes. Attempts to calculate certain large planar diagrams are known in the literature [7] but it seems to us that the choice of diagrams there is rather arbitrary, and the replacement of a propagator by Gaussian expressions seems to be a bad approximation. We believe that a more careful study of this problem is necessary.

Let us consider any large planar diagram (fig. 5). For a moment we shall abandon the rather complicated Feynman rules of fig. 1, replacing the vertices by simple local ϕ^3 or ϕ^4 interactions.

We immediately face two problems:

(i) how to find a convenient parametrization scheme to indicate a point of the graph in the plane, in terms of two parameters σ and τ;

(ii) how to arrive at Gaussian integrands, in order to be able to do the integrations.

These two problems can be solved simultaneously by going to light-cone co-ordinates: we write [8]

$$p^{\pm} = \frac{1}{\sqrt{2}}(p^3 \pm p^0), \qquad \tilde{p} = (p^1, p^2),$$

$$x^{\pm} = \frac{1}{\sqrt{2}}(x^3 \pm x^0), \qquad \tilde{x} = (x^1, x^2). \tag{4.1}$$

Although the gauge particles are massless, we shall consider the slightly more general case of arbitrary masses. The propagators are then

$$\frac{1}{(2\pi)^4 i} \frac{1}{(\tilde{p}^2 + 2p^+p^- + m^2 - i\epsilon)}, \tag{4.2}$$

(for sign conventions, see ref. [5]). We go over to a mixed momentum coordinate representation: at each vertex $V_{(\alpha)}$ we perform an integration over its time co-ordi-

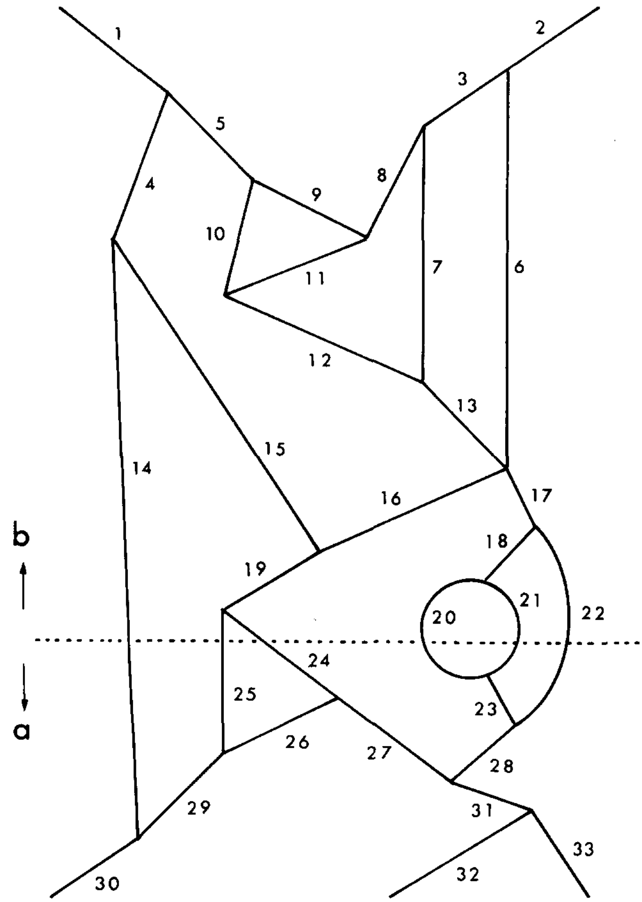


Fig. 5. Example of a planar diagram, divided into two regions a and b (see text).

nate $x^+_{(\alpha)}$, and to each window $F^{(i)}$ of the graph corresponds an integration over the momenta $\tilde{p}_{(i)}$ and $p^+_{(i)}$ (always directed anti-clockwise). In terms of these variables the propagator is the Fourier transform of (4.2) with respect to p^-,

$$\frac{1}{(2\pi)^3 2|p^+|}\,\theta\,(x^+p^+)\,\exp - i\,\frac{x^+}{2p^+}\,(m^2+\tilde{p}^2)\,. \tag{4.3}$$

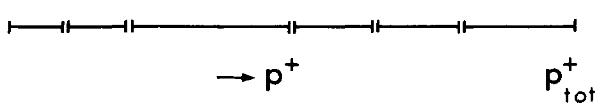


Fig. 6. The components p^+ of the momenta of the propagators that cross the dotted line in fig. 5.

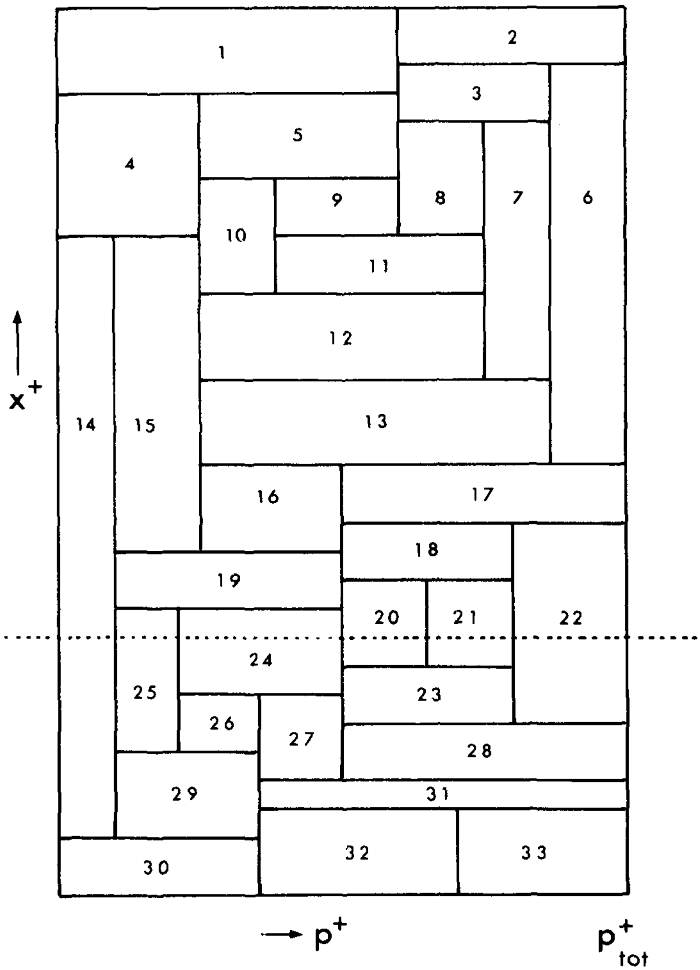


Fig. 7. A new representation of the same diagram. The blocks here correspond to the propagators in fig. 5 and have been numbered accordingly.

Here x^+ is the time difference between the two end points of the line, and $(\tilde{p}, p^+)$ are the difference of the momenta $(\tilde{p}_{(i)}, p^+_{(i)})$ circulating in the windows at both sides of the line. Note that the propagator (4.3) is Gaussian in the transverse momenta $\tilde{p}$.

The parametrization problem can be solved by exploiting the famous θ function in (4.3). For simplicity we shall assume that all external lines with positive p^+ lie next to each other in the plane†. If we divide the set of vertices into: (a) those with $x^+_{(\alpha)} < a$, and (b) those with $x^+_{(\alpha)} > a$, then all lines * going from (a) to (b) have

† If this condition is not fulfilled the resulting plane of fig. 7 will get several "sheets".

* If we want to keep the diagram planar while dividing it into blobs (a) and (b), then we must expect lines going from (a) through (b) back to (a), etc. But it is easy to convince oneself that in those cases the diagram is zero as a consequence of the θ functions.

positive p^+. Now imagine a horizontal line with length p^+_{total} and divide it into segments, each corresponding to a propagator going from (a) to (b), and with a length equal to the (positive) value of p^+ in that propagator (fig. 6). If we now vary the number a, then this line sweeps out a surface with constant width, in which the propagators correspond to blocks; loops in the original diagrams now correspond to vertical lines, and vertices are now horizontal lines. See fig. 7 in which we numbered the blocks corresponding to the propagators in fig. 5. We see that the variables p^+ and x^+ are suitable co-ordinates. The integration in the transverse momenta (or co-ordinates) is Gaussian. Summing and integrating over all possible topologies in the p^+x^+ plane is equivalent to performing the remaining p^+x^+ integrations and the summations over the diagrams.

It is convenient at this point to perform a Wick rotation.

$$i x^+ = \tau \, . \tag{4.4}$$

The factor i in the exponent (4.3) now disappears, together with the factors i at each vertex:

$$(2\pi)^3 \, i \, \lambda \, \mathrm{d} x^+ \to (2\pi)^3 \, \lambda \, \mathrm{d}\, \tau , \tag{4.5}$$

and all amplitudes become real, Gaussian integrals (the θ function in (4.3) now becomes $\theta\ (\tau p^+)$, defining the new regions of integrations).

5. Comparison with the dual string

Instead of considering the transverse momenta $\tilde{p}$, we could study the diagrams in transverse coordinate space. Then we would have a transverse variable $\tilde{x}$ at each vertex of fig. 5, or at each horizontal line in fig. 7. The propagator is also Gaussian in terms of the $\tilde{x}$. The integrand is (after the Wick rotation)

$$C \exp \left\{ - \sum_{ij} \left| \frac{\Delta p^+_{ij}}{2\,(\tau_i - \tau_j)} \right| (\tilde{x}_i - \tilde{x}_j)^2 \right\} . \tag{5.1}$$

where C is independent of the transverse variables, and the summation is performed over all pairs of adjacent horizontal lines in fig. 7. Δp^+_{ij} stands for the width of the block between i and j.

This is to be compared with (the essential part of) the functional integrand for the quantized string:

$$C \exp - \int \mathrm{d}\sigma \, \mathrm{d}\tau \, \left[\left(\frac{\partial \tilde{x}}{\partial \sigma} \right)^2 + \left(\frac{\partial \tilde{x}}{\partial \tau} \right)^2 \right] , \tag{5.2}$$

where $\tilde{x}\ (\sigma, \tau)$ is now a continuous variable on a similar rectangular surface [9]. The difference between (5.1) and (5.2) is profound. The first difference is that in eq. (5.1) we have a partition of the dual surface into meshes, and secondly in (5.1) one must also integrate over all longitudinal variables and sum over all diagrams. This integration and summation together correspond to the summation over all partitions into meshes. The detailed structure of the meshes will depend on the initial Feynman rules, and from those it will probably depend whether (5.1) can be approximated by (5.2) in any way. If so, then the dual string will be an approximate solution of the dynamical equations of our gauge model.

6. A Hamiltonian formalism

Attempts to attain more understanding of the peculiarities of planar diagram field theory have failed until now. There exists, however, a Hamiltonian for this system that might be useful. For simplicity, we confine ourselves to the planar diagrams of ϕ^3 theory (again defined by means of a certain $N \to \infty$ limit). A representation of states $|\psi\rangle$ in a Hilbert space is defined as a set of structures like in fig. 6: a number of "particles" is sitting on a line segment with length p^+_{total}. They have coordinates p_i^+, $i = 1, \ldots, r; r = 0, 1, \ldots$. A transverse loop integration momentum $\tilde{p}_i$ is assigned to each particle (the particles in fig. 6 actually correspond to loops in the original diagram). We put $\tilde{p}_{\mathrm{total}} = 0$, so that $\tilde{p} = 0$ on the boundaries at the left and at the right.

A Wick rotation is not necessary here, so we can take x^+ to be real. The x^+ axis is divided into small segments $x_0^+, x_1^+, \ldots, x_n^+$, with

$$x_{k+1}^+ - x_k^+ = \epsilon . \tag{6.1}$$

Now we write the amplitude formally as

$$A = {}_{x_n^+}\langle \mathrm{out} \,|\, \mathrm{e}^{-i\epsilon H} \,|\, \psi\rangle_{x_{n-1}^+} \langle \psi \,|\, \mathrm{e}^{-i\epsilon H} \,|\, \psi\rangle_{x_{n-2}^+} \langle \psi \,|\, \ldots \quad \ldots |\mathrm{in}\rangle_{x_0^+} , \tag{6.2}$$

where summation and integration over the intermediate states is understood. We now construct the Hamiltonian H that will yield the sum of all planar diagrams. Expand

$$\mathrm{e}^{-i\epsilon H} = 1 - i\epsilon H = 1 - i\epsilon\,(H_0 + H_1), \tag{6.3}$$

where H_0 will be taken to be diagonal in the above-defined representation. If no vertex occurs between x_k^+ and x_{k+1}^+ then only H_0 contributes to

$${}_{x_{k+1}^+}\langle \psi \,|\, \mathrm{e}^{-i\epsilon H} \,|\, \psi\rangle_{x_k^+} .$$

Taking

$$H_0 = \sum_i \frac{m_i^2 + (\tilde{p}_i - \tilde{p}_{i-1})^2}{2\,(p_i^+ - p_{i-1}^+)} , \tag{6.4}$$

we get the correct exponential parts of the propagators (compare (4.3)).

At the vertices (horizontal lines in fig. 7), our particles are created or annihilated. Here H_1 is in action. Let us define operators $a^\dagger(\tilde{p}, p^+)$ and $a(\tilde{p}, p^+)$, with

$$[a^\dagger(\tilde{p}, p^+), a(\tilde{k}, k^+)] = \delta^2(\tilde{p} - \tilde{k})\,\delta(p^+ - k^+), \tag{6.5}$$

creating respectively annihilating particles. We can then take $H_1 = V + V^\dagger$, with

$$V = -(2\pi)^3\lambda(16\pi^3)^{-\frac{3}{2}} \int \mathrm{d}p^+ \int \mathrm{d}^2\tilde{p}\, \frac{a(\tilde{p}, p^+)}{\sqrt{(p_r^+ - p_l^+)(p_r^+ - p^+)(p^+ - p_l^+)}} \tag{6.6}$$

where λ is the coupling constant; p_r^+ and p_l^+ are the coordinates of the closest neighbours at the right and at the left of the point p^+.

Substituting this interaction Hamiltonian into (6.3) and (6.2), we find exactly the Feynman rules for planar diagrams: the square root of the width of each block in fig. 7 always occurs twice, thus giving rise to the required factor $1/p^+$ in the propagator (4.3).

In our gauge theory model, a similar Hamiltonian will describe one quark and one antiquark in interaction. If our theory is to describe hadrons, then its spectrum should come out to be discrete.

7. Conclusion

We are still far away from a satisfactory theory for bound quarks. But, guided by the topological structure of the dual theories, we are led to the planar diagram field theory, in terms of which our problem can easily be formulated: if the eigenstates of a certain Hamiltonian crystallize into a discrete spectrum, despite the fact that the zeroth order Hamiltonian is continuous, then the original particles will condensate into a string that keeps quarks together.

As for baryons, the situation is even more complicated. The Han-Nambu theory clearly suggests $N = 3$. In that case we can raise or lower indices in the following way:

$$\lambda_i \to \lambda^{ij} = \epsilon^{ijk}\lambda_k = -\lambda^{ji}. \tag{7.1}$$

Taking p_i, n_i and λ^{ij} as our elementary fermions we can again consider the $N \to \infty$ limit. The λ quark will then sit in the middle of a string with p and/or n quarks at its ends: we have a string with Σ or Λ baryons! Similarly protons, neutrons and all other baryons can be constructed.

It will be clear that in the case of baryons the $1/N$ expansion is extremely delicate.

If calculations will be possible at all in this theory, then the dual coupling constant will be calculable and of order $\frac{1}{3}$.

References

[1] G. 't Hooft, Marseille Conf., 19–23 June 1972, unpublished;
H.D. Politzer, Phys. Rev. Letters 30 (1973) 1346;
D.J. Gross and F. Wilczek, Phys. Rev. Letters 30 (1973) 1343;
D.J. Gross and F. Wilczek, NAL preprint, Princeton preprint (1973);
S. Coleman and D.J. Gross, Princeton preprint (1973);
G. 't Hooft, Nucl. Phys. B61 (1973) 455; B62 (1973) 444.
[2] M.Y. Han and Y. Nambu, Phys. Rev. 139B (1965) 1006.
[3] D. Amati and M. Testa, CERN preprint TH. 1770 (1973).
[4] G. 't Hooft and M. Veltman, Nucl. Phys. B50 (1972) 318.
[5] G. 't Hooft and M. Veltman, "DIAGRAMMAR", CERN report 73–9 (1973).
[6] K.G. Wilson, Lecture given in Orsay (August 1973), unpublished.
[7] H.B. Nielsen and P. Olesen, Phys. Letters 32B (1970) 203;
B. Sakita and M.A. Virasoro, Phys. Rev. Letters 24 (1970) 1146.
[8] S.-J. Chang and S.-K. Ma, Phys. Rev. 180 (1969) 1506.
[9] J.-L. Gervais and B. Sakita, Phys. Rev. Letters 30 (1973) 716;
S. Mandelstam, Nucl. Phys. B64 (1973) 205.