

IS ASYMPTOTIC FREEDOM ENOUGH?

Gerard 't HOOFT

Institute for Theoretical Physics, NL-3508 TA Utrecht, The Netherlands

Received 13 January 1982

Quantum field theories with strong interactions are usually required to be not only renormalizable but also asymptotically free, in order to avoid diseases such as the Landau ghost. In this paper we suggest an even more restrictive requirement: "asymptotic convergence", which means that at high energies it must be possible to formulate a convergent resummation procedure for the perturbation expansion. Such a convergent resummation technique exists in QCD in the infinite color limit ($N \rightarrow \infty$). We give an outline of a proof of this statement, and a brief account of possible consequences of our asymptotic convergence condition on model building.

History has taught us that progress in our understanding of elementary particles comes partly from experimental observations and partly from purely theoretical requirements concerning self-consistency of the mathematical structure of a model. In the late 60's for example, there was no experimental indication that field theories had to be renormalizable. Yet theorists insisted that they cannot understand their models otherwise and as we know now, their persistence has been extremely successful. The next example is quantum-chromodynamics (QCD), which only then became a respectable model for theorists when it was understood to be asymptotically free: its short distance structure is described by only weakly interacting particles. It is generally felt that this property is sufficient for the theory to be well defined mathematically. For instance one considers an ultraviolet cutoff defined by replacing space-time by a lattice with lattice constant a . If the coupling constant g is chosen to be given by [1]

$$g^{-2} = \beta_0 \log(1/\Lambda_q a) + \kappa \log \log(1/\Lambda_q a), \quad (1)$$

where β_0 and κ are computable renormalization group coefficients [2] and Λ_q is a fixed constant (lattice- Λ -parameter), then the limit $a \rightarrow 0$ appears to exist. One reason why one feels confident with this assumption is that it is true to *all* orders in perturbation expansion, whereas only β_0 and κ are explicitly needed.

They correspond to the one- and two-loop renormalization group coefficients, respectively.

Up to now, however, this assumption has not been proven to be correct beyond perturbation expansion, and due to the complexity of the theory no proof is really in sight. One might therefore speculate that, against all former expectations, the above assumption is wrong. For instance, it could be that the ratio of the proton mass-squared to the string constant depends critically on the details of the lattice action, the nature of the lattice (cubic, triangular, and so on). This situation could be compared with attempts in the 60's to attach unambiguous meaning to non-renormalizable weak-interaction theories such as the four-fermion interaction. Nowadays it is generally accepted that renormalizable weak-interaction theories are superior and one is ready to pay the price of additional degrees of freedom: the W-particle, the Z-particle and the Higgs boson.

Similarly, QCD may have to be extended with more degrees of freedom. The point is that according to eq. (1), the coupling constant g does tend to zero if a goes to zero, but only logarithmically. This is much too slow to enable us to prove uniqueness of the theory [3]. So one might suspect that the theory is not unique.

In four space-time dimensions there are no super-renormalizable field theories with a ground state, be-

cause the only super-renormalizable self interaction, of the form ϕ^3 , is not bounded. At best we can have renormalizability and asymptotic freedom. But we can imagine that at higher energies more and more fields participate in the interaction, such that the parameter β_0 in eq. (1) effectively increases. Since β_0 is proportional to a Casimir operator, which is N in $SU(N)$, theories, we consider the possibility that at higher energies more and more "unification" takes place, from small groups to large groups, and at infinite energy N tends to infinity. Only then g^2 can go to zero faster than logarithmically. Thus, the situation could be similar to that of the usual weak and electromagnetic models, whose high-energy behavior is mild because of extra fields. The extra fields are frozen out at the lower end of the energy scale.

The above philosophy can only possibly make sense if the $N \rightarrow \infty$ limit of $SU(N)$ gauge theories is free from the diseases of QCD. As is well known [4], the limit appears to exist if $\tilde{g}^2 = g^2 N$ is kept fixed, and in the limit $N \rightarrow \infty$ perturbation expansion is described by a set of planar Feynman diagrams (diagrams that can be drawn on a sheet of paper without crossings). In four dimensions this is a highly complicated system that does not seem to be soluble analytically. Also, in terms of \tilde{g}^2 , the theory is again logarithmically asymptotically free.

What we wish to report here is that the planar expansion in \tilde{g}^2 can be summed. Not analytically, but formally, and we argue that there is a unique interpretation that can be attached to the expansion. We have actually proved finiteness of a closely related but simpler theory: a field of $N \times N$ Lorentz scalars with the lagrangian

$$\mathcal{L} = \text{tr} \left(-\frac{1}{2} (\partial_\mu \phi)^2 - \frac{1}{2} m^2 \phi^2 + \frac{1}{4} \lambda \phi^4 \right). \quad (2)$$

Here tr stands for trace, m is a finite mass, and λ indeed has the "wrong" sign. In the limit $N \rightarrow \infty$, $\tilde{\lambda} = \lambda N$ fixed, the vacuum is nevertheless stable because the instantons [5] that describe the decay of the vacuum have infinite action. The most important difference between this theory and the gauge theory is the finite mass m , which is needed as an infrared cutoff. In a gauge theory such a cutoff can probably be achieved by considering the system to be enclosed in a finite box. The "thermodynamic" limit where the volume of the box goes to infinity then still has to be considered. We do not expect that our philosophi-

cal arguments can be affected by such difficulties. The other differences in the Feynman rules are unessential. Our proof of the finiteness of the theory described by eq. (2) in the $N \rightarrow \infty$ limit will be given elsewhere. Let us briefly outline how it goes. The planar diagrams are considered to be composed of possibly divergent subgraphs with four legs or less, which are linked together into larger structures which themselves only contain convergent integrals.

We first postulate that the divergent graphs with four external lines, when summed, are bounded by logarithmically decreasing expressions as suggested by the renormalization group. Then the convergent structures into which these are grouped are bounded by all convergent diagrams with the divergent subgraphs replaced by bare vertices and propagators. The bulk of our proof consists of showing that the series expansion of convergent diagrams has a finite radius of convergence. For this we needed planarity. Since at very high energies the effective coupling can become arbitrarily small, one expects that at low energies λ will be allowed to leave the (fairly small) circle of convergence to some extent. For a certain set of λ values then the six- and more point dressed vertices are uniquely determined if one knows the four-point and self-energy diagrams. These in turn are related to the higher dressed vertices through Dyson-Schwinger equations. The set of closed equations one gets this way determines all amplitudes uniquely if the mass m is positive and sufficiently large. A lower bound for the radius of convergence for the not-divergent planar diagrams is found by observing that the total number of planar diagrams with L loops does not exceed C^L where C is some finite number [6], and each planar diagram with L loops and a limited number of external lines, when L exceeds a certain number, necessarily contains loops with less than six propagators (both statements are not true if N is finite [7]). These loops are bounded by a set of tree diagrams of a certain nature so that diagrams with fewer loops are obtained. By induction one then finds a constant C' so that each diagram does not exceed $(C')^L$ times a complete tree diagram. The most difficult part of this proof is to avoid that at the intermediate steps of this algorithm diagrams appear that are actually ultraviolet or infrared divergent. In our algorithm the infrared divergences could easily be controlled. The ultraviolet divergences are controlled by requiring that the degree of conver-

gence Z for each subgraph at any stage satisfies

$$Z \geq \beta(e - 3), \tag{3}$$

where e is the number of external lines of that subgraph and β some number ($2/3$ for the scalar theory, $1/2$ for gauge theories). Triangular subgraphs can be avoided entirely.

Obviously, planarity was crucial for our proof, and it is conceivable that QCD with $N = 3$ can never be made as healthy as this model. We should mention, however, that progress on constructive finite- N scalar field theories in four dimensions is made [8], and it seems that the methods of ref. [8] are more powerful than ours, so perhaps their methods can simplify our present proof or even overthrow our suspicion that, at infinite energy, N has to tend to infinity. As yet this has not occurred.

There are several ways in which one can imagine "tumbling" from high N to lower N to take place. We were able to construct a model in which ordinary scalar Higgs fields cause repeated symmetry breakdown $SU(N) \rightarrow SU(N - 1) \rightarrow \dots \rightarrow SU(3)$. The renormalizable interactions are postulated to be super-symmetric [9]. We have one $SU(N)$ gauge super-field and M matter super-fields in the elementary representation. The model is asymptotically free (as long as $M < 3N$). We checked explicitly that the required renormalization of the couplings is proportional to the couplings themselves. The scalar field self interaction is of the form

$$-V_4 = \frac{1}{8}g^2 \sum_a \left(\sum_{j=1}^M (A_j^* T^a B_j - B_j^* T^a A_j) \right)^2, \tag{4}$$

where T^a are the generalized Gell-Mann matrices for $SU(N)$. This can be transformed into

$$V_4 = \frac{1}{8}g^2 \sum_a \left(\sum_j (P_j^* T^a P_j - Q_j^* T^a Q_j)^2 \right), \tag{5}$$

where now P_j and Q_j are the scalar fields. The mass terms are now chosen entirely non-supersymmetric:

$$V_2 = \sum_{j=1}^M (-\mu_{jP}^2 P_j^* P_j + \mu_{jQ}^2 Q_j^* Q_j), \tag{6}$$

where μ_{jP}^2 and μ_{jQ}^2 form some hierarchy of scales. We choose μ_{jQ}^2 large enough such that

$$\langle Q_j \rangle = 0. \tag{7}$$

To find $\langle P_j \rangle$ we write

$$\begin{aligned} V &= V_4 + V_2 \\ &= \frac{1}{4}g^2 \sum_{ij} |P_i^* P_j - \kappa P^* P \delta_{ij} - (2/g^2) \tilde{\mu}_j^2 \delta_{ij}|^2 \\ &\quad + \text{constant}. \end{aligned} \tag{8}$$

Here κ is such that

$$\kappa N(2 - \kappa M) = 1, \quad \mu_{jP}^2 = \tilde{\mu}_j^2 - \kappa \sum \tilde{\mu}^2. \tag{9}$$

Note that the $SU(N)$ indices were suppressed. We find that for all j either

$$|\langle P_j \rangle|^2 - \frac{1}{N} \sum_i |\langle P_i \rangle|^2 = \tilde{\mu}_j^2 \quad \text{or} \quad \langle P_j \rangle = 0, \tag{10}$$

Further

$$\langle P_i \rangle^* \langle P_j \rangle = 0 \quad \text{if } i \neq j. \tag{11}$$

We see that a hierarchy of $SU(k)$ groups may emerge, if the μ_{jP}^2 are well chosen. Similar mass hierarchies could successively decouple the unwanted fermions.

Of course the above model is not obviously "natural" in any sense but the requirement of mathematical viability may be stronger than that of naturalness.

Other models can be envisaged in which larger subgroups $SU(4)$, $SU(5)$, ... split off and become confining at energy scales Λ_4 , Λ_5 , ..., larger than the QCD scale Λ_3 .

Besides its possible significance in model building the reader might ignore all that and consider our proof for planar theories just as a small step towards a convergent calculational procedure for QCD.

References

- [1] K. Symanzik, in: Proc. VIth Intern. Conf. on Mathematical physics (Berlin, August 1981).
- [2] D.R.T. Jones, Nucl. Phys. B75 (1974) 531; W.E. Caswell, Phys. Rev. Lett. 33 (1974) 244.
- [3] G. 't Hooft, in: The whys of subnuclear physics (Erice, 1977), ed. A. Zichichi (Plenum, New York) p. 943.
- [4] G. 't Hooft, Nucl. Phys. B72 (1974) 461.
- [5] S. Coleman, in: The whys of subnuclear physics (Erice, 1977), ed. A. Zichichi (Plenum, New York) p. 805.
- [6] E. Brézin et al., Commun. Math. Phys. 59 (1978) 35.
- [7] A.E. Jacobs, Phys. Rev. D23 (1981) 1760.
- [8] C. de Calan and V. Rivasseau, Commun. Math. Phys. 82 (1981) 69.
- [9] P. Fayet and S. Ferrara, Phys. Rep. 32C (1977) 250.