Ref.TH.2901-CERN

VIRTUAL STATIONARY POINTS

W. Nahm

CERN - Geneva

A B S T R A C T

Using ideas from the Morse theory, a type of limiting field configurations is introduced which is expected to yield non-perturbative contributions to coupling constant expansions of topologically non-trivial quantum field theories. Examples are given for two-dimensional non-linear σ models and four-dimensional gauge fields. For the $O(3)$ σ model the complete set is determined. The results are compared with the exact S matrices in two dimensions.

Ref.TH.2901-CERN

16 July 1980

1. - INTRODUCTION

Two-dimensional non-linear σ models [for short $O(N)_2$, etc.] and four-dimensional gauge models have Euclidean functional integrals which get non-perturbative contributions from finite action solutions of the classical equations of motion. Both minima of the action [e.g., the multi-instantons of $O(3)_2$ and of pure gauge theories] and saddle points [as in $O(N)_2$ for $N > 3$] have been extensively discussed ^{1),2)}. For these considerations, saddle points of the finite Morse index i are as relevant as minima, as they can be transformed into minima by fixing i collective co-ordinates.

In $O(3)_2$ the (multi-) instantons and the corresponding anti-instantons are the only Euclidean classical solutions with finite action ³⁾. Taken alone, however, their contribution violates the clustering properties ⁴⁾. To restore them, additional configurations with both instantons and anti-instantons must contribute. But to avoid double counting such configurations should not be put in by hand, because typical instanton-anti-instanton configurations are already included as quantum fluctuations around the minimum.

In this paper a new set of non-perturbative contributions will be identified which has the qualitative features necessary for restoring the cluster properties. They are defined in Section 2 and will be called virtual stationary points. Section 3 contains a description of all virtual stationary points of $O(3)_2$ and partial results for other models. In Section 4 these results are compared with the exact S matrices in two dimensions. Section 5 contains some final remarks.

The present paper just contains basic ideas and results. A more extensive discussion will be given elsewhere. For the mathematical concepts and terminology, consult Ref. 5).

2. - WHAT ARE VIRTUAL STATIONARY POINTS ?

If functional integrals are written as Borel transforms ⁶⁾, i.e.,

$$\int \exp(-A(\varphi)/g) [d\varphi] = \int \exp(-b/g) \rho(b) db, \quad (1)$$

a non-perturbative contribution to $\rho(b)$ vanishes below some $b = b_k$, such that its contribution to the functional integral is proportional to $\exp(-b_k/g)$. In the simplest case the space of fields φ consists of disconnected sectors. If the N^{th} sector has b_N as lower bound of $A(\varphi)$, it will contribute proportionally to $\exp(-b_N/g)$. Instantons are configurations which realize the lower bound. But for some models the bound cannot be realized.

Take, e.g., the Georgi-Glashow model of an $SU(2)$ gauge field plus a Higgs in the adjoint representation, for which

$$A(\varphi, A_\mu) = \int \left(\frac{1}{2} F_{\mu\nu}^2 + \frac{1}{2} (D_\mu \varphi)^2 + \lambda (\varphi^2 - \rho^2)^2 \right) d^4x = A_{\text{gauge}} + A_{\text{kin}} + A_{\text{pot}}. \quad (2)$$

This model has the same topological sectors as the pure gauge model, but a scaling argument shows that for non-zero winding number N there are no minima ⁷⁾. Indeed

$$\varphi'(x) = \varphi(x/\ell), \quad (3)$$

$$A'_\mu(x) = \ell^{-1} A_\mu(x/\ell) \quad (4)$$

yields

$$A(\varphi', A'_\mu) = A_{\text{gauge}} + \ell^2 A_{\text{kin}} + \ell^4 A_{\text{pot}}. \quad (5)$$

Obviously the lower bound of $A(\varphi, A_\mu)$ in the N^{th} sector is the minimum b_N of A_{gauge} . But the bound can only be reached in the limit $\ell \rightarrow 0$. The limiting configuration can be described as a (multi-) instanton of infinitesimal size, plus a Higgs of vanishing length at the instanton position and of constant length outside. Such a limiting configuration will be called a virtual minimum. As for the pure gauge theory, the N^{th} sector yields a contribution $\exp(-b_N/g)$ to the functional integral, where now b_N/g may be interpreted as the action of the virtual minima. Note that the action densities of the virtual minima are well-defined distributions and that the energy momentum tensor vanishes identically.

If one adds suitable higher derivatives with a small coefficient to the Georgi-Glashow Lagrangian, one finds minimal configurations with a small but finite instanton size. Such transformations of virtual stationary points to real ones might be useful for calculating quantum corrections.

In order to generalize the concept of virtual minima to include virtual saddle points, let us go back to the functional integral of Eq. (1). The opening up of a new, disconnected part of the space of fields φ is only the most obvious case of a non-perturbative contribution. This case corresponds to a very crude change of the topology of the field space at the level $A(\varphi) = b_N$. But solutions of the classical equations of motion which are not minima also indicate a change in topology at the corresponding level b_k . Once this is realized, the definition of general virtual stationary points is obvious. Consider for any field theory the spaces

$$M(b) = \{ \varphi \mid A(\varphi) \leq b \}. \quad (6)$$

For increasing b , the homotopy type of $M(b)$ will change at certain values b_k . If for arbitrarily small ε the homotopy type of $M(b_k)$ does not agree with that of $M(b_k - \varepsilon)$, one has a real stationary point. If it does not agree with that of $M(b_k + \varepsilon)$, we call the responsible limiting configurations virtual stationary points. To localize these limiting configurations as limiting points of φ space, one has to find contractable sets U with characteristic functions χ_U , such that $M(b)$ and

$$M(b, \varepsilon \chi_U) = \{ \varphi \mid A(\varphi) \leq b + \varepsilon \chi_U(\varphi) \} \quad (7)$$

have a different homotopy type [minus possible components not connected to $M(b)$] for arbitrarily small ε . From sequences U_n of such sets which shrink to zero, one may select sequences of fields φ_n which allow a description of virtual stationary points as formal limits, like Eqs. (3)-(4) for $\ell \rightarrow 0$.

The definitions given above certainly lack in mathematical rigour, and one can easily make up pathological cases, but for physically interesting Lagrangians they seem to be sufficient. In particular, all the examples discussed later on have well-behaved limits of the action density and of the densities of conserved quantities. The definitions trivially work for virtual minima, where $M(b_0) = \emptyset$, $M(b_0 + \varepsilon) \neq \emptyset$. In this case, the description as limits can just be read off from $M(b_0 + \varepsilon)$ for $\varepsilon \rightarrow 0$.

As for real stationary points one should define for virtual ones the Morse index (= number of negative fluctuation modes) and the nullity (= number of zero modes). For the Morse index i one may imitate a result valid for real stationary points and define it at the lowest number for which

$$\pi_i(M(b_k + \varepsilon), M(b_k)) \neq 1. \quad (8)$$

Similarly, the number z_c of zero modes which generate compact orbits may be defined by writing the minimal dimension d of a space to which $M(b_k + \varepsilon)/M(b_k)$ can be retracted as

$$d = i + z_c. \quad (9)$$

Schrader has worked on an analytic description of Morse indices and zero modes of virtual saddle points, which would be much more convenient to use than the topological descriptions⁸⁾.

3. - EXAMPLES

After compactification at Euclidean infinity the field space of $O(3)_2$ is a space of maps from one two-sphere to another. We denote the spheres by S_z^2 and S_w^2 and parametrize them by the complex numbers

$$z = x + iy \quad (10)$$

and w (including infinity). The action of a map $W : z \rightarrow w$ has the form

$$A(W) = \frac{1}{\pi} \int \left(\left| \frac{\partial W}{\partial z} \right|^2 + \left| \frac{\partial W}{\partial \bar{z}} \right|^2 \right) (1 + |W|^2)^{-2} dx dy, \quad (11)$$

and is bounded below by

$$S(W) = \frac{1}{\pi} \int \left| \left| \frac{\partial W}{\partial z} \right|^2 - \left| \frac{\partial W}{\partial \bar{z}} \right|^2 \right| (1 + |W|^2)^{-2} dx dy, \quad (12)$$

which is the area of S_W^2 covered by the image of S_Z^2 , with the corresponding multiplicity. The action of the instanton

$$W = Z \quad (13)$$

has been normalized to 1.

Consider now only the sector of instanton number zero. The following intuitive arguments show that $M(b)$ for $b \leq 2$ is retractable to $M(0)$, but not for $b > 2$.

For an even instanton number each part of S_W^2 has to be covered an even number of times by the image of S_W^2 . If S_W^2 is to be covered completely, $A(W)$ has to be at least equal to 2, as it is not less than the covered area. But even $A(W) = 2$ is not sufficient, as the lower bound $A(W) = S(W)$ is only realized for instantons and anti-instantons, and to cover S_W^2 once by an instanton and once by an anti-instanton, one would need S_W^2 twice. Consequently, all elements of $M(2)$ will leave some part of S_W^2 uncovered. It is easily visualized that each such map $z \rightarrow w(z)$ can be retracted to the constant map $z \rightarrow w(0)$ by expanding the uncovered region, in a way which depends continuously on the map. The space of those constant maps is just $M(0)$.

To see what happens for $M(2+\epsilon)$ consider only the maps with

$$W(0) = 0. \quad (14)$$

They form the space $\Omega^2 S^2$, the iterated loop space of S^2 . Now

$$\pi_1(\Omega^2 S^2) = \pi_3(S^2) = \mathbb{Z}, \quad (15)$$

which shows that there is a non-retractible loop in $\Omega^2 S^2$, given by the generator of $\pi_3(S^2)$. This generator is represented by the Hopf map

$$S^3 = SU(2) \rightarrow SU(2)/U(1) = S^2. \quad (16)$$

Translating back, this map may be deformed to a non-retractable loop $t \rightarrow M(2+\epsilon)$, $0 \leq t \leq 1$, namely

$$W(z, t) = \left\{ \begin{array}{l} h(3tz) \Theta(1-|z|) + h(3t/\bar{z}) \Theta(|z|-1) \\ \exp(6\pi i t) h(z) \Theta(1-|z|) + h(1/\bar{z}) \Theta(|z|-1) \\ h(3(1-t)z) \Theta(1-|z|) + h(3(1-t)/\bar{z}) \Theta(|z|-1) \end{array} \right\} \text{ for } \left\{ \begin{array}{l} 0 \leq t \leq \frac{1}{3} \\ \frac{1}{3} \leq t \leq \frac{2}{3} \\ \frac{2}{3} \leq t \leq 1 \end{array} \right.$$

(17)

where

$$h(z) = k(\varepsilon) z + \frac{z}{1-|z|}$$

(18)

and

$$k(\varepsilon) \gg \varepsilon^{-1}.$$

(19)

Looking at $t = \frac{1}{2}$ one finds a virtual saddle point consisting of an instanton and an anti-instanton of infinitesimal sizes with a relative angle π between the internal $U(1)$ orientations. The Morse index is 1 by construction and the zero modes are given by the conformal transformations of S^2 .

Considering suitable generators of

$$\pi_2(\Omega^2 S^2) = \pi_1(S^1) = \mathbb{Z}_2,$$

(20)

one finds similar virtual saddle points of the Morse index 2 and action $A(W) = 4$. But this method soon reaches its limits, as the homotopy groups of spheres are notoriously complicated.

It is much easier to try to retract $M(b)$ in such a way that each element moves towards lower action, and to identify all obstructions. In general, such obstructions are given by the real stationary points, for which the gradient of the action vanishes, but also by the virtual stationary points. On approaching a virtual stationary point, the gradient of the action tends to zero everywhere.

For $O(3)_2$ the energy momentum density must tend to zero, too³⁾. Thus the virtual stationary points have to consist of holomorphic and anti-holomorphic pieces. These pieces only can touch in regions where the configuration approaches a constant. This agrees with the structure of the lowest virtual saddle point considered above.

We conclude that the general virtual saddle point of $O(3)_2$ is obtained by iterated insertion of infinitesimal (multi-) instantons or anti-instantons into other real or virtual stationary points. The infinitesimal size is necessary to yield a constant field for the transition region. It is to be measured relative to the size of the configuration into which the insertion takes place. One may, e.g., insert an instanton of size ϵ^2 into an anti-instanton of size ϵ . For instanton number N the virtual saddle points of $O(3)_2$ have the actions

$$b_{Nk} = N + 2k, \quad k = 1, 2, 3, \dots \quad (21)$$

The argumentation developed above may well be generalizable to other non-linear σ models and even to four-dimensional gauge models, as it depends only on the fact that different local solutions can only be joined in regions where they are approximately constant. It is a reasonable conjecture that all these models only have virtual saddle points constructed by inserting classical solutions with infinitesimal scale parameter into other real or virtual stationary points.

For a gauge theory with gauge group G one is interested in the topology of the space of gauge equivalent sets of potentials. This space is homotopic to the space of maps $S^3 \rightarrow G$ with base point⁹⁾, as can be demonstrated easily in a hyperaxial gauge, where no Gribov copies exist. In such a gauge only the gauge potential at infinity matters, as for a given behaviour at infinity one has a topologically trivial affine space. Now the behaviour at infinity is just given by a map $S^3 \rightarrow G$ with a base point. The space of such maps can also be written as the three-fold loop space $\Omega^3 G$.

Consider now the sector of zero instanton number for a gauge field with $G = SU(2)$. The resulting space of maps $S^3 \rightarrow S^3$ is not too different from the space of maps $S^2 \rightarrow S^2$. Thus we may try to construct virtual saddle points in the same way as for $O(3)_2$. In particular one has

$$\pi_1(\Omega^3 S^3) = \pi_4(S^3) = \mathbb{Z}_2, \quad (22)$$

and the generator of this homotopy group is easily constructed by a Freudenthal suspension of the generator of $\pi_3(S^2)$. If one normalizes the action such that the action of one instanton is 1, this generator can be realized in $M(b)$ for $b > 2$, as follows : use a hyper-axial gauge for which $A_1 = 0$ everywhere, $A_2 = 0$ for $x^1 = 0$. Deform the vacuum in such a way that one obtains an instanton at position $(d, 0, 0, 0)$ and an anti-instanton at position $(-d, 0, 0, 0)$, with scale parameters small compared to d . Then at $x^1 = -\infty$ one has one complete covering of $SU(2)$ and at $x^1 = +\infty$ another. The two coverings come together at one group element g_0 for $x^1 = 0$. Now rotate the $SU(2)$ orientation of the instanton by 2π , keeping g_0 fixed, and finally let the instanton-anti-instanton pair annihilate again. The corresponding cycle in $\Omega^3 S^3$ is essentially a suspension of the $\Omega^2 S^2$ cycle of Eq. (17). If the instanton sizes are small enough, it can be realized in $M(2+\epsilon)$ for any ϵ . The configuration is certainly an obstruction against a retraction of $M(2+\epsilon)$ to $M(2)$, as instanton-anti-instanton pairs with an angle π between their $SU(2)$ orientations repel each other. If the angle is 0, it is well known that they attract each other ¹⁰⁾.

It would be nice to have a proof that $M(b)$ for $b \leq 2$ can be retracted to $M(0)$, which only contains the vacuum, but such a proof seems to be harder to find than for $O(3)_2$. More generally, one would like to know if Eq. (21) also applies to gauge theories, but, of course, this problem is even harder.

4. - A CHECK : EXACT S MATRICES

Our results can be checked by a comparison with the exact S matrices of scale invariant two-dimensional models without particle production. Apart from a kinematic factor the two-body S matrices have to be meromorphic functions of the rapidity difference θ alone, where

$$\theta = 2 \log \frac{\sqrt{s} + \sqrt{s - 4m^2}}{2m} = \log \frac{s}{m^2} + \sum_{k=1}^{\infty} c_k \left(\frac{m^2}{s}\right)^k. \quad (23)$$

The functions possess summable asymptotic expansions in terms of $1/\theta$. For $O(3)_2$ the expansion even has a finite radius of convergence, as one has a function which is regular at infinity ¹¹⁾.

Because of the scale invariance, $\log(s/m^2)$ can only appear in the combination

$$\log \frac{s}{m^2} = \log \frac{s}{\mu^2} + \int \frac{2dg'}{\beta(g')} = \log \frac{s}{\mu^2} - \frac{2}{\beta_2 g} - \frac{2\beta_3}{\beta_2^2} \log g + \dots \quad (24)$$

Consequently, the S matrix has a perturbation expansion which is a multiple power series in terms of g , $g^2 \log g$, $g^2 \log(s/\mu^2)$ and

$$\frac{\mu^2}{s} g^{2\beta_3/\beta_2^2} \exp(2/\beta_2 g). \quad (25)$$

Powers of $g^2 \log g$ are not unusual for theories with infra-red problems ^{12),13)}. Powers of the expression (25) are expected from the analysis of renormalons ¹⁴⁾.

The interesting point is that there are no further terms of the form $\exp(-b_k/g)$. Consequently, the contributions of real and virtual stationary points have to fit into the renormalon expansion. One obtains the following remarkable result :

for scale invariant two-dimensional models without particle production the actions of all real and virtual stationary points have to be multiples of the renormalon action $-2/\beta_2 g$.

In Section 3 we have only found multiples of the one-instanton action anyhow, which for $O(N)_2$ turns out to be $-2(N-2)/\beta_2 g$, as expected. The result indicates a possible geometric meaning of the renormalons, but for the moment the connection between renormalons and instantons still is completely obscure.

5. - OPEN QUESTIONS

As we have seen, an analysis of the topology of the function spaces of non-linear σ models and gauge models shows that a certain type of limiting field configurations should yield non-perturbative contributions to the functional integrals. These configurations typically contain instantons of infinitesimal size.

The calculations of quantum corrections to their contributions is an open question. One might expect problems, as in expressions like $\exp(-b_k/g)$ one usually takes the coupling constant g at the instanton scale parameter, which makes no sense, if this parameter is infinitesimal. However, the expansion of the explicit S matrices shows that the relevant scale is given by the momenta of the scattering particles, whereas the instanton size does not enter. A reproduction of the $O(3)_2$ S matrix by a generalization of perturbation theory should lead to a better understanding of these problems.

A technical problem is our use of the compactification of the Euclidean space at infinity. For the moment one cannot even decide if this problem is spurious or important.

All our considerations become irrelevant if one uses the $1/N$ expansion. In fact, the $O(N)_2$ S matrix is even a regular function of N at $N = \infty$, which leaves no room for separate instanton contributions.

I am indebted to R. Schrader for extensive discussions of the properties of the virtual instantons.

REFERENCES

- 1) E. Corrigan and P. Goddard - "Some Aspects of Instantons" (Review), DAMTP 79-18 (1979).
- 2) A.M. Din and W.J. Zakrzewski - Nuclear Phys. B168 (1980) 173.
- 3) W. Garber, S. Ruijsenaars and E. Seiler - Ann.Phys. 119 (1979) 305.
- 4) B. Berg - "Recent Results on the Instanton Approximation", DESY 79/58 (1979).
- 5) J. Milnor - "Morse Theory", Princeton University Press (1963).
- 6) R. Balian, G. Parisi and A. Voros - in "Feynman Path Integrals", Proceedings, Marseille (1978), Springer, p. 337.
- 7) Y.M. Cho - Phys.Letters 81B (1979) 25.
- 8) R. Schrader - Private communication.
- 9) I.M. Singer - Commun.Math.Phys. 60 (1978) 7.
- 10) D. Förster - Phys.Letters 66B (1977) 279.
- 11) A.B. Zamolodchikov and A.B. Zamolodchikov - Nuclear Phys. B133 (1978) 525.
- 12) K. Symanzik - Commun.Math.Phys. 45 (1975) 79.
- 13) W. Buchmüller and E. Remiddi - Nuclear Phys. B162 (1980) 250.
- 14) G. Parisi - Nuclear Phys. B150 (1979) 163.