Ref.TH.3484-CERN

ON FALSE PERTURBATIVE MINIMA OF EFFECTIVE POTENTIALS

Robert D.C. Miller
CERN -- Geneva

ABSTRACT

A perturbatively evaluated effective potential can possess false minima, artefacts of the method approximation method itself. I show that within pure perturbation theory itself it seems almost impossible to conclude whether or not a perturbative minimum actually approximates a true minimum. Renormalization group (R.G.) methods can provide answers on the falseness.

To illustrate my general remarks I study massless ϕ^4 to a two-loop level. I show that different R.G. methods to those of Coleman and Weinberg must be employed to establish the falseness of a one-loop minimum which this theory possesses.

1. INTRODUCTION

If we minimize an effective potential and find a non-zero VEV, then the theory under consideration is said to be spontaneously broken. When expanded about the new vacuum the structure of the theory is generally radically changed from its original unbroken form. Almost all effective potential calculations to date are performed using perturbation theory [recently non-perturbative (lattice) methods have been undergoing development, see for example Callaway and Maloof 1982, Callaway 1982, Ginsparg and Levine 1982]. With a perturbative approach there is a quite worrying aspect, the question is not simply how reliable is the result but whether the indicated spontaneous breaking actually exists. The approximate nature of perturbation theory can generate *false* (i.e. non-existent) minima to effective potentials. This question is the topic addressed here.

The first part of this paper considers what might be useful criteria in determining whether or not a sequence of approximate minima (generated by perturbation series) actually converges to a true minimum of the effective potential. This is a complex issue and I do not attempt to provide any absolute rigour. Rather, I will present four illustrative examples which will show that most intuitive ideas on what might be useful criteria, fail. The examples will indicate that i) perfectly respectable sequences can converge to false minima, and within perturbation theory we would never be able to realize this. ii) A true minimum can be the limit of a sequence of approximate minima generated only by every alternate order in perturbation theory, the approximate minima being non-existent at the other alternative orders. iii) It is always possible that perturbation series will never uncover all minima as in the case where a countably infinite number of them exist. Furthermore no finite order in perturbation theory can assure us that we have uncovered the absolute minimum, it is always possible a lower minimum occurs at higher order still. Finally, iv) it is easy to construct examples where a true minimum exists while no trace of this will be detected in perturbation theory and this is *not* related to $\exp(-1/g^2)$ arguments! It is reasonable to conclude that *confident* use of the effective potential formalism can only come from a non-perturbative analysis. However, as originally pointed out by Coleman and Weinberg (1973) the renormalization the renormalization group (R.G.) can in certain cases establish that a minimum is false (although it cannot prove that a minimum is actually true, at least to my knowledge).

In the second part of this paper, I revisit massless ϕ^4 theory which possesses a false minimum. This is a nice example as it illustrates some of my previous remarks. The zero loop effective potential reveals no breaking. At a one-loop level a non-zero minimum is generated (Coleman and Weinberg 1973); I will take this to the two-loop level and show that the one-loop minimum disappears. In the light

of my general remarks this is seen *not* to provide evidence that the minimum is false. However, I will display the false nature of this minimum by showing it violates the infra-red structure of ϕ^4 and this is done using the R.G. in a different way to that employed by Coleman and Weinberg. (I will show that R.G. improvement of the one-loop effective potential *still* possesses the false minimum; this is in contradiction to a conclusion reached by Coleman and Weinberg.)

In conclusion, it is worth while mentioning that the possible "non-existence" of ϕ^4 (see for example Wilson 1971, Schrader 1976, Baker and Kincaid 1979; see Callaway and Maloof 1982 for a more complete list) does not concern us here as I use it only as a model example. Furthermore, I have only dealt with perturbation theory: the recent work of Callaway and Maloof (1982) apparently provides a non-perturbative proof of the false nature of the discussed ϕ^4 minimum; however, that work also possesses some rather disturbing conclusions as to the usefulness of perturbative effective potential evaluation.

2. GENERAL REMARKS ON FALSE EXTREMA

At a primary level minimizing an effective potential (V) reduces to a search for non-zero real (i.e. $\in \mathbb{R}$) zeros of a given function (possibly of more than one variable).

Normal perturbation theory (for V) in coupling constants can always be re-summed so as to give a perturbation series in terms of the variable(s) $\ln(\sigma^2/\mu^2) = \ell_\sigma$ where σ is (are) the VEV(s) of the problem and μ the subtraction point. Each order of perturbation theory provides a polynomial approximation in ℓ_σ to the effective potential^{*)}.

At this level the issue of false minima becomes the same as the issue of false zeros to a power series. In this section we study the problem from this view point. The aim is to provide four simple examples which will display how it is easy to invalidate the most intuitive criteria for deciding whether or not a zero is false.

For purposes of simplicity, we consider a convergent power series in a single variable x . When truncated at order x^n the truncated series will be denoted by S_n and thus $S(x)$ will be the limit of the sequence $\{S_n\}$. We can loosely define a *false zero* of $S(x)$ as a non-zero, real zero which occurs in at least one of the S_n and which does *not* approximate a *true* zero of S .

A first example will show that false zeros exist. Let

$$f(x) = \sum_{n=0}^{\infty} (-x)^n \quad (1)$$

*) In that we may write $V = V_0 v(\ell_\sigma)$ where V_0 is the tree level effective potential.

The first non-constant approximation to f is given by $f_1(x) = 1 - x$. $f_1(x)$ does have a real zero i.e. $x = +1$. However, we know that $\{f_n\}$ converges to $f(x) = 1/(1+x)$ which has no zeros. Thus $x = +1$ is a false zero of f . Of course, since we *know* the whole series one can provide plenty of reasons why the false zero arose. The radius of convergence of (1) is $r_c = 1$. Since $x = 1$ does not lie within the circle of convergence we would expect this to be a false zero, as it is. Furthermore, one may argue that the approximation $f_1(x) = 1 - x$ is only valid for $x \ll 1$ thus $x = 1$ is obviously a meaningless zero. Another point might be that if we go to the next non-trivial order, $f_2(x) = 1 - x + x^2$, we find no approximate real zeros to f and consequently $x = 1$, which appeared at first order as an approximate real zero, may be discarded forthwith as a false zero. Now let us reflect upon what happens if the *whole* series is not known, moreover as in perturbation theory we might assume we are limited to very few terms of the series. Firstly, we could *not* conclude whether the approximate zero lay inside, on, or outside the circle of convergence. The radius of convergence is a property of the *whole* series and can never be gleaned from any finite number of terms. The first obvious reason for rejecting $x = 1$ as an approximation to a true zero is then not justified, i.e. we cannot conclude it is a false zero in this way. The intuitive argument that " $f_1(x)$ only approximates f for $x \ll 1$, hence $x = 1$, could hardly be trusted", is also invalid, we will see this in a forthcoming example. The third piece of intuition is put into disarray when we observe that $f_n(x)$ has $x = 1$ as a zero for *all* odd n ! Thus, just because an approximate zero disappears at one level of approximation does not guarantee it will not reoccur an infinite number of times in higher order. One may then counter this point by observing that f_n for all even n do not have any real zeros, so "obviously" a criterion for a false zero is that it is not stable at each order; again this is false intuition as will be seen in our next example.

Consider the series

$$g(x) = 1 + 3 \sum_{n=1}^{\infty} (-x)^n \quad (2)$$

This series sums to $(1-2x)/(1+x)$ and has $r_c = 1$. It provides an example where at all even order g_n there is *no* approximate real zero. However, at each odd order g_n a real zero does exist and the sequence of these zeros converges fairly rapidly to a true zero, $x = \frac{1}{2}$: viz $z_1 = 0.333$, $z_3 = 0.443$, $z_5 = 0.482$, $z_7 = 0.495 \dots$ where $g_n(z_n) = 0$ for n odd. Thus we see that without knowledge of r_c in our first example it is difficult to argue $x = 1$ is not a valid approximation to a true zero.

Next we consider a disturbing example. Let

$$h(x) = x \left[1 + \sum_{n=1}^{\infty} \frac{(-x^2)^n}{(2n+1)!} \right] \quad (3)$$

This series sums to $\sin x$. This function has an infinite number of true real zeros at $x = \pm n\pi$, $n = 0, 1, 2, \dots$. Furthermore, (3) has a radius of convergence $r_c \rightarrow \infty$. We are discussing the existence of non-zero real zeros. Approximate zeros of this type occur first in $h_1(x) = x - \frac{x^3}{3!}$, which has real non-zero zeros at $x = \pm\sqrt{6}$. Assuming we have no knowledge of the radius of convergence, the intuitive argument of earlier would claim these are nonsensical zeros since $h_1(x)$ is only meant as a good approximation when $x \ll 1$. However, it is clear that $x = \pm\sqrt{6}$ are in fact first order approximations to the true zeros $\pm\pi$! This example is one in which at each order, $h_n(x)$ will uncover first-order approximations to two new real, true zeros. In the context of minima for perturbative effective potential we can thus never be sure then that we are expanding around the true vacuum; it is always possible that a new minimum of lower energy may be uncovered as we go to higher orders. Furthermore, we will never know what order is "high enough" to be sure that we have expanded about the true vacuum.

The conclusion from this would be that the location of the absolute minimum is strictly a non-perturbative question, despite the fact that a perturbative effective potential is calculable to any finite order.

Our final example provides an illustration where the true zero is outside the "perturbative domain" and "can never be revealed by perturbative theory". Let

$$I(x) = 1 + 3 \sum_{n=1}^{\infty} (-x)^n \quad (4)$$

This is just $g(x) + 1$ and sums to $(2-x)/(1+x)$, with $r_c = 1$. $g(x)$ has a true zero at $x = 2$ which lies outside the circle of convergence, hence we would expect it to be a zero non-recoverable from the series of (4). In fact, one finds that I_n for even n have no real zeros. Furthermore, I_n for odd n do have real zeros and a sequence is generated; $z_1 = 0.667$, $z_3 = 0.796$, $z_5 = 0.852$, $z_7 = 0.884$, $z_9 = 0.904$, $z_{11} = 0.918$, etc. It is easy to show this sequence is bounded above by 1 and presumably this is the limit of this sequence. Thus $x = 2$ is not recovered as a true zero and more surprising is the fact that a perfectly respectable sequence of approximate zeros is generated in which each term lies within the circle of convergence, yet converges to a point on the circle of convergence which is not a zero. However, we must again remember that in perturbation series we never know the radius of convergence and consequently intuition would accept this false zero as a true zero; once more intuition fails.

As a final comment we make some remarks on false minima. In the context of what has been discussed above, i.e. essentially false extrema, the question of whether or not one has a false minimum is answered only when one has confirmed that a true zero exists. The true zero will be a limit of a sequence of approximate zeros. One may then hopefully study the sequence of derivatives $\{S'_n(z_n)\}$ to

determine whether or not the minimum is false. Again one has problems. Even if every term in this sequence is positive (hence each approximate extremum was a minimum) the limit of the sequence may be zero, in which case the limit may not be a minimum. One can imagine many other pathological cases, for example an alternating sequence converging to zero.

3. MASSLESS ϕ^4 REVISITED

Massless ϕ^4 is a simple example which possesses a false minimum. It will be used to illustrate some of the features mentioned in Section 2. Also we will look at *how* the minimum is proved to be false by use of the R.G. My approach will contrast the original method of Coleman and Weinberg (1973) which depends upon R.G. improvement of the one-loop effective potential. (In fact I will show that R.G. improvement of V does not eliminate the false minimum as originally claimed.)

The two-loop effective potential for this theory has been calculated by Lee and Sciaccaluga (1975). I redefine the subtraction point (μ) (hence present results in a variant of m.s.) to obtain

$$V(\phi_{cl}) = \frac{\lambda}{4!} \phi_{cl}^4 S(\phi_{cl}) \quad (5)$$

where $S(\phi_{cl})$ is a perturbation series

$$S(\phi_{cl}) = 1 + \alpha \frac{3}{2} l_\phi + \alpha^2 \frac{3}{56} (84 l_\phi^2 - 176 l_\phi - 137) \quad (6)$$

Here ϕ_{cl} is the classical field ($\phi_{cl} = \delta W[J]/\delta J$) and I have introduced shorthand notations

$$\alpha = \frac{\lambda}{16\pi^2} \quad (7)$$

[like Stevenson's (1981) couplant] and

$$l_\phi = L_n(\phi_{cl}^2/\mu^2) \quad (8)$$

One can always resum the $S(\phi_{cl})$ perturbation series. At u loop order,

$$S(\phi_{cl}) = \sum_{m=0}^u (\alpha l_\phi)^m C_m \quad (9)$$

C_m is just a pure perturbation series in α , i.e.

$$C_m = \sum_{n=0}^{u-m} \alpha^n r_n \quad (10)$$

r_n just plain numerical coefficients. Resumming (6) for example gives

$$S(\phi_{cl}) = (1 + \alpha \cdot 0 - \alpha^2 \frac{411}{56}) + (\alpha l_\phi) (\frac{3}{2} - \alpha \frac{66}{7}) + (\alpha l_\phi)^2 (\frac{9}{2}) \quad (11)$$

Let us consider a standard perturbative analysis of this. Spontaneous breaking is indicated by a non-zero (minimal) solution to

$$\left. \frac{dV(\phi_{cl})}{d\phi_{cl}} \right|_{\phi_{cl}=\sigma} = 0 \quad (12)$$

This means, in terms of $S(\ell_\sigma)$, that the non-zero solutions must satisfy

$$2S(\ell_\sigma) + \frac{\partial S(\ell_\sigma)}{\partial \ell_\sigma} = 0 \quad (13)$$

At a two-loop level this implies

$$\left[2 + \alpha \frac{3}{2} - \alpha^2 \frac{675}{28} \right] + (\alpha \ell_\sigma) \left[3 - \alpha \frac{69}{7} \right] + (\alpha \ell_\sigma)^2 [9] = 0 \quad (14)$$

Had we only worked to a one-loop level we would have obtained the equation

$$\alpha \ell_\sigma = -\frac{2}{3} - \frac{\alpha}{2} \quad (15)$$

which is a minimum of the one-loop $S(\ell_\sigma)$ (hence one-loop V). Coleman and Weinberg also provided intuitive arguments why this is a false minimum. In view of Section 2, we can see that their intuitive argument is hard to validate. Intuitively from (11) we expect perturbation theory to make sense when

$$|\alpha \ell_\sigma| \ll 1 \quad (16a)$$

and

$$\alpha \ll 1 \quad (16b)$$

Since we may choose α as small as we wish by varying μ we see (15) implies that for such α

$$|\alpha \ell_\sigma| \approx \frac{2}{3} \quad (17)$$

although not $\ll 1$ it is < 1 , so the solution in (15) is not utter nonsense at this stage^{*)}. Furthermore, Section 2 illustrated that (16) may not be a totally realistic requirement as we have no idea of the radius of convergence in (14) (i.e. ≤ 1 may be sufficient, we do not know).

We now proceed to the two-loop level. The minimization condition has lead to (14). This is just a quadratic equation in $\alpha \ell_\sigma$. The discriminant of this equation is numerically found to be

$$\Delta \doteq -63 (1 + 1.80\alpha - 15.3\alpha^2) \quad (18)$$

*) Because Coleman and Weinberg chose a different expansion parameter, (17) in their case looked nonsensical, hence intuitive falseness.

If we look towards the $(\alpha\ell_0)^0$ coefficient in (14), it appears to make perturbative sense only for $\alpha \lesssim 0.1$. Thus, for sensible two-loop perturbation theory Δ is negative and (14) has no real solution. In particular the one-loop minimum has disappeared. As stressed in Section 2 this does *not* provide any evidence in favour of rejecting the one-loop minimum as false. This is rather surprising since for example, suppose Mahanthappa and Sher (1980) had found that the one-loop minimum, which predicts the Coleman-Weinberg Higgs mass, disappeared at a two-loop level. One would then naturally be very dubious in accepting the one-loop prediction as physical but I would argue that such a doubt is not justified. Of course, this is only a hypothetical situation (Mahanthappa and Sher confirmed an improved approximation of the minimum at the two-loop level).

So much for pure perturbation theory. We now show what happens when it is combined with the R.G.

First I begin with a brief review of the R.G. for ϕ^4 . The field and coupling constant both undergo renormalization

$$\begin{aligned}\phi_{c1,0} &= Z_\phi \phi_{c1} \\ \lambda_0 &= Z_\lambda \lambda\end{aligned}\tag{19}$$

I define R.G. functions γ and β by

$$\gamma = \frac{\mu}{\phi_{c1}} \frac{d\phi_{c1}}{d\mu} = -\frac{\mu}{Z_\phi} \frac{dZ_\phi}{d\mu}\tag{20}$$

and

$$\beta = \frac{\mu}{\alpha} \frac{d\alpha}{d\mu}\tag{21}$$

The calculation of γ and β at a one-loop level is trivial, the results are

$$\gamma = 0\tag{22}$$

$$\beta = 3\alpha\tag{23}$$

(These are scheme-independent results.) From (21) we see

$$\ln(\mu/\mu') = \int_{\alpha'}^{\alpha} \frac{dx}{x\beta(x)}\tag{24}$$

where α' denotes $\alpha(\mu')$ and α , $\alpha(\mu)$. Since β is positive near $x = 0$ [see (23)], (24) may be expanded asymptotically for both μ and $\mu' \rightarrow 0$, to obtain

$$\alpha = \frac{\alpha'}{1 + \frac{3}{2}\alpha' \ln(\mu'^2/\mu^2)}\tag{25}$$

If we denote by ϕ'_{c1} , $\phi_{c1}(\mu')$ then to R.G. one-loop order (20) and (22) tell us that

$$\phi'_{c1} = \phi_{c1} \quad (26)$$

(This result is of particular importance later.)

Finally, imagine μ' fixed in (24) and then allow $\mu \rightarrow \infty$. The left-hand side of (24) then $\rightarrow +\infty$ and assuming the only zero of $\beta(x)$ is at $x = 0$ we must conclude $\alpha \rightarrow +\infty$, i.e.

$$\begin{aligned} \alpha(\mu) &\rightarrow +\infty \\ \text{as } \mu &\rightarrow +\infty \end{aligned} \quad (27)$$

(25) and (27) reflect the standard behaviour of infra-red free theories in the absence of non-zero infra-red fixed points.

We can now provide our simple proof on the false nature of the one-loop minimum. We can satisfy (16b) by allowing $\mu \rightarrow 0$. Let us do this for fixed μ' . In this limit (25) tells us that

$$\alpha(\mu) \sim \frac{1}{-\frac{3}{2} \ln \mu^2 + \text{small term}} \quad (28)$$

(as $\mu \rightarrow 0$ the $\ln$ will be negative).

However, for the *same* limit, i.e. $\mu \rightarrow 0$ and σ fixed [such a procedure is valid because of (26)], the one-loop minimum, satisfying (15) tells us that

$$\alpha(\mu) \sim \frac{1}{+\frac{3}{2} \ln \mu^2 + \text{small term}} \quad (29)$$

One would like to feel that a sign error has been made in obtaining (29) but this is not the case. We thus see that the one-loop minimum leads to an incorrect infra-red behaviour for $\alpha(\mu)$. This establishes its unphysical, i.e. false, nature.

I will close this section by showing that R.G. improvement of the one-loop effective potential does not lead to the disappearance of the one-loop perturbative minimum [this is contrary to the original claims of Coleman and Weinberg (1973)].

First we must agree upon what is meant by R.G. improved perturbation theory. If an object $\Gamma(P)$ is known perturbatively then sensible use of that perturbative result will restrict us to a certain domain of parameter space from which we may choose P . This domain of definition for renormalized parameters p may be extended by using the R.G. In other words the R.G. may be viewed as a method of analytic continuation for perturbation theory.

Coleman and Weinberg chose to R.G. improve the perturbative result for the fourth partial derivative of V w.r.t. the classical field, $V^{(4)}$. We will directly

R.G. improve V itself for this is just as easy (their motivation was that $V^{(4)}$ is dimensionless hence may be expressed in terms of λ and ϕ_{c1}/μ , i.e. two variables instead of three).

The principal observation is that the effective potential must be R.G. invariant as is readily seen by considering the momentum expansion of the generating functional for one particle irreducible amplitudes. Using the notation already introduced we immediately conclude

$$V(\mu', \phi'_{c1}, \alpha') = V(\mu, \phi_{c1}, \alpha) \quad (30)$$

I now adopt the convention that the point (μ, ϕ_{c1}, α) belongs to the perturbative domain of parameter space. The point $(\mu', \phi'_{c1}, \alpha')$ will be an arbitrary point in the R.G. extended domain, which of course contains the perturbative domain. For this reason although we *cannot* generally use the perturbative result for $V(\mu', \phi'_{c1}, \alpha')$ we *can* do so for $V(\mu, \phi_{c1}, \alpha)$. Hence, at the one-loop level

$$V(\mu', \phi'_{c1}, \alpha') = \frac{\lambda}{4!} \phi_{c1}^4 (1 + \alpha \frac{3}{2} \ell_\phi) \quad (31)$$

This is required to be expressed in terms of μ' , ϕ'_{c1} and α' ; this we can do using (25) and (26). To locate the extremum of (31), it is sufficient to make all ϕ'_{c1} dependence explicit. (25) shows α has no ϕ_{c1} dependence when expressed in terms of α' and μ' , thus (26) alone determines the ϕ'_{c1} dependence of (31) which is trivially

$$V(\mu', \phi'_{c1}, \alpha') = \frac{\lambda}{4!} \phi_{c1}'^4 (1 + \alpha' \frac{3}{2} \ell_\phi') \quad (32)$$

Consequently, the minimum remains with us

$$\alpha' \ell_\phi' (\sigma'^2/\mu'^2) = -\frac{2}{3} - \frac{\alpha'}{2}$$

or using (25)

$$\alpha' \ell_\phi' = -\frac{2}{3} - \frac{\alpha'}{2} [1 + 4 \ell_\phi' (\mu'^2/\mu^2)] \quad (33)$$

The minimum has been shown to persist upon R.G. improvement (albeit it is shifted in location) and this is a simple consequence of the fact that at a one-loop level $\gamma = 0$. (This is not simply a contradiction of the Coleman-Weinberg conclusion, since it is straightforward to locate their error, but this is not the point here.)

4. CONCLUDING REMARKS

In this paper I address the question as to whether or not (within pure perturbation theory) it is possible to confidently conclude that an approximate minimum to an effective potential actually represents an approximation to a true minimum and that it is not just an artefact of the approximation method itself. A number of examples seem to indicate that this is not possible, one must go beyond pure perturbation theory for definite answers.

Massless ϕ^4 theory is revisited (at the two-loop level) to illustrate some of the points made, for this theory possesses a false minimum. I show that to this level perturbation theory really tells us nothing of the falseness of the one-loop minimum. It is shown that simple R.G. arguments are sufficient to establish the false nature of the minimum, one does not have to take a purely non-perturbative approach. Here, my arguments differ from Coleman and Weinberg for, regarding their approach, I show that R.G. improvement of the one-loop effective potential does not eliminate the false minimum as was originally claimed.

Purely non-perturbative methods (Callaway and Maloof 1982) seem to verify the falseness of the one-loop minimum. However, they could equally pose a threat to the usefulness of perturbation theory when it comes to effective potential analyses.

Acknowledgements

I would like to thank A. Petermann for a reading of the manuscript.

REFERENCES

- Baker G.A., Kincaid J., 1979, Phys. Rev. Lett. 42, 1431.
- Callaway D.J.E., 1982 (July), "Non-perturbative analyses of the effective potential in the $O(N)$ model", ANL-HEP-PR-82-30.
- Callaway D.J.E., Maloof D.J., 1982 (April), "Effective potential of lattice ϕ^4 theory", ANL-HEP-PR-82-19.
- Coleman S., Weinberg E., 1973, Phys. Rev. D 7, 1888.
- Ginsparg P., Levine H., 1982 (November), "Monte Carlo evaluation of the effective action", HUTP-82/A034.
- Lee S.Y., Sciaccaluga A.M., 1975, Nucl. Phys. B96, 435.
- Mahanthappa K.T., Sher M.A., 1980, Phys. Rev. D 22, 1711.
- Schrader R., 1976, Phys. Rev. B 14, 172.
- Stevenson P.M., 1981, Phys. Rev. D 23, 2916.
- Wilson K.G., 1971, Phys. Rev. B 4, 3184.