

INTRODUCTION TO THE BACKGROUND FIELD METHOD*

BY L. F. ABBOTT**

CERN, Geneva

(Received July 20, 1981)

The background field approach to calculations in gauge field theories is presented. Conventional functional techniques are reviewed and the background field method is introduced. Feynman rules and renormalization are discussed and, as an example, the Yang-Mills β function is computed.

PACS numbers: 11.10.Np, 11.10.Gh

1. Introduction

The background field method is a technique for quantizing gauge field theories without losing explicit gauge invariance. It makes gauge theories easier to understand and greatly simplifies computations. In this review I will present the formalism of this method and show how it is applied to gauge theory calculations.

The background field method was introduced by DeWitt [1, 2] in a formalism which was applicable to one-loop processes. The extension to multi-loop calculations, which involved a reformulation of the method, was first made by 't Hooft [3] and then discussed in more detail by DeWitt [4], Boulware [5] and by myself [6]. It is this extension, valid to all orders of perturbation theory, which I will present here. The background field method is used extensively in gravity [7] and supergravity [8] theories. In addition, it has been used to derive light-particle effective field theories from grand unified models [9], to compute the Yang-Mills β function up to two loops [6] and to perform calculations in lattice gauge theories [10]. In all of these applications, the great simplifications introduced by the method play a key role.

Any formulation of a gauge field theory begins with a gauge invariant Lagrangian. However, in order to quantize the theory a gauge must be chosen. In the conventional formulation, this means that the Lagrangian you actually use to derive Feynman rules and perform calculations, consisting of the classical Lagrangian plus gauge-fixing and ghost terms, is not gauge invariant. Of course, any physical quantity calculated will be

* Presented at the XXI Cracow School of Theoretical Physics, Paszówka, May 29 — June 9, 1981.

** Address: Physics Department, Brandeis University, Waltham, MA 02254, USA.

gauge invariant and independent of the particular gauge chosen, but quantities with no direct physical interpretation like off-shell Green's functions or divergent counterterms will not be gauge invariant. Green's functions in the conventional formulation do not directly reflect the underlying gauge invariance of the theory but rather obey complicated Slavnov-Taylor identities [11] resulting from BRS invariance [12]. In the background field approach, one arranges things so that explicit gauge invariance, present in the original Lagrangian, is still present once gauge-fixing and ghost terms have been added. As a result, in this formalism, Green's functions obey the naive Ward identities of gauge invariance and even unphysical quantities like divergent counterterms take a gauge invariant form.

To show what the retention of explicit gauge invariance means let me present a simple "proof" of the renormalizability of pure Yang-Mills theory. Suppose by power counting arguments you have convinced yourself that all of the divergences in Yang-Mills theory can be cancelled by local counterterms of dimension four. You would then list all such terms like

$$(\partial_\mu A_\nu)^2, [A_\mu, A_\nu]^2, F_{\mu\nu}^2, \dots \quad (1.1)$$

Next, using gauge invariance, you might argue that the only term on this list which can actually appear is $F_{\mu\nu}^2$ since it alone is gauge invariant. Then, since $F_{\mu\nu}^2$ is also the bare Lagrangian you would find that the theory is renormalizable. Well, this is a nice argument, but the problem is that in the conventional formulation of gauge theories it is completely wrong. The reason for this is that the counterterms, being unphysical, do not take the gauge invariant form $F_{\mu\nu}^2$. However, in the background field approach the divergences *are* restricted by the requirement of gauge invariance, and the above proof becomes correct. In fact, arguments exactly like this are used to prove the absence of divergences in various orders for gravity [7] and supergravity [8] theories. These proofs of course rely on the background field method for their validity.

This introduction consists of two main parts. In the first part I will develop the background field formalism. This formalism relies heavily on conventional functional methods in field theory. Therefore I will begin by reviewing standard functional techniques, introducing the generating functionals for disconnected, connected and one-particle-irreducible graphs. Then I will introduce the background field method and show how it is related to the conventional approach. In the second part I will discuss how to apply the method to actual calculations. I will give the Feynman rules, show how the renormalization programme works and, as an explicit example, describe the calculation of the Yang-Mills β function in the background field method.

2. The background field formalism

2.1. Review of functional methods

In any field theory, the ultimate goal of one's calculations is the S matrix. The S matrix can be obtained from the Green's functions of the theory by LSZ reduction [13]. In the functional approach, the Green's functions are determined by taking functional derivatives

with respect to the source function J of the generating functional

$$Z[J] = \int \delta Q \exp i[S[Q] + J \cdot Q]. \quad (2.1.1)$$

In this formula, Q is a field with classical action S and

$$J \cdot Q \equiv \int d^4x (JQ), \quad (2.1.2)$$

$Z[J]$ is determined by performing the functional integral over all configurations of the field Q . The Green's functions are defined by

$$\langle 0|T\{Q \dots Q\}|0\rangle = \int \delta Q(Q \dots Q) \exp iS[Q] = \left(\frac{1}{i} \frac{\delta}{\delta J}\right)^n Z[J] \Big|_{J=0}. \quad (2.1.3)$$

In the discussion here, I will suppress all indices and arguments of Q and J . The meaning of the various formulae will be clarified by writing them in diagrammatic form in figures. In Sections 2.1 and 2.2 I will not be discussing gauge theories so Q might be taken to be a scalar field. The complications introduced by a gauge theory will be discussed in Section 2.3.

The Green's functions defined in Eq. (2.1.3) are the disconnected Green's functions, that is, they contain completely disjoint pieces. The relation between disconnected and connected Green's functions is illustrated in Fig. 1. The disconnected pieces of Green's

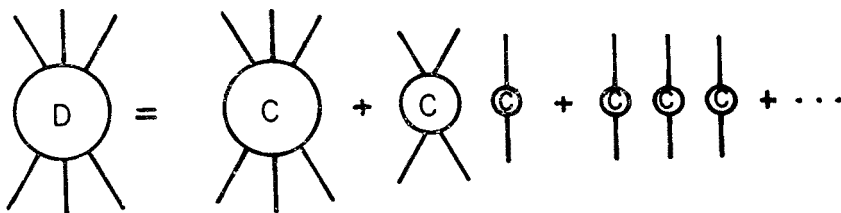


Fig. 1. The relation between connected and disconnected graphs. D = disconnected, C = connected

functions do not contribute to the S matrix. Therefore, it is better to work only with connected Green's functions. These are generated by taking J derivatives of

$$W[J] = -i \ln Z[J]. \quad (2.1.4)$$

At first sight it may seem surprising that the disconnected pieces can be removed just by taking the logarithm in Eq. (2.1.4). Let me show how this works by constructing the first few Green's functions from (2.1.4). Taking a derivative of (2.1.4) with respect to J we find from Eq. (2.1.3)

$$\frac{\delta W}{\delta J} = \frac{\langle 0|Q|0\rangle}{\langle 0|0\rangle}, \quad (2.1.5)$$

which is just the normalized vacuum expectation value of Q in the presence of the source J . Taking a second derivative gives,

$$\frac{1}{i} \frac{\delta^2 W}{\delta J^2} = \left[\frac{\langle 0|T\{QQ\}|0\rangle}{\langle 0|0\rangle} - \left(\frac{\langle 0|Q|0\rangle}{\langle 0|0\rangle} \right)^2 \right]. \quad (2.1.6)$$

The term in brackets is just the full two-point Green's function minus its disconnected part. The factor of $1/i$ is just part of the definition of the connected two-point function, so the

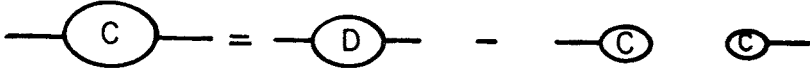


Fig. 2. The relation between the connected (C) and disconnected (D) two-point functions

second derivative of W is the connected two-point function as promised. Equation (2.1.6) is displayed graphically in Fig. 2. Now let us take one more J derivative,

$$\begin{aligned} \left(\frac{1}{i} \right)^2 \frac{\delta^3 W}{\delta J^3} &= \frac{\langle 0|T\{QQQ\}|0\rangle}{\langle 0|0\rangle} - 3 \frac{\langle 0|T\{QQ\}|0\rangle \langle 0|Q|0\rangle}{\langle 0|0\rangle^2} + 2 \left(\frac{\langle 0|Q|0\rangle}{\langle 0|0\rangle} \right)^3 \\ &= \frac{\langle 0|T\{QQQ\}|0\rangle}{\langle 0|0\rangle} - 3 \left(\frac{1}{i} \frac{\delta^2 W}{\delta J^2} \right) \frac{\langle 0|Q|0\rangle}{\langle 0|0\rangle} - \left(\frac{\langle 0|Q|0\rangle}{\langle 0|0\rangle} \right)^3. \end{aligned} \quad (2.1.7)$$

To see the significance of this Equation, we draw it out graphically in Fig. 3. Clearly the third derivative of W is just the connected three-point Green's function, all disconnected pieces have been correctly removed. I will not bother to prove that W generates all the

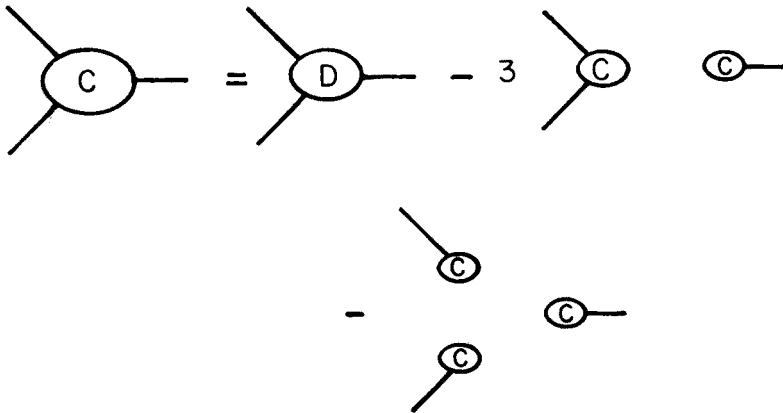


Fig. 3. The relation between connected (C) and disconnected (D) three-point functions

connected Green's functions. I hope my derivation of the first three shows you how things work. It is not too difficult to complete the proof for all Green's functions inductively.

The connected Green's functions can further be simplified by expressing them in terms of one-particle-irreducible pieces as is shown in Fig. 4. A diagram is called one-particle

reducible if it can be split into two disjoint pieces by cutting a single internal line. It is more economical to compute only one-particle-irreducible (1PI) graphs and then string these together into trees of 1PI parts than it is to compute all the connected Green's functions

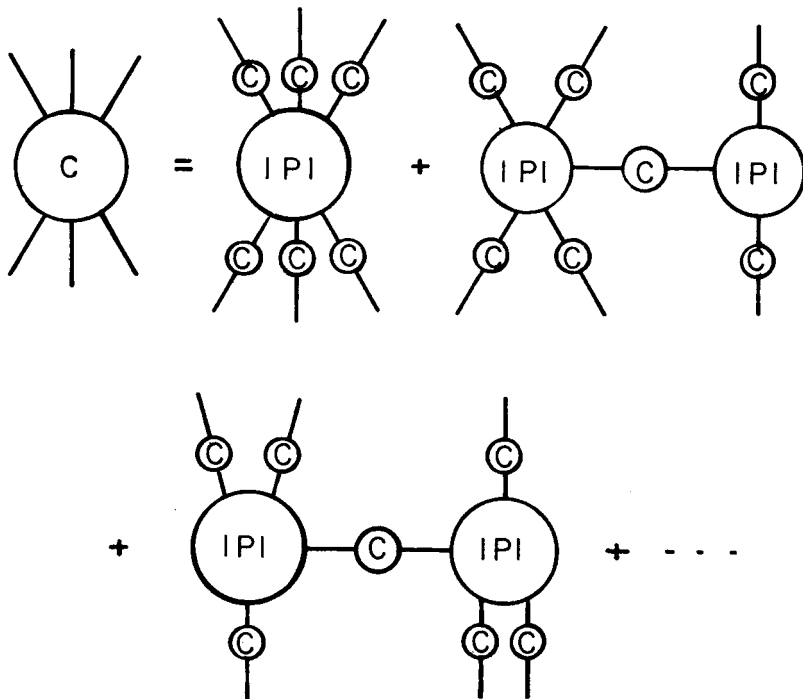


Fig. 4. The relation between connected (C) and 1PI graphs

directly. The 1PI Green's functions are generated by a functional called the effective action. It is defined as

$$\Gamma[\bar{Q}] = W[J] - J \cdot \bar{Q}, \quad (2.1.8)$$

where

$$\bar{Q} \equiv \frac{\delta W}{\delta J}. \quad (2.1.9)$$

One should be careful to distinguish between the original field variable Q , and \bar{Q} which from Eqs. (2.1.9) and (2.1.5) can be seen to be the vacuum expectation value of Q in the presence of the source J . It is quite remarkable that the simple transformation of Eqs. (2.1.8) and (2.1.9) is enough to assure that the \bar{Q} derivatives of $\Gamma[\bar{Q}]$ will be the 1PI Green's functions. Again I will not prove that this is so for all Green's functions, but just consider the first few cases [14]. The first derivative of Eq. (2.1.8) gives

$$\frac{\delta \Gamma}{\delta \bar{Q}} = -J. \quad (1.2.10)$$

This is just the quantum-mechanical field equation for \bar{Q} which replaces the classical field equation $\delta S/\delta Q = -J$ in the quantized theory. Taking a derivative of (2.1.10) we find

$$\frac{\delta^2 \Gamma}{\delta \bar{Q}^2} = -\frac{\delta J}{\delta \bar{Q}}. \quad (2.1.11)$$

However, recall that

$$\bar{Q} = \frac{\delta W}{\delta J} \quad (2.1.12)$$

so that

$$-\frac{\delta J}{\delta \bar{Q}} = \left[-\frac{\delta \bar{Q}}{\delta J} \right]^{-1} = \left[-\frac{\delta^2 W}{\delta J^2} \right]^{-1} = iD^{-1}, \quad (2.1.13)$$

where D is just the full propagator. Thus we find that the second derivative of Γ is the inverse propagator (up to a factor of i which is conventionally absorbed into the definition of the 1PI Green's functions)

$$\frac{1}{i} \frac{\delta^2 \Gamma}{\delta \bar{Q}^2} = D^{-1}. \quad (2.1.14)$$

It is easy to see why the 1PI two-point function is the inverse propagator. Write Eq. (2.1.14) as

$$D \frac{1}{i} \frac{\delta^2 \Gamma}{\delta \bar{Q}^2} D = D. \quad (2.1.15)$$

As shown in Fig. 5, this equation just means that the full propagator is obtained from the 1PI two-point function by dressing the two external legs with propagators. Equation (2.1.13) can also be used to derive the identity

$$\frac{\delta}{\delta \bar{Q}} = \frac{\delta J}{\delta \bar{Q}} \frac{\delta}{\delta J} = D^{-1} \frac{1}{i} \frac{\delta}{\delta J}. \quad (2.1.16)$$

This equation is really the reason why the effective action generates 1PI graphs. When operating on W , $\delta/\delta J$ adds an external line to a Green's function. Then, by (2.1.16), $\delta/\delta \bar{Q}$

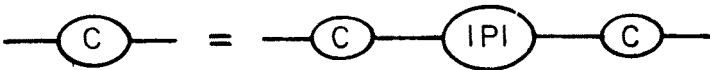


Fig. 5. The relation between connected (C) and 1PI two-point functions

when operating on Γ adds an external line and removes the propagator from this line. This continual amputation of the external propagators is what keeps the diagrams 1PI. For example, let us consider the third derivative of the effective action. From Eqs. (2.1.11)

and (2.1.13) we find that

$$\frac{\delta^3 \Gamma}{\delta \bar{Q}^3} = D^{-1} \frac{1}{i} \frac{\delta}{\delta J} \left[-\frac{\delta^2 W}{\delta J^2} \right]^{-1} = D^{-1} \frac{1}{i} \frac{\frac{\delta^3 W}{\delta J^3}}{\left(\frac{\delta^2 W}{\delta J^2} \right)^2} = i D^{-3} \frac{\delta^3 W}{\delta J^3} \quad (2.1.17)$$

or equivalently

$$\left(\frac{1}{i} \right)^2 \frac{\delta^3 W}{\delta J^3} = i D^3 \frac{\delta^3 \Gamma}{\delta \bar{Q}^3}. \quad (2.1.18)$$

This equation, expressed graphically in Fig. 6, just states that the connected three-point

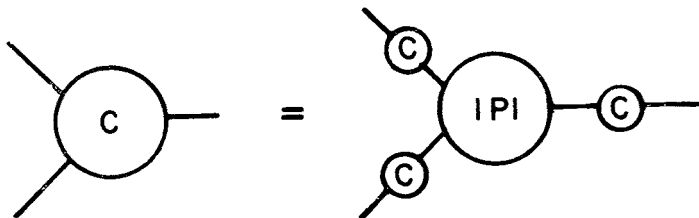


Fig. 6. The relation between connected (C) and IPI three-point functions

function is obtained from the IPI three-point function by adding propagators to the external legs. Finally, consider

$$\begin{aligned} \frac{\delta^4 \Gamma}{\delta \bar{Q}^4} &= D^{-1} \frac{1}{i} \frac{\delta}{\delta J} \left[\frac{\delta^3 W}{\delta J^3} \left(\frac{\delta^2 W}{\delta J^2} \right)^{-3} \right] = D^{-4} \frac{\delta^4 W}{\delta J^4} + 3i D^{-5} \left(\frac{\delta^3 W}{\delta J^3} \right)^2 \\ &= D^{-4} \frac{\delta^4 W}{\delta J^4} - 3i D \left(\frac{\delta^3 \Gamma}{\delta \bar{Q}^3} \right)^2. \end{aligned} \quad (2.1.19)$$

This can also be written as

$$\left(\frac{1}{i} \right)^3 \frac{\delta^4 W}{\delta J^4} = i D^4 \frac{\delta^4 \Gamma}{\delta \bar{Q}^4} + 3 \left(i \frac{\delta^3 \Gamma}{\delta \bar{Q}^3} \right) D \left(i \frac{\delta^3 \Gamma}{\delta \bar{Q}^3} \right) D^4. \quad (2.1.20)$$

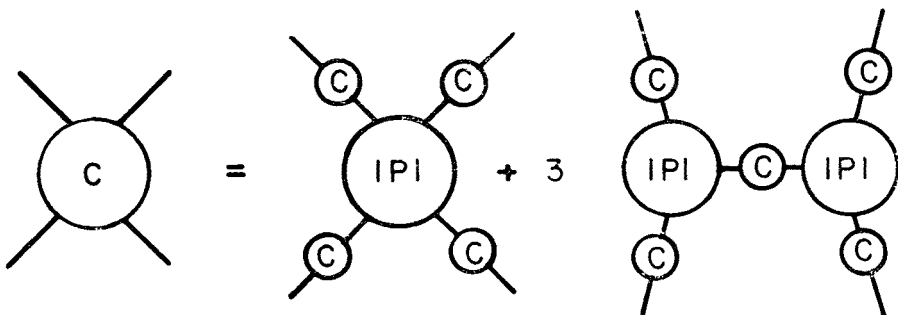


Fig. 7. The relation between connected (C) and IPI four-point functions

This equation is drawn out in Fig. 7 and clearly shows how the full connected four-point function is correctly constructed from its IPI parts.

2.2. The background field method

We have seen in the last Section that an important quantity to calculate in a field theory is the effective action, $\Gamma[\bar{Q}]$. Once this is known, the S matrix can be constructed by stringing together trees of 1PI parts to generate the full connected Green's functions, then amputating external propagators, putting all momenta on shell and adding appropriate external wave-function factors. The background field method is a convenient way of computing the effective action. I will begin by considering non-gauge theories. For these theories, the background field method is identical to the conventional "field-shifting" method which has frequently been used to compute the effective action and, in particular, the effective potential [15].

Recall that the generating functional for disconnected graphs was

$$Z[J] = \int \delta Q \exp i[S[Q] + J \cdot Q]. \quad (2.2.1)$$

Let us define an analogous quantity in which we write the classical action S as a function of the field Q plus an arbitrary background field ϕ . Thus

$$\tilde{Z}[J, \phi] = \int \delta Q \exp i[S[Q + \phi] + J \cdot Q], \quad (2.2.2)$$

\tilde{Z} depends both on the conventional source J and on the background field ϕ which can be thought of as an alternate source. Also, by analogy with the conventional generator of connected graphs,

$$W[J] = -i \ln Z[J] \quad (2.2.3)$$

we define

$$\tilde{W}[J, \phi] = -i \ln \tilde{Z}[J, \phi]. \quad (2.2.4)$$

In the conventional approach we defined

$$\bar{Q} = \frac{\delta W}{\delta J} \quad (2.2.5)$$

so here we write

$$\tilde{Q} = \frac{\delta \tilde{W}}{\delta J}. \quad (2.2.6)$$

Finally, to replace the conventional effective action

$$\Gamma[\bar{Q}] = W[J] - J \cdot \bar{Q} \quad (2.2.7)$$

we define the background field effective action

$$\tilde{\Gamma}[\tilde{Q}, \phi] = \tilde{W}[J, \phi] - J \cdot \tilde{Q}. \quad (2.2.8)$$

To see what the point of all these definitions is, we shift the variable of integration in Eq. (2.2.2), $Q \rightarrow Q - \phi$. This allows us to relate the conventional and background field generating functions. One immediately finds

$$\tilde{Z}[J, \phi] = Z[J] \exp -iJ \cdot \phi. \quad (2.2.9)$$

Taking a logarithm of this equation gives

$$\tilde{W}[J, \phi] = W[J] - J \cdot \phi. \quad (2.2.10)$$

Next, differentiate Eq. (2.2.10) with respect to J . Recalling the definitions (2.2.5) and (2.2.6), we obtain the relation

$$\tilde{Q} = \bar{Q} - \phi. \quad (2.2.11)$$

Finally, from Eqs. (2.2.8) and (2.2.7) we see that

$$\tilde{\Gamma}[\tilde{Q}, \phi] = W[J] - J \cdot \phi - J \cdot \bar{Q} + J \cdot \phi = \Gamma[\bar{Q}]. \quad (2.2.12)$$

But from Eq. (2.2.11) $\bar{Q} = \tilde{Q} + \phi$ so

$$\tilde{\Gamma}[\tilde{Q}, \phi] = \Gamma[\tilde{Q} + \phi]. \quad (2.2.13)$$

This is our main result. As a special case of (2.2.13), take $\tilde{Q} = 0$ so that

$$\tilde{\Gamma}[0, \phi] = \Gamma[\phi]. \quad (2.2.14)$$

This equation means that the effective action can be determined by computing $\tilde{\Gamma}[0, \phi]$.

The background field effective action $\tilde{\Gamma}[\tilde{Q}, \phi]$ is just a conventional effective action computed in the presence of the background field ϕ . It therefore consists of all 1PI graphs contributing to Green's functions. Recall that the 1PI Green's functions are generated by taking derivatives of the effective action. In this case, the derivatives of $\tilde{\Gamma}[\tilde{Q}, \phi]$ with respect to \tilde{Q} would generate 1PI Green's functions in the presence of the background field ϕ . Now $\tilde{\Gamma}[0, \phi]$ has no dependence on \tilde{Q} so it generates no graphs with external lines. Instead, $\tilde{\Gamma}[0, \phi]$ is the sum of all 1PI vacuum graphs in the presence of the ϕ field. This is one advantage of the background field approach: it allows you to calculate the effective action by summing only vacuum graphs (graphs with no external lines). Equation (2.2.14) then assures you of getting the correct effective action with this method.

There are two quite different approaches to the calculation of $\tilde{\Gamma}[0, \phi]$. The first is to treat the ϕ field exactly. One sums all 1PI vacuum graphs using the shifted Lagrangian $S[Q + \phi]$ to generate Feynman rules. One must use the exact Q field propagator in the presence of the background field ϕ . Then one computes the usual ϕ field diagrams with vertices that depend on ϕ . Of course, this approach is only possible for very simple background fields. For example, it is used to generate the effective potential for scalar field theories [15, 16] when ϕ is taken to be a constant, and has been used to calculate the effective action for covariantly constant gauge fields [17].

The second approach, which is the one used in this paper, is to treat the background field ϕ perturbatively. Thus, we consider 1PI vacuum graphs of the Q field with ϕ fields appearing as external lines. In this approach the background field is arbitrary and in fact does not have to be specified. Instead, we proceed as follows. We use the action $S[Q + \phi]$ to generate Feynman rules. We take from the part of S quadratic in Q the ordinary Q field propagator. Then, the non-quadratic parts of S generate interactions. Interactions among the Q fields are used inside diagrams while interactions involving Q fields and ϕ fields are used to generate external lines. In this way, any of the 1PI Green's functions can be

computed. For example, imagine that $S[Q+\phi]$ contained a term aQ^3 and another term $bQ^2\phi$. Suppose you wanted to calculate the 1PI three-point function for this theory. One graph contributing to this Green's function would be that of Fig. 8. In this graph all the internal lines would be ordinary Q field propagators obtained from the part of S quadratic in Q . The two three-point vertices joining three internal lines would be given by the interaction term aQ^3 . However, the three-point vertices in this diagram which connect two internal

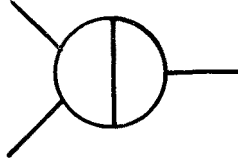


Fig. 8. A sample graph for background field calculation

lines to one external line would be given by the interaction $bQ^2\phi$. Thus in the background field approach you end up calculating diagrams identical to those you would have calculated in a conventional approach. The only difference is that in the background field approach the Feynman rules for vertices inside diagrams may differ from those for vertices connected to external lines. These Feynman rules for gauge theories are given in Section 3.1.

2.3. Gauge theories and the background field gauge

The discussion of the last Section was concerned with non-gauge theories. For gauge theories there is the important distinction that one must choose a gauge. In place of Eq. (2.2.1) defining the generating functional, one has for a gauge theory

$$Z[J] = \int \delta Q \det \left[\frac{\delta G^a}{\delta \omega^b} \right] \exp i \left[S[Q] - \frac{1}{2\alpha} G \cdot G + J \cdot Q \right], \quad (2.3.1)$$

where Q is now a gauge field Q_μ^a . In this equation S is the gauge field action

$$S = -\frac{1}{4} \int d^4x (F_{\mu\nu}^a)^2 \quad (2.3.2)$$

with

$$F_{\mu\nu}^a = \partial_\mu Q_\nu^a - \partial_\nu Q_\mu^a + g f^{abc} Q_\mu^b Q_\nu^c \quad (2.3.3)$$

for a group with structure constants f^{abc} . In addition,

$$J \cdot Q \equiv \int d^4x J_\mu^a Q_\mu^a \quad (2.3.4)$$

and

$$G \cdot G \equiv \int d^4x G^a G^a. \quad (2.3.5)$$

G^a is the gauge-fixing term: for example, $G^a = \partial_\mu Q_\mu^a$ is a common choice. Finally, $\delta G^a / \delta \omega^b$ is the derivative of G^a with respect to a gauge transformation

$$\delta Q_\mu^a = -f^{abc} \omega^b Q_\mu^c + \frac{1}{g} \partial_\mu \omega^a. \quad (2.3.6)$$

The background field generating functional, (2.2.2), is, for a gauge theory,

$$\tilde{Z}[J, A] = \int \delta Q \det \left[\frac{\delta \tilde{G}^a}{\delta \omega^b} \right] \exp i \left[S[Q+A] - \frac{1}{2\alpha} \tilde{G} \cdot \tilde{G} + J \cdot Q \right], \quad (2.3.7)$$

where $\delta \tilde{G}^a / \delta \omega^b$ is the derivative of the gauge-fixing term under the infinitesimal gauge transformation

$$\delta Q_\mu^a = -f^{abc} \omega^b (Q_\mu^c + A_\mu^c) + \frac{1}{g} \partial_\mu \omega^a. \quad (2.3.8)$$

We can now define the various quantities of Section 2.2 and relate them. Recall that the relation between background field quantities and conventional quantities in Section 2.2 was derived by shifting the integration variable in the functional integral. With this in mind, it is straightforward to derive in analogy with Eq. (2.2.14)

$$\tilde{\Gamma}[0, A] = \Gamma[A] \quad (2.3.9)$$

with the following provision. If $\tilde{\Gamma}[0, A]$ on the left-hand side of Eq. (2.3.9) is calculated using a gauge-fixing term $\tilde{G}^a = \tilde{G}^a(Q, A)$, then it will be equal to the conventional effective action $\Gamma[\bar{Q}]$ calculated with the gauge-fixing term $G^a = \tilde{G}^a(Q-A, A)$ and evaluated at $\bar{Q} = A$. The gauge $G^a = \tilde{G}^a(Q-A, A)$ may be an unusual gauge for the conventional approach. Nevertheless, Eq. (2.3.9) and the gauge independence of physical quantities assures us that the background field method, although it will give different Green's functions than the conventional approach in a conventional gauge, will give us the same S matrix. Thus it is a valid method for performing gauge theory calculations.

As was stated in the introduction, the great advantage of the background field method is that it retains explicit gauge invariance. What this means is that there exists a choice of gauge fixing term \tilde{G}^a for which the background field effective action $\tilde{\Gamma}[0; A]$ is a gauge invariant functional of A . As a result its form is severely restricted. This gauge choice is

$$\tilde{G}^a = \partial_\mu Q_\mu^a + g f^{abc} A_\mu^b Q_\mu^c. \quad (2.3.10)$$

It is easy to see that with this choice of gauge, the background field generating functional (2.3.7) is invariant under the transformations

$$\delta A_\mu^a = -f^{abc} \omega^b A_\mu^c + \frac{1}{g} \partial_\mu \omega^a, \quad (2.3.11)$$

$$\delta J_\mu^a = -f^{abc} \omega^b J_\mu^c. \quad (2.3.12)$$

To show this, make the change of integration variables in Eq. (2.3.7)

$$Q_\mu^a \rightarrow Q_\mu^a - f^{abc} \omega^b Q_\mu^c. \quad (2.3.13)$$

Equations (2.3.12) and (2.2.13) represent an adjoint group rotation so the term $J \cdot Q$ in (2.3.7) is clearly invariant. Adding (2.3.11) and (2.3.13) we find

$$\delta(Q_\mu^a + A_\mu^a) = -f^{abc} \omega^b (Q_\mu^c + A_\mu^c) + \frac{1}{g} \partial_\mu \omega^a. \quad (2.3.14)$$

This is just a gauge transformation on the field variable ($Q_\mu^a + A_\mu^a$) so the action $S[Q+A]$ is also invariant. Finally, (2.3.11) is a gauge transformation on A while (2.3.13) is an adjoint rotation of Q . The gauge fixing term \tilde{G}^a of Eq. (2.3.10) is just the background field covariant derivative of Q . By definition, this is invariant under such transformations. Likewise, it can easily be shown that the determinant factor in (2.3.7) is invariant. Thus \tilde{Z} is invariant under (2.3.11) and (2.3.12).

Since \tilde{Q} is just the conjugate variable to J , it follows immediately from the invariance of \tilde{Z} that $\tilde{\Gamma}[\tilde{Q}, A]$ is invariant under

$$\delta A_\mu^a = -f^{abc}\omega^b A_\mu^c + \frac{1}{g}\partial_\mu\omega^a, \quad (2.3.15)$$

$$\delta\tilde{Q}_\mu^a = -f^{abc}\omega^b\tilde{Q}_\mu^c. \quad (2.3.16)$$

Note that Eq. (2.3.16) is a homogeneous transformation so $\tilde{\Gamma}[0, A]$ is invariant under the transformation (2.3.15) alone. Since this is just an ordinary gauge transformation of A , we see that $\tilde{\Gamma}[0, A]$ is a gauge invariant functional of A . As a result, 1PI Green's functions generated by differentiating $\tilde{\Gamma}[0, A]$ with respect to A will obey the naive Ward identities of gauge invariance as stated in the introduction. Thus our goal has been accomplished. What remains is to see how the effective action $\tilde{\Gamma}[0, A]$ is computed.

3. Background field calculations

3.1. Feynman rules

The background field effective action $\tilde{\Gamma}[0, A]$ generates 1PI Green's functions. These are calculated using Feynman rules derived from the shifted action $S[Q+A]$, the gauge fixing term (2.3.10) and the determinant factor in Eq. (2.3.7) which is written in terms of an anti-commuting scalar ghost field in the usual way. Vertices involving Q fields are used inside diagrams while vertices involving A fields are used for external lines. All propagators appearing in these 1PI graphs are Q field propagators. This is essential since the A field propagator is undefined because A field gauge invariance has not been broken. Since only 1PI graphs are being considered, vertices involving only one Q field will never contribute and so can be ignored.

The complete Feynman rules are given in Fig. 9. The lines terminating in the letter A denote external lines. The procedure for calculating an S matrix element from these Feynman rules is the following. Write down all 1PI graphs for a given Green's function. Apply the Feynman rules of Fig. 9 using the vertices with lines terminating in A 's for external lines and the plain lines for internal lines. Once the 1PI Green's function has been computed the hard work is over, renormalization has been carried out and the simple Ward identities can be checked. At this point, you must choose a gauge for the background field A . It does not have to be related to the original background field gauge choice. The advantage of this method is that the fixing of a gauge for the A field could be postponed until this point. Now, the A field propagator is defined and using it to connect 1PI parts, trees are built up to generate the connected Green's functions. Finally, the S matrix is determined by LSZ reduction.

Fig. 9

$$\begin{aligned}
 & \text{Diagram 1: } a, \mu \text{ (wavy), } b, \nu \text{ (dashed), } k \text{ (momentum)} \rightarrow \frac{-i\delta_{ab}}{k^2 + i\epsilon} \left[g_{\mu\nu} - \frac{k_\mu k_\nu}{k^2} (1 - \alpha) \right] \\
 & \text{Diagram 2: } a \text{ (dashed), } b \text{ (dashed), } k \text{ (momentum)} \rightarrow \frac{i\delta_{ab}}{k^2 + i\epsilon} \\
 & \text{Diagram 3: } a, \mu \text{ (wavy), } b, \nu \text{ (dashed), } c, \lambda \text{ (dashed), } p, q, r \text{ (momenta)} \rightarrow g^{abc} \left[g_{\mu\lambda} \left(p - r - \frac{1}{\alpha} q \right)_\nu \right. \\
 & \quad \left. + g_{\nu\lambda} (r - q)_\mu + g_{\mu\nu} \left(q - p + \frac{1}{\alpha} r \right)_\lambda \right] \\
 & \text{Diagram 4: } a, \mu \text{ (wavy), } b, \nu \text{ (dashed), } c, \lambda \text{ (dashed), } p, q, r \text{ (momenta)} \rightarrow g^{abc} \left[g_{\mu\lambda} (p - r)_\nu + g_{\nu\lambda} (r - q)_\mu \right. \\
 & \quad \left. + g_{\mu\nu} (q - p)_\lambda \right] \\
 & \text{Diagram 5: } a, \mu \text{ (wavy), } b, \nu \text{ (dashed), } c, \lambda \text{ (dashed), } d, \rho \text{ (dashed), } p, q, r \text{ (momenta)} \rightarrow \left. \begin{aligned} & -ig^2 [f_{abx} f_{xcd} (g_{\mu\lambda} g_{\nu\rho} - g_{\mu\rho} g_{\nu\lambda} - g_{\mu\nu} g_{\lambda\rho}) \\ & + f_{adx} f_{xbc} (g_{\mu\nu} g_{\lambda\rho} - g_{\mu\lambda} g_{\nu\rho} - \frac{1}{\alpha} g_{\mu\nu} g_{\nu\lambda}) \\ & + f_{acx} f_{xbd} (g_{\mu\nu} g_{\lambda\rho} - g_{\mu\rho} g_{\nu\lambda})] \end{aligned} \right\} \\
 & \text{Diagram 6: } a, \mu \text{ (wavy), } b, \nu \text{ (dashed), } c, \lambda \text{ (dashed), } d, \rho \text{ (dashed), } p, q, r \text{ (momenta)} \rightarrow \left. \begin{aligned} & -ig^2 [f_{abx} f_{xcd} (g_{\mu\lambda} g_{\nu\rho} - g_{\mu\rho} g_{\nu\lambda}) \\ & + f_{adx} f_{xbc} (g_{\mu\nu} g_{\lambda\rho} - g_{\mu\lambda} g_{\nu\rho}) \\ & + f_{acx} f_{xbd} (g_{\mu\nu} g_{\lambda\rho} - g_{\mu\rho} g_{\nu\lambda})] \end{aligned} \right\}
 \end{aligned}$$

Fig. 9 - continued

$$\begin{aligned}
 & \text{Diagram 7: } a, \mu \text{ (wavy), } b, \nu \text{ (dashed), } c, \lambda \text{ (dashed), } d, \rho \text{ (dashed), } p, q, r \text{ (momenta)} \rightarrow \left. \begin{aligned} & -ig^2 [f_{abx} f_{xcd} (g_{\mu\lambda} g_{\nu\rho} - g_{\mu\rho} g_{\nu\lambda} + \frac{1}{\alpha} g_{\mu\nu} g_{\lambda\rho}) \\ & + f_{adx} f_{xbc} (g_{\mu\nu} g_{\lambda\rho} - g_{\mu\lambda} g_{\nu\rho} - \frac{1}{\alpha} g_{\mu\rho} g_{\nu\lambda}) \\ & + f_{acx} f_{xbd} (g_{\mu\nu} g_{\lambda\rho} - g_{\mu\rho} g_{\nu\lambda})] \end{aligned} \right\} \\
 & \text{Diagram 8: } a, \mu \text{ (wavy), } b, \nu \text{ (dashed), } c, \mu \text{ (wavy), } p, q \text{ (momenta)} \rightarrow \zeta_\mu - g f_{acb} p_\mu \\
 & \text{Diagram 9: } a, \mu \text{ (wavy), } b, \nu \text{ (dashed), } c, \mu \text{ (wavy), } d, \nu \text{ (dashed), } p, q \text{ (momenta)} \rightarrow \left. \begin{aligned} & -ig^2 f_{acx} f_{xdb} g_{\mu\nu} \\ & -ig^2 g_{\mu\nu} (f_{acx} f_{xdb} \\ & + f_{adx} f_{xcb}) \end{aligned} \right\}
 \end{aligned}$$

Fig. 9. Feynman rules for background field calculations in Yang-Mills theory. Wavy lines are Q field propagators, wavy lines terminating in an A represent external gauge particles and dashed lines represent ghost propagators

3.2. Renormalization

During the calculation of the effective action, $\tilde{\Gamma}[0, A]$, divergences will be encountered which must be renormalized. This is done by relating bare and renormalized quantities by

$$\begin{aligned}(A_\mu)_0 &= Z_A^{1/2} A_\mu, \\ g_0 &= Z_g g, \\ \alpha_0 &= Z_\alpha \alpha.\end{aligned}\tag{3.2.1}$$

These field, coupling constant and gauge fixing parameter renormalizations are sufficient to renormalize the theory. There are two more fields, the Q field and the ghost field, which do not have to be renormalized. This is because they only appear inside loops. If these fields were normalized we would find that the renormalization factors appearing at the vertices would cancel those of the propagator, so renormalization of these fields is irrelevant.

Because explicit gauge invariance is retained in the background field method, the renormalization factors Z_g and Z_A are, in fact, related. The infinities appearing in the gauge-invariant effective action $\tilde{\Gamma}[0, A]$ must take the gauge invariant form of a divergent constant times $(F_{\mu\nu}^a)^2$. Now, according to (3.2.1), $F_{\mu\nu}^a$ is renormalized by

$$(F_{\mu\nu}^a)_0 = Z_A^{1/2} [\partial_\mu A_\nu^a - \partial_\nu A_\mu^a + Z_g Z_A^{1/2} g f^{abc} A_\mu^b A_\nu^c].\tag{3.2.2}$$

This will only take on the gauge covariant form of a constant times $F_{\mu\nu}^a$ if

$$Z_g = Z_A^{-1/2}.\tag{3.2.3}$$

This is the relation between charge and background field renormalizations in the background field method.

The easiest way to renormalize a Yang-Mills theory is to use dimensional regularization [18] and minimal subtraction [19]. In this scheme we perform all loop momentum integrals in $4-2\epsilon$ dimensions and write the renormalization constants as sums over poles in ϵ such as

$$Z_A = 1 + \sum_{n=1}^{\infty} \frac{Z_A^{(n)}}{\epsilon^n}.\tag{3.2.4}$$

In general, if we calculate to r loops we will get contributions in Eq. (3.2.4) to $1/\epsilon, 1/\epsilon^2, \dots, 1/\epsilon^r$.

In the dimensional regularization scheme, the bare coupling constant is not dimensionless. This is because the action in $4-2\epsilon$ dimensions involves terms like

$$\int d^{4-2\epsilon} x (\partial_\mu A_\nu)^2\tag{3.2.5}$$

and

$$\int d^{4-2\epsilon} x g_0^2 (A_\mu A_\nu)^4.\tag{3.2.6}$$

From (3.2.5) we find that if the action is to be dimensionless, A must have dimensions of $(\text{mass})^{1-\epsilon}$. Then Eq. (3.2.6) tells us that g_0 has dimensions $(\text{mass})^\epsilon$. If we still wish to use

a dimensionless renormalized coupling constant, then we must introduce an arbitrary mass parameter μ and replace the coupling constant renormalization in Eq. (3.2.1) by

$$g_0 = Z_g \mu^\epsilon g. \quad (3.2.7)$$

The introduction of this μ parameter is the reason why Yang-Mills theory is not characterized by a dimensionless coupling constant, but rather by a mass parameter which is a particular combination of μ and g usually called Λ . Since μ is arbitrary, we must require that g_0 be independent of μ which gives [20]

$$\mu \frac{\partial g_0}{\partial \mu} = 0 = Z_g \mu^\epsilon \left[\epsilon g + g \mu \frac{\partial \ln Z_g}{\partial \mu} + \mu \frac{\partial g}{\partial \mu} \right]. \quad (3.2.8)$$

Now, by definition,

$$\beta = \mu \frac{\partial g}{\partial \mu} \quad (3.2.9)$$

so we find that

$$\beta = -\epsilon g - g \mu \frac{\partial \ln Z_g}{\partial \mu}. \quad (3.2.10)$$

Finally, using the chain rule

$$\mu \frac{\partial}{\partial \mu} = \mu \frac{\partial g}{\partial \mu} \frac{\partial}{\partial g} = \beta \frac{\partial}{\partial g} \quad (3.2.11)$$

(3.2.10) can be written in the simpler form

$$\beta = -\epsilon g - g \beta \frac{\partial \ln Z_g}{\partial g}. \quad (3.2.12)$$

From Eq. (3.2.3) this can also be written, in the background field gauge, as

$$\beta = -\epsilon g + \frac{1}{2} g \beta \frac{\partial \ln Z_A}{\partial g}. \quad (3.2.13)$$

Now Z_A , from Eq. (3.2.4), is a sum of powers of $1/\epsilon$. If the β function is to be finite, as it must be, these various powers must conspire to cancel in (3.2.13). From this one can derive important relations [20, 6] between the coefficients $Z_A^{(n)}$. The only finite term in β can come when the $-\epsilon g$ term of β hits the $1/\epsilon$ term in Z_A on the right-hand side of (3.2.13). Thus, taking the limit $\epsilon \rightarrow 0$ we find that

$$\beta = -\frac{1}{2} g^2 \frac{\partial}{\partial g} Z_A^{(1)}. \quad (3.2.14)$$

Thus, the β function can be determined from a knowledge of the coefficient of the $1/\epsilon$ term in the background field renormalization factor. This is calculated in the next section.

3.3. Calculation of the Yang-Mills β function

A good example of the tremendous simplifications introduced by the background field method is provided by a computation of the Yang-Mills β function. Because the divergences in this method are forced to take a gauge invariant form the β function can be determined from the background field renormalization factor Z_A through Eq. (3.2.14). Z_A can be obtained just by calculating loop corrections to the gauge field propagator [6].

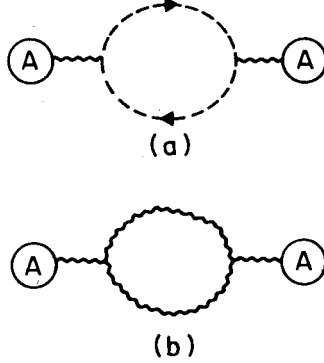


Fig. 10. Graphs for a one-loop calculation of the Yang-Mills β function

No vertex corrections need to be considered. This is in contrast to the conventional approach [21, 22] where the gauge propagator, ghost propagator and gauge-ghost-ghost vertex all must be computed.

For example, the one-loop Yang-Mills β function can be determined just by calculating the two graphs of Fig. 10. The first graph, Fig. 10a, gives for the $1/\epsilon$ pole (remembering a factor -1 for the ghost loop)

$$\frac{ig^2 C_A \delta^{ab}}{(4\pi)^2} \left(\frac{1}{3\epsilon} \right) [g_{\mu\nu} k^2 - k_\mu k_\nu] \quad (3.3.1)$$

while Fig. 10b gives

$$\frac{ig^2 C_A \delta^{ab}}{(4\pi)^2} \left(\frac{10}{3\epsilon} \right) [g_{\mu\nu} k^2 - k_\mu k_\nu], \quad (3.3.2)$$

where C_A is defined by

$$f^{abc} f^{bcd} = C_A \delta^{ad} \quad (3.3.3)$$

and is N for $SU(N)$. Adding (3.3.1) and (3.3.2) together, we find that the divergences are cancelled by defining

$$Z_A = 1 + \frac{11C_A}{3\epsilon} \frac{g^2}{(4\pi)^2}. \quad (3.3.4)$$

When substituted into Eq. (3.2.14) this gives the well-known result [21]

$$\beta = -\frac{11C_A}{3} \frac{g^3}{(4\pi)^2}. \quad (3.3.5)$$

Fig. 11 - continued

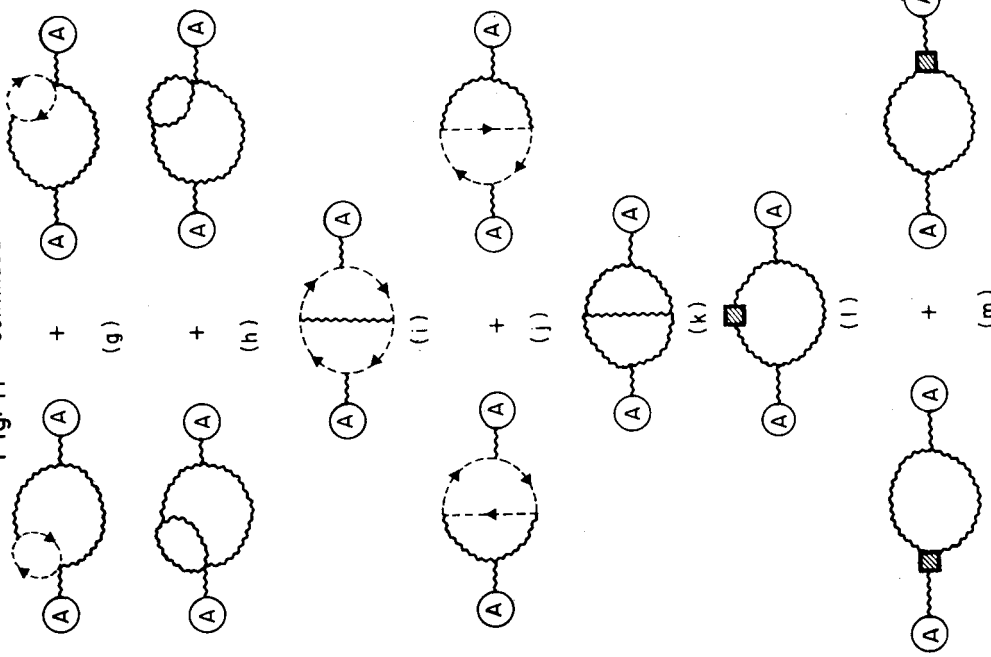


Fig. 11

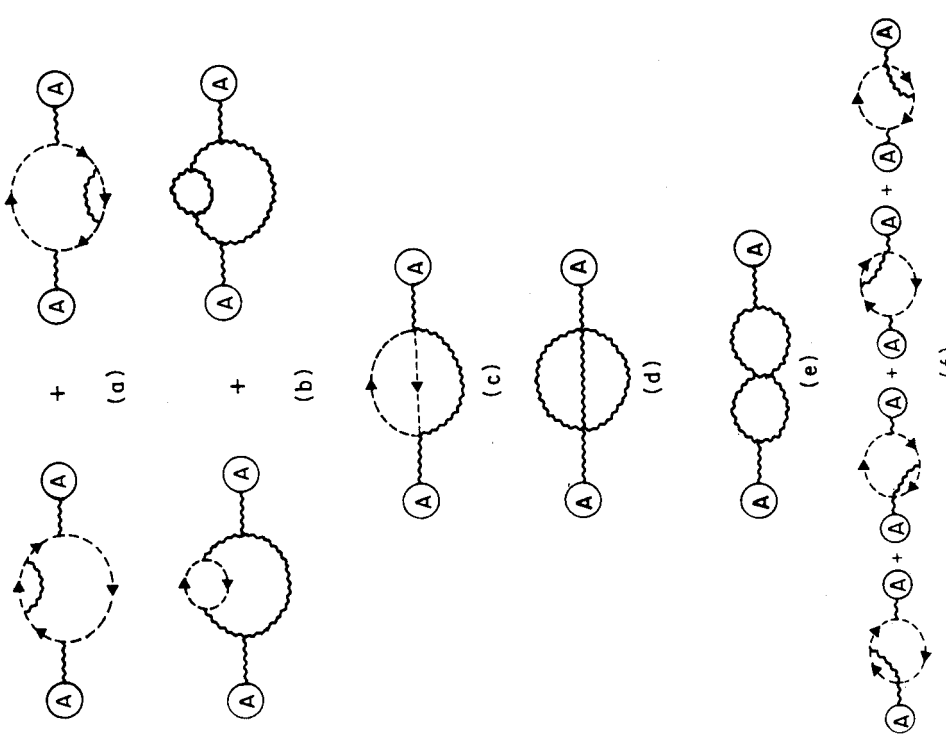


Fig. 11. Graphs for a two-loop calculation of the Yang-Mills β function. Boxes represent gauge-fixing term insertions resulting from renormalization of the gauge-fixing parameter. Details of this calculation are given in Ref. [6]

I have also calculated the two-loop contribution to the Yang-Mills β function [6] from the graphs of Fig. 11. Here, the background field method results in an enormous reduction in labour over conventional methods [22].

4. Conclusions

The background field formalism now exists for calculations in gauge theories to arbitrary numbers of loops. The method can be used for general discussions about the structure and particularly about the divergences of gauge theories, and for specific calculations. In both of these applications, the retention of explicit gauge invariance made possible by the background field gauge should be extremely useful.

I am very grateful to the organizers of the Cracow School for their hospitality.

REFERENCES

- [1] B. S. DeWitt, *Phys. Rev.* **162**, 1195 (1967); also in *Dynamic Theory of Groups and Fields*, Gordon and Breach 1965.
- [2] J. Honerkamp, *Nucl. Phys.* **B36**, 130 (1971); *Nucl. Phys.* **B48**, 269 (1972); R. Kallosh, *Nucl. Phys.* **B78**, 293 (1974); I. Ya. Arefieva, L. D. Faddeev, A. A. Slavnov, *Theor. Mat. Fiz.* **21**, 311 (1974); S. Sarkar, *Nucl. Phys.* **B82**, 447 (1974); S. Sarkar, H. Strubbe, *Nucl. Phys.* **B90**, 45 (1975); H. Kluberg, J. B. Zuber, *Phys. Rev.* **D12**, 482 (1975); G. 't Hooft, *Nucl. Phys.* **B62**, 444 (1973); M. Grisaru, P. van Nieuwenhuizen, C. C. Wu, *Phys. Rev.* **D12**, 3203 (1975).
- [3] G. 't Hooft, Acta Universitatis Wratislavenensis No 38 XII, Winter School of Theoretical Physics in Karpacz; *Functional and Probabilistic Methods in Quantum Field Theory* Vol. 1 (1975).
- [4] B. S. DeWitt, *A Gauge Invariant Effective Action*, Santa Barbara preprint NSF-ITP-80-31 (1980) to appear in *Quantum Gravity II*, ed. C. Isham, R. Penrose and S. Sciama.
- [5] D. Boulware, *Gauge Dependence of the Effective Action*, University of Washington preprint RLO-1388-822 (1980).
- [6] L. F. Abbott, *Nucl. Phys.* **B185**, 189 (1981).
- [7] G. 't Hooft, *Nucl. Phys.* **B62**, 444 (1973); G. 't Hooft, M. Veltman, *Ann. Inst. Henri Poincaré* **20**, 69 (1974).
- [8] See for example: S. Deser, J. Kay, K. Stelle, *Phys. Rev. Lett.* **38**, 527 (1977).
- [9] S. Weinberg, *Phys. Lett.* **91B**, 51 (1980).
- [10] R. Dashen, D. Gross, *Phys. Rev.* **D23**, 2340 (1981).
- [11] For a review see: B. Lee in *Methods in Field Theory*, ed. C. R. Balian, J. Zinn-Justin, North Holland 1976.
- [12] C. Becchi, A. Rouet, R. Stora, *Commun. Math. Phys.* **42**, 127 (1975).
- [13] See for example: J. Bjorken, S. Drell, *Relativistic Quantum Fields*, McGraw-Hill, New York 1965.
- [14] For a similar discussion see: B. Zumino in *Lectures on Elementary Particles and Quantum Field Theory*, ed. S. Deser, M. Grisaru, H. Pendleton, MIT Press, Cambridge 1970.
- [15] See for example: R. Jackiw, *Phys. Rev.* **D9**, 1686 (1974).
- [16] S. Coleman, E. Weinberg, *Phys. Rev.* **D7**, 1888 (1973).
- [17] I. A. Batalin, G. K. Savvidy, Yerevan Physics Institute preprint ERFI 78-299 (1978).
- [18] G. 't Hooft, M. Veltman, *Nucl. Phys.* **B44**, 189 (1972).
- [19] G. 't Hooft, *Nucl. Phys.* **B61**, 455 (1973).
- [20] D. Gross in *Methods in Field Theory*, ed. C. R. Balian, J. Zinn-Justin, North Holland 1976.
- [21] H. D. Politzer, *Phys. Rev. Lett.* **30**, 1346 (1973); D. Gross, F. Wilczek, *Phys. Rev. Lett.* **30**, 1343 (1973).
- [22] W. E. Caswell, *Phys. Rev. Lett.* **33**, 244 (1974); D. R. T. Jones, *Nucl. Phys.* **B75**, 531 (1974).