THE REGULARISATION AND

RENORMALISATION OF GAUGE THEORIES

Thesis by
Brian J. Warr

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California.

1987
(Submitted May 21, 1987)

To Bette, James and Mark

ACKNOWLEDGEMENTS

The person who has most influenced how I think is my advisor, and I hope friend, John Preskill. Whatever I do in the future, I shall be grateful for the privilege of his inspiration.

ABSTRACT

The method of effective Lagrangian flow provides the most physically illuminating discussion of renormalisation theory. At distance scales much larger than some physical cutoff, the physics is described by a small number of parameters, which can be identified purely by dimensional analysis. For scalar theories a rigorous yet simple proof of renormalisability, based on this concept, was given by Polchinski, and this work forms the bedrock of this thesis.

For gauge theories there is the extra issue of the unitarity of the renormalised S-matrix, which can only be guaranteed by proving renormalised Ward identities, and this is what we carry out for all cases of interest in $d = 4$. In particular we cover the case of $N = 1$ super Yang-Mills.

We prove that the cancellation of anomalies at the one-loop level is a sufficient as well as necessary condition for a theory to be perturbatively quantisable, and hence that there are no higher-loop anomalies.

Table of Contents

INTRODUCTION

In this thesis I shall discuss, using the method of effective Lagrangians, the regularisation and renormalisation of Yang-Mills theories. The main advantage of the technique is that the different cases of symmetric, broken, chiral and supersymmetric gauge theories can all be analysed in the same way. This is the first such unified approach.

The principle result will be that perturbatively there are no more anomalies other than the well-documented one-loop anomaly, and so cancellation of the one-loop anomaly becomes a *sufficient* as well as *necessary* condition for Yang-Mills to be acceptable as a quantum field theory.

The Physical Motivation for the Study

A spontaneously broken gauge theory is at present the only known consistent theory of massive vector bosons. The consistency was first proven by 't Hooft and Veltman [1], using their brilliant invention of dimensional regularisation, and this seminal work led to the explosion of interest in gauge theories in the 1970's. This activity culminated in the highly successful 'Standard Model' of strong, electromagnetic and weak interactions, based on the gauge group $SU(3)_c \times SU(2) \times U(1)$, and this model has now been tested up to energy scales approaching 1 TeV. (For a review, see Ref. [2].)

However, the Standard Model is not a *fundamental* theory, since it does not include gravity, and it is current theoretical prejudice that it does not even survive unscathed for more than about a single order of magnitude in energy above the present experiments. One of the strongest arguments that we have for the emergence, in up-coming experiments, of new physics is the so-called 'hierarchy problem.' Here it is noted that the presence of scalars in the theory is *unnatural.* What is meant by this is that generically in a field theory scalar particles tend to have a mass of order of the physical cutoff (say the Planck mass, $m_{\rm Pl}$) and so one must 'fine tune' the parameters very carefully in order to make them light.

One particularly attractive solution to this problem involves an $N = 1$ supersymmetric, chiral Yang-Mills model. The idea is that the chiral fermions can be kept rigorously massless by the gauge symmetry, and so their scalar super-partners would be kept massless by the supersymmetry. Thus the natural mass scale for the scalars will be determined in this scenario by the mass scale of supersymmetry breaking, which can plausibly be much lower than $m_{\rm Pl}$. (For a fascinating speculation on this subject, see Ref. [3].)

Now the above idea presupposes that chiral and/or supersymmetric Yang-Mills theories are renormalisable and unitary. This was *not* proven by 't Hooft and is in fact a very difficult question to discuss within dimensional regularisation, or the other main scheme, namely BPHZ subtraction. It is clearly very important to elucidate the conditions under which these theories are so consistent, not just for the exotic case above, but as a basic input to all future phenomenology.

Difficulties with Conventional Renormalisation Schemes

The main goal of *any* scheme is to prove the existence of renormalised Ward identities, so that we are guaranteed to have a renormalised and unitary S-matrix.

In dimensional regularisation the *regularised* Ward identities automatically hold exactly (in the non-chiral case), and from there 't Hooft and Veltman manage to construct the *renormalised* Ward identities by induction. This exactness of the regularised identities is crucial to their argument, and comes about because the naïve manipulations on the functional integral, which formally lead to the identities, can be *translated* into the manipulations of shifting integrals and using Lorentz vector algebra in the Feynman diagram expansion. This *diagrammatic* analysis can be rigorously justified using the axioms of dimensional regularisation, as explained in the excellent review by Collins [4].

However for chiral gauge theories, i.e., theories where the fermions are strictly massless due to the gauge symmetry, and for supersymmetric theories, the Lorentz algebra fails and the regularised identities are no longer exact. Of course one can compute the one-loop obstruction, and one can indeed demand that the fermions are in a representation such that this anomaly vanishes, but this is to no avail: the regularised identities are still not exact, and so the renormalisation proof cannot proceed[f1].

In BPHZ subtraction scheme there is no problem with the Lorentz algebra, since we stick to exactly four dimensions, but now we must face highly non-trivial combinatorics. Even for the simplest case, namely a completely broken, non-chiral theory where all the particles are massive, the proof that overlapping divergences really do cancel is somewhat complicated. There is no readily apparent reason *within* the BPHZ language as to why the cancellations have to occur, and hence one must simply rely on the towering strength of Zimmermann and Lowenstein [7].

The situation is in fact even worse for chiral gauge theories because of the *masslessness* of some of the particles, (and of course this is precisely the case one must address for the anomaly question). The presence of infrared divergences prevents one from Taylor expanding the Feynman integrands in the usual way about the origin of momenta, so one must choose some non-zero euclidean momentum instead (or put in a fictitious mass). The technique is not then manifestly gauge invariant, and the recovery of the Ward identities becomes very difficult to exhibit [8].

Moreover, for the case of supersymmetry we have the problem that in most gauges the BPHZ scheme cannot even be defined! This is because there are dimensionless fields in the theory whose propagators vary as the inverse fourth power of momentum. Now the intertwining of the infrared and ultraviolet divergences becomes inextricable, since the Taylor expansion of integrands does not exist about *any* point.

---

f_1 For chiral gauge theories Costa et al. [5] make a bold attempt to prove the 'Sufficiency Theorem' for one-loop anomalies using dimensional regularisation. However their manipulations on the functional integral seem hard to justify, and moreover they end up with a strange list of conditions for the theorem to hold. In the supersymmetric case, dimensional regularisation is discussed in Ref. [6].

Appreciating these difficulties, then, we are motivated to pursue a framework of regularisation and renormalisation that is more intuitive and physical.

Outline of the Thesis

In the first chapter, I introduce Wilson's brilliantly insightful concept of the 'flowing effective Lagrangian' [9], and extensively review the renormalisation theorems derived in this picture by Polchinski [10].

In the second chapter, I discuss the problem of unitarity for field theories with vector particles, for which the only known solution utilises the 'gauge principle,' and thus set up the task of proving the renormalised Ward identities. The technique for this involves a new regulator, based on higher covariant derivatives, where it is the mass scale of the high dimension operators that provides the running cutoff. I note here that previous attempts [11] to regulate along these lines have been incorrect, and I explain why. Using the new regulator, and the results of Chapter I, the renormalised Ward identities are derived very simply, and I conclude with a proof that these identities do indeed entail unitarity. For clarity, I consider in this chapter only pure Yang-Mills and Yang-Mills coupled to non-chiral matter.

In the third chapter, I investigate the issue of perturbative anomalies. The new regulator in the second chapter turns out to be too difficult to generalise to chiral and supersymmetric Yang-Mills, so I first describe a modified form that can be used for both these cases. Here its use in the effective Lagrangian is a little more involved, but is totally straightforward.

CHAPTER I

THE METHOD OF EFFECTIVE LAGRANGIANS

The outstanding concept in field theory is that of effective Lagrangian flows, pioneered by Wilson [9], and it is upon this that I focus my attention. One abandons the idea that a theory should be valid all the way down to zero distance; rather one supposes there is a high energy cutoff Λ_0 above which the physics can be something totally different. (These days there is even a popular candidate as to what that 'something' is.) Below the cutoff the theory is determined by a very general Lagrangian, where the infinitely many coupling constants, by the criterion of naturalness, are simply of order of Λ_0 raised to the appropriate power.

One now considers the physics at an energy E far below Λ_0. We could compute the Green's functions directly, using the Lagrangian at Λ_0, but this involves performing large loop integrals. These integrals actually bring *all* the coupling constants into play, and so *prima facie* the theory has, in principle, no predictive capability.

However, we have not yet used the stipulation that $E \gg \Lambda_0$. To illuminate the relevance of this, imagine smoothly lowering the cutoff to $\Lambda_R \gtrsim E$. Of course to keep the physics, that is, the path integral, fixed we must make compensating changes in the strength of the couplings. This defines the 'running' or 'effective' Lagrangian. Within the (very large) space of functionals we find that the effective Lagrangian flows in the infrared towards a fixed surface of very low dimension $\mathbf{D}$, e.g., three in the case of Z_2-symmetric $\lambda\phi^4$ theory, so that the low energy Green's functions are in fact defined completely by only $\mathbf{D}$ parameters. (Actually for Λ_R/Λ_0 non-zero the surface is not quite sharp, rather it has a 'thickness,' and the physical computations have an accuracy limited by $O(\Lambda_R/\Lambda_0)^2$.)

What has happened is that of all the original couplings at Λ_0, only $\mathbf{D}$ independent combinations of them have entered into physical quantities.

Now for any point on the fixed surface at Λ_R there is a flow towards it from a wide variety of initial Lagrangians at Λ_0, and these are all equivalent as far as the low energy Green's functions are concerned. In particular there is a flow from Λ_0 where the only non-zero couplings are those of the dimension ≤ 4 operators, the counting of which supplies the value of **D**. Since we can simply choose to use this special Lagrangian, known familiarly as the 'bare Lagrangian,' instead of the true effective Lagrangian at Λ_0 for the low energy computations, we see automatically that 'power-counting renormalisable' theories are indeed renormalisable, and that nature, at low energy, is *bound* to be describable by such theories. Wilson's ideas have thus taught us the physical meaning of renormalisation.

A rigorous proof of these properties of the effective Lagrangian flow for scalar field theories in perturbation was given in a beautiful paper by Polchinski [10]. His analysis is astonishingly simple, and completely illuminating. In this chapter we review Polchinski's work, and extend the results so that they may be applied to the later chapters on Yang-Mills and supersymmetry.

1. A Toy Model

Why don't we include a four-Fermi interaction, $\lambda(\bar{\psi}\psi)^2$ in QED, or at any rate insist that $\lambda^{-\frac{1}{2}} \geq 100$Gev? The reason is not aesthetics; it is rather that the effect of such a term can be completely absorbed into the renormalisable couplings when looking at low energy processes. This fact is not at all obvious when looking at the Feynman diagrams. For example, consider the overlapping divergence diagram of Fig. 1. How can one be sure that its presence does not induce the necessity of $\log p^2$ counterterms?

The ideal way to approach this problem is the study of the effective Lagrangian flow. Since the cutoff Λ_0 is kept finite, albeit very large compared to the physical scale Λ_R, all the couplings in the effective Lagrangian can be made arbitrarily small *at all scales* . This means that the anomalous dimension of any operator can also be made arbitrarily small, so that in perturbation theory it is the trivially computable canonical dimension which acts as the discriminant for 'relevant' and 'irrelevant' parts.

The essential behaviour of how the couplings of 'relevant' and 'irrelevant' operators run is captured in Polchinski's toy equations [10] (which he then elevates to the full field theory case).

He considers a system described by two couplings, g_4 and g_6, one dimensionless and the other of dimension -2, for which the flow equations have the generic form:

$$\Lambda\frac{\partial g_4}{\partial\Lambda} = \beta_4(g_4, \Lambda^2 g_6) \tag{1.1a}$$

$$\Lambda\frac{\partial g_6}{\partial\Lambda} = \Lambda^2\beta_6(g_4, \Lambda^2 g_6)\ . \tag{1.1b}$$

Defining the dimensionless variables $\lambda_4 = g_4,\ \lambda_6 = \Lambda^2 g_6$, we rewrite these equations as

$$\Lambda\frac{\partial\lambda_4}{\partial\Lambda} = \beta_4(\lambda_4, \lambda_6) \tag{1.2a}$$

$$\Lambda\frac{\partial\lambda_6}{\partial\Lambda} = 2\lambda_6 + \beta_6(\lambda_4, \lambda_6)\ . \tag{1.2b}$$

What we would like to do is to study the nature of solutions to (2) as we run the scale Λ down from a high value to a much smaller value. To do this it is instructive to consider two nearby flows, so that we can observe the changes at low energy due to given changes at high energy. Suppose we have a solution such that at some initial high scale Λ_0 the couplings are λ_4^0 and λ_6^0. Let me denote this by

$$\lambda_i(\Lambda) = \bar{\lambda}_i(\Lambda, \Lambda_0, \lambda_4^0, \lambda_6^0) \tag{1.3a}$$

where

$$\lambda_i(\Lambda_0) = \bar{\lambda}_i(\Lambda_0, \Lambda_0, \lambda_4^0, \lambda_6^0) = \lambda_i^0\ , \tag{1.3b}$$

and the index i takes on the values $4, 6$. A nearby solution may start at initial scale $\Lambda_0 + \delta\Lambda_0$, with couplings $\lambda_4^0 + \delta\lambda_4^0$ and $\lambda_6^0 + \delta\lambda_6^0$, and this is

$$\lambda_i(\Lambda) = \bar{\lambda}_i(\Lambda, \Lambda_0, \lambda_4^0, \lambda_6^0) + \epsilon_i(\Lambda) \tag{1.4a}$$

where

$$\epsilon_i(\Lambda) = \Lambda_0 \frac{\partial \bar{\lambda}_i}{\partial \Lambda_0} \frac{\delta \Lambda_0}{\Lambda_0} + \frac{\partial \bar{\lambda}_i}{\partial \lambda_4^0} \delta \lambda_4^0 + \frac{\partial \bar{\lambda}_i}{\partial \lambda_6^0} \delta \lambda_6^0 \,. \tag{1.4b}$$

The initial condition at $\Lambda = \Lambda_0 + \delta\Lambda_0$ then tells us that

$$0 \;=\; \Lambda \frac{\partial \bar{\lambda}_i}{\partial \Lambda}\bigg|_{\Lambda=\Lambda_0} \;+\; \Lambda_0 \frac{\partial \bar{\lambda}_i}{\partial \Lambda_0}\bigg|_{\Lambda=\Lambda_0} \tag{1.5a}$$

and

$$\delta_{ij} \;=\; \frac{\partial \bar{\lambda}_i}{\partial \lambda_j^0}\bigg|_{\Lambda=\Lambda_0} \,. \tag{1.5b}$$

Now the increments $\epsilon_i(\Lambda)$ satisfy the linearized form of equations (2), namely,

$$\Lambda \frac{\partial \epsilon_4}{\partial \Lambda} = \overline{\frac{\partial \beta_4}{\partial \lambda_4}} \epsilon_4 + \overline{\frac{\partial \beta_4}{\partial \lambda_6}} \epsilon_6 \tag{1.6a}$$

$$\Lambda \frac{\partial \epsilon_6}{\partial \Lambda} = \overline{\frac{\partial \beta_6}{\partial \lambda_4}} \epsilon_4 + (2 + \overline{\frac{\partial \beta_6}{\partial \lambda_6}}) \epsilon_6 \;, \tag{1.6b}$$

and the key point here is the presence of the '2' in the coefficient of ϵ_6. This '2' came of course just from the canonical dimension of the coupling g_6, and we see that if it is dominant then the vector $(\epsilon_4\,,\epsilon_6)$ will approach the *tangent vector* to the surface $\lambda = \bar{\lambda}_i$ for small Λ. The deviation from the tangent vanishes as the *square* of Λ/Λ_0.

Thus we expect that as we let the initial couplings vary in their two-dimensional space, the couplings at a much lower scale Λ_R will run near a one-dimensional subspace. In particular, given $\lambda_4(\Lambda_R)$ then $\lambda_6(\Lambda_R)$ will already be known, to an accuracy of $(\Lambda_R/\Lambda_0)^2$. This is illustrated in Fig. 2.

To prove this, first let the initial couplings vary along an arbitrary curve, that is, set $\lambda_6^0 = f(\lambda_4^0)$, and hence $\delta\lambda_6^0 = f'(\lambda_4^0)\delta\lambda_4^0$. Thus in (4b),

$$\epsilon_i(\Lambda) = \Lambda_0 \frac{\partial \bar{\lambda}_i}{\partial \Lambda_0} \frac{\delta \Lambda_0}{\Lambda_0} + \frac{D \bar{\lambda}_i}{D \lambda_4^0} \delta \lambda_4^0 \qquad \text{where} \quad \frac{D}{D\lambda_4^0} \equiv \frac{\partial}{\partial \lambda_4^0} + \frac{df}{d\lambda_4^0} \frac{\partial}{\partial \lambda_6^0} \,. \tag{1.7}$$

Now let me stipulate that at Λ_R the two flows have the same value for λ_4. This defines a relation between $\delta\Lambda_0$ and $\delta\lambda_4^0$, namely,

$$0 = \Lambda_0 \frac{\partial \bar{\lambda}_4}{\partial \Lambda_0}(\Lambda_\mathrm{R}, \Lambda_0, \lambda_4^0, f(\lambda_4^0)) \frac{\delta\Lambda_0}{\Lambda_0} + \frac{D\bar{\lambda}_4}{D\lambda_4^0}(\Lambda_\mathrm{R}, \Lambda_0, \lambda_4^0, f(\lambda_4^0))\delta\lambda_4^0 \ , \tag{1.8}$$

and so now the increment ϵ_6 is

$$\epsilon_6(\Lambda) = \left[\Lambda_0 \frac{\partial \bar{\lambda}_6(\Lambda)}{\partial \Lambda_0} - \frac{D\bar{\lambda}_6(\Lambda)}{D\lambda_4^0}\left(\frac{D\bar{\lambda}_4(\Lambda_\mathrm{R})}{D\lambda_4^0}\right)^{-1} \Lambda_0 \frac{\partial \bar{\lambda}_4(\Lambda_\mathrm{R})}{\partial \Lambda_0}\right] \frac{\delta\Lambda_0}{\Lambda_0} \ . \tag{1.9}$$

The equation of motion (6b) for $\epsilon_6(\Lambda)$ is coupled to that of $\epsilon_4(\Lambda)$ and this makes it tricky to solve. However, we can define a new quantity $V_6(\Lambda)\delta\Lambda_0/\Lambda_0$ which coincides with $\epsilon_6(\Lambda)$ at the interesting point $\Lambda = \Lambda_\mathrm{R}$, and its equation of motion is uncoupled. We have

$$V_6(\Lambda) = \Lambda_0 \frac{\partial \bar{\lambda}_6}{\partial \lambda_0} - \frac{D\bar{\lambda}_6}{D\lambda_4^0}\left(\frac{D\bar{\lambda}_4}{D\lambda_4^0}\right)^{-1} \Lambda_0 \frac{\partial \bar{\lambda}_4}{\partial \Lambda_0} \ , \tag{1.10}$$

where the arguments of all the functions are now Λ. After some elementary calculus we derive

$$\Lambda \frac{\partial V_6}{\partial \Lambda} = V_6 \left[2 + \overline{\frac{\partial \beta_6}{\partial \lambda_6}} - \overline{\frac{\partial \beta_4}{\partial \lambda_6}}\left(\frac{D\bar{\lambda}_4}{D\lambda_4^0}\right)^{-1} \frac{D\bar{\lambda}_6}{D\lambda_4^0}\right] \ . \tag{1.11}$$

Now provided the non-linear terms (the anomalous dimensions of V_6) are kept small, there is an approximate solution, namely,

$$V_6(\Lambda_\mathrm{R}) \simeq V_6(\Lambda_0)(\frac{\Lambda_\mathrm{R}}{\Lambda_0})^2 \ , \tag{1.12}$$

where from the initial conditions (5)

$$V_6(\Lambda_0) = -\Lambda \frac{\partial \bar{\lambda}_6}{\partial \Lambda}\bigg|_{\Lambda=\Lambda_0} + \frac{df}{d\lambda_4^0} \Lambda \frac{\partial \bar{\lambda}_4}{\partial \Lambda}\bigg|_{\Lambda=\Lambda_0} \tag{1.13a}$$

$$= -(2f(\lambda_4^0) + \beta_6(\lambda_4^0, f(\lambda_4^0)) + \frac{df}{d\lambda_4^0}\beta_4(\lambda_4^0, f(\lambda_4^0)) \ . \tag{1.13b}$$

The dependence of $V_6(\Lambda_0)$ on Λ_0 is determined by the dependence, defined by (8), of λ_4^0 on Λ_0. If we can bound it to vary slowly, as we shall for its analogue in the field theory, then $V_6(\Lambda_R)$ tends to zero as $\Lambda_0 \to \infty$. But remember we have

$$V_6(\Lambda_R) = \epsilon_6(\Lambda_R) = \Lambda_0 \frac{d}{d\Lambda_0}\bar{\lambda}_6(\Lambda_R, \Lambda_0, \lambda_4^0, f(\lambda_4^0))\Big|_{\bar{\lambda}_4(\Lambda_R)\text{fixed}} \ . \tag{1.14}$$

Thus integrating (14), $\bar{\lambda}_6(\Lambda_R)$ is approaching a definite limit as $\Lambda_0 \to \infty$, and this behaviour is sketched in Fig. 3. All we have to do now is show that this limiting value of $\bar{\lambda}_6(\Lambda_R)$ is *independent* of the choice of initial curve $f(\cdot)$.

In considering a family of curves $f_t(\cdot)$, we define

$$W_6(\Lambda, t) = \frac{d}{dt}\lambda_6(\Lambda, \Lambda_0, \lambda_4^0, \lambda_6^0)\Big|_{\lambda_4(\Lambda)\text{fixed}} \ , \tag{1.15}$$

and a similar analysis reveals that $W_6(\Lambda_R)$ also vanishes as the square of Λ_R/Λ_0 .

This completes the argument that the solutions to the flow equations (2) run toward a fixed surface of dimension one, and this indeed is the analogous statement of renormalisability: λ_4^0 can be found (as a possibly divergent function of Λ_0) such that at the lower scale Λ_R the couplings λ_4^R and λ_6^R are finite as Λ_0 is taken to infinity. Moreover, the initial condition function $f(\cdot)$ has been absorbed entirely into the relation between λ_4^0 and λ_4^R, and does not affect the 'irrelevant' variable λ_6^R once the 'relevant' variable λ_4^R has been chosen.

2. The Effective Lagrangian Flow Equations for $\lambda\phi^4$ Theory

There are many different forms of the effective Lagrangian flow equations, which in Wilson's language correspond to the different relative rates of 'partial integration' [9] of high momentum modes in the functional integral. A nice feature of Polchinski's analysis is that it is insensitive to so many details that it clearly will work no matter which set of equations is chosen. In particular, all the flows possess the *same* fixed point.

I write the euclidean partition function with momentum cutoff Λ in the following general way:

$$Z[J,\Lambda] = \int D\phi \, \exp \frac{1}{\hbar}\left[S - \frac{1}{2}\int \frac{d^4p}{(2\pi)^4} P^{-1}\phi_p\phi_{-p} + \int \frac{d^4p}{(2\pi)^4} AJ_p\phi_{-p} + B\right] , \quad (1.16)$$

where $P^{-1}(p^2,\Lambda^2)$ is the inverse of the cutoff propagator, S is the interaction (a functional of the field ϕ) and B may be some functional of the source J. The partition function is to be independent of Λ, so its Λ-derivative must vanish. We have

$$\hbar\Lambda \frac{d\log Z}{d\Lambda} = \langle \Lambda \frac{\partial S}{\partial \Lambda}\rangle - \frac{1}{2}\int \frac{d^4p}{(2\pi)^4}\Lambda \frac{\partial P^{-1}}{\partial \Lambda}\langle \phi_p\phi_{-p}\rangle + \int \frac{d^4p}{(2\pi)^4}\Lambda\frac{\partial A}{\partial \Lambda}J_p\langle\phi_{-p}\rangle + \Lambda\frac{\partial B}{\partial \Lambda} ,$$

$$(1.17)$$

where $\langle X\rangle$ is the average of the quantity X with respect to the probability measure $d\mu \equiv D\phi \, \exp \mathbf{S}/\hbar$.

The tactic now is to find some identity which contains among other things the terms $\langle\phi_p\phi_{-p}\rangle$ and $J_p\langle\phi_{-p}\rangle$. Using the rules for integrating by parts[f1], we have

$$0 = \int D\phi \frac{\delta}{\delta\phi_q}\left(\frac{1}{2}\frac{\delta}{\delta\phi_{-q}} + P^{-1}\phi_q - AJ_q\right)\exp\frac{1}{\hbar}\mathbf{S} . \quad (1.18)$$

Writing out the RHS of (18), and then adding in the same thing with $q \rightarrow -q$, this reads

$$0 = \langle \hbar\frac{\delta^2 S}{\delta\phi_q\delta\phi_{-q}} + \frac{\delta S}{\delta\phi_q}\frac{\delta S}{\delta\phi_{-q}}\rangle - P^{-2}\langle\phi_q\phi_{-q}\rangle$$

$$+ \hbar P^{-1}\delta^4(0) - A^2 J_q J_{-q} + P^{-1}A(J_q\langle\phi_{-q}\rangle + J_{-q}\langle\phi_q\rangle) . \quad (1.19)$$

Multiplying (1.19) throughout by $-\frac{1}{2}\int\frac{d^4q}{(2\pi)^4}\Lambda\frac{\partial P}{\partial\Lambda}$ we see from (1.17) that $\Lambda\frac{\partial}{\partial\Lambda}\log Z$ indeed vanishes if we *choose*

$$A = \beta P^{-1} \quad (1.20a)$$

---

f_1 Treating the path integral perturbatively, that is, as the generator of Feynman diagrams, equation (1.18) holds. For this and other functional manipulations, see Ref.[11].

$$B = -\frac{1}{2}\int \frac{d^4q}{(2\pi)^4}\left[\hbar\delta^4(0)\log P + \beta^2 P^{-1} J_q J_{-q}\right] + \gamma \qquad (\beta, \gamma \ \ \Lambda - \text{independent}) \tag{1.20b}$$

$$\Lambda\frac{\partial S}{\partial\Lambda} = -\frac{1}{2}\int \frac{d^4q}{(2\pi)^4}\Lambda\frac{\partial P}{\partial\Lambda}\left[\hbar\frac{\delta^2 S}{\delta\phi_q\delta\phi_{-q}} + \frac{\delta S}{\delta\phi_q}\frac{\delta S}{\delta\phi_{-q}}\right] . \tag{1.20c}$$

These equations are of the simplest non-trivial form to describe the running Lagrangian. The derivation makes it completely clear that there are many possibilities, for we could have used identities involving four or more functional derivatives.

Expanding the interaction S in powers of the field ϕ we can rewrite the flow in terms of dimensionless coefficient functions:

$$S(\phi,\Lambda) = \sum_{m=1}^{\infty}\frac{1}{(2m)!}\int\frac{d^4p_1\ldots d^4p_{2m}}{(2\pi)^{8m-4}\Lambda^{2m-4}}\delta^4(\textstyle\sum p)A_{2m}(p_1,\ldots,p_{2m};\Lambda)\phi(p_1)\ldots\phi(p_{2m}) , \tag{1.21}$$

where we respect the $\phi \to -\phi$ symmetry (it is preserved by (1.20c)) and ignore the field-independent piece, since it does not affect functional averages. The vertices A_{2m} then evolve according to (1.20c) as

$$\begin{aligned}\Lambda^{2m-4}\Lambda\frac{\partial}{\partial\Lambda}\frac{1}{\Lambda^{2m-4}}&A_{2m}(p_1,\ldots,p_{2m};\Lambda) = \\ &-\sum_{l=1}^{m} Q(P,\Lambda)\,A_{2l}(p_1,\ldots,p_{2l-1},P;\Lambda)\,A_{2m+2-2l}(p_{2l},\ldots,p_{2m},-P;\Lambda) \\ &+\text{ perms } - \frac{\hbar}{2}\int\frac{d^4p}{(2\pi\Lambda)^4}Q(p,\Lambda)\,A_{2m+2}(p_1,\ldots,p_{2m},p,-p;\Lambda) ,\end{aligned} \tag{1.22}$$

where $P = \sum_{i=1}^{2l-1} p_i$, and

$$Q = \Lambda^3\frac{\partial}{\partial\Lambda}P(p^2,\Lambda^2) . \tag{1.23}$$

These evolution equations can be written for simplicity in a graphical notation. Denoting A_{2m} by a $2m$-legged vertex and Q by a straight line we have, ignoring all numerical factors, the graphs in Fig. 4. The closed loop there indicates an integral over momentum, $\int\frac{d^4p}{(2\pi\Lambda)^4}$. Equation (1.22) is the analogue in the field theory of equation (1.2) of the toy system. Here there are three 'relevant' variables as opposed to one, namely the coefficients of the three dimension ≤ 4 operators $\phi^2, (\partial^x_\mu \phi)^2$, and ϕ^4 , i.e.,

$$\rho_1 = -\Lambda^2 A_2(0,0;\Lambda) \tag{1.24a}$$

$$\rho_2 = -\frac{1}{8}\Lambda^2 \frac{\partial^2 A_2}{\partial p^2}(p,-p;\Lambda)\big|_{p=0} \tag{1.24b}$$

$$\rho_3 = -A_4(0,0,0,0;\Lambda) \ , \tag{1.24c}$$

and all the other variables, denoted collectively by $\vec{\mathbf{I}}$, are irrelevant. Following the same idea as before, we now make the analogous definitions to Eqs. (1.10) and (1.15).

$$V(\phi,\Lambda,\Lambda_0,\rho_b,t) = \Lambda_0 \frac{d}{d\Lambda_0} S(\phi,\Lambda,\Lambda_0,\rho_b^0,\vec{\mathbf{I}}^0)\big|_{\rho_b \text{ fixed}} \tag{1.25a}$$

$$W(\phi,\Lambda,\Lambda_0,\rho_b,t) = \frac{d}{dt} S(\phi,\Lambda,\Lambda_0,\rho_b^0,\vec{\mathbf{I}}^0)\big|_{\rho_b \text{ fixed}} \tag{1.25b}$$

where $\vec{\mathbf{I}}^0 = \vec{f}_t(\rho_b^0)$, and

$$\rho_b^0 = \rho_b^0(\Lambda,\Lambda_0,\rho_a,[\vec{f}_t]) \ . \tag{1.25c}$$

Just as in Eq. (1.11), we consider the Λ-derivative of V and W in order to show that they tend to zero as Λ_0 is taken to infinity. We have

$$V = \Lambda_0 \frac{\partial S}{\partial \Lambda_0} - \sum_{b=1,2,3} B_b \Lambda_0 \frac{\partial \rho_b}{\partial \Lambda_0} \ , \tag{1.26}$$

where

$$B_b = \sum_{a=1,2,3} \frac{DS}{D\rho_a^0} \left(\frac{D\rho}{D\rho^0}\right)^{-1}_{ab} \ , \tag{1.27}$$

and the capitalized partial derivative is defined analogously to that in (1.7). The calculus yields

$$\Lambda \frac{\partial V}{\partial \Lambda} = M[V] - \sum_{b=1,2,3} B_b M_b[V] \tag{1.28a}$$

where

$$M[\,\cdot\,] = -\frac{1}{2}\int \frac{d^4p}{(2\pi\Lambda)^4}\Lambda^2 Q(p^2,\Lambda^2)\left\{2\frac{\delta S}{\delta\phi_{-p}}\frac{\delta}{\delta\phi_p} + \frac{\delta^2}{\delta\phi_p\delta\phi_{-p}}\right\}[\,\cdot\,]\,, \qquad (1.28b)$$

and $M_b[V]$ is the coefficient of ϕ^2, $(\partial^x_\mu\phi)^2$, ϕ^4 in $M[V]$.

Similarly for W we find

$$\Lambda\frac{dW}{d\Lambda} = M[W] - \sum_{b=1,2,3} B_b M_b[W]\,. \qquad (1.29)$$

Expanding the unknown quantities V and B_b as a power series in ϕ,

$$V(\Lambda) = \sum_{m=1}^{\infty}\frac{1}{(2m)!}\int\frac{d^4p_1\ldots d^4p_{2m}}{(2\pi)^{8m-4}\Lambda^{2m-4}}\delta^4(\textstyle\sum p)V_{2m}(\{p\};\Lambda)\phi(p_1)\ldots\phi(p_{2m}) \qquad (1.30)$$

$$B_b(\Lambda) = \sum_{m=1}^{\infty}\frac{1}{(2m)!}\int\frac{d^4p_1\ldots d^4p_{2m}}{(2\pi)^{8m-4}\Lambda^{2\delta_{b1}+2m-4}}\delta^4(\textstyle\sum p)B_{b,2m}(\{p\};\Lambda)\phi(p_1)\ldots\phi(p_{2m})\,, \qquad (1.31)$$

then equation (1.28a) thus contains vertices of three types, A, B and V, and it is easiest to read it off in its graphical form. Using the notation of a $2m$-legged vertex with dashed-line legs to represent B_{2m} and double-line legs to represent V_{2m}, we can rewrite (1.28a) as Fig. 5.

We now have the equations of motion for the A- and V- vertices. It only remains to find that of the B-vertices for the analysis to begin. (The equation of motion of the W-vertices is identical to that of the V-vertices, so I won't rewrite it.) The operator M is linear, and we find

$$\Lambda\frac{\partial B_b}{\partial\Lambda} = M[B_b] - \sum_a B_a M_a[B_b]\,. \qquad (1.32)$$

The graphical form is given in Fig. 6. Figures 4, 5, and 6 are the key to the procedure. Their schematic representation in fact includes all the information that we require; in particular, the numerical factors and signs that we have suppressed play no role in the arguments to follow.

3. Description of the Flow in Perturbation

Assume for the moment there exists a flow such that at Λ_0 the interaction takes on the simple form,

$$S(\phi, \Lambda_0) = -\int d^4x \left[\frac{1}{2}\rho_1^0\phi^2 + \frac{1}{2}\rho_2^0(\partial_\mu^x\phi)^2 + \frac{1}{24}\rho_3^0\phi^4\right] , \tag{1.33}$$

i.e., $\vec{f_t}(\cdot)$ is chosen to be identically zero, and down at the scale $\Lambda_{\rm R}$ the relevant parameters are given by

$$\rho_1^{\rm R} = 0 \tag{1.34a}$$

$$\rho_2^{\rm R} = 0 \tag{1.34b}$$

$$\rho_3^{\rm R} = \lambda^{\rm R} . \tag{1.34c}$$

Of course the full interaction at $\Lambda_{\rm R}$ is certainly not of the form (1.33). Remember that S here is the interaction only; it does not contain the inverse propagator terms in the full action.

Now even for a finite value of Λ_0, there is a non-trivial part in this assumption, namely that $\rho_b^0$and $\rho_b^{\rm R}$ are *both* taken to be strictly finite. It could have happened after all that the flow down from (1.33) led to an interaction at $\Lambda_{\rm R}$ containing the term $\int \frac{d^4p}{(2\pi)^4} \log p^2/L_{\rm R}^2\, \phi_p\phi_{-p}$, so that $\rho_2^{\rm R}$ would in fact be infinite. This would be really problematic because a finite change in the cutoff (i.e., a finite 'blocking' of the original variables) would be changing the interaction from a local to a non-local form.

In perturbation we shall simply demonstrate that this does not happen, and indeed the assumption is correct. Here we see the first benefit of the effective Lagrangian method, for in other schemes the possibility of non-local counterterms is a very hard question to deal with.

By 'perturbation' I mean that I consider ρ_b^0, and hence every quantity, to be a power series in $\lambda^{\rm R}$. We shall show that order by order in $\lambda^{\rm R}$ we can find ρ_b^0 such that the flow is as described[f2] .

The importance of this result is that we will have tied down the two ends of the flow. All that would be left are the irrelevant parameters $\vec{\mathbf{I}}^{\mathrm{R}}$, which we discuss using the flow for $V(\phi,\Lambda)$. Note that (1.34) involves no loss of generality. If I just say that ρ_1^{R} and ρ_2^{R} are finite numbers then I can always set them to zero with a finite wavefunction and mass renormalisation.

Clearly we need bounds on the A-vertices, since we are trying to prove a finiteness property, and it is easiest if we consider the arguments of the vertex functions to be restricted in range. This can be accomplished if the cutoff propagator p is of compact support. Values of the vertex functions for momenta outside this range would then be unphysical, and thus ignorable. For definiteness let me choose for the cutoff propagator,

$$P = \frac{1}{p^2+m^2} K(p^2/\Lambda^2) \tag{1.35a}$$

$$K(x) = \begin{cases} 1 & \text{if } x \leq 1 \\ (1+\mathrm{e}^{\frac{1}{1-x}})\mathrm{e}^{\frac{1}{x-4}}\mathrm{e}^{\frac{1}{3}} & \text{if } 1 < x < 4 \\ 0 & \text{if } x \geq 4 \end{cases} . \tag{1.35b}$$

A simplifying feature of this choice is that K is infinitely differentiable, and so there exist constants C and D_n such that

$$\int \frac{d^4p}{(2\pi\Lambda)^4} \, |Q| \, < \, C \tag{1.36a}$$

$$\| \frac{\partial^n Q}{\partial p^n} \| \, < \, D_n \Lambda^{-n} \; , \tag{1.36b}$$

---

f_2 The perturbation series makes sense for values of λ^{R} sufficiently small such that all the couplings at all the scales are small. The measure of size is $\Lambda_{\mathrm{R}}/\Lambda_0$, and so none of the analysis applies to the strict continuum limit, $\Lambda_0 = \infty$.

where the 'double norm' is defined by $\|f(p_i,\Lambda)\| \equiv \max_{p_i^2\le 4\Lambda^2} |f(p_i,\Lambda)|$, and Q is defined by (1.23). From (1.22) we have directly

$$\|\Lambda^{2m-4}\Lambda\frac{\partial}{\partial\Lambda}A_{2m}\| \le \sum_{l=1}^{m}\frac{1}{2}\binom{2m}{2l-1}D_0\|A_{2l}\|\cdot\|A_{2m+2-2l}\|+\frac{1}{2}\hbar C\|A_{2m+2}\| \ , \tag{1.37}$$

or, schematically, for the rth order in λ^{R},

$$\|\Lambda^{2m-4}\Lambda\frac{\partial}{\partial\Lambda}A^{(r)}_{2m}\| \lesssim \sum_{l=1}^{m}\sum_{t=0}^{r}\|A^{t}_{2l}\|\cdot\|A^{(r-t)}_{2m+2-2l}\| \ + \ \|A^{(r)}_{2m+2}\| \ , \tag{1.38}$$

where the symbol $\lesssim$ indicates the suppression of numerical factors. Taking derivatives of (1.22) with respect to momenta we can similarly derive

$$\|\Lambda^{2m-4}\Lambda\frac{\partial}{\partial\Lambda}\partial^n_p A^{(r)}_{2m}\| \lesssim \sum_{l=1}^{m}\sum_{t=0}^{r}\sum_{n_1+n_2\le n}\|\partial^{n_1}_p A^{(t)}_{2l}\|\cdot\|\partial^{n_2}_p A^{(r-t)}_{2m+2-2l}\| + \|\partial^n_p A^{(r)}_{2m+2}\| \ . \tag{1.39}$$

I now prove that I can find ρ^0_b order by order in λ^{R} such that all the vertices and their derivatives are strictly finite. This is enough to prove the above assumption, equations (1.33) and (1.34), for then we'll have a finite functional relation $\rho^{\mathrm{R}}_b = \rho^{\mathrm{R}}_b(\rho^0_a)$ which is invertible, due to

$$\frac{\partial\rho^{\mathrm{R}}_b}{\partial\rho^0_a} = \delta_{ba} + O(\lambda^{\mathrm{R}}) \ . \tag{1.40}$$

Proceeding inductively, suppose the hypothesis holds at order $r-1$. I.e.,

$$\|\partial^n A^{(s)}_{2m}(p_1,\ldots,p_{2m};\Lambda)\| < \infty \qquad\qquad \text{all}\, p,\ s\le r-1 \ . \tag{1.41}$$

It certainly holds at zeroth order, where $\rho^{(0)}_b = 0$ and so $S^{(0)}(\phi,\Lambda) = 0$. (This is just the statement that at zeroth order in the coupling there is no interaction, and the theory contains only the free cutoff propagator.) At rth order I note that for m sufficiently large $A^{(r)}_{2m}$ is identically zero, since the linking together of r 4-vertices can yield at most a $2r+2$-vertex. So the induction hypothesis holds at rth order for $m \geq r+2$ and now we proceed downwards in m. Given that the hypothesis holds for $m \geq k+1$ then (1.39) tells us that

$$\|\Lambda^{2k-4}\Lambda\frac{\partial}{\partial\Lambda}\frac{1}{\Lambda^{2k-4}}\partial^n_p A^{(r)}_{2k}\| < \infty \; . \tag{1.41}$$

Integrating down from Λ_0 to Λ we see that $\|A^{(r)}_{2k}(\Lambda)\|$ is indeed finite provided its value at Λ_0 is chosen to be finite, which it is. We can go all the way down to $m=1$, and this proves the hypothesis.

4. The Renormalisation Theorem

Using the 'tying down' of the previous section we're now ready to study the renormalisation flow of $V(\phi,\Lambda)$. Since the equation of motion (1.28a), c.f. Fig. 5, involves vertices of all types A,B and V, we'll need bounds on types A and B which can then be fed in to get the desired bounds on type V.

From (1.22), or rather its estimated form (1.39), we will prove that

$$\|\partial^n_p A^{(r)}_{2m}(\Lambda)\| \leq \Lambda^{-n} P(\log\frac{\Lambda_0}{\Lambda_R}) \; , \tag{1.43}$$

where $P(\cdot)$ is some polynomial with positive coefficients. This of course is our prized statement that $S(\phi,\Lambda)$ flows according to its canonical dimensions, excepting for possible logarithmic deviations (which are kept small by powers of the coupling λ^{R}). From (1.32), c.f. Fig. 6, and (1.43) we then prove an estimate for the B-vertices, namely,

$$\|\partial^n_p B^{(r)}_{b,2m}(\Lambda)\| \leq \Lambda^{-n} P(\log\frac{\Lambda_0}{\Lambda_R}) \; . \tag{1.44}$$

Finally the *coup de grâce.* From (1.28a) (1.43) and (1.44) we prove

$$\|\partial^n_p V^{(r)}_{2m}(\Lambda)\| \leq (\frac{\Lambda}{\Lambda_0})^2 \, \Lambda^{-n} P(\log\frac{\Lambda_0}{\Lambda_R}) \; . \tag{1.45}$$

So V vanishes as the inverse-square of Λ_0, just as in the toy system, c.f.(1.12). This indeed is according to the canonical dimension, for the leading irrelevant parameters are ϕ^6 and $\partial_x^2\phi^4$. From the defining equation (1.25a) the interaction $S(\phi, \Lambda_{\rm R})$ therefore has a definite limit as $\Lambda_0 \to \infty$, which it approaches as Λ_0^{-2}.

This last property now holds for the low energy Green's functions, i.e., those with external momenta satisfying $p_i^2 \leq \Lambda_{\rm R}^2$. We imagine calculating them using the interaction at $\Lambda_{\rm R}$, for the whole point of the flow was that we could calculate them at any scale we wish. At rth order $G^{(r)}(\Lambda_0, \Lambda_{\rm R}, \lambda^{\rm R})$ is given by a finite number of terms of the form,

$$\int d^{4l}p\, A_{2m_1}^{(r_1)}(\{p\}, \Lambda_{\rm R}) \ldots A_{2m_k}^{(r_k)}(\{p\}, \Lambda_{\rm R})\, P(p_1, \Lambda_{\rm R}) \ldots P(p_n, \Lambda_{\rm R}) \ , \qquad (1.46)$$

and these integrals are over a finite range and therefore converge for $m^2 > 0$.

For the massless case note that the vertex functions A_{2m} have a smooth limit as $m^2 \to 0$. This is simply because, with the choice (1.35) of cutoff propagator, Q has a smooth limit, as have the bounding constants C and D_n. Thus all we have to do is choose unexceptional momenta[f3] for the Green's functions, and the terms (1.46) are all finite. No infrared regularisation is required !

Finally for the case of 'Spontaneous Symmetry Breaking' we have $m^2 < 0$. This does not adversely affect the flow analysis for $S(\phi, \Lambda)$ provided $\Lambda^2 > |m^2|$, because Q only has support for $\Lambda^2 \leq p^2 \leq 4\Lambda^2$. Thus the bounds (1.36) on Q still hold. To obtain the Green's functions we take the interaction and kinetic terms at $\Lambda_{\rm R}$ and *shift* the field ϕ by some finite constant. The shifted vertex functions are still finite, and now the propagator has a positive mass-square so that we can once again perform the integrals in (1.46).

We see that the effective Lagrangian approach is by far the cleanest way to separate the ultraviolet and infrared divergences, and to show that the low-energy phenomenon of SSB is entirely separate from the issue of renormalisation. This will be very important in our study of gauge theories.

---

f_3 The momenta in the argument of a Green's function are said to be 'unexceptional' if they are all euclidean and no partial sum vanishes.

The proof of estimates (1.43), (1.44) and (1.45) is given below. It is lifted entirely from the paper by Polchinski, and I reproduce it because it will be important later to understand the mechanics of the argument. The idea is very similar in nature to the discussion in Section 2, that is, we have induction in the order of the coupling, λ^{R}, and the number of external lines, $2m$. For a given order we go downwards in m (for the same reason as before) and then we go from one order to the next.

A nice extra feature of the argument is that we see directly the naturalness problem for light scalars, as mentioned in the Introduction. For the renormalised trajectory that we have described is not 'typical.' Normally the mass parameter ρ_2 is of order Λ_0^2, whatever the scale, so in any real model we would expect the scalars all to be heavy and have a mass comparable to the physical cutoff.

Proof of bound on A-vertices, Eq.(1.43)

The bound clearly holds at zeroth order, and at any order r for $m \geq r+2$, since as stated above $A^0_{2m} \equiv 0$ and $A^{(r)}_{2m} = 0$ for $m \geq r+2$. Assume that it holds at order $r-1$, and at order r for $m \geq k+1$. The RHS of (1.39) for $m = n$ is then bounded by hypothesis and we have

$$||\Lambda^{2k-4}\Lambda\frac{\partial}{\partial\Lambda}\frac{1}{\Lambda^{2k-4}}\partial^n_p A^{(r)}_{2k}|| \leq \Lambda^{-n}\, P(\log\frac{\Lambda_0}{\Lambda_{\mathrm{R}}}) \,. \tag{1.47}$$

For $k \geq 3$ the initial value is zero so integrating down from Λ_0 to Λ,

$$\begin{aligned} ||\partial^n_p A^{(r)}_{2k}|| &\leq \Lambda^{-n} P(\log\frac{\Lambda_0}{\Lambda_{\mathrm{R}}}) \int_{\Lambda}^{\Lambda_0} \frac{d\Gamma}{\Gamma}(\frac{\Lambda}{\Gamma})^{2k+n-4} \\ &\leq \Lambda^{-n} P(\log\frac{\Lambda_0}{\Lambda_{\mathrm{R}}}) \qquad , k \geq 3 \,. \end{aligned} \tag{1.48}$$

This carries to $k = 2, n = 1$, but for $k = 2, n = 0$ there is an initial value which is as yet unknown. But by virtue of the 'tying down' we do know the end point value at zero momentum, namely $A^{(r)}_4(0,0,0,0;\Lambda_{\mathrm{R}}) = -\delta^{r1}$. Now from (1.48) we obtain

(1.47) for $k=2, n=0$, i.e.,

$$\left|\Lambda\frac{\partial}{\partial\Lambda}A_4^{(r)}(0,0,0,0;\Lambda_{\rm R})\right| \leq P(\log\frac{\Lambda_0}{\Lambda_{\rm R}}) \ , \tag{1.49}$$

whence integrating *up* from $\Lambda_{\rm R}$ to Λ we derive

$$\left|A_4^{(r)}(0,0,0,0;\Lambda)\right| \leq P(\log\frac{\Lambda_0}{\Lambda_{\rm R}}) \ . \tag{1.50}$$

Now, since $A_4^{(r)}(p_1,p_2,p_3,p_4;\Lambda)$ is reconstructible, by Taylor's theorem, from $A_4^{(r)}(0,0,0,0;\Lambda)$ and $\partial^2/\partial p_i^\mu \partial p_j^\nu\, A_4^{(r)}(p_1,p_2,p_3,p_4;\Lambda)$, for both of which the bound holds, then the bound holds for it also. For $k=1$ we proceed analogously. For $k=1, n=4$ there is no initial value so we integrate down from Λ_0; for $n=2$ and $n=0$ we know the zero momentum end point value and we integrate up from $\Lambda_{\rm R}$. This allows a full reconstruction of $A_2^{(r)}(p,-p;\Lambda)$, thus completing the induction.

The naturalness problem is seen by considering the integral of (47) for $k=1, n=0$ down from Λ_0. This would put the bound on $A_2(\Lambda)$ to be $O(\Lambda_0/\Lambda)^2$, so that to get the much smaller value of $O(1)$ the initial conditions have to be very finely tuned.

Proof of bound on B-vertices, Eq.(1.44)

Just as above, we need to know information about the initial condition and the end-point. From definition (1.27) we have directly

$$B_{b,2}^{(r)}(0,0;\Lambda) = -\delta^{r0}\delta_{b1} \tag{1.51a}$$

$$\frac{\partial^2}{\partial p^\mu \partial p^\nu}B_{b,2}^{(r)}(p,-p;\Lambda)\big|_{p=0} = -\frac{1}{2}\delta^{r0}\delta^{\mu\nu}\delta_{b2} \tag{1.51b}$$

$$B_{b,4}^{(r)}(0,0,0,0;\Lambda) = -\delta^{r0}\delta_{b3} \ , \tag{1.51c}$$

and from our choice of $\vec{f}_t(\cdot)=0$ the initial condition is

$$B_{b,2m}(p_1,\ldots,p_{2m};\Lambda_0) = -\delta_{b1}\delta_{m1} - \delta_{b2}\delta_{m1}\frac{p_1^2}{\Lambda_0^2} - \delta_{b3}\delta_{m2} \ . \tag{1.52}$$

Consideration of the evolution equation (1.32) at zeroth order now yields

$$B^{(0)}_{b,2m}(p_1,\dots,p_{2m};\Lambda) = -\delta_{b1}\delta_{m1} - \delta_{b2}\delta_{m1}\frac{p_1^2}{\Lambda^2} - \delta_{b3}\delta_{m2}\ , \tag{1.53}$$

so the bound holds for $r=0$. It also holds at any order r for $m \geq r+3$, for then $B^{(r)}_{b,2m}$ vanishes, as can be seen by plugging in the vanishing of $A^{(r)}_{2m}|_{m\geq r+2}$ and the initial condition into (1.32).

Assume that the bound holds up to order $r-1$ and at order r for $m \geq k+1$. Estimating (1.32) analogously to (1.39),

$$\begin{aligned}
&\|\Lambda^{2k-4+2\delta_{b1}}\Lambda\frac{\partial}{\partial\Lambda}\frac{1}{\Lambda^{2k-4+2\delta_{b1}}}\partial_p^n B^{(r)}_{b,2k}\| \lesssim \\
&\quad \sum_{l=1}^{k}\sum_{t=0}^{r}\sum_{n_1+n_2\leq n}\|\partial_p^{n_1}A^{(t)}_{2l}\|\cdot\|\partial_p^{n_2}B^{(r-t)}_{b,2k+2-2l}\| + \|\partial_p^n B^{(r)}_{b,2k+2}\| \\
&\quad + \sum_{t=0}^{r}\|\partial_p^n B^{(t)}_{1,2k}\|\cdot\|B^{(r-t)}_{b,4}\| + \sum_{t=0}^{r}\|\partial_p^n B^{(t)}_{2,2k}\|\cdot\|\Lambda^2\partial_p^2 B^{(r-t)}_{b,4}\| \\
&\quad + \sum_{t=0}^{r}\|\partial_p^n B^{(t)}_{3,2k}\|\cdot\|B^{(r-t)}_{b,6}\|\ ,
\end{aligned} \tag{1.54}$$

then we see that the induction cannot automatically continue due to the presence of unknown order r terms on the RHS. These are the last three terms for $t=0$ and the third-to-last term for $t=r$. The former problem is not present if $k \geq 3$, due to (1.53), and the latter problem only involves $b=3$. Thus for $k\geq 3, b=1,2$ the induction *can* proceed and we have

$$\|\Lambda^{2k-4+2\delta_{b1}}\Lambda\frac{\partial}{\partial\Lambda}\frac{1}{\Lambda^{2k-4+2\delta_{b1}}}\partial_p^n B^{(r)}_{b,2k}\| \leq \Lambda^{-n}\,P(\log\frac{\Lambda_0}{\Lambda_R})\quad k\geq 3,\ b=1,2. \tag{1.55}$$

Integrating down from Λ_0 to Λ,

$$\|\partial_p^n B^{(r)}_{b,2k}\| \leq \Lambda^{-n}\,P(\log\frac{\Lambda_0}{\Lambda_R})\ . \tag{1.56}$$

Going back to (1.54) we can feed this result in to remove the problem for $b=3$ and so the induction step is fully established down to $n=3$.

The bound on $B^{(r)}_{b,6}(\Lambda)$ now allows (1.55) to be derived for $n = 2$. We can integrate down since we know the initial condition (1.52), and so the induction is carried. The last domino falls since now (1.55) can be derived for $n = 1$. We can integrate down for $p \geq 2$, and finally the $n = 1, p = 0$ bound is provided by (1.51a).

Proof of bound on V-vertices, Eq.(1.45)

To obtain the initial condition note that for the choice $\vec{f_t} = 0$ we have, c.f. (1.13a),

$$V(\phi,\Lambda,\Lambda_0,\rho^0)\big|_{\Lambda=\Lambda_0} = -\Lambda\frac{\partial}{\partial\Lambda}S(\phi,\Lambda,\Lambda_0)\big|_{\Lambda=\Lambda_0}\ , \tag{1.57}$$

whereupon equation (1.47) evaluated at Λ_0 yields

$$\|\partial_p^n V_{2m}^{(r)}(\Lambda_0)\| \leq \Lambda_0^{-n}\, P(\log\frac{\Lambda_0}{\Lambda_R})\ . \tag{1.58}$$

Just as for the A-vertices, $V^{(0)}_{2m} = 0$, and $V^{(r)}_{2m}\big|_{m\geq r+2} = 0$. Assume once again that the bound holds up to order $r-1$ and at order r for $m \geq k+1$. The estimated form of the evolution equation (1.28a) is

$$\begin{aligned}\|\Lambda^{2k-4}\Lambda\frac{\partial}{\partial\Lambda}\frac{1}{\Lambda^{2k-4}}\partial_p^n V_{2m}^{(r)}\| \lesssim{}& \sum_{l=1}^{k}\sum_{t=0}^{r}\sum_{n_1+n_2\leq n}\|\partial_p^{n_1}A_{2l}^{(r)}\|\cdot\|\partial_p^{n_2}V_{2k+2-2l}^{(r-t)}\| + \|\partial_p^n V_{2k+2}^{(r)}\| \\ &+\sum_{t=0}^{r}\|\partial_p^n B_{1,2k}^{(t)}\|\cdot\|V_4^{(r-t)}\| + \sum_{t=0}^{r}\|\partial_p^n B_{2,2k}^{(t)}\|\cdot\|\Lambda^2\partial_p^2 V_4^{(r-t)}\| \\ &+\sum_{t=0}^{r}\|\partial_p^n B_{3,2k}^{(t)}\|\cdot\|V_6^{(r-t)}\|\ ,\end{aligned} \tag{1.59}$$

and as in the previous case the induction is not immediate. But the bothersome $t = 0$ terms do not arise for $k \geq 3$ and so

$$\|\Lambda^{2k-4}\Lambda\frac{\partial}{\partial\Lambda}\frac{1}{\Lambda^{2k-4}}\partial_p^n V_{2k}^{(r)}\| \leq (\frac{\Lambda}{\Lambda_0})^2\,\Lambda^{-n}P(\log\frac{\Lambda_0}{\Lambda_R}) \qquad k\geq 3\,. \tag{1.60}$$

Integrating down from Λ_0 and using (1.58),

$$\|\partial_p^n V_{2k}^{(r)}\| \leq \Lambda^{-n}(\frac{\Lambda}{\Lambda_0})^2 P(\log\frac{\Lambda_0}{\Lambda_R})\int_{\Lambda}^{\Lambda_0}\frac{d\Gamma}{\Gamma}(\frac{\Lambda}{\Gamma})^{2k+n-6} + \Lambda^{-p}(\frac{\Lambda}{\Lambda_0})^{2k+n-4}P(\log\frac{\Lambda_0}{\Lambda_R})$$

$$\leq \Lambda^{-n}(\frac{\Lambda}{\Lambda_0})^2 P(\log\frac{\Lambda_0}{\Lambda_R}) , \tag{1.61}$$

and the induction step is established down to $k \geq 3$. The bound on V_6 gives us (1.60) for $k = 2$, and the integration down is possible for $n \geq 2$. For $k = 2, n = 0$ the end point value is known, $V_4(0,0,0,0;\Lambda) \equiv 0$, and the induction carries. The $k = 1$ bound is then analogously established, using the end point values for $n = 2$ and $n = 0$.

This completes the proof of the renormalisability of $\lambda\phi^4$ theory, in the conventional sense: the interaction (1.33), containing only coupling constant, mass, and wavefunction renormalisations, has been shown to yield finite Green's functions in the limit $\Lambda_0 \to \infty$. We have not yet addressed the possible effects of the initial condition function $\vec{f}_t(\cdot)$, set to zero in the above, and it is to this and other generalisations that I now turn.

5. General Results from the Method of Effective Lagrangians

The Polchinski analysis, presented in detail in the last two sections, has the considerable merit of being very flexible. In this way it truly captures the spirit of the effective Lagrangian idea. Using the *identical* reasoning I shall show that a wide class of initial condition functions can be absorbed (as in the toy system), and moreover that the different forms of the flow equations do indeed give the same results. Then I shall discuss the extension to theories with many fields, be they bosons or fermions, scalars, spinors, or vectors.

The Initial Condition Function

Instead of the simple interaction (1.33) suppose we allow at scale Λ_0 the presence of higher dimension operators. For canonical dimension D the coefficient is of the form

$$\frac{1}{\Lambda_0^D} G_t(\rho_1^0/\Lambda_0^2, \rho_2^0, \rho_3^0) . \tag{1.62}$$

We must check that we still flow down to Λ_R without generating non-local terms, and that the estimates on the A, B and V-vertices still hold. Also we must check that

W vanishes as $\Lambda_0 \to \infty$, to verify that the fixed surface at $\Lambda_{\rm R}$ is indeed independent of the initial conditions.

Now the inductive arguments that we used, based on going downwards in m for a given order r, relied on only two properties: namely that $A^{(r)}_{2m}$, and hence $B^{(r)}_{2m}$ and $V^{(r)}_{2m}$, vanished for sufficiently large m, and that $A^{(0)}_{2m} = 0$. I simply demand that we keep these properties. Thus we consider initial conditions where the operators that appear have a finite power of the field ϕ, with coefficients that go to zero when $\rho^0_b = 0$.

The 'tying down' argument of Section I.3 now goes over word for word. In the proof of the A-vertices bound, the integrated equation (1.48) now contains the initial values (1.62). But this is no problem because we can still bound ρ^0_b by integrating upwards, and then we feed these bounds back in.

For the B-vertices we have to make some restrictions in order to let the induction continue. An initial term $\rho^0_3\,\phi^8/\Lambda_0^4$, for example, would be disastrous for then $B^{(0)}_{3,8}(\Lambda) \neq 0$ and equation (1.53) could *not* be established for $n = 4$. It is somewhat contrary to the philosophy to have to preclude this term, but it causes manifestly a break in the proof. I shall take the attitude that an improved proof could cover this case, but I will not need this in the sequel. Rather for simplicity, all the subsequent analysis will use the Polchinski arguments only. It is sufficient if in (1.62) we say that $G_t(\ldots)$ is a function of ρ^0_3 only, with $dG_t/d\rho^0_3$ vanishing at zeroth order in ρ^0_3 for $m \geq 3$.

For example, the terms $(\rho^0_3)^2\phi^8/\Lambda_0^4$ and $\rho^0_3\,\partial_x^{2N}\phi^4/\Lambda_0^{2N}$ are acceptable. (This is good because these are the sorts of terms that I'll introduce in the regularisation of gauge theories; see Section II.2.) Now it's still true that $B^{(0)}_{1,2m}$ and $B^{(0)}_{2,2m}$ vanish for $m \geq 2$ and $B^{(0)}_{3,2m}$ vanishes for $m \geq 3$, and the argument now holds. The integrated equation (1.56) will now contain initial values, admittedly, but these are controlled by the bound on ρ^0_b. For the V-vertices note that (1.57) must be changed to include the presence of $G_t(\ldots)$, as in the toy equations (1.13a). However the extra terms are again already bounded, and so the initial condition still satisfies (1.58). Naturally $V^{(0)}_{2m}$ still vanishes, and the induction can proceed exactly as it can for the B-vertices.

Finally, for the W-vertices it is guaranteed by the form of $G_t(\rho_3^0)$ that

$$\|\partial_p^n W_{2m}^{(r)}(\Lambda_0)\| \le \Lambda_0^{-p}\, P(\log\frac{\Lambda_0}{\Lambda_{\rm R}}) \tag{1.63}$$

$$W_{2m}^{(0)} = 0 \ , \tag{1.64}$$

and these were just the properties required to bound the V-vertices. Since the equations of motion are the same, we can conclude the same bound, i.e.,

$$\|\partial_p^n W_{2m}^{(r)}(\Lambda)\| \le (\frac{\Lambda}{\Lambda_0})^2 \Lambda^{-p}\, P(\log\frac{\Lambda_0}{\Lambda_{\rm R}}) \ , \tag{1.65}$$

and this completes the proof of the absorption of the initial condition function.

Arbitrariness in the Flow Equations

Returning to the derivation in Section I.2 of the flow equations (1.20), suppose we had employed, say, quartic functional derivatives. In our graphical notation this would have introduced terms like those in Fig. 7, but it would still be true that $A_{2m}^{(0)} = 0$ and $A_2^{(0)} = 0$ for large enough m. As should by now be clear, this means that there is no obstruction to the Polchinski arguments, and the flow still runs towards a 3-dimensional fixed surface.

Now for completeness, we should check the physical equivalence of different forms of the flow equations.

Suppose $(\rho_b^0)^{\mathbf{1}}$ are the initial relevant parameters which flow down using form **1** of the flow equations to the fixed surface defined by **1**. The Green's functions are Λ_0-finite, and depend only on $\Lambda_{\rm R}$, $(\rho_b^{\rm R})^{\mathbf{1}}$ and 'functionally' on form **1**. I shall denote this by $G(\Lambda_{\rm R}, (\rho_b^{\rm R})^{\mathbf{1}}, \mathbf{1})$.

Now the renormalisation condition is normally made on the Green's functions, say the 2-pt and 4-pt functions, and as an example we could have[f4]

$$G^{(2)}(p^2) = g_1 \ \text{ at } p^2 = 0 \tag{1.66a}$$

f_4 These conditions are appropriate to a massive theory; in the massless case we choose the 'subtraction point' to be at some euclidean momentum.

$$\frac{\partial}{\partial p^2} G^{(2)}(p^2) = g_2 \ \text{ at } p^2 = 0 \tag{1.66b}$$

$$G^{(4)}(p_1, p_2, p_3, p_4) = g_3 \ \text{ at } p_i = 0 \ . \tag{1.66c}$$

Rewriting the Green's functions in terms of the conditions $\vec{g}$ we get $\tilde{G}(\Lambda_{\rm R}, \vec{g}, \mathbf{1})$. Now, repeating the above for form **2** of the flow equations we get $\tilde{G}(\Lambda_{\rm R}, \vec{g}, \mathbf{2})$, and what we wish to establish is

$$\tilde{G}(\Lambda_{\rm R}, \vec{g}, \mathbf{1}) = \tilde{G}(\Lambda_{\rm R}, \vec{g}, \mathbf{2}) \ . \tag{1.67}$$

To show this, take $(\rho_b^0)^{\mathbf{1}}$ and flow down using form **2**. Since Green's functions calculated from $(\rho_b^0)^{\mathbf{1}}$ are Λ_0-finite, it must be that the vertex functions at $\Lambda_{\rm R}$ are also Λ_0-finite, i.e., we've reached the fixed surface defined by form **2**. Suppose the relevant parameters are now $(\rho_b^{\rm R})^{\mathbf{1,2}}$, then we have

$$G(\Lambda_{\rm R}, (\rho_b^{\rm R})^{\mathbf{1}}, \mathbf{1}) = G(\Lambda_{\rm R}, (\rho_b^{\rm R})^{\mathbf{1,2}}, \mathbf{2}) \ , \tag{1.68}$$

and we see that the two flow forms correspond to two different *schemes* of renormalisation, related by a finite redefinition of the parameters. By definition of the conditions $\vec{g}$ we have

$$G(\Lambda_{\rm R}, (\rho_b^{\rm R})^{\mathbf{1}}, \mathbf{1}) = \tilde{G}(\Lambda_{\rm R}, \vec{g}, \mathbf{1}) \tag{1.69}$$

and

$$G(\Lambda_{\rm R}, (\rho_b^{\rm R})^{\mathbf{1,2}}, \mathbf{2}) = \tilde{G}(\Lambda_{\rm R}, \vec{g}, \mathbf{2}) \ , \tag{1.70}$$

and so indeed (1.67) holds.

This construction proves the equivalence: the renormalised Green's functions are completely insensitive to the arbitrariness in the flow equations.

General 'Power-Counting Renormalisable' Theories

The above framework now enables us to prove that *all* field theories are defined in the infrared by a fixed surface of dimension equal to the number of naïvely relevant operators.

We can see how the generalisation goes by considering first a theory with two scalar fields ϕ and χ. The vertex functions now have a label denoting how many legs there are of a given field. In particular, equation (1.22), c.f. Fig. 4, is replaced by Fig. 8. There are seven relevant operators, assuming a $\phi \to -\phi, \chi \to -\chi$ symmetry, namely kinetic and mass terms for both fields and three quartic couplings $\phi^4, \chi^4, \phi^2\chi^2$. Thus the index on the B-vertices now runs over seven values. (This restriction is not necessary, of course; we could perfectly well have no symmetry at all and have fifteen relevant operators.) The analogues of (1.28a) and (1.32) are lengthy to write down, even in the graphical notation, but they are completely straightforward.

Now imagine running through the entire analysis of Sections I.3 and I.4. At the scale Λ_{R} the relevant parameters are chosen to be

$$\rho^{\mathrm{R}}_{1\phi} = \rho^{\mathrm{R}}_{1\chi} = 0 \qquad (1.71a)$$

$$\rho^{\mathrm{R}}_{2\phi} = \rho^{\mathrm{R}}_{2\chi} = 0 \qquad (1.71b)$$

$$\rho^{\mathrm{R}}_{\phi^4} = \lambda^{\mathrm{R}} \qquad (1.71c)$$

$$\rho^{\mathrm{R}}_{\chi^4} = \kappa^{\mathrm{R}} \qquad (1.71d)$$

$$\rho^{\mathrm{R}}_{\phi^2\chi^2} = \pi^{\mathrm{R}} \;, \qquad (1.71e)$$

where the renormalised couplings may be considered either dependent or independent, but of course small. Once again we have the crucial properties that $A^{(0,0,0)}_{2m,2n} = 0$ and $A^{(r_1,r_2,r_3)}_{2m,2n} = 0$ for large enough m, n and so we can perform the inductive arguments successively in λ^{R} then κ^{R} then π^{R}, provided that the initial condition function satisfies the criteria mentioned above.

Now it is obvious that we can extend to any number of fields. There is no problem with global internal symmetries, since these are preserved by Λ-evolution. In this case each leg of the vertex functions can be interpreted as a representation, and the internal lines as a matrix. Down at the scale Λ_R only the coefficients of symmetric relevant operators need be specified. We do not have to confine ourselves to scalar fields; spinors and vectors are admissible since Λ-evolution also respects global euclidean invariance.

Finally I remark that the statistics of any field does not matter here. The difference between fermions and bosons lies only in the *signs* of various terms in $\Lambda\partial S/\partial\Lambda$ and these disappear in the procedure of taking bounds.

6. The Renormalisation of Composite Operators

To conclude the discussion of effective Lagrangian flow I describe how it can be used to deal with operator insertions. Apart from being a natural addendum, we'll need this information later on when we study Ward identities.

Consider the renormalisation of an operator O_d of dimension d. The partition function at Λ_0, with an insertion of O_d, is

$$Z_{op}[J,\Lambda_0] = \int D\phi\, O_d(x) \mathrm{exp}\, \frac{1}{\hbar}\left\{S(\Lambda_0) - \frac{1}{2}\int \frac{d^4p}{(2\pi)^4} P^{-1}(\Lambda_0)\phi_p\phi_{-p} + \int \frac{d^4p}{(2\pi)^4} J_p\phi_{-p}\right\} \tag{1.72}$$

but this is a difficult object to allow to flow down in scale. Instead, we couple the operator to a source $\epsilon(x)$, of dimension $4-d$, and raise it into the exponent. Thus we consider a new interaction S_{Total} which has an expansion in powers of $\epsilon(x)$:

$$S_{\mathrm{Total}}(\Lambda) = S_{(0)}(\Lambda) + \epsilon S_{(1)}(\Lambda) + \epsilon^2 S_{(2)}(\Lambda) + \dots\ , \tag{1.73}$$

where $S_{(0)}$ is just our original interaction S, since it remains uncontaminated by the insertion, and $S_{(1)}$ is the 'running operator insertion.' We expect, from the dominance of canonical dimensions, $S_{(1)}$ to flow in the infrared toward a fixed surface defined by *its* relevant operators, which are those of the same symmetry as O_d and having dimension $\leq d$. One might think that we have in fact proved this already, just by substituting S_{Total} for S in the preceding sections, but actually we have not, because S_{Total} does not satisfy all the conditions listed. In particular it does not vanish at zeroth order in λ^{R}.

However there's not much more work, and we follow exactly the same tactics as before. If ρ_b and ξ_α are respectively the relevant parameters for S and $S_{(1)}$, which I choose for simplicity to have initial conditions $\rho_b = \rho_b^0, \xi_\alpha = \xi_\alpha^0$ and all irrelevant parameters zero, then define

$$\begin{aligned} V_{(1)}(\Lambda) &= \Lambda_0 \frac{d}{d\Lambda_0} S_{(1)}(\Lambda)\big|_{\rho_b,\xi_\alpha \text{ fixed}} \\ &= \Lambda_0 \frac{\partial S_{(1)}}{\partial \Lambda_0} - \sum_b C_b \Lambda_0 \frac{\partial \rho_b}{\partial \Lambda_0} - \sum_\alpha D_\alpha \Lambda_0 \frac{\partial \xi_\alpha}{\partial \Lambda_0} , \end{aligned} \tag{1.74}$$

where

$$C_b = \sum_a \left[\sum_{\alpha,\beta} \frac{\partial S_{(1)}}{\partial \xi_\alpha^0} \left(\frac{\partial \xi}{\partial \xi^0}\right)^{-1}_{\alpha\beta} \frac{\partial \xi_\beta}{\partial \rho_a^0} - \frac{\partial S_{(1)}}{\partial \rho_a^0} \right] \left(\frac{\partial \rho}{\partial \rho^0}\right)^{-1}_{ab} \tag{1.75a}$$

$$D_\alpha = \sum_\beta \frac{\partial S_{(1)}}{\partial \xi_\beta^0} \left(\frac{\partial \xi}{\partial \xi_0}\right)^{-1}_{\beta\alpha} . \tag{1.75b}$$

$V_{(1)}(\Lambda)$ is the quantity which we wish to show is driven to zero as $\Lambda_0 \to \infty$. Using the equation of motion (1.20c) now interpreted for S_{Total} we find

$$\Lambda \frac{\partial S_{(1)}}{\partial \Lambda} = M[S_{(1)}] \tag{1.76}$$

$$\Lambda \frac{\partial V_{(1)}}{\partial \Lambda} = M[V_{(1)}] + \sum_b C_b M_b[V] - \sum_\alpha D_\alpha M_\alpha[V_{(1)}] - N[S_{(1)}; V] + \sum_\alpha D_\alpha N_\alpha[S_{(1)}; V] \tag{1.77}$$

$$\Lambda \frac{\partial C_b}{\partial \Lambda} = M[C_b] - \sum_a C_a M_a[B_b] - \sum_\alpha D_\alpha M_\alpha[C_b] + N[S_{(1)}; B_b] - \sum_\alpha D_\alpha N_\alpha[S_{(1)}; B_b] \tag{1.78}$$

$$\Lambda \frac{\partial D_\alpha}{\partial \Lambda} = M[D_\alpha] - \sum_\beta D_\beta M_\beta[D_\alpha] , \tag{1.79}$$

where $M[\cdot]$ and $M_b[\cdot]$ are as defined in (1.28b), and $M_\alpha[V_{(1)}]$ are the coefficients of the $S_{(1)}$-relevant operators in $M[V_{(1)}]$. The new functional operator $N[S_{(1)};X]$ is given by

$$N[S_{(1)};X] \equiv \int \frac{d^4q}{(2\pi\Lambda)^4}\Lambda^2\, Q\, \frac{\delta X}{\delta\phi_q}\frac{\delta S_{(1)}}{\delta\phi_{-q}} \ , \tag{1.80}$$

and $N_\alpha[S_{(1)};X]$ is the analogous restriction.

Equations (1.76)-(1.79) are similar in nature to those governing the V and B vertices, (1.28a) and (1.32). The vertices of type $S_{(1)}$, D and C can sequentially be bounded order by order in λ^{R} with the same bound as on the A and B vertices, and then $V_{(1)}$ can be bounded identically to V. Note that there is *no* obstruction here due to the non-vanishing of $S^{(0)}_{(1)}$.

As an example, in $\lambda\phi^4$ theory consider the renormalisation of $\phi^4(x)$. There are five relevant operators, namely $\phi^4, \phi\partial^2\phi, \partial_\mu\phi\partial_\mu\phi, \phi^2$ and $\mathbf{1}$, and so at Λ_0 the insertion is generically of the form

$$\phi^4_{\mathrm{ren}}(x) \;=\; \xi^0_1\,\phi^4(x) + \xi^0_2\,\phi\partial^2\phi(x) + \xi^0_3\,\partial_\mu\phi\partial_\mu\phi(x) + \xi^0_4\Lambda_0^2\,\phi^2(x) + \xi^0_5\Lambda_0^4\,\mathbf{1} \ , \tag{1.81}$$

where the renormalisation constants ξ^0_i are of order $P(\log\Lambda_0/\Lambda_{\mathrm{R}})$. This phenomenon, that the renormalisation of $\phi^4(x)$ is not simply multiplicative, is known as 'operator mixing.'

7. Summary

The method of effective Lagrangians makes the theory of renormalisation very clear, in fact it reduces it to dimensional analysis. The crucial idea is that the continuum limit $\Lambda_0 = \infty$ need not actually be taken.

Once Λ_0 is considered to be finite, the field theory can have both small bare couplings and small renormalised couplings, so that in such a regime the perturbation series is well defined. Since the theory is kept arbitrarily close to being free, the anomalous dimensions of all operators are small. Thus the splitting of operators into 'relevant' and 'irrelevant' can truly be decided naïvely, that is by the canonical, or free-field, dimensions, and this is the key point. Polchinski's proof of these statements is appropriately simple. The essential quality that a finite cutoff Λ_0 introduces is that all momentum integrals can be trivially estimated, and these estimates are indeed enough. Nothing more than this enters the discussion; we neither cite Weinberg's theorem nor combinatorial folk-lore! For reference, I list the main results in perturbation theory established in this chapter.

(i) There exists a 'renormalised interaction at Λ_0,' $S(\Lambda_0)$, which yields Λ_0-finite Green's functions. In general it includes all the naïvely relevant operators, with coefficients given by Λ_0 raised to the canonical power multiplied by a polynomial of $\log \Lambda_0/\Lambda_R$.

(ii) There are many different renormalised interactions at Λ_0 which yield the same renormalised Green's functions. These correspond to the different choices of initial condition functions.

(iii) In the usual language of 'subtracting infinities,' suppose an interaction at Λ_0 has been constructed, as a power series in $\hbar$, to make the 2-pt and 4-pt Green's functions Λ_0-finite up to nth loop order. Then automatically *all* the Green's functions are n-loop finite.

To see this, imagine flowing down in scale this 'partially renormalised' interaction. Down at Λ_R the theory *still lies on the fixed surface*, since the flow is attracted to it regardless of starting point, it's just that the relevant parameters are not now Λ_0-finite. Shifting the relevant parameters is all that's required to achieve full renormalisation.

(iv) The properties (i), (ii), and (iii) also hold for composite operator renormalisation. In general an operator will mix with all those of the same symmetry with the same or lower canonical dimension.

CHAPTER II
NON-CHIRAL GAUGE THEORIES

For a field theory to be physically acceptable the S-matrix must be unitary. This is no problem for scalars and spinors, provided they obey the spin-statistics rule, but for vector models this is a very serious constraint [1]. Consider first massive vectors. The propagator is of the form

$$P_{\mu\nu}(k) = \frac{\rho_{\mu\nu}(k)}{k^2 + m^2} , \tag{2.1}$$

where $\rho_{\mu\nu}$ is a Lorentz tensor. In the rest frame of the particle there are the three states, corresponding to its having spin one; thus $\rho_{\mu\nu}(k^2 = -m^2) = \mathrm{diag}(0,1,1,1)$, and so a possible candidate for $P_{\mu\nu}(k)$ is

$$P_{\mu\nu}(k) = \left[g_{\mu\nu} + \frac{k_\mu k_\nu}{\alpha m^2 - (1-\alpha)k^2} \right] \frac{1}{k^2 + m^2} . \tag{2.2}$$

However for $\alpha = 1$ the propagator behaves as k^0, and not k^{-2}, for large k^2. This in fact causes a breakdown of the renormalisation theorems in Chapter I, because the estimates (1.36) which were crucial to the argument don't hold. Moreover, for $\alpha \neq 1$ there is an extra pole which corresponds to a scalar particle of mass-square $\frac{-\alpha m^2}{(1-\alpha)}$ with residue $\frac{\alpha}{(1-\alpha)}$. If the mass-square is positive the residue is negative, so we have a 'ghost,' and if the residue is positive the mass-square is negative, which represents a 'tachyon'. The candidate propagator (2.2) is therefore *prima facie* unacceptable.

It's easy to see that in fact there is no propagator which avoids these problems. In general, we write

$$P_{\mu\nu}(k) = A(k^2) g_{\mu\nu} + B(k^2) k_\mu k_\nu , \tag{2.3}$$

where we demand $P_{\mu\nu}$ to behave as k^{-2} as $k^2 \to \infty$, and its inverse to be a polynomial in momenta (so that it can be derived from a local action). This implies

$$A(k^2) \sim k^{-2} \tag{2.4a}$$

$$B(k^2) \sim k^{-4} \,. \tag{2.4b}$$

Taking the poles in B all to have positive mass-square, we find that B is essentially a sum of terms of the form

$$\begin{aligned} &\frac{\alpha_{ij}}{(k^2+m_i^2)(k^2+m_j^2)} \\ &= \frac{\alpha_{ij}}{m_j^2 - m_i^2}\left(\frac{1}{k^2+m_i^2} - \frac{1}{k^2+m_j^2}\right) \,, \end{aligned} \tag{2.5}$$

and so the sum of the residues is zero. Thus we're obliged to have ghosts unless $B = 0$. But if B vanishes then there is a ghost from the timelike component of $g_{\mu\nu}$, and so we're stuck.

Of course, this had to happen. In a manifestly Lorentz-covariant formalism we must only use representations of the Lorentz group - but a 4-vector has one more degree of freedom than we require, and due to the signature of the spacetime this appears as a ghost or a tachyon. (For the exceptional case of $\alpha = 1$ above, the degree of freedom is removed by setting $\partial_\mu A_\mu = 0$, but this is generically spoiled by local interactions.) For massless vectors the situation is at least as bad, for now there are only two physical states. We really must keep unitarity, locality, and the existence of the vacuum (the absence of tachyons). Demanding, for simplicity, that also the Lorentz invariance is indeed manifest, then we *must* deal with ghosts and/or non-manifest renormalisability.

The only known acceptable solutions to the dilemma are gauge theories, for which one proves that the S-matrix is indeed unitary when restricted to a 'physical' subspace of the full Hilbert space. This 'decoupling' of the physical and unphysical parts crucially relies on the Ward identities, and it is to these that I now devote this chapter.

1.The Naïve Ward identities

The classical Lagrangian for pure Yang-Mills with a Lorentz-covariant gauge fixing is

$$L = -\frac{1}{4}F^a_{\mu\nu}F^a_{\mu\nu} + \frac{1}{2\alpha}\left[f(\frac{\partial^2}{m^2})\partial_\mu A^a_\mu\right]^2 + i\partial_\mu\bar{\eta}^a(D_\mu\eta)^a \ , \tag{2.6}$$

where $D_\mu = \partial_\mu - ig[A_\mu,\ ]$, $F_{\mu\nu} = -\frac{1}{ig}[D_\mu, D_\nu]$, and η is the Faddeev-Popov ghost. The weight function $f(\cdot)$ is arbitrary at this point, and m is just some (physically irrelevant) scale. The BRS variations [12] which leave the action invariant are

$$\delta A^a_\mu = -\frac{1}{g}(D_\mu\eta)^a\,\varepsilon \tag{2.7a}$$

$$\delta\eta^a = \frac{1}{2}f^{abc}\eta^b\eta^c\,\varepsilon \tag{2.7b}$$

$$\delta\bar{\eta}^a = \frac{-i}{\alpha g}f^2(\frac{\partial^2}{m^2})\partial_\mu A^a_\mu\,\varepsilon \ , \tag{2.7c}$$

where ε is a Grassman parameter, and it is this invariance which we wish to reflect at the quantum level.

Naïvely, we write down the functional integral, $Z[J_\mu,\bar{\zeta},\zeta]$, where the sources J_μ, $\bar{\zeta}$ and ζ are coupled respectively to A_μ,η and $\bar{\eta}$, and perform a BRS transformation on the integration variables. This generates an infinite set of identities, from which we derive the canonical ones by differentiating with respect to ζ and setting ζ and $\bar{\zeta}$ to zero. Taking the Jacobian to be unity, this procedure yields

$$0 = \int DA_\mu D\bar{\eta}D\eta\left[-\bar{\eta}^a(y)\int d^4x J^b_\mu(D_\mu\eta)^b(x) + \frac{i}{\alpha}f^2(\frac{\partial^2}{m^2})\partial_\mu A^a_\mu(y)\right]\exp\int(-\frac{L}{\hbar}+J_\mu A_\mu). \tag{2.8}$$

We can make a simplification of this *formal* identity. Since the integrand of Z is an exponential linear in $\bar{\eta}$, then performing the integral over $\bar{\eta}$ results in a 'delta functional.' Insertion into Z of that operator which is the argument of the delta-functional would thus give zero, and so we have

$$0 = \int DA_\mu D\bar{\eta} D\eta \left[-\bar{\eta}^a(y)\partial_\mu (D_\mu \eta)^b(x) + i\hbar \delta^{ab}\delta^4(x-y) \right] \exp\int(-\frac{L}{\hbar} + J_\mu A_\mu). \quad (2.9)$$

Now splitting up J_μ into transverse and longitudinal parts, we can combine (2.8) and (2.9) to get

$$0 = igf^{bcd}\int d^4x J_\mu^{b\mathrm{tr}}(x)\frac{\delta G^{ad}(y,x)}{\delta J_\mu^c(x)} + \hbar\partial^{-2}\partial_\mu J_\mu^a(y) + \frac{1}{\alpha}f^2(\frac{\partial^2}{m^2})\partial_\mu^y \frac{\delta W}{\delta J_\mu^a(y)}, \quad (2.10)$$

where $W = \log Z$, $G^{ad}(y,x) = \langle \bar{\eta}^a(y)\eta^d(x)\rangle_J$, and A_μ^a has been replaced by $\delta/\delta J_\mu^a$.

It is worthwhile to pause and ask what this complicated looking identity (2.10) really means. The identity relates Green's functions with a single longitudinal gauge boson to Green's functions with a single ghost insertion, and essentially it says that the unphysical contributions to any process from the longitudinal gauge bosons and the ghosts actually cancel. This is why the Ward identity is so important.

What we wish to do, then, is to properly define the terms in (2.10) via some regularisation scheme, and then show that a corresponding identity holds as a statement about renormalised quantities. Let us simply accept for now that this would indeed be sufficient to prove unitarity.

There are many conceivable methods for regularisation, and a prudent choice would be one for which proving renormalised equations is simple. The cleanest route from regularisation to renormalisation follows the flow of effective Lagrangians, where the regulator is essentially a mass scale, and so it is this technique that I'd like to use.

Now as I've described it so far, the running mass scale has been a momentum cutoff, but it's clear from the physical idea of the flow that any mass scale regulator should work. I look for an alternative to the momentum cutoff, not because there's anything intrinsically wrong with it, but rather for a technical reason: the breakage of BRS invariance that it induces persists to all loop orders.

What I mean by this is the following: suppose one tries to derive, as a first step, a regularised Ward identity, where no counterterms for the Λ_0-divergences are yet added. The momentum cutoff causes a breaking, and we must compensate for this by adding an explicitly gauge-non-invariant operator to the action. This operator, however, can only be written as an *infinite* series in $\hbar$, not a finite one.

This is bad because in our framework renormalisation theory is just dimensional analysis. What we expect from this is that the renormalised interaction at Λ_0 contains all the relevant couplings, and this result is obviously not enough to deduce a renormalised Ward identity. However, one would hope that together with an exact regularised Ward identity we could proceed. But any such prospective argument is *bound* to be inductive in the order of the couplings (or $\hbar$), and so it's disastrous if the input, the regularised Ward identity, requires an infinite expansion in $\hbar$ before we can begin.

In searching for a mass scale regulator which respects the regularised Ward identities, an immediate candidate to study is 'higher covariant derivatives.' However there are a couple of tricky points with this regulator, which, remarkably, have eluded many authors.

First and foremost, the regulator must actually regulate, that is, render all the Feynman diagrams finite. Previous work [11] has failed right here, due to a misunderstanding of the fact that the inclusion of higher covariant derivatives into the action only *mollifies* the structure of the $n \geq 2$ -loop divergences rather than dismantles it. The combination of higher covariant derivatives with a one-loop regulator (like Pauli-Villars particles) must in fact be constructed with some care, and I discuss this in the next section.

Secondly, the regulating particles require the presence of a pre-regulator, in order to make sense of adding Feynman diagrams together. The regulator would then be said to be successful if the Green's functions are finite as the pre-regulator is removed. This is not mere pedantry. Let's say that the pre-regulator is a momentum cutoff: somewhat insidiously the cutoff will contribute non-vanishing terms in the derivation of the regularised Ward identity, and we must make sure that these terms can be controlled.

2. The Method of Higher Covariant Derivatives

Consider the following manifestly Lorentz invariant, higher derivative Lagrangian of gauge and ghost fields,

$$L = -\frac{1}{4}F^a_{\mu\nu}\left[1+(\frac{D^2}{M_0^2})^{2n}\right]F^a_{\mu\nu} + \frac{1}{2\alpha}\left[f(\frac{\partial^2}{m^2})\partial_\mu A^a_\mu\right]^2 + i\partial_\mu\bar{\eta}^a(D_\mu\eta)^a \ . \qquad (2.11)$$

For simplicity I shall choose the 'gauge-weighting' function $f(\cdot)$ to be given by

$$f(x) = 1 + x^{n+p} \ . \qquad (2.12)$$

We take $p \geq 0$ so that the longitudinal part of the propagator decays rapidly compared to the transverse part for large momentum, and the large k behaviour of the propagator is $k^{-(4n+2)}$. In this case I say that the gauge field is *of type* $4n+2$. The large k behaviour for the ghosts is k^{-2}, and so the ghost field is of type 2.

I note here that this particular gauge, call it the 'transverse Lorentz gauge,' is chosen to make the discussion of regularisation and renormalisation the most tractable; we will, in the end, generalise to other gauges.

Now covariantly couple the gauge bosons in (2.11) to regulator fields Φ_t of type t, whose dimensions are again made correct by powers of M_0. What we wish to do is find a new action for which, by some clever arrangement, the Green's functions are finite. The higher covariant regulator, for our special gauge, would then be in place with M_0 as the regulation scale.

To analyse the possibility of this, we compute the *superficial* degree of divergence, D_Γ, of the Feynman diagrams. (The subscript refers to the fact that we're measuring the divergence with respect to the pre-regulating cutoff, Γ.) The leading momentum behaviour of the propagators and vertices in the theory are shown in Fig. 9, where for simplicity I ignore self-couplings of the regulator fields, e.g. Φ_t^4, since such couplings don't materially affect the results.

Denoting by V_j, G, and $U_{t,i}$ the number of those vertices in Fig. 9 appearing in a given Feynman diagram, the superficial degree is

$$D_\Gamma = 4L - (4n+2)I - I_g - \sum_t tI_t + \sum_j (4n+4-j)V_j + G + \sum_{t,i}(t-i)U_{t,i} \ , \quad (2.13)$$

where I, I_g, I_t, are respectively the number of gauge, ghost, and regulator lines, and L is the number of loops. The topological relations are

$$L = I + I_g + \sum_t I_t + 1 - \sum_j V_j - G - \sum_{t,i} U_{t,i} \quad (2.14)$$

$$E = -2I + \sum_j jV_j + G + \sum_{t,i} iU_{t,i} \quad (2.15a)$$

$$E_g = -2I_g + 2G \quad (2.15b)$$

$$E_t = -2I_t + 2\sum_i U_{t,i} \ , \quad (2.15c)$$

where E, E_g, E_t are the numbers of external lines. The arithmetic gives us,

$$D_\Gamma = 4 - 4n(L-1) - E - (2n+1)E_g - \sum_t (2n+2-\frac{1}{2}t)E_t \ . \quad (2.16)$$

We see that by choosing n large enough compared to the largest value of t in the system, the superficial degree is non-negative only for $L = 1$, $E = 2, 3, 4$ and $E_g = E_t = 0$. ($E = 1$ is ignorable, for the tadpole graphs vanish by symmetric integration, which is allowed due to the presence of the pre-regulating momentum cutoff Γ.) Thus the only superficial Γ-divergences come from one-loop graphs with 2,3 or 4 external gauge legs. This is a nice feature, for with this choice there is only a small number of divergences to take care of. Alternatively, we could include regulators of the same type as the gauge field, whereupon there would also be one-loop divergences from graphs with 2 or 4 external regulator lines. Either way, we must find an example of an action where all these divergences do in fact cancel.

Now the *raison d'être* for the inclusion of higher derivatives is to ensure that no new superficial divergences arise in loop orders ≥ 2, for from (2.16) this beneficial property only obtains if the type of the gauge field is at least six . However by covariance these must be accompanied by higher couplings, which make the calculation of even the one-loop graphs frightfully tedious. It is tempting therefore to try to use regulator fields with the identical propagators and vertices to those of the gauge field, so that it is simple to relate the one-loop divergences from the regulator and gauge cycles. One would then hope to find an algebraic condition, analogous to the Pauli-Villars condition, for cancellation. Such Pauli-Villars particles would have an action of the form

$$S_{PV} = \frac{1}{2}\int d^4x d^4y \Phi^a_\mu(x)\left[\frac{\delta^2 S_{YM}}{\delta A^a_\mu(x)\delta A^b_\nu(y)} + cM_0^2\delta^4(x-y)\delta_{\mu\nu}\delta^{ab}\right]\Phi^b_\nu(y)\ , \qquad (2.17)$$

where S_{YM} is the gauge field part of the action in (2.11).

Now as noted above, for a Pauli-Villars regulator field there are divergent diagrams where it appears as external. To see the relevance of this, note that the three diagrams of Fig. 10 have the same divergence, up to a sign.

If $10(i)$ and $10(ii)$ together have canceling divergence, which is trivial to arrange, then one is left with $10(iii)$ uncanceled. This will rear its ugly head in, for example, the two-loop graphs of Fig. 11, which shows the contributions of the same topology, from the gauge field and Pauli-Villars field to the gauge two-point function. The sum is not finite, unfortunately: if one considers the left-hand sub-graph then $11(i)$ and $11(ii)$ sum to a finite result but $11(iii)$ and $11(iv)$ do not. This is known colloquially as the problem of 'overlapping divergences'.

We cannot simply ignore the one-loop graphs with external regulator lines. These must be made finite, too. However, it's not so easy. If we add in another Pauli-Villars field coupled to the first, then we have the one-loop graph where this new field appears as external. So we add another, and another, and the need for ever more regulator fields causes them to proliferate *ad infinitum.*

Abandoning this approach, we go back to considering the types of the regulator fields all to be smaller than that of the gauge field. Then there is no problem with the graphs with external regulators, but the price we pay is that we cannot keep both the properties of manifest gauge invariance *and* the simplicity of the divergence cancellations.

One way to arrange for the cancellation conditions to be relatively trivial is to take the Pauli-Villars action (2.17) and remove some powers of ordinary derivatives even-handedly from the inverse propagators and vertices. This can retain sufficiently well the one-loop divergence structure, while also reducing the type of the regulating fields. But clearly this breaks the gauge invariance, and so the main issue with this tactic is how to *recover* the Ward identity. It is a scheme based on these lines that I discuss in the chapters below on chiral gauge theories and supersymmetry.

Here in this discussion, however, we shall maintain gauge invariance, and accept the fact that cancellation will be tricky to exhibit. In many ways, actually, this procedure is conceptually easier.

The technique is essentially to write down all the conceivable invariant actions for regulator fields of type $t < 4n + 2$ and compute each one's contribution to the various one-loop Γ-divergences. These actions will contain various parameters, and what we must show is that there is enough play in them for cancellation to occur. To see the likelihood of success, let us make a count.

There are ten topologically distinct, divergent one-loop graphs, and they are shown in Fig. 12. Fortunately, we do not have to cancel the divergences in each graph separately. Simply canceling the divergences in the one-loop Green's functions will be enough. This is because all Green's functions can be generated from their skeleton expansions, which are made trivial here due to the fact that the divergences are either 'disjoint' or 'nested'.

I list here the various possible divergences that can occur, consistent with Lorentz invariance and Bose symmetry, in the Green's functions. For the 2-point function, c.f. Fig. 12 $(i), (ii)$, we may have both quadratic and logarithmic Γ-divergences,

$$1. \quad g^2 \int_\Gamma \frac{d^4k}{(2\pi)^4} k^2 g_{\mu\nu}\, \delta^{ab} \tag{2.18a}$$

$$2. \quad g^2 \int_\Gamma \frac{d^4k}{(2\pi)^4} (p^2 g_{\mu\nu} - p_\mu p_\nu)\, \delta^{ab} \tag{2.18b}$$

$$3. \quad g^2 \int_\Gamma \frac{d^4k}{(2\pi)^4} (p^2 g_{\mu\nu} + 2p_\mu p_\nu)\, \delta^{ab} \tag{2.18c}$$

For the 3-point function, c.f. Fig. 12 (iii),(iv),(v), we may only have

$$4. \quad g^3 \int_\Gamma \frac{d^4k}{(2\pi)^4} f^{abc}\ [g_{\mu\nu}(p-q)_\rho + g_{\mu\rho}(r-p)_\nu + g_{\nu\rho}(q-r)_\mu] \tag{2.18d}$$

and for the 4-point function, c.f. Fig. 12 (vi)-(x), we may have

$$5. \quad g^4 \int_\Gamma \frac{d^4k}{(2\pi)^4} \mathrm{Str}\, T^a T^b T^c T^d\, (g_{\mu\nu} g_{\rho\sigma} + g_{\mu\rho} g_{\nu\sigma} + g_{\mu\sigma} g_{\nu\rho}) \tag{2.18e}$$

$$\begin{aligned} 6. \quad g^4 \int_\Gamma \frac{d^4k}{(2\pi)^4} \big[& f^{abe} f^{cde} (g_{\mu\sigma} g_{\nu\rho} - g_{\mu\rho} g_{\nu\sigma}) + f^{ade} f^{cbe} (g_{\mu\nu} g_{rho\sigma} - g_{\mu\rho} g_{\nu\sigma}) \\ & + f^{ace} f^{bde} (g_{\mu\sigma} g_{\nu\rho} - g_{\mu\nu} g_{\rho\sigma}) \big] \, . \end{aligned} \tag{2.18f}$$

Now, to estimate the freedom in the parameters, consider for the moment a real scalar regulator Φ of type $2r$ in representation $\mathbf{R}$ of the gauge group. The most general Lagrangian quadratic in Φ is

$$-\frac{1}{4}\Phi\,[D^4 + \alpha g D_\mu F_{\mu\nu} D_\nu + \beta g [D_\mu, F_{\mu\nu}] D_\nu + \gamma g^2 F_{\mu\nu} F_{\mu\nu}]\, \frac{D^{2r-4}}{M_0^{2r-2}} \Phi + \text{ h.c. } + \dots \tag{2.19}$$

where α, β, and γ are arbitrary real constants and the ellipses denote operators which cause no contribution to the one-loop divergences, e.g., $\Phi F^r \Phi$.

The Feynman rules are obtained from (2.19) by functional differentiation with respect to A^a_μ. The vertices thus depend polynomially on the type r, and since we're only interested in a maximum number of *four* external gauge lines the maximum power of r that appears is four. In computing the one-loop divergences, therefore, their coefficients must be of the form

$$\{1, r, r^2, r^3, r^4\} \times \text{powers of } \alpha,\ \beta \,\text{and}\, \gamma$$

and it turns out (see Appendix 1) that the combinations that actually appear are r, r^2, r^3, r^4 and $(\alpha + 2\gamma)$. Thus the inclusion of several real scalars of various types and representations provides for us five tunable parameters, namely,

$$\sum r, \sum r^2, \sum r^3, \sum r^4 \text{ and } \sum(\alpha + 2\gamma)\ , \tag{2.20}$$

where in this notation the sums imply a plus sign for bosonic fields and a minus sign for fermions. So if we have $n^{(b)}_{r,\mathbf{R}}$ bosons of type $2r$ in representation $\mathbf{R}$, and $n^{(f)}_{r,\mathbf{R}}$ fermions, then the symbol $\sum r$ denotes $\sum_{r,\mathbf{R}} (n^{(b)}_{r,\mathbf{R}} - n^{(f)}_{r,\mathbf{R}})\, r$.

Of these five parameters the first four are constrained to be integers while the other one can take on any real value, and clearly the latter is the more useful. To increase this number, as we must in fact do, we can add in vector fields and spinors.

It's not obvious that our tactic will work, but indeed we do wind up with enough parameters, as I prove in Appendix 1. According to the results presented there, the following Lagrangian represents a higher covariant derivative regulator:

$$L_{\text{HCD}} = -\frac{1}{4} F^a_{\mu\nu} \left[1 + (\frac{D^2}{M_0^2})^{2n}\right] F^a_{\mu\nu} + \frac{1}{2\alpha}\left[f(\frac{\partial^2}{m^2})\partial_\mu A^a_\mu\right]^2 + i\partial_\mu \bar{\eta}^a (D_\mu \eta)^a$$

$$\sum -\frac{1}{4}\Phi \left[D^4 + \alpha_r g D_\mu F_{mu\nu} D_\nu + \beta_r g [D_\mu, F_{\mu\nu}] D_\nu + \gamma_r g^2 F_{\mu\nu} F_{\mu\nu}\right] \frac{D^{2r-4}}{M_0^{2r-2}} \Phi +$$

$$\sum -\frac{1}{2} \bar{V}_\mu [\{D^4 + \epsilon_r g D_\rho F_{\rho\sigma} D_\sigma + \eta_r g [D_\rho, F_{\rho\sigma}] D_\sigma + \kappa_r g^2 F_{\rho\sigma} F_{\rho\sigma}\} g_{\mu\nu} +$$

$$+\{(\lambda_r - 1)D_\mu D_\nu + \pi_r F_{\mu\nu}\}D^2]\frac{D^{2r-4}}{M_0^{2r-2}}V_\nu + \text{ h.c.} + \dots \ , \qquad (2.21)$$

where V_μ is a complex vector, and the 'regulator constants' $\alpha_r, \beta_r \dots \pi_r$ satisfy certain algebraic equations. The ellipses denote operators involving the regulator fields that don't contribute to regulation, namely operators of lower dimension or those containing too many powers of the field strength F.

One might wonder whether these constants depend on the gauge parameter α, or on the precise choice (2.12) of the gauge-weighting function. In fact there is no dependence, since an insertion of the longitudinal part of the gauge propagator automatically makes any one-loop graph Γ-finite.

Let me summarize the key ideas in the construction of the regulated Lagrangian (2.21). Higher derivatives are put in so that the superficial degree D_Γ is negative for Feynman diagrams with more than one loop. This requires using at least four extra derivatives *and* a special gauge choice where the longitudinal part of the propagator decays rapidly. Regulating particles are then chosen with a smaller 'type' (that is number of higher derivatives) than the gauge particles, in order that the graphs with external *regulator* lines converge. This surmounts the overlapping divergence problem, which is the main obstacle to the attempt [11] to regularise using ordinary Pauli-Villars particles.

With the new approach the only divergences come from the one-loop graphs with external *gauge* lines, and the question that we faced was whether there existed any such collection of regulating particles that could actually achieve cancellation. There being no principle one way or the other, this question was settled (in the affirmative!) by direct calculation. In the next section I use the regulated Lagrangian (2.21) to derive the regularised Ward identities. From there I shall construct the renormalised Ward identities, which are at the heart of the proof of unitarity.

3. The Regularised Ward Identity

As I mentioned at the end of Section II.1, the regularised Ward identities cannot be immediately written down just from the Lagrangian (2.21); the breakage due to the pre-regulator must be taken into account. For analytical simplicity, I select a *smooth* momentum cutoff, by including factors of $\left[1+(\partial^2/\Gamma^2)^{\mathbf{P}}\right]$ in the kinetic terms. To save writing out once again the rather lengthy list of operators in the Lagrangian, let me establish some notation. I say L_{HCD} is a sum of kinetic terms for each field and the rest, namely the interaction, S_{HCD}:

$$L_{\mathrm{HCD}} = -\frac{1}{2}A^a_\mu P^{-1}_{\mu\nu} A^a_\nu + i\bar{\eta}^a P^{-1}\eta^a - \sum \frac{1}{2}\Phi P^{-1}_{\Phi_r}\Phi - \sum \bar{V}_\mu (P_{V_r})^{-1}_{\mu\nu} V_\nu + S_{\mathrm{HCD}} \quad (2.22)$$

where the inverse propagators in k-space are given by

$$P^{-1}_{\mu\nu} = (k^2 g_{\mu\nu} - k_\mu k_\nu)(1+\frac{k^{4n}}{M_0^{4n}}) + \frac{1}{\alpha}f^2(\frac{k^2}{m^2})k_\mu k_\nu \quad (2.23a)$$

$$P^{-1} = k^2 \quad (2.23b)$$

$$P^{-1}_{\Phi_r} = \frac{k^{2r}}{M_0^{2r-2}} + \dots \quad (2.23c)$$

$$(P_{V_r})^{-1}_{\mu\nu} = \frac{k^{2r}}{M_0^{2r-2}}(g_{\mu\nu} + \lambda_r \frac{k_\mu k_\nu}{k^2}) + \dots \quad (2.23d)$$

Here the ellipses denote possible (irrelevant) terms of lower powers of k^2, in correspondence with those in (2.21). Now to form the fully regulated and pre-regulated Lagrangian, L_{pre}, I replace the inverse propagators (2.23) by,

$$P^{-1}_{\mu\nu}[\Gamma] = (k^2 g_{\mu\nu} - k_\mu k_\nu)\left[1+\frac{k^{4n}}{M_0^{4n}}(1+\frac{k^{2\mathbf{P}}}{\Gamma^{2\mathbf{P}}})\right] + \frac{1}{\alpha}f^2(\frac{k^2}{m^2})k_\mu k_\nu \quad (2.24a)$$

$$P^{-1}[\Gamma] = k^2(1+\frac{k^{2\mathbf{P}}}{\Gamma^{2\mathbf{P}}}) \quad (2.24b)$$

$$P^{-1}_{\Phi_r}[\Gamma] = \frac{k^{2r}}{M_0^{2r-2}}(1+\frac{k^{2\mathbf{P}}}{\Gamma^{2\mathbf{P}}}) + \dots \quad (2.24c)$$

$$(P_{V_r})^{-1}_{\mu\nu}[\Gamma] = \frac{k^{2r}}{M_0^{2r-2}}(g_{\mu\nu} + \lambda_r \frac{k_\mu k_\nu}{k^2})(1+\frac{k^{2\mathbf{P}}}{\Gamma^{2\mathbf{P}}}) + \dots \quad (2.24d)$$

and so the complete functional integral for our study is,

$$Z_{\text{pre}}[J_\mu, \zeta, \bar\zeta] = \int d\mu_{\text{pre}} \exp \int (J_\mu A_\mu + \bar\zeta\eta + \bar\eta\zeta) \tag{2.25a}$$

$$d\mu_{\text{pre}} = DA_\mu D\bar\eta D\eta \left(\prod D\Phi DV\right) \exp \int -\frac{L_{\text{pre}}}{\hbar}\,. \tag{2.25b}$$

Note that in (2.24a) the longitudinal part of $P^{-1}_{\mu\nu}[\Gamma]$ does not carry a Γ-cutoff factor; none is necessary, due to the large power of momentum already there. However, my choice is merely a convenience for the later arguments presented in Appendix 3, and one certainly could include the Γ-cutoff factor if desired. (See the comments after equation (2.30).)

To obtain the regularised Ward identities we perform the same manipulations on the functional integral as before: that is, a BRS transformation on the fields, followed by differentiation with respect to the ghost source ζ and setting ζ and $\bar\zeta$ to zero. I define the BRS variation of the regulator fields to be just a gauge variation, $\delta\Phi_j = i(T^a)_{jk}\eta^a\varepsilon$.

The Jacobian of the transformation is indeed unity (see Appendix 2), and so we have in correspondence to (2.8),

$$0 = \left\langle -\bar\eta^a(y)\int d^4x J^b_\mu (D_\mu\eta)^b(x) + \frac{i}{\alpha} f^2\left(\frac{\partial^2}{m^2}\right)\partial_\mu A^a_\mu(y) + \bar\eta^a(y)\int d^4x\,\delta L_{\text{pre}}\right\rangle_J \tag{2.26}$$

where δL_{pre} is the BRS variation of the pre-regulating terms (2.24), and, as a reminder, the function $f(\cdot)$ is given by the choice (2.12) of gauge-fixing term.

Now δL_{pre} contains precisely the same number of extra derivatives that is used in the propagator as the momentum cutoff, so we might worry that they are actually *infinite*, that is, not even pre-regulated. This would certainly invalidate equation (2.26), but if all are finite, then it stands as an exact identity.

To find out we compute, for each of the operator insertions in δL_{pre}, the superficial degree, D_∞^{op} , with which the inserted diagrams truly diverge. For the diagrams to be finite in this sense, that is, pre-regulated, we require $D_\infty^{\mathrm{op}} < 0$.

As a warm-up we check that the uninserted diagrams, for that matter, are pre-regulated. The same arithmetic as in the derivation of (2.16) yields

$$D_\infty = D_\Gamma - 2\mathbf{P}\left(I + I_g + \sum_t I_t\right) , \tag{2.27}$$

whereupon D_∞ is strictly negative if $\mathbf{P} > 2$. This also holds for the $A_\mu \eta$ insertion in the first term of (2.26).

Now the δL_{pre} insertions are shown in Fig. 13 with their leading momentum behaviour. For 13(i), 13(ii) and 13(iii), D_∞^{op} is indeed negative except for the tadpole graphs of Fig. 14, and these tadpoles vanish by tracelessness of the group generators. To be more precise, we *define* the graphs using the pre-pre-regulator Δ of Appendix 2, and then they become unambiguously zero. Fig. 14(iv) looks dangerous but in fact is not. The momentum at the vertex must be contracted with the longitudinal part of the gauge propagator, thus lowering the degree by $4p$ and making it negative. Thus for all terms in (2.26) we have $D_\infty^{\mathrm{op}} < 0$, and this completes the proof of its veracity.

The first two terms in (2.26) are in fact not only pre-regulated but also regulated, that is they are Γ-finite in the limit $\Gamma \gg M_0$. These are the 'good' Ward identity terms. However the contribution from δL_{pre}, written as an asymptotic expansion in Γ, has, by power counting, quadratic and logarithmic divergences, as well as finite parts and inverse powers. But since the RHS of (2.26) is identically zero for all finite Γ, the coefficients of the divergences must in fact vanish. Thus despite appearances, the cutoff can be taken to infinity.

The Γ-finite parts from δL_{pre} are the potential anomalies. We require them to be zero. Actually what I shall show, in Appendix 3, is that they can be *canceled* by adding to the Lagrangian a gauge-non-invariant, dimension four, local polynomial of gA_μ and derivatives. Moreover, this compensating operator is *linear* in $\hbar$, unlike the infinite series in $\hbar$ required when using a pure momentum cutoff regulator. As I intimated before, this is going to be crucial in the discussion of renormalisation.

The reason why it's easy to prove that such a compensator exists is that only the one-loop diagrams of the operator insertions need be considered (this is the magic of higher covariant derivatives). Even then the finite parts need not actually be calculated, for they satisfy the Wess-Zumino consistency condition[f4] and can thus be classified by some elementary algebra.

Repeating the steps to (2.26) using the improved Lagrangian L_{reg}, related to L_{pre} by the addition of the compensating operator, we derive

$$0 = \int d\mu_{\text{reg}} \left[-\bar{\eta}(y) \int d^4x J^b_\mu (D_\mu \eta)^b + \frac{i}{\alpha} f^2 (\frac{\partial^2}{m^2}) \partial_\mu A^a_\mu(y) \right] \exp \int J_\mu A_\mu + O(\frac{1}{\Gamma^2}) \tag{2.28a}$$

$$d\mu_{\text{reg}} = DA_\mu D\bar{\eta} D\eta (\prod D\Phi DV) \exp \int -\frac{L_{\text{reg}}}{\hbar} \; . \tag{2.28b}$$

Corresponding to the formal identity (2.9), we also have

$$0 = \int d\mu_{\text{reg}} \left[-\bar{\eta}^a(y) \partial_\mu (D_\mu \eta)^b(x) - \bar{\eta}^a(y) \partial^2 (\frac{\partial^2}{\Gamma^2})^{\mathbf{P}} \eta^b(x) + i\hbar \delta^{ab} \delta^4(x-y) \right] \exp \int J_\mu A_\mu \; , \tag{2.29}$$

where the extra term from the momentum cutoff vanishes as Γ is taken to infinity. Combining these two, we finally obtain the regularised Ward identity, c.f. equation

f_4 Using the pre-regulator given in the text, this is not strictly true : there are some terms in δL_{pre} which do not satisfy WZ. However, if I take the scale of momentum cutoff for the ghosts, call it Ω, to be different from that of the gauge and regulator particles, Γ, then the troublesome terms vanish in the limit $\Gamma \gg \Omega \gg M_0$. This is explained in Appendix 3. Note that this refinement does not affect the rest of the discussion.

(2.10):

$$0 = igf^{bcd}\int d^4x J^{btr}_\mu(x)\frac{\delta G^{ad}_{\rm reg}(y,x)}{\delta J^c_\mu} + \hbar\partial^{-2}\partial_\mu J^a_\mu(y) + \frac{1}{\alpha}f^2(\frac{\partial^2}{m^2})\partial^y_\mu\frac{\delta W_{\rm reg}}{\delta J^a_\mu(y)} + O(\frac{1}{\Gamma^2}) \ , \tag{2.30}$$

where $W_{\rm reg} = \log Z_{\rm reg}$ and $G^{ad}_{\rm reg}(y,x) = \langle\bar\eta^a(y)\eta^d(x)\rangle^{\rm reg}_J$.

Note that while (2.30) was derived using a particular smooth pre-regulator Γ, this was only for simplicity of the analysis of the breaking terms $\delta L_{\rm pre}$, given in Appendix 3. If Γ is now replaced by *any* other momentum cutoff at, say, Λ, then the Green's functions in (2.30) only suffer a change of order,

$$O(\frac{M_0}{\Gamma})^2 + O(\frac{M_0}{\Lambda})^2 \ , \tag{2.31}$$

since the Feynman integrals are really regulated at M_0. This becomes negligible in the limit $\Gamma, \Lambda \gg M_0$, and so our specification of (2.24) was really only a technical device. The regularised Ward identity holds no matter what is the nature of the pre-regulator.

In order now to address renormalisation, we rewrite these results in the language of Chapter I. The Lagrangian $L_{\rm reg}$ is the sum of a *propagating part* and an *interaction*, $S_{\rm reg}$, where the inverse propagators are given in (2.24), and the interaction is the sum of the 'higher covariant derivative' interaction, $S_{\rm HCD}$, and the compensating operator. Let me write this latter as

$$\begin{aligned} S_{\rm reg} \ = & -\frac{1}{2}gf^{abc}(\partial_\mu A_\nu - \partial_\nu A_\mu)^a A^b_\mu A^c_\nu + \frac{1}{4}g^2 f^{abc}f^{ade}A^b_\mu A^c_\nu A^d_\mu A^e_\nu \\ & + gf^{abc}\partial_\mu\bar\eta^a A^b_\mu \eta^c + S_R[A_\mu, \{\Phi\}, M_0, g] + \hbar P_4[gA_\mu] \ , \end{aligned} \tag{2.32}$$

where S_R contains the higher derivative parts and the regulator field interactions in $S_{\rm HCD}$, and $\hbar P_4$ is the compensating operator.

Now consider a new interaction, $\tilde{S}_{\rm reg}$, related to $S_{\rm reg}$ by a change in the coupling constant and wavefunction normalisation. We have

$$\tilde{S}_{\rm reg} = -\frac{1}{4}(Z_\gamma - 1)(\partial_\mu A^a_\nu - \partial_\nu A^a_\mu)^2 - \frac{1}{2}Z_\gamma \tilde{g} f^{abc}(\partial_\mu A_\nu - \partial_\nu A_\mu)^a A^b_\mu A^c_\nu$$

$$+ \frac{1}{4}Z_\gamma \tilde{g}^2 f^{abc} f^{ade} A^b_\mu A^c_\nu A^d_\mu A^e_\nu + i(Z_\eta - 1)\partial_\mu \bar{\eta}^a \partial_\mu \eta^a$$

$$+ Z_\eta \tilde{g} f^{abc} \partial_\mu \bar{\eta}^a A^b_\mu \eta^c + S_R[A_\mu, \{\Phi\}, M_0, \tilde{g}] + \hbar P_4[\tilde{g} A_\mu] \ , \tag{2.33}$$

where Z_γ, Z_η, and $\tilde{g}$ are constants. Clearly $\tilde{L}_{\rm reg}$ is regulated and pre-regulated in exactly the same way as is $L_{\rm reg}$, and so we can immediately write down the new regularised Ward identity,

$$0 = iZ_\eta \tilde{g} f^{bcd} \int d^4x J^{b{\rm tr}}_\mu \frac{\partial \tilde{G}^{ad}_{\rm reg}(y,x)}{\partial J^c_\mu(y)} + \hbar \partial^{-2} \partial_\mu J^a_\mu(y) + \frac{1}{\alpha} f^2(\frac{\partial^2}{m^2}) \partial_\mu \frac{\delta \tilde{W}_{\rm reg}}{\delta J^a_\mu(y)} + O(\frac{1}{\Gamma^2}) \ . \tag{2.34}$$

Taking Z_γ, Z_η, and $\tilde{g}$ to be functions of the regulation scale M_0, some reference scale $M_{\rm R}$, and a coupling $g_{\rm R}$, the question now to be asked is whether the constants can be chosen order by order in $\hbar$ such that the Green's functions have finite limits as M_0 tends to infinity. Equation (2.34) would then be the *renormalised* Ward identity.

Parenthetically, we might wonder why we'll need $\tilde{g}$ to be different from g. After all, the regulated action (2.21) is BRS invariant, so why should the counterterms needed for renormalisation not also be BRS invariant? Whence we'd only need Z_γ and Z_η.

The reason is that the regulated Ward identity (2.30) does not just relate Green's functions to themselves, as in an $O(N)$-sigma model for example, rather it relates Green's functions to operator insertions. All that it will guarantee, in fact, is the existence of a new, *renormalised* BRS invariance.

4. The Renormalised Ward Identity

The Polchinski analysis for renormalisation involved a *sharp* momentum-cutoff *regulator.* This sharpness was technically very useful, for it reduced all integrals to integrals over a finite domain, and therefore to be trivially estimable. However, in our method of higher covariant derivatives, the effective cutoff at the regulation scale M_0 is very smooth and, moreover, works by a 'conspiracy' of cancellations. Thus all the integrals in the flow equations have in fact an infinite range, or a range all the way up to the pre-regulator Γ, and the Polchinski scheme of estimation does not appear to work: in the simple bounds of Chapter I all information that M_0 is the true regulation scale is lost.

One might persevere with the method of effective Lagrangians by looking for a more intricate means to bound the flowing vertex functions, in particular one that takes the Γ-cancellations automatically into account. But this would be technically very difficult, and hence contrary to our goal of keeping the analysis easy to follow. The strategy that we're going to employ is to make maximum use of the regularised Ward identity. The crude Polchinski bounds will actually be enough, for this identity will characterise and restrict the possible counterterms.

In order to directly apply the Polchinski results, let us make a modification to our pre-regulation procedure. The smooth momentum cutoff Γ in (2.24) is replaced with a sharp momentum cutoff at NM_0. (We take N to be finite for the moment and then we shall investigate the limit $N \to \infty$.) Now as we let the regulation scale M_0 flow down to the much smaller scale M_{R}, the pre-regulator flows down with it to scale NM_{R}, (instead of remaining at some high value), and this protects all the integrals. Thus the inverse propagators in (2.24) now run, and at scale M become,

$$P^{-1}_{\mu\nu}(k,M,N) = \left[(k^2 g_{\mu\nu} - k_\mu k_\nu)(1+\frac{k^{4n}}{M^{4n}}) + \frac{1}{\alpha}k_\mu k_\nu f^2(\frac{k^2}{m^2})\right] K^{-1}(k^2/(NM)^2) \tag{2.35a}$$

$$P^{-1}(k,M,N) = k^2 K^{-1}(k^2/(NM)^2) \tag{2.35b}$$

$$P_{\Phi_r}^{-1}(k, M, N) = \left[\frac{k^{2r}}{M^{2r-2}} + \ldots\right] K^{-1}(k^2/(NM)^2) \tag{2.35c}$$

$$(P_{V_r})_{\mu\nu}^{-1}(k, M, N) = \left[\frac{k^{2r}}{M^{2r-2}}(g_{\mu\nu} + \lambda_r \frac{k_\mu k_\nu}{k^2}) + \ldots\right] K^{-1}(k^2/(NM)^2) \ , \tag{2.35d}$$

where $K(\cdot)$ is the cutoff factor defined in (1.35). There are now bounds on the propagators analogous to (1.36), with the running scale Λ replaced by the running scale M,

$$\int \frac{d^4p}{(2\pi M)^4} |Q| \ < \ C(N) \tag{2.36a}$$

$$\| \frac{\partial^n Q}{\partial p^n} \|_{NM} \ < \ D_n(N) M^{-n} \ , \tag{2.36b}$$

where the 'double norm' here is interpreted over the range $p_i^2 \leq 4(NM)^2$.

The Polchinski method of effective Lagrangians can now be applied. From the summary in Section I.7 we immediately have the result that the renormalised interaction at $M = M_0$ may be written in the form

$$\begin{aligned} S(M_0,N) \ = \ \int d^4x \Big\{ & \rho_1^0 \, A_\mu^2 + \rho_2^0 \, (\partial_\mu A_\mu)^2 + \rho_3^0 \, (\partial_\mu A_\nu - \partial_\nu A_\mu)^2 \\ & + \rho_4^0 \, [A_\mu, A_\nu](\partial_\mu A_\nu - \partial_\nu A_\mu) + \rho_5^0 \, \{A_\mu, A_\nu\}(\partial_\mu A_\nu + \partial_\nu A_\mu) + \rho_6^0 \, [A_\mu, A_\nu][A_\mu, A_\nu] \\ & + \rho_7^0 \, \{A_\mu, A_\nu\}\{A_\mu, A_\nu\} + \rho_8^0 \, (A_\mu^2)^2 + \rho_9^0 \, \bar\eta\eta + \rho_{10}^0 \, \partial_\mu\bar\eta\partial_\nu\eta + \rho_{11}^0 \, \partial_\mu\bar\eta[A_\mu, \eta] \\ & + \rho_{12}^0 \, \partial_\mu\bar\eta\{A_\mu, \eta\} + \rho_{13} \, [A_\mu, \bar\eta][A_\mu, \eta] + \rho_{14} \, \{A_\mu, \bar\eta\}\{A_\mu, \eta\} + \rho_{15} \, [\bar\eta, \bar\eta][\eta, \eta] \\ & + \text{relevant operators for regulator fields} \, . \Big\} \end{aligned} \tag{2.37}$$

The Green's functions from (2.37) are M_0-finite for some choice of the bare couplings ρ_b^0, and generically we have ρ_1^0 and ρ_9^0 being of order $M_0^2 P_N(\log M_0/M_{\rm R})$ and the rest of them being of order $P_N(\log M_0/M_{\rm R})$. We may also add irrelevant operators to (2.37) provided they satisfy the criteria for the initial condition function, c.f. Section I.5.

Note that I place a subscript 'N' on the polynomials of logarithms to remind us that their coefficients may, for all we know so far, depend on N.

Now we need a much stronger result. We'd like to show that $S(M_0, N)$ is in fact of the much more restricted form of $\tilde{S}_{\rm reg}$, with the Green's functions finite not only as $M_0 \to \infty$, but also as $N \to \infty$. The Ward identity (2.34), for $\Gamma \gg M_0 \gg M_{\rm R}, N \gg 1$ would then hold as the renormalised Ward identity, and the pre-regulating scale would have dropped out completely.

The first step in restricting $S(M_0, N)$ comes from noting that with our choice of gauge (2.12), the operators $\bar{\eta}\eta, AA\bar{\eta}\eta$, and $\bar{\eta}\bar{\eta}\eta\eta$ are in fact *irrelevant.* Thus we can stipulate their coefficients, ρ_9^0, ρ_{13}^0, ρ_{14}^0 and ρ_{15}^0, to be zero at the high scale.

This is because the longitudinal and transverse parts of the gauge propagator (2.35b) have different bounding coefficients, in particular,

$$\int \frac{d^4p}{(2\pi M)^4} |Q_{\rm Long}| < C'(N)(\frac{m^2}{M^2})^{2n+2p} \tag{2.38a}$$

$$\|\frac{\partial^n Q_{\rm Long}}{\partial p^n}\|_{NM} < D'_n(N) M^{-n}(\frac{m^2}{M^2})^{2n+2p} \ , \tag{2.38b}$$

where the extra factors relative to (2.36) come from the gauge-fixing function. Thus we have the new bounds,

$$|A_{\bar{\eta}\eta}(0,0;M,N)| \leq (\frac{m^2}{M^2})^{2n+2p} P_N(\log\frac{M_0}{M_{\rm R}}) \tag{2.39a}$$

$$|A_{\bar{\eta}\bar{\eta}\eta\eta}(0,0,0,0;M,N)| \leq (\frac{m^2}{M^2})^{2n+2p} P_N(\log\frac{M_0}{M_{\rm R}}) \tag{2.39b}$$

and

$$|A_{AA\bar{\eta}\eta}(0,0,0,0;M,N)| \leq (\frac{m^2}{M^2})^{2n+2p} P_N(\log\frac{M_0}{M_{\rm R}}) \ . \tag{2.39c}$$

To see why this happens, consider the generation in the effective Lagrangian of a quartic ghost coupling from two $A\bar{\eta}\eta$ couplings, as in Fig. 15.

By Lorentz invariance the $A\bar{\eta}\eta$ operator contains at least one derivative. If in Fig. 15 the derivative acts on an external ghost line, then this makes no contribution to the zero momentum value of the vertex, (so we ignore this), and if it acts on the internal gauge line, then it projects out Q_{Long}. It is the stronger bound (2.38) on Q_{Long} which allows us to derive, in the same inductive way as before, the stronger bound (2.39b) on the $\bar{\eta}\bar{\eta}\eta\eta$ vertex.

The same consideration applies for the generation of a zero momentum $\bar{\eta}\eta$ vertex and $AA\bar{\eta}\eta$ vertex, and this is enough to prove irrelevance.

Actually we can get more for our money, for in this gauge we can also show that the renormalisation constant of $\partial_\mu\bar{\eta}A_\mu\eta(x)$ is M_0-finite, instead of the expected $\rho^0_{11} \sim \rho^0_{12} \sim P(\log M_0/M_{\mathrm{R}})$. To establish this, imagine computing the 1PI Green's function $\Gamma_{A_\mu\bar{\eta}\eta}(p_1,p_2,p_3)$ with the interaction at M_0. The tree term is just a *linear* combination of the momenta p_i with coefficients ρ^0_{11} and ρ^0_{12}, and in fact this is the *only* contribution of this form. In Feynman diagrams with loops, e.g., Fig. 16, the derivative in the $A\bar{\eta}\eta$ coupling again must act either externally or on the longitudinal gauge propagator, and hence the loop contributions to $\Gamma_{A_\mu\bar{\eta}\eta}$ are the more complicated functions $p_{1\mu}p_2\cdot p_3F(p_i)$, $p_{1\mu}\left[m^2G(p_i)\right]^{2n+2p}$, etc. The unspecified functions of momenta here, whatever they are, have dimension -2 and arise from the loop integrals.

But the sum of the trees and loops is certainly M_0-finite, since (2.37) is the renormalised interaction, and this is true over a whole range of values of p_i. Thus the coefficients of the linear term, namely ρ^0_{11} and ρ^0_{12} , are separately M_0-finite, as claimed.

With the identical reasoning it can also be shown that the operator $f^{abc}A^b_\mu\eta^c(x)$, which we must study since its insertion into Green's functions appears in the identity that we're trying to prove, is an M_0-*finite* operator. Now, Section I.6 on the renormalisation of composite operators would tell us to expect

$$f^{abc}(A^b_\mu\eta^c)_{\mathrm{ren}}(x) = \xi^0_1\, f^{abc}A^b_\mu\eta^c(x) + \xi^0_2\,\partial_\mu\eta^a(x) \ , \tag{2.40}$$

where ξ_1^0 and ξ_2^0 are of order $P_N(\log M_0/M_{\rm R})$. In fact, ξ_1^0 is M_0-finite, and the value of ξ_2^0 is unimportant since in (2.34) we take a transverse projection. To see this we consider the inserted 1PI Green's function $\Gamma^{\rm op}_{A_\mu \eta}(p_1,p_2)$, and as above, the trees and loops must be separately M_0-finite.

All the machinery is now in place for an inductive proof of the renormalised Ward identity, where the induction is in the order of $\hbar$.

By a suitable choice of the relevant parameters $\rho_b^{\rm R}$ at the low scale, the renormalised interaction $S(M_0,N)$ can be written as the sum of $S_{\rm reg}$, c.f. (2.32), plus terms of order at least one in $\hbar$. This is simply because at *tree level* we don't need any renormalisation. Let me denote this fact as

$$S(M_0,N) = \tilde{S}_{\rm reg}^{(0,\infty)} + O(\hbar^1) \ , \tag{2.41}$$

where the superscript indicates that the renormalisation constants $Z_\gamma, Z_\eta,$ and$\tilde{g}$ are as yet of zeroth order in $\hbar$, so that $\tilde{S}_{\rm reg}$ coincides with $S_{\rm reg}$, and that the constants are finite as $N \to \infty$.

Now since $\tilde{S}_{\rm reg}^{(0,\infty)}$ is regularised at M_0, its M_0-divergences are independent of N. Thus the renormalisation parameters ρ_b^0 are N-finite at the next order, that is $O(\hbar^1)$, and so, making no further assumption about the renormalised interaction, we can choose $Z_\gamma^{(1,\infty)}$, $Z_\eta^{(1,\infty)}$ and $(Z_\eta \tilde{g})^{(1,\infty)}$ to be equal respectively to the actual values of ρ_3^0, ρ_{10}^0 and ρ_{11}^0 in first order, and write

$$S(M_0,N) = \tilde{S}_{\rm reg}^{(1,\infty)} + O(\hbar^1) \ . \tag{2.42}$$

Now we must find out what are the unknown $O(\hbar)$ terms in $S(M_0,N)$. To do this we define a new interaction at M_0 by *dropping* these unknown $O(\hbar^1)$ terms,

$$\bar{S}(M_0,\infty) = \tilde{S}_{\rm reg}^{(1,\infty)} \ . \tag{2.43}$$

Of course the Green's functions from $\bar{S}(M_0,\infty)$ may not be M_0-finite, but they are N-finite and they satisfy the regularised Ward identity *exactly*, up to terms $O(N^{-2})$. To analyse the possible M_0-divergences it's most convenient to consider the interaction scaled down to $M_{\rm R}$, for which the difference between (2.42) and (2.43) is simply a *change in position* on the fixed surface, as described in Section I.7 (*iii*). The coordinate shifts are

$$\rho_b^{\rm R} \to \rho_b^{\rm R} + \delta\rho_b^{\rm R} \tag{2.44a}$$

$$\vec{\rm I}^{\rm R}(\rho_b^{\rm R}) \to \vec{\rm I}^{\rm R}(\rho_b^{\rm R} + \delta\rho_b^{\rm R}) \ , \tag{2.44b}$$

where $\delta\rho_b^{\rm R} = O(\hbar^1)$ except for $\delta\rho_3^{\rm R}$, $\delta\rho_{10}^{\rm R}$ and $\delta\rho_{11}^{\rm R}$, which vanish by construction. The shift is *a priori* M_0-divergent and N-divergent, but it is nonetheless small and treatable perturbatively.

We now characterise this shift by looking at the Green's functions. Since we're going to be counting powers of $\hbar$, let's be clear on how to do that: the power P of $\hbar$ of a Feynman diagram with E external legs is related to the loop order L by $L \leq P + 1 - E$. In the following I change the normalisation of the Green's functions by a factor $\hbar^{1-E}$ so that now $L \leq P$.

Consider the longitudinal part of the gauge 2-point function computed from the shifted interaction at $M_{\rm R}$, namely, $\bar{S}(M_{\rm R}, \infty)$. (Remember we can compute the Green's functions using the interaction at any scale we wish.) The M_0-divergences at first order in $\hbar$ come purely from $\delta\rho_1^{\rm R}$ and $\delta\rho_2^{\rm R}$, since the shifts in the other parameters only contribute to the Green's function by forming loops and therefore they only enter in higher order. However the regularised Ward identity for $\bar{W}[J]$ states that

$$0 = \delta^4(x-y)\delta^{ab} + \frac{1}{\alpha\hbar} f^2(\frac{\partial^2}{m^2})\partial_\mu^x \partial_\nu^y \frac{\delta \bar{W}[J]}{\delta J_\mu^a(x) \delta J_\nu^b(y)}\bigg|_{J=0} + O(\frac{1}{N^2}) \ , \tag{2.45}$$

and so the *full* longitudinal Green's function from $\bar{W}$ is given purely by the tree-level contribution, and in particular is M_0-finite. This implies that $\delta\rho_1^{\rm R}$ and $\delta\rho_2^{\rm R}$ are, despite our general expectations, M_0-finite at $O(\hbar^1)$.

For the gauge 3-point function the Ward identity reads

$$i(Z_\eta\tilde{g})^{(1)}\left[f^{bde}(g_{\mu\nu}-\frac{\partial_\mu^y\partial_\nu^y}{\partial^2})\frac{\delta\bar{G}^{ae}[J;x,y]}{\delta J_\mu^d(y)\delta J_\rho^c(z)}\bigg|_{J=0}+\begin{matrix}\mathrm{b}\sim\mathrm{c}\\ \mathrm{y}\sim\mathrm{z}\\ \nu\sim\rho\end{matrix}\right]$$
$$=\frac{-1}{\alpha}f^2(\frac{\partial^2}{m^2})\partial_\mu^x\frac{\delta\bar{W}[J]}{\delta J_\mu^a(x)\delta J_\nu^b(y)\delta J_\rho^c(z)}\bigg|_{J=0}+O(\frac{1}{N^2})\ ,\quad(2.46)$$

and this is depicted in Fig. 17. Since the operator insertion $f^{abc}A_\mu^b\eta_c$ is M_0-finite for the full interaction $S(M_0)$, c.f. (2.40), we note that ξ_1 is also M_0-finite for $\bar{S}(M_0,\infty)$ at $O(\hbar^1)$. Moreover, $(Z_\eta\tilde{g})^{(1)}$ is M_0-finite, due to the M_0-finiteness of ρ_{11}^0. Thus, the inserted Green's function on the LHS of (2.46) is also M_0-finite at $O(\hbar^1)$, since given the M_0-finiteness of $\delta\rho_1^{(1)}$, $\delta\rho_2^{(1)}$ and $\delta\rho_3^{(1)}$ the M_0-divergences in the parameter shifts can again only contribute by forming loops.

Identity (2.46) therefore tells us that the gauge 3-point function from $\bar{W}$, multiplied by various factors of momenta, is already M_0-finite at $O(\hbar^1)$. Since the putative M_0-divergence could only have been proportional to a linear function of momenta, this says that the gauge 3-point function itself is in fact M_0-finite at $O(\hbar^1)$, and hence so are $\delta\rho_4^{\mathrm{R}}$ and $\delta\rho_5^{\mathrm{R}}$.

The identical argument works for the 4-point function, and thus for $\delta\rho_6^{\mathrm{R}}$, $\delta\rho_7^{\mathrm{R}}$ and $\delta\rho_8^{\mathrm{R}}$. Now everything in the generic form (2.37) has been taken care of[f5] : the M_0-divergences in $\delta\rho_b^{\mathrm{R}}$ are all $O(\hbar^2)$.

Going back from $\bar{S}(M_0,\infty)$ to $S(M_0,N)$, we have demonstrated by this argument that (2.42) can really be written

$$S(M_0,N)=\tilde{S}_{\mathrm{reg}}^{(1,\infty)}+O(\hbar^2)\ .\qquad(2.47)$$

The induction is now clear. From (2.41) we have derived (2.47), and by repeated

---

f_5 Actually there's one small addendum necessary. We must show that the dimension four part of the regulator interaction is also of gauge invariant form at $O(\hbar^1)$. This is proved in exactly the same manner, alluded to in the following section, as for non-chiral matter coupled to the gauge field.

application of the reasoning we can deduce that for any n,

$$S(M_0, N) = \tilde{S}_{\text{reg}}^{(n,\infty)} + O(\hbar^{n+1}) \; . \tag{2.48}$$

Thus we have shown that any BRS-breaking terms or N-divergences can be hidden away beyond an arbitrarily high order in $\hbar$. This completes the proof that we can construct, order by order in $\hbar$, a renormalised interaction of the form (2.33), which yields Green's functions finite in the limit $M_0, N \to \infty$, satisfying the renormalised Ward identity (2.34).

5. Non-chiral Matter Couplings

In this chapter I have so far been discussing pure Yang-Mills theories. However it's a trivial extension, with the techniques developed, to consider coupling in non-chiral matter. For example,

$$L_{\text{matter}} = D_\mu \bar{\phi} D_\mu \phi + \mu^2 \bar{\phi}\phi + \lambda(\bar{\phi}\phi)^2 + i\bar{\psi}(\not{D} + m)\psi \; . \tag{2.49}$$

The extra regulator fields required to cancel the one-loop divergences of the matter fields are of the standard Pauli-Villars type. Adding these in, as well as L_{HCD} from (2.21), we now have a BRS-invariant, fully regulated Lagrangian for the non-chiral gauge theory. We now wish to be able to find wavefunction renormalisations Z_γ, Z_η, Z_ϕ and Z_ψ, mass renormalisations Z_μ and Z_m, and coupling constants $\tilde{g}$ and $\tilde{\lambda}$ such that the regulating scale M_0 can be taken to infinity.

We proceed as before. We derive the regularised Ward identity, where now the functionals have sources for the matter fields as well as for the gauge field, and the identity restricts the possible counterterms that can appear. As explained in Chapter I, the possibility of SSB makes no difference to our renormalisation argument, and thus we have proved the renormalised Ward identities for an arbitrary non-chiral gauge theory, in both the symmetric and broken phases.

6. Unitarity

Let me outline what we've achieved so far. Throughout the discussion we have been considering a manifestly Lorentz covariant formalism, in order to guarantee automatically that the renormalised S-matrix satisfies the laws of Special Relativity. However, the generic results of renormalisation theory gave us an interaction (2.37) for which the S-matrix was not *unitary*, due to the non-trivial coupling of physical and unphysical modes, and this was the problem that we had to overcome.

A big improvement was obtained by considering a special class of gauge-fixings. For the transverse Lorentz gauge we managed to find a BRS-invariant regulator (2.21), from which we proved the regularised Ward identities (2.30). These identities restricted the nature of the counterterms, and we showed that the renormalised interaction was in fact of the form (2.33), with concomitant renormalised Ward identity, (2.34).

The relevance of this is that, broadly speaking, the Ward identity describes the underlying gauge invariance of the theory. Since the identity survives renormalisation, this means that the gauge invariance is retained at the full quantum level. The renormalised S-matrix[f6] should therefore be the same in all gauges.

Now we can find a gauge with manifest Lorentz invariance, and also a gauge with manifest unitarity, so the (unique) S-matrix must have both of these benevolent properties! What we wish to do is construct a proof, in this spirit, of the unitarity of our already Lorentz-invariant, renormalised S-matrix.

Let me consider a continuous path in the space of gauge conditions, parametrised by $\bar{M}$, from the transverse Lorentz gauge to the 'axial' gauge. The latter is manifestly unitary due to the manifest decoupling of the Faddeev-Popov ghosts. Now, for a point on this path which corresponds to a 'transverse gauge,' i.e., a gauge where the

f_6 Strictly speaking, this discussion should refer to the case of spontaneous symmetry breaking, so that the S-matrix exists at all. For clarity, I shall use the *language* of the pure gauge theory, but this affords no loss of generality.

transverse part of the propagator dominates at large momentum, I shall show using the renormalised Ward identity that the renormalised S-matrix satisfies

$$\frac{d}{d\bar{M}}\mathbf{S} = 0 . \tag{2.50}$$

Thus the renormalised S-matrix is indeed the same throughout the region $\mathbf{R}$ of transverse gauges. If we could reach the axial gauge maintaining 'transverse-ness,' that is, staying within $\mathbf{R}$, then we'd be finished. However, the axial gauge is *not* itself transverse, and we cannot directly prove (2.50) at that point. In particular, we don't have at our disposal a renormalised Ward identity for this gauge.

To obviate this problem I shall find a path in $\mathbf{R}$ which comes 'sufficiently close' to the axial gauge, and argue that the axial gauge limit can actually be taken. Let me explain this point. I choose a convenient family of gauge-fixing conditions,

$$g^a[A_\mu] = X\,\partial_\mu A_\mu^a + iXY\,t_\mu A_\mu^a \ , \tag{2.51a}$$

with new gauge-fixing and ghost terms replacing those in the Lagrangian (2.11),

$$L_{\text{gf}} = \frac{1}{2\alpha}\left[f(\frac{\partial^2}{m^2})g^a[A_\mu]\right]^2 + i\bar{\eta}^a\frac{\delta g^a}{\delta A_\mu^b}(D_\mu\eta)^b \ . \tag{2.51b}$$

In (2.51a) t_μ is some fixed 4-vector, say, $(1,0,0,0)$, $f(\cdot)$ is the gauge-weighting function in (2.12), and X, Y are factors, depending on m, $\bar{M}$ and ∂^2, which are yet to be chosen. This gives us a gauge field propagator of the form:

$$P_{\mu\nu} = \frac{1}{k^2(1+k^{4n}/M_0^{4n})}\left[(g_{\mu\nu}-\frac{k_\mu k_\nu}{k^2}) + B\frac{k_\mu k_\nu}{k^2} + C\frac{k_\mu t_\nu}{k^2} + D\frac{t_\mu k_\nu}{k^2}\right] \tag{2.52a}$$

where

$$B = \frac{Y^2(k\cdot t)^2 + Y^2t^2k^2 + \alpha k^2(1+k^{4n}/M_0^{4n})f^{-2}(k^2/m^2)X^{-2}}{k^4 - Y^2(k\cdot t)^2} \tag{2.52b}$$

$$C = -\frac{Yk^2}{k^2 + k\cdot t} \tag{2.52c}$$

$$D = \frac{Yk^2}{k^2 - k\cdot t} \quad . \tag{2.52d}$$

Now the properties that we require are: (a) eq.(2.52) reduces to the transverse Lorentz gauge for $\bar{M} = m$; (b) we reduce to the axial gauge for $\bar{M} = \infty$; and (c) there is a range of values for the path parameter such that the non-transverse terms B, C and D vanish as $k^2 \to \infty$ and are already suppressed by powers of m/M_0 for $k^2 \sim M_0^2$. (I shall call this range the 'transverse range'.) We can achieve these properties by choosing, for example,

$$X = (\frac{m}{\bar{M}})^{3/4} \tag{2.53a}$$

$$Y = \frac{\bar{M}}{m}(1 - \frac{m}{\bar{M}})\frac{m}{1+\partial^4/\bar{M}^4} \quad , \tag{2.53b}$$

giving us a transverse range,

$$m \leq \bar{M} \ll (\frac{M_0}{m})^{4/5} \quad . \tag{2.53c}$$

(Note that there is no problem associated with the fact that Y is in general non-local, for the action at $\bar{M} = m$ is certainly local, and we are going to prove that the S-matrix is the same for all $\bar{M}$.)

Property (c) is the key here. It gives us a range of $\bar{M}$ for which the gauge condition is in region $\mathbf{R}$ and the renormalised Ward identity can be derived (see below). Moreover, this range extends up to infinity if M_0 is taken to infinity first. Thus, by comparing (b) and (c), we see that if we can reverse the order of limits for the scales, keeping the S-matrix the same, then the gauge condition leaves region $\mathbf{R}$

and reaches the axial gauge. Our discussion of this reversal will then complete the argument for unitarity.

Now first we must prove (2.50). I claim that with the interpolating gauge choice (2.51), and with $\bar{M}$ in the transverse range, we still have a renormalised interaction of the form (2.33), despite the breaking of Lorentz invariance. To see this, imagine repeating the entire argument from (2.11) to (2.48). The regulation method (2.21) certainly holds, and so does the regularised Ward identity (2.30). For renormalisation, the bounds (2.36) on the propagators are valid, and, moreover, we still have the important stronger bounds, c.f.(2.38), on the non-transverse parts of the gauge propagator:

$$\int \frac{d^4p}{(2\pi M)^4}|Q_{\text{non-trans}}| < C(\frac{\bar{M}}{M})^5 \tag{2.54a}$$

$$\|\frac{\partial^n Q_{\text{non-trans}}}{\partial p^n}\|_{NM} < D_n M^{-n}(\frac{\bar{M}}{M})^5 \,. \tag{2.54b}$$

Here the non-transverse parts include the longitudinal part and the Lorentz breaking part. The power '5' comes from my choice of Y. The bounds (2.54) once again allow us to argue that various unwanted operators are in fact irrelevant, in particular all Lorentz breaking operators. The rest of the argument to (2.48) now proceeds unimpeded.

I now sketch the proof that for all $\bar{M}$ satisfying (2.53c), the renormalised S-matrix is the same. Consider the generating functional $Z_{\text{ren}}[J,\bar{M}]$ of renormalised Green's functions in the gauge defined by (2.51). Suppose we now change $\bar{M}$ in (2.51) by an infinitesimal amount $\delta\bar{M}$, keeping everything else in the action at M_0 fixed. This changes the Lagrangian at M_0 by

$$\delta L_{\text{gf}} = \delta\bar{M}\left[\frac{1}{\alpha}g^a\frac{dg^a}{d\bar{M}} + i\bar{\eta}^a\frac{d}{d\bar{M}}\frac{\delta g^a}{\delta A^b_\mu}(D_\mu\eta)^b\right] \,, \tag{2.55}$$

where we *don't* change the pre-regulating part of the action. This can remain the same for all $\bar{M}$.

To first order in $\delta\bar{M}$, the new generating functional, call it $Z_{\rm reg}^{\delta M}[J,\bar{M}]$, differs from $Z_{\rm ren}[J,\bar{M}]$ by an operator insertion of $\delta L_{\rm gf}$, i.e.,

$$Z_{\rm reg}^{\delta M} - Z_{\rm ren} = \delta\bar{M}\langle \int d^4x \delta L_{\rm gf}(x)\rangle_J \ . \tag{2.56}$$

One can check that this operator insertion is indeed regularised, that is, finite with respect to the pre-regulator Γ, so that (2.56) makes sense. Of course the insertion may not be renormalised, that is M_0-finite, but that won't matter.

Now, I can derive a new regularised Ward identity, exactly analogous to our original one (2.30), by including in the functional $Z_{\rm ren}$ a source K coupled to $dg/d\bar{M}$. Performing a BRS transformation on the fields, differentiating with respect to the ghost source ζ and also K, and setting $\zeta = \bar{\zeta} = K = 0$, yields

$$0 = \langle i\bar{\eta}^c(z)\frac{dg^a(y)}{d\bar{M}}\int d^4x J_\mu^b (D_\mu\eta)^b(x) + \frac{1}{\alpha}g_c(z)\frac{dg^a(y)}{d\bar{M}} + i\bar{\eta}^c(z)\frac{d}{d\bar{M}}\frac{\delta g^a}{\delta A_\mu^b}(D_\mu\eta)^b(y)\rangle_J \ . \tag{2.57}$$

Setting now $a=c$ and summing, and $y=z$ and integrating, we can combine (2.55), (2.56) and (2.57) and obtain, to first order in $\delta\bar{M}$,

$$Z_{\rm reg}^{\delta M}[J,\bar{M}] = \left\langle \exp - i\delta\bar{M}\int d^4x d^4y J_\mu^b (D_\mu\eta)^b(x)\bar{\eta}^a\frac{dg^a}{d\bar{M}}(y)\right\rangle_J \ . \tag{2.58}$$

Equation (2.58) says that the only difference between $Z_{\rm reg}^{\delta M}$ and $Z_{\rm ren}$ is a change in what the source J_μ is coupled to. In the former it is coupled not just to $A_\mu^b(x)$ but rather to

$$A_\mu^b(x) - i\delta\bar{M}\int d^4y (D_\mu\eta)^b(x)\bar{\eta}^a\frac{dg^a}{d\bar{M}}(y) \ . \tag{2.59}$$

Taking the source to be transverse, the extra term in (2.59) is at least quadratic in the fields, whereupon it has no effect when the Green's functions are amputated to form S-matrix elements. (That's because an insertion of a non-linear function of the fields does not have a pole singularity.) Thus the S-matrix computed from $Z_{\text{reg}}^{\delta M}$ is equal to that computed from Z_{ren}. I shall denote this by

$$\mathbf{S}_{\text{reg}}^{\delta M}(\bar{M}) = \mathbf{S}_{\text{ren}}(\bar{M}) . \tag{2.60}$$

Now, what is the relation between $\mathbf{S}_{\text{reg}}^{\delta M}$ and $\mathbf{S}_{\text{ren}}(\bar{M}+\delta\bar{M})$? They are derived from actions with the same gauge-fixing, but the values of the renormalisation constants Z_γ, Z_η and $\tilde{g}$, c.f. (2.33), may be different. (After all, we have *not* claimed that the Green's functions from $Z_{\text{reg}}^{\delta M}[J,\bar{M}]$ are renormalised.) However, $\mathbf{S}_{\text{reg}}^{\delta M}(\bar{M})$ *is* renormalised, since it is equal to $\mathbf{S}_{\text{ren}}(\bar{M})$, and so it must be equal to $\mathbf{S}_{\text{ren}}(\bar{M}+\delta\bar{M})$ when we choose the same renormalisation prescription. Calling this prescription $\vec{g}$, we have therefore proved

$$\mathbf{S}_{\text{ren}}(\bar{M}+\delta\bar{M},\vec{g}) = \mathbf{S}_{\text{ren}}(\bar{M},\vec{g}) = \mathbf{S}_{\text{ren}}(\vec{g}) . \tag{2.61}$$

The taking of the axial gauge limit, $\bar{M}\to\infty$, is now the very last thing that we must consider. To do this we go all the back to the most general ideas about the flow of effective Lagrangians.

For *any* choice of gauge, be it a 'renormalisable' one or otherwise, we can put in a cutoff M_0 and calculate low energy Green's functions and S-matrix elements. The M_0-divergences can be hidden in a choice of the bare parameters, and the only difference between the 'renormalisable' and 'non-renormalisable' cases in this regard is whether the number of parameters is finite or infinite.

Now, for a given renormalisation prescription, which I shall denote by $\{\vec{g},\vec{h}\}$, where $\vec{g}$ represents the parameters in (2.61) and $\vec{h}$ describe the rest, the dependence of this general S-matrix on the gauge is completely smooth. So in particular for the $\bar{M}$-gauges we have an S-matrix,

$$\mathbf{S}_{\text{general}}(\bar{M},M_0,\vec{g},\vec{h}) , \tag{2.62}$$

where one fact, at least, that we do know about $\mathbf{S}_{\text{general}}$ is that it possesses a unique, smooth limit as $\bar{M}, M_0 \to \infty$. Thus the limits $M_0 \to \infty$ and $\bar{M} \to \infty$ are interchangeable for $\mathbf{S}_{\text{general}}$. Now $\mathbf{S}_{\text{general}}(\infty, \infty, \vec{g}, \vec{h})$ is unitary because for $\bar{M}$ going to infinity before M_0 we recover the axial gauge. Of course there's a lot of information that we don't know from this argument, for example, whether or not $\mathbf{S}_{\text{general}}$ is Lorentz-invariant, and how many parameters $\{\vec{g}, \vec{h}\}$ are needed to specify it.

But now we can compare with our previous results. If we study $\mathbf{S}_{\text{general}}$ along the section

$$\bar{M} = (\frac{M_0}{m})^{\frac{1}{2}} m \,, \tag{2.63}$$

then for M_0 large the transverse condition (2.53c) is satisfied. In the limit $M_0 \to \infty$ we must therefore recover the S-matrix in (2.61), and so

$$\mathbf{S}_{\text{general}}(\infty, \infty, \vec{g}, \vec{h}) = \mathbf{S}_{\text{ren}}(\vec{g}) \,. \tag{2.64}$$

This, at last, completes the proof that $\mathbf{S}_{\text{ren}}(\vec{g})$ is unitary.

7. Summary

Perturbative renormalisability is a general property of field theories defined by a physical mass scale cutoff, Λ_0, as evinced by considering the 'method of effective Lagrangians'. This powerful concept was first described by Wilson, and a rigorous but amazingly simple proof of its properties was provided by Polchinski. The analysis is presented in Chapter 1, and the main result is that for *any* theory we can find a (non-unique) interaction $S(\Lambda_0)$ from which the calculated Green's functions are Λ_0-finite. $S(\Lambda_0)$ contains the relevant operators, with Λ_0-divergent coefficients known as the 'bare couplings,' and it can also contain irrelevant operators (multiplied by the appropriate power of Λ_0) with *arbitrary* coefficients. The presence of the irrelevant operators affects only the relationships between the bare and renormalised couplings, and does not affect the (physical) relationships among the renormalised Green's functions themselves.

In the above context, gauge theories are special in that, essentially, we require some of the relevant operators *not to appear* in $S(\Lambda_0)$. For example, we must preclude a bare mass term for the gauge boson. This is because only when $S(\Lambda_0)$ takes on a certain restricted form can the S-matrix be unitary. However, the generic results of renormalisation theory give us a bare mass term of order Λ_0^2 times a polynomial of $\log(\Lambda_0/E_{\mathrm{phys}})$, so clearly we need to characterise more finely the nature of the effective Lagrangian flow.

The key to our discussion was the new regulator that we presented. This had the properties of being physical, so that the method of effective Lagrangians could be applied; and gauge-invariant, so that *regularised* Ward identity could be derived. Essentially what the regularised Ward identity does is to relate operators of different naïve scaling properties. In particular the unwanted relevant operators are related to operators that are truly irrelevant. It is this which provides the necessary restrictions on the form of the 'bare interaction,' and this allows us to obtain the *renormalised* Ward identity.

The final step of the argument is to link the renormalised Ward identity to unitarity. The identity can be proved only in a certain class $\mathbf{R}$ of gauges, namely the 'transverse gauges,' see Eq. (2.12), and it tells us that within this class the S-matrix is unique. Now $\mathbf{R}$ does *not* include the manifestly unitary gauge $\mathbf{U}$, but we argue that $\mathbf{U}$ can be approached 'sufficiently closely' by a sequence of $\mathbf{R}$-gauges so that the S-matrix is the same for it also. By this chain of reasoning we conclude that non-chiral Yang-Mills theories really do have a renormalised, Lorentz invariant and unitary S-matrix.

Appendix 1. 'Construction of the HCD regulator'

In this appendix I shall simply present the results for the one-loop Γ-divergences contributed by the various regulator fields. It will not, fortunately, be necessary to compute the contributions from the gauge and ghost particles (with one exception), since the regulator contribution contains many tunable parameters. These parameters, as I say, can be chosen to provide a net cancellation of divergences.

Now, the Lagrangians for real scalar and complex vector regulator particles are, c.f.(2.21),

$$L_{\text{scalar}} = -\frac{1}{4}\Phi\left[D^4 + \alpha_r g D_\mu F_{\mu\nu} D_\nu + \beta_r g[D_\mu, F_{\mu\nu}]D_\nu + \gamma_r g^2 F_{\mu\nu}F_{\mu\nu}\right]\frac{D^{2r-4}}{M_0^{2r-2}}\Phi$$
$$+ \text{h.c.} + \ldots \qquad (A1.1a)$$

$$L_{\text{vector}} = -\frac{1}{2}\bar{V}_\mu\Big[\{D^4 + \epsilon_r g D_\rho F_{\rho\sigma} D_\sigma + \eta_r g[D_\rho, F_{\rho\sigma}]D_\sigma + \kappa_r g^2 F_{\rho\sigma}F_{\rho\sigma}\}g_{\mu\nu} +$$
$$+ \{(\lambda_r - 1)D_\mu D_\nu + \pi_r F_{\mu\nu}\}D^2\Big]\frac{D^{2r-4}}{M_0^{2r-2}}V_\nu + \text{h.c.} + \ldots\ , \qquad (A1.1b)$$

and the contributions from the complex vectors to the divergences listed in Eq. (2.18) are given below.

1. $-4r$ $\qquad (A1.2a)$

2. $$\Big[\pi_r^2(1+\lambda_r^{-1}) + 2\pi_r(1-\lambda_r^{-1}) + 2(\epsilon_r + 2\kappa_r)(3+\lambda_r^{-1})$$
$$-\frac{1}{3}(6r^2 - 8r + 14) + \frac{1}{6}(6r^2 - 12r + 14)(\lambda_r + \lambda_r^{-1})\Big] \qquad (A1.2b)$$

3. $\frac{2}{3}(r-1) - \frac{2}{3}r\lambda_r^{-1} + \frac{1}{3}(\lambda_r + \lambda_r^{-1})$ $\qquad (A1.2c)$

4. $$\Big[\pi_r^2(1+\lambda_r^{-1}) + 2\pi_r(1-\lambda_r^{-1}) + 2(\epsilon_r + 2\kappa_r)(3+\lambda_r^{-1}) + \frac{1}{6}(-3r^3 + 7r - 12)$$

$$+\frac{1}{2}\lambda_r^{-1}(r^3-4r^2+3r)+(\lambda_r+\lambda_r^{-1})(r^2-2r+1)\Big] \qquad (A1.2d)$$

5. $0 \qquad (A1.2e)$

6. $$\Big[2(\epsilon_r+2\kappa_r)(3+\lambda_r^{-1})+\pi_r^2(1+\lambda_r^{-1})+\frac{1}{2}\pi_r(4-3\lambda_r-\lambda_r^{-1})-\frac{1}{6}(\lambda_r^2+\lambda_r^{-2})$$
$$+(\lambda_r+\lambda_r^{-1})(r^2-2r+\frac{5}{3})-\frac{10}{3}r^2+4r-3\Big] . \qquad (A1.2f)$$

In (A1.2) the values listed for divergences 1, 2, 3, 4 and 6 should all be multiplied by $T(\mathbf{R})$, which is the normalisation constant for the group generators, i.e. $\mathrm{Tr}T^aT^b = T(\mathbf{R})\delta^{ab}$, where $T(\mathbf{Fund}) \equiv \frac{1}{2}$. The values for the real scalar contributions can be obtained from (A1.2) by setting $\lambda_r = 1, \pi_r = 0$, replacing $\epsilon_r, \eta_r, \kappa_r$ by α_r, β_r, γ_r, and dividing by 8.

Now we see that divergences 2,3,4 and 6 can be taken care of using our freedom in the parameters $\alpha_r, \beta_r, \ldots, \pi_r$, but this is not so for divergences 1 and 5. To be sure of their cancellation, we must in fact compute the gauge and ghost contributions, call them d_1 and d_5 respectively. Now the Lagrangian for the gauge and ghost fields is, c.f. (2.11),

$$L=-\frac{1}{4}F^a_{\mu\nu}\left[1+(\frac{D^2}{M_0^2})^{2n}\right]F^a_{\mu\nu}+\frac{1}{2\alpha}\left[f(\frac{\partial^2}{m^2})\partial_\mu A^a_\mu\right]^2+i\partial_\mu\bar{\eta}^a(D_\mu\eta)^a \ , \qquad (A1.3)$$

and d_1 is found to be equal to $-T(\mathbf{Adj})\big[1 + \frac{3}{2}(3n^2+2n)\big]$, with d_5 equal to zero. Clearly divergence 5 is no problem, and d_1 can be canceled by real scalars in the adjoint representation, satisfying

$$\sum_r r = -2-3(3n^2+2n) \ . \qquad (A1.4)$$

To cancel d_2, d_3, d_4 and d_6 we need only consider complex vectors in the adjoint. Eq. (A1.2) provides us with no less than nine independent real parameters and one can trivially find combinations that exhibit cancellation.

Appendix 2. 'The Jacobian of the BRS transformation'

We wish to show that the Jacobian of the BRS transformation (2.7) is trivial. Consider the functional integral

$$I = \int DA_\mu D\bar{\eta} D\eta \,\exp - L_\Gamma(A_\mu, \bar{\eta}, \eta)\ . \tag{A2.1}$$

The subscript Γ denotes that the integral is defined by its expansion in Feynman graphs, which have been pre-regulated with some *smooth* cutoff. Performing a BRS transformation on the integration variables, suppose the Jacobian is

$$J = 1 + \Psi(A_\mu, \bar{\eta}, \eta)\,\varepsilon\ , \tag{A2.2}$$

where ε is the Grassman parameter. This means that Ψ must satisfy

$$0 = \int DA_\mu D\bar{\eta} D\eta \left[\Psi L_\Gamma - \frac{1}{g}\frac{\delta L_\Gamma}{\delta A^a_\mu}(D_\mu\eta)^a + \frac{1}{2} f^{abc}\frac{\delta L_\Gamma}{\delta \eta^a}\eta^b\eta^c - \frac{i}{\alpha g}\frac{\delta L_\Gamma}{\delta \bar{\eta}^a} f^2(\frac{\partial^2}{m^2})\partial_\mu A^a_\mu\right]\ . \tag{A2.3}$$

Unfortunately equation (A2.3) is *formal*, even given the presence of the pre-regulator. The reason is simply that the variation of L_Γ will include the variation of the inverse propagators that carry the pre-regulating cutoff factor, and the insertion of these may actually be infinite. To solve this problem I invoke a 'pre-pre-regulator,' Δ : instead of the local BRS variations (2.7) I consider a *smeared* version,

$$\delta A^a_\mu = -\frac{1}{g}\exp - \frac{\partial^2}{\Delta^2}(D_\mu\eta)^a\,\varepsilon \tag{A2.4a}$$

$$\delta\eta^a = \frac{1}{2} f^{abc}\exp - \frac{\partial^2}{\Delta^2}\eta^b\eta^c\,\varepsilon \tag{A2.4b}$$

$$\delta\bar{\eta}^a = \frac{-i}{\alpha g}\exp - \frac{\partial^2}{\Delta^2} f^2(\frac{\partial^2}{m^2})\partial_\mu A^a_\mu\,\varepsilon\ . \tag{A2.4c}$$

The analogue of (A2.3) now stands exactly. Now, integrating by parts the second term with respect to A^a_μ, the third term with respect to η^a, and the fourth term with respect to $\bar{\eta}^a$, we see that they all vanish. Thus for the pre-pre-regulated transformation Ψ must be zero, that is, the Jacobian must indeed be trivial. (Note that the pre-pre-regulator drops out of all the identities generated in Chapter II in the limit $\Delta \gg \Gamma$, so this discussion does not affect the derivation of the regularised and renormalised Ward identities.)

Appendix 3. 'The Compensating Operator for Non-Chiral Gauge Theories'

The insertions of $\delta L_{\rm pre}$ are shown in Fig. 13. Here we are interested in the Γ-finite parts of the one-loop inserted diagrams.

For Fig. 13(*ii*) the only Γ-finite parts come from closing up the regulator lines into a regulator cycle with external gauge legs. Now if the regulator field had been 'quadratic,' that is, coupled only to the gauge field with no self-couplings, then this regulator cycle would have represented the BRS variation of a functional of A_μ, namely, the functional determinant obtained by integration. Thus a second BRS variation would have yielded zero. This is known as the Wess-Zumino consistency condition [13], and it would have applied separately to the Γ-divergent and Γ-finite parts.

Even though our regulator regulators do have self-couplings, and are also coupled to each other, it's true nevertheless that the Γ-finite parts satisfy WZ. To see this, consider a new Lagrangian for the regulators related to the original (2.21) by simply dropping the self-couplings and mutual couplings. The new cycles satisfy WZ by construction, and they exactly reproduce the regulator cycles of interest in their Γ-divergent and Γ-finite parts.

Now written in terms of an operator G, the Γ-finite part is just a ghost field η^a times some dimension four polynomial of gA_μ and derivatives. Thus the WZ condition is a purely *algebraic* constraint. Becchi et al. [12] analysed this condition, and proved that such an operator can be written as a linear combination of the ABBJ anomaly and the BRS variation of a local polynomial of A_μ, i.e.,

$$G = a\, g^2 \varepsilon_{\mu\nu\rho\sigma}\, \mathrm{tr}\left[\eta\partial_\mu(A_\nu\partial_\rho A_\sigma + \frac{1}{2} g A_\nu A_\rho A_\sigma)\right] + b\, \delta_{\rm BRS} P_4[gA_\mu]\ , \qquad (A3.1)$$

where a and b are numerical coefficients.

The coefficient of the anomaly must be zero, since there is no source of the parity violating anti-symmetric tensor $\varepsilon_{\mu\nu\rho\sigma}$, and so the Γ-finite parts of Fig. 13(*ii*) can indeed be canceled by a compensating operator added to the Lagrangian. Similar remarks go for the gauge cycle from Fig. 13(i) and the ghost cycle from Fig. 13(iii).

Now Fig. 13(i) and Fig. 13(iv) contain other one-loop graphs which cannot be dealt with in this way. These are shown in Fig. 18. To solve this problem, I change the scale of momentum cutoff for the ghosts, so that that the pre-regulating term for the ghosts is now, c.f.(2.24),

$$i\partial_\mu \bar{\eta}^a \left[1 + (\frac{\partial^2}{\Omega^2})^{2\mathbf{Q}}\right] \partial_\mu \eta^a \ . \tag{A3.2}$$

Taking $\Gamma \gg \Omega \gg M_0$, the contribution from Fig. 18(i) now dies like a power of $(\Omega/\Gamma)^{2\mathbf{P}}$. For Fig. 18($ii$) the relevant momentum integrals are of the form

$$\int_0^\infty \frac{k^{2\mathbf{Q}}/\Omega^{2\mathbf{Q}}}{(1 + k^{2\mathbf{P}}/\Gamma^{2\mathbf{P}})^s (1 + k^{2\mathbf{Q}}/\Omega^{2\mathbf{Q}})} \, k^t \, dk \ , \tag{A3.3}$$

where the exponents s and t can take on the values $s = 2, 3, 4 \,\text{and}\, t = 1, -1$. For $t = 1$ we have contributions of order

$$\Gamma^2 O(1) + \Omega^2 O(\frac{1}{\mathbf{Q}}) + O(\frac{\Omega}{\Gamma})^{2\mathbf{P}} \ , \tag{A3.4a}$$

and for $t = -1$ we have

$$\log(\frac{\Gamma}{\Omega}) O(1) + O(\frac{1}{\mathbf{Q}}) + O(\frac{\Omega}{\Gamma})^{2\mathbf{P}} \ . \tag{A3.4b}$$

The quadratic Γ and Ω divergences can be ignored, as stated in the text. We see that there is a finite piece even in our new limit, $\Gamma \gg \Omega \gg M_0$, but this vanishes as $\mathbf{Q} \gg 1$. (This last limit can indeed be taken because the full inserted Green's function, to any loop order, contains no positive powers of $\mathbf{Q}$.)

This completes the proof that there exists a compensating operator which exactly cancels the finite parts of δL_{pre}. The operator is obviously proportional to $\hbar$, since it is a one-loop object.

CHAPTER III
PERTURBATIVE ANOMALIES

In Chapter II we limited our discussion of Yang-Mills theories to the non-chiral, non-supersymmetric case. From the work of 't Hooft and Veltman using dimensional regularisation, [1], it was already known that these theories are physically acceptable, but we managed at least to construct a much simpler proof than the original by considering the method of effective Lagrangians. The strategy was to find a *physical* regulator, that is to say a mass-scale cutoff, that preserved BRS invariance. This allowed us to derive the *regularised* Ward identities, and these could be used to restrict the form of the operators appearing in the bare Lagrangian. Order by order in $\hbar$ we showed that the low energy Green's functions do indeed satisfy the *renormalised* Ward identities, and therefore that these theories are unitary.

Now we turn to theories which include *chiral* fermions, whose bare masses are constrained to be zero by the gauge symmetry. Here there is the fascinating phenomenon that at the quantum level the gauge symmetry may become anomalous [14], thus rendering the theory unphysical. Whether or not this happens at *one-loop* depends on the representation $\mathbf{R}$ of the chiral fermions, and for one-loop anomaly freedom we must demand,

$$\mathrm{tr}\ T^a_{\mathrm{R}}\{T^b_{\mathrm{R}}, T^c_{\mathrm{R}}\} = 0 . \tag{3.1}$$

This 'cancellation condition' for the one-loop ABBJ anomaly is *necessary* for renormalisability and unitarity, and we ask the question in this chapter whether it is also *sufficient*, that is to say that there no more obstructions to quantisation at higher loop orders.

As I explained in the Introduction, this question is very hard to discuss in the context of dimensional regularisation or BPHZ, and it is clear that if we wish to pursue the method of effective Lagrangians we also are faced with a serious difficulty. The mass scale in any physical regulator is bound to be associated with a *breaking* of the chiral gauge symmetry, even when the ABBJ anomaly cancels, so there exists *no* 'higher covariant derivative' regulator analogous to the one in Chapter II, for this case. How, then, are we to proceed?

The idea that I present here is that of a physical regulator with not one, but *two* mass scales of regularisation. The higher one M_χ will control the gauge breaking operators, while the lower one M_0 will only enter through the higher covariant derivatives. If we satisfy the ABBJ condition (3.1), the corrections to the regularised Ward identity will be bounded by $O(M_0^2/M_\chi^2)$, and we'll be able to find renormalisation constants such that M_χ and M_0 can be taken to infinity in the order

$$M_\chi \gg M_0 \gg E_{\text{phys}} \,. \tag{3.2}$$

It is in this 'sequential' limit that the renormalised Ward identity will hold, thereby proving that there are indeed no higher-loop anomalies.

This new scheme of regularisation also has the merit that it can be directly applied to $N = 1$ supersymmetry. In the language of Chapter II, the problem here is that a manifestly supersymmetric formalism necessitates the presence of dimensionless fields. This causes there to be divergences in Green's functions with an arbitrary number of these fields as external. Thus the number of different divergences that one must regulate is actually *infinite*, and this renders the gauge-invariant method of Chapter II rather unworkable.

However, regularisation is easy to achieve if we allow ourselves to *break* the gauge invariance. Once again the gauge breaking is controlled by having two mass scales, with the breaking being reduced to zero as we take one scale to infinity before the other. We prove that supersymmetry is a good symmetry in QFT, the only anomalies being the natural supersymmetric extension of the ABBJ anomaly.

1. Regularisation of χGT

In this section we concentrate on the *regularisation* of chiral gauge theories. Once we've obtained the regularised Ward identity, the path to proving renormalisability and unitarity will follow the well-trodden route described in Chapter II.

We would still like to use higher covariant derivatives for the gauge field, since it was this which allowed us to consider only one-loop objects, and so we shall try to regulate by employing the following higher derivative action:

$$L = -\frac{1}{4}F^a_{\mu\nu}\left[1+(\frac{D^2}{M_0^2})^{2n}\right]F^a_{\mu\nu}+\frac{1}{2\alpha}\left[f(\frac{\partial^2}{m^2})\partial_\mu A^a_\mu\right]^2+i\partial_\mu\bar{\eta}^a(D_\mu\eta)^a+i\bar{\psi}\not{D}P_+\psi \ . \quad (3.3)$$

In (3.3) I have, for simplicity, written just one chiral fermion field. This is to be considered as a *representative* of any chiral fermion content that one might choose, and of course in the end we'll demand that we satisfy the condition (3.1).

I divide up the one-loop divergences of (3.3) into three classes: the field running around the loop may be a gauge boson, a ghost, or a chiral fermion. Now, massive regulator fields that are added to (3.3) to cancel the *chiral fermion loop* Γ-divergences will *break* the chiral gauge invariance. It does no good, therefore, to have their mass scale equal to M_0, for then the regulation scale and the gauge breaking scale would be the same. The regulated Ward identities would then be very hard to prove, in exact analogy to the case of using a momentum cutoff regulator for pure Yang-Mills.

Thus I shall take the chiral-regulator fields to have a mass scale M_χ much larger than M_0. This will allow me to separate off the gauge breaking contributions to the Ward identity, and then bound them to be arbitrarily small.

The regulator fields for the gauge and ghost loops could of course be chosen as in Chapter II. Instead I shall find a new *gauge-non-invariant* action to cancel the gauge and ghost Γ-divergences, with the feature that, unlike before, no explicit calculation of Γ-divergences is required. (The motivation for this is the ability to generalise the discussion to supersymmetry.) Like the chiral regulator fields, the mass scale for the gauge and ghost regulators will be M_χ, not M_0.

For the chiral fermion and ghost cycles we choose ordinary Pauli-Villars regulators. The actions for these fields are

$$L_{\mathrm{f-reg}} = i\bar{\Psi}(\not{\partial} - ig\not{A}P_+ + i\kappa M_\chi)\Psi \quad (3.4a)$$

$$L_{\text{gh-reg}} = i\partial_\mu \bar{c}^a[(D_\mu c)^a + \lambda^2 M_\chi^2 c^a] \ , \tag{3.4b}$$

and so if we have f_α chiral regulators of mass-square $\kappa_\alpha^2 M_\chi^2$, and g_β ghost-regulators of mass-square $\lambda_\beta^2 M_\chi^2$, then the Γ-divergences from the chiral fermions and ghosts are canceled when

$$0 = \sum_\alpha f_\alpha + 1 \tag{3.5a}$$

$$0 = \sum_\alpha f_\alpha \kappa_\alpha^2 \tag{3.5b}$$

$$0 = \sum_\beta g_\beta + 1 \tag{3.5c}$$

$$0 = \sum_\beta g_\beta \lambda_\beta^2 \ . \tag{3.5d}$$

For the gauge cycle it's not so easy, since the Pauli-Villars regulator fields in (2.17) are not acceptable. The problem, remember, is that we want *all* the regulator fields to have a lower *type* than the gauge field, c.f. (2.16), in order to have no overlapping divergences.

However, from (2.17) we can *construct* a new, gauge-non-invariant action, for which the type is lower and the conditions for Γ-divergence cancellation are still as simple as (3.5).

First, take (2.17) and drop all but the highest dimension operators in the action. This yields

$$S'_{PV} = -\frac{1}{8}\int d^4x d^4y \Phi_\mu^a \left[\frac{\delta^2}{\delta A_\mu^a \delta A_\nu^b}(F_{\rho\sigma}^c \frac{D^{4n}}{M_0^{4n}} F_{\rho\sigma}^c) + \frac{1}{\alpha} f^2 (\frac{\partial^2}{m^2}) \partial_\mu \partial_\nu \delta^{ab} \right] \Phi_\nu^b \ . \tag{3.6}$$

Now expand S'_{PV} in powers of A_μ, and integrate by parts so that a generic term in the series for S'_{PV} is put in the form

$$\int d^4x \frac{1}{M_0^{4n}} \Phi_\mu^a \Big\{ \alpha\ A's\ ,\ \beta\ \partial's \Big\} \Big\{ \gamma\ \partial's \Big\} \Phi_\nu^b\ , \tag{3.7}$$

where by the notation I mean that only γ of the $4n + 2 - \alpha$ derivatives act on Φ_ν^b

and the rest act on the gauge fields. The point to note is that the terms for which $\alpha + \beta \geq 5$ do not contribute to the Γ-divergences of the Φ-cycle, so we drop these terms also.

The next step is to reduce γ in the surviving terms by $4n + 2 - 2m$, by removing $2n + 1 - 2m$ powers of (∂^2/M_0^2). The resulting action is of type $2m$, and the one-loop Γ-divergences are clearly related to those of S_{PV}, i.e., to those of the gauge field. Finally, we replace M_0 by the gauge-breaking scale M_χ, and add in a mass term for the regulator field with mass-square ζM_χ^2. I shall denote these gauge-regulator fields of type $2m$ as Φ_μ^{2m}, and their action as $L_{\mathrm{g-reg}}^{2m}$.

It is quite straightforward, now, to work out the cancellation conditions for the Γ-divergences. The divergences from the gauge cycle are identical to those from S_{PV} or S_{PV}^{prime}, and so we have only to relate S'_{PV} to $L_{\mathrm{g-reg}}^{2m}$.

For the diagrams with external momenta $\{p_\mu\}$ set to zero the Γ-divergences from S'_{PV} and $L_{\mathrm{g-reg}}^{2m}$ are manifestly the same, by the construction. This is because all that's different in the Feynman integrand, at large loop momentum k, is an *equal* change in the power of k^2 in the numerator as in the denominator. When the external momenta are not set to zero, the change is in powers of $(k + p)^2$, and so Taylor expanding in p tells us that the Γ-divergences *linear* in p depend *linearly* on the type $2m$. Similarly,the Γ-divergences quadratic in p depend quadratically on $2m$, and for higher powers of p the graphs converge.

Remembering that the type of S'_{PV} is $4n + 2$, we can therefore cancel the gauge cycle divergences using a set of gauge-regulators, including an integral number a_{2m} of type $2m$, provided

$$0 = \sum_m a_{2m} + 1 \tag{3.8a}$$

$$0 = \sum_m m a_{2m} + (2n + 1) \tag{3.8b}$$

$$0 = \sum_m m^2 a_{2m} + (2n + 1)^2 \ . \tag{3.8c}$$

It is easy to find a set of integers a_{2m} such that (3.8) is satisfied. In summary, our fully regulated action for a chiral gauge theory is,

$$L_{\rm HCD} = -\frac{1}{4}F^a_{\mu\nu}\left[1+(\frac{D^2}{M_0^2})^{2n}\right]F^a_{\mu\nu} + \frac{1}{2\alpha}\left[f(\frac{\partial^2}{m^2})\partial_\mu A^a_\mu\right]^2 + i\partial_\mu\bar{\eta}^a(D_\mu\eta)^a$$

$$+ i\bar{\psi}\not{D}P_+\psi + \sum L^{2m}_{\rm g-reg} + \sum L_{\rm f-reg} + \sum L_{\rm gh-reg} \, . \qquad (3.9)$$

As in Section II.2, the regulator works due to the combination of higher covariant derivatives with regulator fields of lower type than the gauge field. Here, however, we have two scales of regularisation, with M_0 associated with operators that respect the gauge invariance, and M_χ with those that break it. The Green's functions from (16) are Γ-finite, but of course still have quadratic and logarithmic divergences in M_0 and M_χ.

2. The Regularised Ward Identity for χGT

The derivation of the regularised Ward identities proceeds analogously as in Section II.3. We couple the gauge, ghost, and chiral fields to sources J_μ, ζ, and ξ respectively, put the smooth pre-regulating terms into the action (3.9), and consider the response of the functional integral to a BRS transformation. This yields, c.f. (2.26),

$$0 = \Big\langle -\bar{\eta}^a(y)\int d^4x J^b_\mu(D_\mu\eta)^b(x) + \frac{i}{\alpha}f^2(\frac{\partial^2}{m^2})\partial_\mu A^a_\mu(y)$$
$$+ ig\bar{\eta}^a(y)\int d^4x \mathrm{tr}(\bar{\xi}P_+\eta^bT^b\psi - \bar{\psi}P_+\eta^bT^b\xi) + \bar{\eta}^a(y)\int d^4x\,\delta L_{\rm pre}\Big\rangle_{J,\bar{\xi},\xi} \qquad (3.10)$$

where $\delta L_{\rm pre}$ now is the BRS variation not only of the momentum cutoff pre-regulator Γ, but also of the intrinsically gauge non-invariant couplings of the regulator fields.

Once again, we must check the validity of (3.10): we make sure that all the terms are strictly finite by computing the superficial degree of true divergence $D^{\rm op}_\infty$. Now we take the pre-regulator off to infinity, but this time the discussion of the Γ-finite parts (the potential anomalies) is a little more involved than before. This is because of the intrinsically gauge-non-invariant couplings, and I go through this fully in Appendix 4.

The essential point, in fact, is that the BRS variation of the regulator cycles is some functional of the gauge field A_μ involving the scale M_χ. The asymptotic expansion in M_χ has leading terms M_χ^2 and 1, multiplied by local polynomials of A_μ and derivatives, of dimension 2 and 4 respectively. Importantly, the remainder is bounded by inverse powers of M_χ.

Now, we can easily remove the quadratic M_χ-divergences from *all* Green's functions, with or without an insertion of δL_{pre}, by appending to (3.5) and (3.8) the conditions,

$$0 = \sum_\alpha f_\alpha \kappa_\alpha^2 \log \kappa_\alpha \tag{3.11a}$$

$$0 = \sum_\beta g_\beta \lambda_\beta^2 \log \lambda_\beta \tag{3.11b}$$

$$0 = \sum_\gamma a_{2m,\gamma}\, \zeta_{2m,\gamma}^2 \ , \tag{3.11c}$$

where in (3.11c) $\zeta_{2m,\gamma}^2 M_\chi^2$ is the mass-square of the γ-th gauge regulator of type $2m$.

Moreover, in the case of a vanishing ABBJ anomaly, the M_χ-finite term in δL_{pre} can be canceled by adding to the action (3.9) a compensating operator $\hbar P_4[gA_\mu]$, so only the inverse powers of M_χ remain. This is because the ABBJ anomaly is the unique possible one-loop obstruction, c.f. (A3.1). Now, this property will hold for operator insertions of δL_{pre} *to all loop orders*, since by (3.11) there is no source of positive powers of M_χ. By this reasoning, we conclude that the compensated action yields the identity

$$0 = g \int d^4x \left[i f^{abc} J_\mu^{b\text{tr}} \frac{\delta G_{\text{reg}}^{ad}}{\delta J_\mu^c} + \text{tr}(\bar{\xi} P_+ T^b \frac{\delta G_{\text{reg}}^{ab}}{\delta \bar{\xi}} - \frac{\delta G_{\text{reg}}^{ab}}{\delta \xi} P_+ T^b \xi) \right]$$

$$+ \hbar\partial^{-2}\partial_\mu J^a_\mu(y) + \frac{1}{\alpha}f^2(\frac{\partial^2}{m^2})\partial^y_\mu \frac{\delta W_{\rm reg}}{\delta J^a_\mu(y)} + O(\frac{1}{\Gamma^2}) + O(\frac{1}{M_\chi^2}) \ . \qquad (3.12)$$

Equation (3.12) is the regularised Ward identity for the chiral gauge theory. It differs from its analogue (2.30) in the non-chiral case in that after Γ has been sent to infinity there are still two scales left, M_χ and M_0. Here M_χ cannot yet be taken to infinity, since the Green's functions in (3.7) contain, of course, logarithmic M_χ-divergences. However, if we can find bare couplings depending on M_χ and M_0 such that the sequential limit (3.2) can be taken, then (3.12) will indeed reproduce the *renormalised* Ward identity, which is the main goal of the discussion.

3. The Renormalised Ward Identity for χGT

The proof of the renormalised Ward identities from the regularised identities (3.12) again follows almost exactly the analysis in Section II.4. There, all we needed to know was that for some choice of relevant operators (not necessarily a gauge-invariant choice) the M_0-divergences could be removed from the low energy Green's functions. This was guaranteed by Polchinski's theorem on effective Lagrangian flow. Combining this fact with the regularised Ward identity gave us the 'one-loop renormalised' Ward identity, and from there the 'two-loop' identity, and so on inductively to all loops.

Here, if I take $M_\chi = XM_0$, then all we'd need to know is that for some choice of relevant operators (not necessarily gauge-invariant) *both* M_0 and X can be taken to infinity.

This again is true for M_0, (by Polchinski), but the effective Lagrangian flow arguments have difficulty dealing with X. However, we note that the X-divergences, which are logarithmic, only enter through regulator cycles. In particular these divergences are disjoint, i.e. non-overlapping, so that once they have been removed at the one-loop level they have been removed altogether. The argument in Section II.4 now goes through, and so the renormalised Ward identities hold.

4. The Adler-Bardeen Theorem

The arguments in this chapter prove, finally, that a vanishing one-loop ABBJ anomaly is a *sufficient* condition for a gauge theory to be renormalisable, Lorentz-invariant and unitary. The key ingredients have been the effective Lagrangian flow and the physical regulator with two mass scales. The technique I've described can naturally be applied to many perturbation theory questions, and in the next section I shall exhibit the application to super Yang-Mills. Here, I conclude with an application of the technique to prove a very famous result, namely the Adler-Bardeen theorem for the axial $U(1)_A$ anomaly.

This states that under some suitable scheme of renormalisation,

$$\partial_\mu j_{\mu 5} = \frac{1}{16} g^2 F^a_{\mu\nu} \tilde{F}^a_{\mu\nu} \, , \tag{3.13}$$

where $j_{\mu 5}$ is the non-gauged axial current of fermions ψ coupled to a gauge field A^a_μ. To prove this we consider the regulated action, c.f. (3.9),

$$L_{\text{HCD}} = -\frac{1}{4} F^a_{\mu\nu} \left[1 + (\frac{D^2}{M_0^2})^{2n} \right] F^a_{\mu\nu} + \frac{1}{2\alpha} \left[f(\frac{\partial^2}{m^2}) \partial_\mu A^a_\mu \right]^2 + i \partial_\mu \bar{\eta}^a (D_\mu \eta)^a$$

$$+ i\bar{\psi} \not{D} \psi + \sum i\bar{\phi}(\not{D} + i\kappa M_\chi)\phi + \sum L^{2m}_{\text{g-reg}} + \sum L_{\text{gh-reg}} \, . \tag{3.14}$$

In (3.14) the ϕ fields are the Pauli-Villars regulators, with mass κM_χ, for the fermions ψ. As in (3.3), the fermion action is to be considered as a *representative*, this time of any action with a global axial symmetry.

Now we know already that this regulated action can be renormalised with wave-function and coupling renormalisations, Z_γ, Z_ψ, and $\tilde{g}$. Including these bare parameters so that the Green's functions are finite, we perform an axial $U(1)_A$ rotation on the fields, given by

$$\delta\psi = i\alpha\gamma_5\psi \tag{3.15a}$$

$$\delta\phi = i\alpha\gamma_5\phi \tag{3.15b}$$

$$\delta A^a_\mu = 0 \ , \tag{3.15c}$$

where we have sources J_μ, $\bar{\xi}$ and ξ coupled respectively to A_μ, ψ and $\bar{\psi}$. This yields the Ward identity

$$0 = \left\langle \partial_\mu j_{\mu 5} + i(\bar{\xi}\gamma_5\psi - \bar{\psi}\gamma_5\xi) - 2\textstyle\sum\kappa M_\chi\bar{\phi}\gamma_5\phi \right\rangle_{J,\bar{\xi},\xi} \tag{3.16}$$

where

$$j_{\mu 5} = Z_\psi\bar{\psi}\gamma_\mu\gamma_5\psi + \textstyle\sum\bar{\phi}\gamma_\mu\gamma_5\phi \ . \tag{3.17}$$

The breaking term due to the Pauli-Villars fields can easily be evaluated using Feynman diagrams [4], and we have

$$\langle\textstyle\sum\kappa M_\chi\bar{\phi}\gamma_5\phi\rangle_{A_\mu} = -\frac{1}{32\pi^2}\tilde{g}^2 F^a_{\mu\nu}\tilde{F}^a_{\mu\nu} + O(\frac{1}{M_\chi^2}) \ ; \tag{3.18}$$

thus in (3.16)

$$0 = \left\langle \partial_\mu j_{\mu 5} + i(\bar{\xi}\gamma_5\psi - \bar{\psi}\gamma_5\xi) - \frac{1}{16\pi^2}\tilde{g}^2 F^a_{\mu\nu}\tilde{F}^a_{\mu\nu} \right\rangle_{J,\bar{\xi},\xi} + O(\frac{1}{M_\chi^2}) \ . \tag{3.19}$$

As in the discussion of chiral gauge theories, we insist that the regularised graphs have no quadratic M_χ-divergences, c.f.(3.11), so the $O(1/M_\chi^2)$ term will drop out as $M_\chi \to \infty$.

Now, before we take the regulation scales to infinity we must renormalise the operator insertions (the uninserted Green's functions having been renormalised already). From effective Lagrangian flow arguments we know that $\partial_\mu j_{\mu 5}$ can be made finite by a subtraction S and a multiplicative constant C. Performing this operation on (3.19), and setting $\bar{\xi} = \xi = 0$ yields

$$C\left\langle \partial_\mu j_{\mu 5} - S \right\rangle_J = C\left\langle \frac{1}{16\pi^2}\tilde{g}^2 F^a_{\mu\nu}\tilde{F}^a_{\mu\nu} - S \right\rangle_J + O(\frac{1}{M_\chi^2}) \ , \tag{3.20}$$

and now the sequential limit $M_\chi \gg M_0 \gg E_{\mathrm{phys}}$ can be taken. The LHS of (3.20) is finite by construction, and hence so is the RHS. This defines the appropriate renormalisation of $F\tilde{F}$, and proves, by a remarkably economical argument, the $U(1)_A$ Adler-Bardeen theorem.

5. Supersymmetry

We are now ready to discuss $N=1, d=4$ Super Yang-Mills, and in this section we shall work with a manifestly supersymmetric formalism through the use of superfields.

As an introduction to the use of superfields in the method of effective Lagrangians, let me write down the flow equations for the theory of a single, self-coupled chiral super-multiplet. The notation is as in Ref. [5]. Splitting the action up into a kinetic term, with momentum cutoff factor $K^{-1}(p^2/\Lambda^2)$, and an interaction $S(\Lambda)$, we have

$$L = \int d^8 z \frac{1}{2} (\phi \quad \bar{\phi}) \begin{pmatrix} \frac{-mD^2}{4\partial^2} & 1 \\ 1 & \frac{-m\bar{D}^2}{4\partial^2} \end{pmatrix} K^{-1}(\frac{\partial^2}{\Lambda^2}) \begin{pmatrix} \phi \\ \bar{\phi} \end{pmatrix} + S(\Lambda) \,, \qquad (3.21a)$$

and the equation of motion is, c.f. (1.20c),

$$\Lambda \frac{\partial S}{\partial \Lambda} = -\frac{1}{2} \mathrm{tr} \int d^8 z \int d^8 z' \, \Lambda \frac{\partial}{\partial \Lambda} K(\frac{\partial^2}{\Lambda^2}) \, \Delta_{\mathrm{GRS}}(z, z') \left[\frac{\delta^2 S}{\delta \vec{\phi}(z) \delta \vec{\phi}(z')} + \frac{\delta S}{\delta \vec{\phi}(z)} \frac{\delta S}{\delta \vec{\phi}(z')} \right] \,.$$

$$(3.21b)$$

Here $\vec{\phi}$ denotes the vector $(\phi, \bar{\phi})$, and the Grisaru-Roček-Siegel propagator is given by

$$\Delta_{\mathrm{GRS}} = \frac{1}{\partial^2 - m^2} \begin{pmatrix} \frac{mD^2}{4\partial^2} & 1 \\ 1 & \frac{m\bar{D}^2}{4\partial^2} \end{pmatrix} \delta^8(z - z') \,. \qquad (3.21c)$$

By power counting, the perturbatively relevant operators are $\int d^4\vartheta \, \bar{\phi}\phi$ and $\int d^2\vartheta \, \phi^r$ for $r = 1, 2, 3$, since ϕ, ϑ and $d\vartheta$ have dimensions 1, $-\frac{1}{2}$ and $\frac{1}{2}$ respectively. Polchinski's reasoning applied to (3.21b) thus tells us that there exist bare couplings for these operators, at the high scale $\Lambda = \Lambda_0$, such that the Green's functions are rendered Λ_0-finite. Moreover, a consideration of the 'D-algebra' shows that the purely chiral operators $\int d^2\vartheta\phi^r$ are in fact Λ_0-finite [5], so the only Λ_0-divergence comes in the wavefunction renormalisation, $Z_{\bar{\phi}\phi}$.

Regularisation of $N = 1$ Super Yang-Mills

The classical action, without gauge fixing, for $N = 1$ super Yang-Mills is

$$L_{N=1} = \int d^2\vartheta \left[\frac{1}{4g^2}\mathrm{tr}W^\alpha W_\alpha + m_{ij}\phi_i\phi_j + g_{ijk}\phi_i\phi_j\phi_k\right] + \mathrm{h.c.} + \int d^4\vartheta\bar{\phi}_i e^{gV}\phi_i \ , \tag{3.22}$$

where ϕ is a chiral superfield *representative* of any matter field content, and V is the vector superfield containing the gauge boson. The non-abelian field strength W_α is given by $W_\alpha = -\frac{1}{4}\bar{D}^2 e^{-gV}D_\alpha e^{gV}$. The 'super-gauge' transformations respected by (3.22) are

$$\phi \longrightarrow e^{i\Lambda}\phi \tag{3.23a}$$

$$\bar{\phi} \longrightarrow \bar{\phi}e^{-i\bar{\Lambda}} \tag{3.23b}$$

$$e^{gV} \longrightarrow e^{i\bar{\Lambda}}e^{gV}e^{-i\Lambda} \ , \tag{3.23c}$$

and we can fix these out in a manifestly supersymmetric way by using the gauge-fixing functions $\bar{D}^2V$ and D^2V. Thus we have the action for the gauge-fixing terms and ghosts [6],

$$L_{\mathrm{gf}} = \frac{1}{8}(D^2V)f(\bar{D}^2V) + (\eta+\bar{\eta})L_{\frac{gV}{2}}\left[(c+\bar{c}) + \coth L_{\frac{gV}{2}}(c-\bar{c})\right] \ , \tag{3.24}$$

where η and c are the ghost chiral superfields, and f is a 'gauge weighting' factor (which can be a function of spacetime derivatives). In (3.24) implied is a trace over the group and an integration $\int d^4\vartheta$. The total action, $L_{N=1}+L_{\rm gf}$, then has a superfield BRS invariance given by [6],

$$\delta V = \xi \, L_{\frac{gV}{2}} \left[(c+\bar{c}) + \coth L_{\frac{gV}{2}} (c-\bar{c}) \right] \tag{3.25a}$$

$$\delta c = -\xi \, c^2 \tag{3.25b}$$

$$\delta \eta = -\frac{\xi}{8g} f \bar{D}^2 D^2 V \tag{3.25c}$$

$$\delta \phi = \xi \, c\phi \ , \tag{3.25d}$$

with analogous variations for $\bar{c}$, $\bar{\eta}$ and $\bar{\phi}$.

It is worth reminding ourselves at this point of the difficulties that standard renormalisation theory faces in dealing with $N=1$ Super Yang-Mills. The superfield V is dimensionless, so the number of primitive divergences is actually infinite. Moreover V is massless, so the infrared and ultraviolet divergences can become entwined in a very complicated way. However, in our *physical* scheme of regularisation and renormalisation these problems dissolve.

Let us now regulate the total action in our familiar way of combining higher covariant derivatives with regulator fields. The higher covariant derivatives in the vector superfield action enter with scale M_0, and the ghost-, matter-, and gauge-regulator fields, whose action is constructed exactly as in (3.9), have mass scale M_χ. Note that once again we must choose the gauge-weighting function $f(\partial^2/m^2)$ such that its inverse decays rapidly at large momentum, in order to have the analogue of the 'transverse gauges,' see Eq. (2.12).

To be explicit, the kinetic term for V in the regulated action is

$$\frac{1}{8} D^\alpha V \left[1 + (\frac{\partial^2}{M_0^2})^{2n} \right] \bar{D}^2 D_\alpha V + \frac{1}{16} V (\bar{D}^2 D^2 + D^2 \bar{D}^2) f V \ , \tag{3.26a}$$

yielding a V-propagator,

$$\left[\frac{1}{1+(\partial^2/M_0^2)^{2n}} P_T + \frac{1}{f}(1-P_T) \right] \partial^{-2} \, \delta^8(z-z') \ , \tag{3.26b}$$

where P_T is the projection operator $-\frac{1}{8\partial^2} D\bar{D}^2 D$. A suitable choice of $f(\cdot)$, therefore, to bring us to a 'super transverse gauge' is

$$f(\frac{\partial^2}{m^2}) = 1 + (\frac{\partial^2}{m^2})^{2n+2p} \qquad , p > 0 \; . \tag{3.26c}$$

Now the gauge propagator is proportional to P_T at large momentum and is of type $4n+2$. At small momentum the propagator tends to ∂^{-2}, so that there's no problem in the infrared: the Green's functions are to be renormalised for *unexceptional* momenta, just as in ordinary gauge theories. Of course we shall renormalise using the method of effective Lagrangians, which separates completely the infrared and ultraviolet divergences.

In summary, the fully regulated, manifestly supersymmetric action for $N = 1$ super Yang-Mills is

$$\begin{aligned} L^{\text{reg}}_{\text{HCD}} &= \text{tr} \int d^4\vartheta \left[\frac{1}{8} V \bar{D}^2 D^2 f V + (\eta + \bar{\eta}) L_{\frac{gV}{2}} (c + \bar{c}) + \coth L_{\frac{gV}{2}} (c - \bar{c})] + \bar{\phi}_i e^{gV} \phi_i \right] \\ &+ \text{tr} \int d^2\vartheta \left[\frac{1}{4g^2} W^\alpha \left\{ 1 + (\frac{\partial^2_{\text{cov}}}{M_0^2})^{2n} \right\} W_\alpha + m_{ij}\phi_i\phi_j + g_{ijk}\phi_i\phi_j\phi_k \right] + \text{h.c.} \\ &+ L_{\text{matter}} + \sum L^{2m}_{\text{g-reg}} + \sum L_{\text{gh-reg}} + \sum L_{\text{matt-reg}} \; , \end{aligned} \tag{3.27}$$

where the ghost- and matter- regulators are simply Pauli-Villars fields, and the gauge-regulators are derived from Pauli-Villars in the manner described for chiral gauge theories in Section III.1.

Inserting the pre-regulating momentum cutoff Γ, the regularised Ward identities can be derived from (3.27), using the BRS variations (3.25), just as before. The discussion of the Γ-divergences and Γ-finite parts, and the existence of a compensating operator then follows exactly that of Appendix 4. Thus in the case of canceling one-loop anomaly we have the identity, c.f. (3.12),

$$0=\left\langle\int d^8z' J(z')\langle\eta(z)L_{\frac{gV}{2}}\left[(c+\bar{c})+\coth L_{\frac{gV}{2}}(c-\bar{c})\right](z')-\frac{1}{8}f\bar{D}^2D^2V(z)\right\rangle_J^{\text{reg}}$$
$$+O(\frac{1}{\Gamma^2})+O(\frac{1}{M_\chi^2})\,.\quad(3.28)$$

Once again anomaly freedom at one-loop will be a sufficient condition for the theory to be renormalisable and unitary. Now, the one-loop anomaly has been studied by various authors [7], and has been completely characterised by algebraic and cohomological arguments. It was found that the unique possible one-loop anomaly is merely the supersymmetric extension of the ordinary ABBJ anomaly, and therefore its cancellation is guaranteed by the ordinary ABBJ cancellation condition, (3.1).

Renormalisation and Unitarity

The proof of the renormalised Ward identities from (3.28) follows the arguments presented above for chiral gauge theories. Despite the infinitude of primitively divergent graphs, we find that in the 'super transverse gauge' (3.26c) there are only *two* divergent bare couplings, namely the linear wavefunction renormalisations of V and ϕ. The mass- and self- couplings of the matter fields, m_{ij} and g_{ijk}, are unrenormalised, just as in a pure matter theory, and so is the gauge coupling g. This last can be seen in Fig. 4, and occurs for the same reason that the ghost-ghost-gauge vertex $\partial_\mu\bar{\eta}[A_\mu,\eta]$ in an ordinary gauge theory is unrenormalised when computed in a transverse gauge. See Section II.4. Thus the renormalised action at the high scale takes on the simple form

$$L\text{ren}=\operatorname{tr}\int d^4\vartheta\left[\frac{1}{8}V\bar{D}^2D^2fV+(\eta+\bar{\eta})L_{\frac{gV}{2}}[(c+\bar{c})+\coth L_{\frac{gV}{2}}(c-\bar{c})]+Z_{\bar{\phi}\phi}\bar{\phi}e^{gV}\phi\right]$$
$$+\operatorname{tr}\int d^2\vartheta\left[\frac{1}{4g^2}Z_VW^\alpha W_\alpha+m_{ij}\phi_i\phi_j+g_{ijk}\phi_i\phi_j\phi_k\right]+\text{h.c.}$$

$$+ \text{ higher derivative terms } + \text{ regulator terms },\tag{3.29}$$

where Z_V and $Z_{\bar{\phi}\phi}$ are logarithmic divergences in M_0 and M_χ. Note that $Z_{\bar{\phi}\phi}$ is actually also finite when $g_{ijk} = 0$.

In other supersymmetric gauges, e.g., the choice $f(\cdot) = 1$, the renormalisation programme is much more difficult to carry out, since V does not renormalise simply as $V \rightarrow Z_V V$. Rather there is a *functional* renormalisation,

$$V \rightarrow \alpha_1 V + \alpha_2 V^2 + \alpha_3 V^3 + \ldots ,\tag{3.30}$$

and there is an *infinite* number of renormalisation constants. A direct proof of the renormalised Ward identities in such a gauge is beyond the techniques I've presented, but fortunately no such direct proof is necessary.

To see this, let us go back to the discussion of unitarity in Section II.6. There we approached a given gauge choice **G**, the 'axial' gauge in fact, as a limit of a one-parameter family of transverse gauges. This showed that the S-matrix computed in **G** is the same as computed in any of the transverse gauges, and therefore that it is renormalisable (as well as unitary). The same argument can be applied here since *any* gauge choice, supersymmetric or not, can be approached as a limit of 'super transverse gauges.' Thus, for $f(\cdot) = 1$ in particular, all but one of the renormalisation constants in (3.30) are gauge artifacts. To conclude, note that we can also reach the 'axial Wess-Zumino gauge,' in which there are neither ghosts nor unphysical auxiliary fields. Therefore, the combination of the renormalisation result (3.29) and the 'limit argument' proves, in one fell swoop, that $N = 1$ super Yang-Mills possesses a renormalised, Lorentz-invariant, supersymmetric and unitary S-matrix.

6. Gravitational Anomalies

In this chapter we have proven that gauge theories, with or without rigid $N = 1$ supersymmetry, are renormalisable, provided they satisfy the ABBJ condition (3.1). Moreover, by power counting there exist no other renormalisable theories in $d = 4$, and so from Wilson's insight we can say that the physical low-energy model which desribes our world *must* be of this class. It's tempting to also conclude that condition (3.1) is the only *a priori* constraint on low energy phenomenology, but this would be too hasty.

There is another restriction in fact, that we must append to (3.1), *if* we demand that our low-energy model be derived from a higher theory that contains *gravity*. This is due to the so-called 'gravitational anomalies.'

Now, this statement at first glance seems a little strange. If it's really true that the low-energy limit of any higher theory is *decoupled* from gravity, up to inverse powers of the cutoff $m_{\rm Pl}$, then surely I can use a 'bare Lagrangian' at the high scale which is also decoupled. But if the bare Lagrangian doesn't contain gravity, there can be no gravitational anomalies; *therefore* it is enough to just satisfy (3.1).

This argument would be correct if we were happy for the bare Lagrangian to contain all the possible relevant operators. However, for gauge theories this is not the case. Rather we demand some of the relevant operators to be absent, in order to guarantee the renormalised Ward identities and hence unitarity. It is quite conceivable that the 'irrelevant' gravity couplings induce *unwanted* relevant couplings, thereby ruining the low-energy gauge symmetry.

In order to see the non-trivial effect of the presence of gravity, we take the action (3.9) and make it 'Generally Covariant'. This is done by introducing the vielbein e^a_μ and the spin connection ω. A review of the relevant differential geometry is given in [15]. We have

$$\begin{aligned}\frac{1}{\sqrt{-g}}L = &-\frac{1}{4}g^{\mu\rho}g^{\nu\sigma}F^a_{\mu\nu}\left[1+(\frac{\nabla^2}{M_0^2})^{2n}\right]F^a_{\rho\sigma}+\frac{1}{2\alpha}\left[f(\frac{g^{\rho\sigma}\partial_\rho\partial_\sigma}{m^2})g^{\mu\nu}\partial_\mu A^a_\nu\right]^2 \\ &+ig^{\mu\nu}\partial_\mu\bar{\eta}^a(\nabla_\mu\eta)^a+i\bar{\psi}\not{\mathcal{D}}P_+\psi+M_0^2\mathcal{R} \\ &+\sum L^{2m}_{\rm g-reg}+\sum L_{\rm f-reg}+\sum L_{\rm gh-reg}+\sum L_{\rm grav-reg}\ , \qquad (3.31a)\end{aligned}$$

where, in the language of exterior derivatives,

$$\nabla = \mathrm{d} + \mathbf{A} + \mathbf{\Gamma} \tag{3.31b}$$

$$\not{D} = \mathrm{d} + \mathbf{A} + \omega \tag{3.31c}$$

and $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ is the metric, $\mathbf{\Gamma}$ is the Christoffel connection, and $\mathcal{R}$ is the Ricci scalar.

In (3.31a) the mass scale of the gauge-invariant and generally covariant terms is M_0, which we can think of as m_{Pl}, and the mass scale of the regulators is $M_\chi \gg M_0$ as in (3.9). The $h_{\mu\nu}$ 'graviton' loops are regulated just like the gauge bosons, i.e., we include higher derivatives into $\mathcal{R}$ and add extra regulating particles, $L_{\mathrm{grav-reg}}$, c.f. Section 1.

Now let us try to derive the regularised Ward identities for BRS invariance exactly as before, (where the BRS variations of the graviton and graviton-regulators are defined to be zero). Once again the possible obstructions G are completely classified at the one-loop level, and must be of the form, c.f. (A3.1),

$$G = \eta \times \{\dim 4\,\mathrm{polynomial\,of}\,A_\mu\,,\ h_{\mu\nu}\,\mathrm{and\,derivatives}\,\}\ , \tag{3.32}$$

such that $\delta_{\mathrm{BRS}} G = 0$. There are now *two* solutions for G, namely the ABBJ anomaly (A3.1), and,

$$G_{\mathrm{grav}} = \varepsilon^{\mu\nu\rho\sigma}\,\mathrm{tr}\,\eta\, R_{\mu\nu}^{\ \ \kappa\lambda} R_{\rho\sigma\kappa\lambda}\ , \tag{3.33}$$

where in (3.33) I've taken the linearised form using $h_{\mu\nu}$ and made it generally covariant [16].

We see that from (3.33) we must demand a representation $\mathbf{R}$ of the chiral fermions such that

$$\mathrm{tr}\ T^a_{\mathbf{R}} = 0\ . \tag{3.34}$$

Of course (3.34) is automatically satisfied for a semi-simple Lie group, and is only a constraint on the $U(1)$ factors.

7. Summary

In this chapter we have characterised the two sources of obstruction to perturbatively quantising a gauge theory. These are the ABBJ and gravitational anomalies, and when absent at the one-loop level they are absent to all loop orders.

The proof of this result made heavy use of a physical, mass-scale regulator that broke the BRS invariance in a *controllable* way. For this we needed, in fact, *two* scales of regularisation, but this caused us no difficulty, and we could still rely on the effective Lagrangian flow arguments of Chapter I. It was the use of these arguments, of course, that allowed us to extend the renormalisation theorems to all loop orders, and allowed us to understand them, for the first time, in a simple and intuitive way. In particular we did not have to worry about overlapping divergences and infrared divergences, which is fortunate since both these issues are very complicated and completely unphysical.

Appendix 4. 'The Compensating Operator for χGT'

This appendix discusses the Γ-finite parts of the gauge-breaking insertion δL_{pre} in (3.10). The aim is to show, as in Chapter II, that there is a compensating operator $\hbar P_4[gA_\mu]$ that can be added to the action (3.9) to cancel the Γ-finite part of δL_{pre}.

The terms in δL_{pre} due to the momentum cutoff pre-regulator can be analysed exactly as in Appendix 3 and so all that remains are the operator insertions due to the intrinsic gauge non-invariance of the regulator couplings. I show these insertions in Fig. 1, with their leading powers of momentum. Now, remember that we need only consider *one-loop* graphs, for if we cancel the Γ-finite parts in those, then they are canceled to *all* loop orders. First, I shall look at those graphs from Fig. 1 where the regulator fields do *not* form a closed cycle. I distinguish between two classes of these graphs. There are those, exemplified in Fig. 2, which contain at least one internal regulator line, and there are those, in Fig. 3, which have all the regulator lines external.

The Fig. 2 graphs are Γ-finite, and hence represent a potential obstruction, but they vanish as a power of M_0/M_χ due to the presence of the heavy propagators. The Fig. 3 graphs don't have heavy propagators, but on the other hand we're only interested in Green's functions where the regulator fields are internal only. Thus in some higher loop diagram the regulator fields in Fig. 3 must close up somewhere, reducing the graphs to the case of Fig. 2, or to the case where the regulator fields form a closed cycle. For the closed cycles of regulator fields we may invoke the Wess-Zumino consistency condition, as we did in Appendix 3, to characterise the Γ-finite parts. By (3.11) we have no quadratic M_χ-divergences, and the M_χ-finite terms can indeed be removed by a compensating operator provided the one-loop anomaly cancels. One may think there is a $\log M_\chi$ term, but actually the logarithm can only appear as $\log M_\chi/\Gamma$, which is not Γ-finite, so this has already been taken care of in (3.10).

This concludes the proof that a compensating operator can be found such that the regularised Ward identity (3.12) holds up to terms $O(M_0/M_\chi)^2$.

CONCLUSION
TOPICS IN RENORMALISATION THEORY

I have been dealing, in this thesis, with questions in the admittedly rather technical field of renormaliation theory. However, the reasoning has had a strong physical motivation, and I hope the presentation has de-mystified the subject somewhat.

Now, there still remain unresolved issues in renormalisation theory, and in this concluding chapter I shall describe some of them.

Equivalence of Different Regulators

A truly basic question is whether or not *all* consistent schemes of regularisation and renormalisation are equivalent. The sheer generality of this problem makes it very hard to answer.

For a start, we can prove that the 'physical' regulator introduced in Chapters II and III are equivalent to an ordinary momentum cutoff. One simply runs the scale M_0 (or M_0 and M_χ) *up* to the pre-regulating scale Γ and then integrates out the regulating particles. This produces an action which (a) is well defined with coefficients of operators depending only on Γ and E_{phys}, and (b) yields renormalised Green's functions satisfying the renormalised Ward identity. This action is of course very complicated, and contains BRS-non-invariant operators of arbitrary dimension and arbitrary order in $\hbar$ (as predicted in Section II.1). Plausibly an argument like this can be used to show the equivalence of all schemes based on a mass scale.

Now, comparing a momentum cutoff with BPHZ, in the former one performs the Feynman integrals and then subtracts the Γ-divergent parts, and in the latter one subtracts the integrand first and then performs the integral. It was the great achievement of Zimmermann [17] to find the right definition of integrand subtraction such that it reproduced the results of using a momentum cutoff and counterterms.

Unfortunately, it gets harder to show that a momentum cutoff is equivalent to dimensional regularisation or point-splitting, for the nature of the regulators seems so very different. Indeed, there has been recent work by Siopsis [18], which suggests that a certain definition of point-splitting allows one to quantise an anomalous gauge theory, which is a result definitely inconsistent with the renormalisation scheme I've presented here.

The Use of 'Dimensional Reduction'

For anomaly-canceled theories like the Standard Model, Chanowitz et al. [19] claim that it is fully satisfactory to use dimensional regularisation with the Dirac algebra in $d = 4$ rather than $d = 4 - \epsilon$. In particular, γ_5 is set to anti-commute with γ_μ. (Of course this presupposes the 'Sufficiency Theorem' established in this thesis.)

This prescription is analogous to using 'dimensional reduction' in an anomaly-canceled supersymmetric theory, (where the supersymmetry algebra is performed in $d = 4$ and the loop integrals are in $d = 4 - \epsilon$). From a computational point of view it is important to know if these prescriptions are indeed valid, for they are by far simpler to use than ordinary dimensional regularisation. However they are difficult to justify, since *strictly* the use of the $d = 4$ algebra is inconsistent [20]. The 'catch-22' is that of course the $d = 4$ algebra is inconsistent, else there would be no anomalies at all !

Finite Theories

It would be pleasing from a pedagogical point of view to classify which quantum field theories are perturbatively finite [21], and to understand how they behave in the continuum limit. In $d = 2$ such a classification may be of interest for string theory [22].

Triviality of $\lambda\phi^4$ and QED

I end with a question that is as old as the hills, or is at any rate pretty old [23]. Does QED exist in the continuum limit ? A greater understanding of the renormalisation group flow may lead us to be able to settle this, where one's naïve intuition at present would give the answer 'No,' because the theory is not 'asymptotically free.' However, there has been much recent speculation [24] that the related theory $\lambda\phi^4$ *does* have a non-trivial phase, so this famous question is still open.

REFERENCES

[1] G.'t Hooft and M. Veltman, 'Diagrammar,' CERN report 73-9.

[2] C. Quigg, 'Gauge Theories of the Strong, Weak and Electromagnetic Interactions,' Benjamin/Cummings (1983).

[3] M. Dine, R. Rohm, N. Seiberg and E. Witten, Phys. Lett. B156 (1985) 55.

[4] J. Collins, 'Renormalisation,' Cambridge University Press (1984).

[5] G. Costa, J. Julve, T. Marinucci and M. Tonin, Nuovo Cimento 38A (1977) 373.

[6] L.V. Avdeev and A.A. Vladimirov, Nucl. Phys. B219 (1983) 262.

[7] W. Zimmermann, Brandeis Lectures (1970), eds. S. Deser, M. Grisaru and H.Pendleton, MIT Press.

J.H. Lowenstein, Phys. Rev. D4 (1971) 2281.

[8] T.E. Clark and J.H. Lowenstein, Nucl. Phys. B113 (1976) 109.

[9] K.G. Wilson and J.G. Kogut, Phys. Reports 12 (1974) 75.

[10] J. Polchinski, Nucl. Phys. B231 (1984) 269.

[11] L.D. Faddeev and A.A. Slavnov, 'Gauge Fields, Introduction to the Quantum Theory,' Benjamin/Cummings (1980).

[12] C. Becchi, A. Rouet and R. Stora, Ann. Phys. 98 (1976) 287.

[13] J. Wess and B. Zumino, Phys. Lett. 37B (1971) 95.

[14] S.L. Adler and W.A. Bardeen, Phys. Rev. 182 (1969) 1517.

[15] L. Alvarez-Gaumé and P. Ginsparg, 'The Structure of Gauge and Gravitational Anomalies,' HUTP-84/A016.

[16] R. Delbourgo and A. Salam, Phys. Lett. 40B (1972) 381.

[17] W. Zimmermann, Comm. Math. Phys. 15 (1969) 208.

[18] G. Siopsis, Caltech Thesis (unpublished).

[19] M. Chanowitz, M. Furman and I. Hinchliffe, Nucl. Phys. B159 (1979) 225.

[20] W. Siegel, Phys. Lett. 94B (1980) 37.

R. van Damme and G.'t Hooft, Phys. Lett. 150B (1985) 133.

[21] O. Piguet and K. Sibold, 'Non-renormalisation Theorems of Chiral Anomalies and Finiteness in Supersymmetric Yang-Mills Theories,' MPI-PAE/PTh 31/86.

[22] L. Alvarez-Gaumé and P. Ginsparg, 'Finiteness of Ricci flat Supersymmetric non-linear σ-models,' HUTP-85/A016.

[23] L.D. Landau, A.A. Abrikosov and I.M. Khalatnikov, Doklady Akad. Nauk SSSR 95 (1954) 1177.

[24] P.M. Stevenson and R. Tarrach, Phys. Lett. B176 (1986) 436.

Figure 1

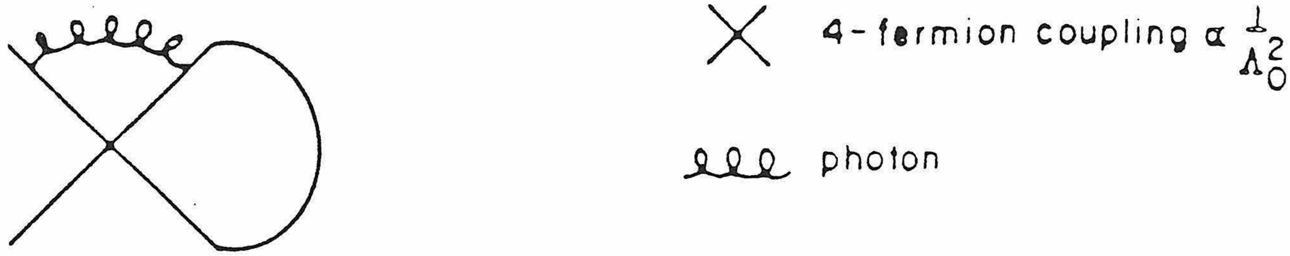


Figure 2

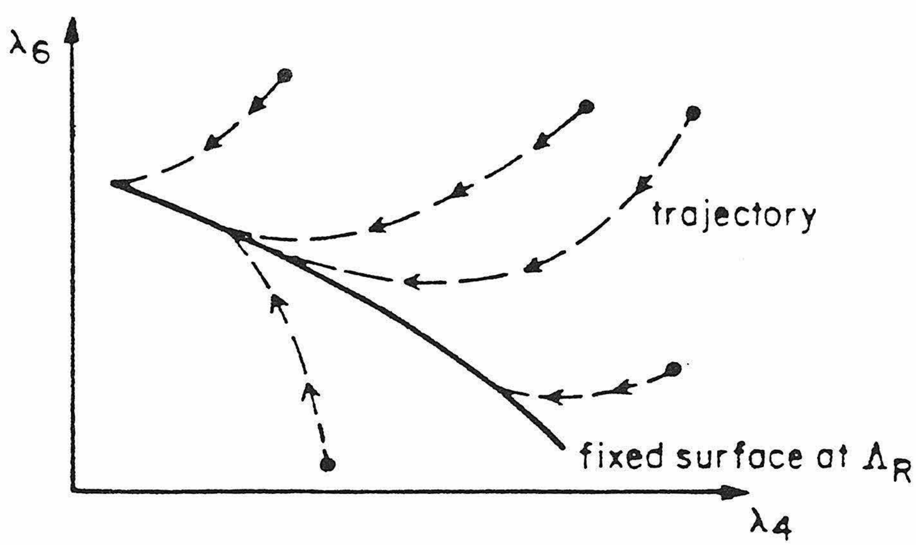


Fig 3

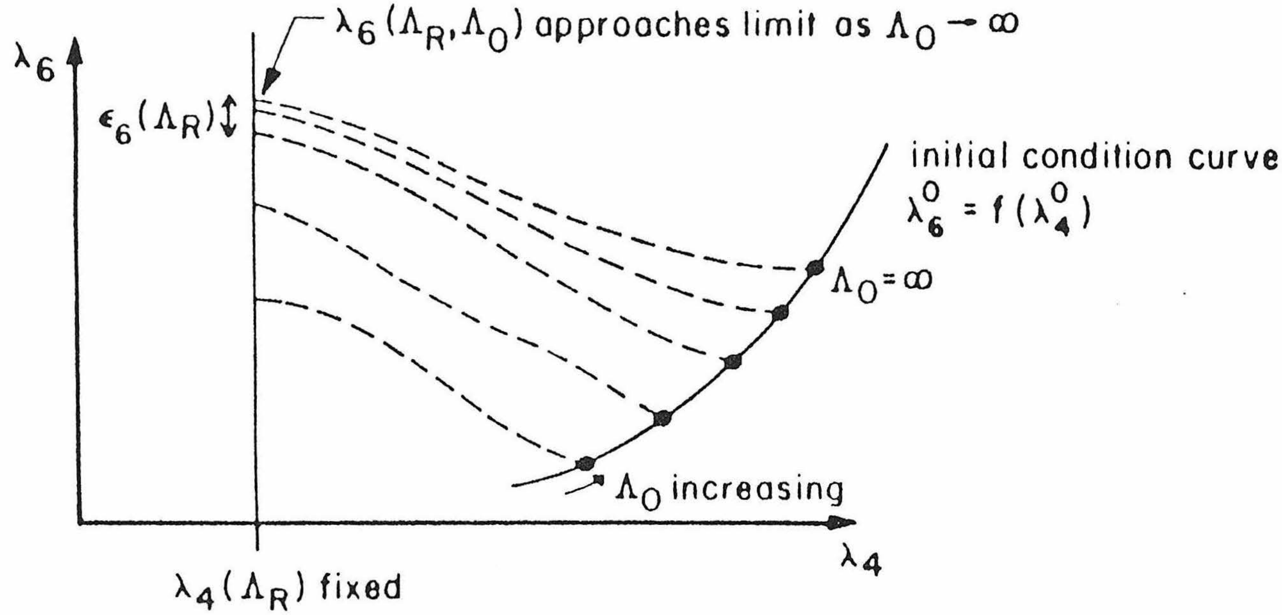

$\lambda_6(\Lambda_R,\Lambda_0)$ approaches limit as $\Lambda_0 \to \infty$
λ_6
$\epsilon_6(\Lambda_R)$
initial condition curve
$\lambda_6^0 = f(\lambda_4^0)$
$\Lambda_0 = \infty$
Λ_0 increasing
λ_4
$\lambda_4(\Lambda_R)$ fixed

Fig 4

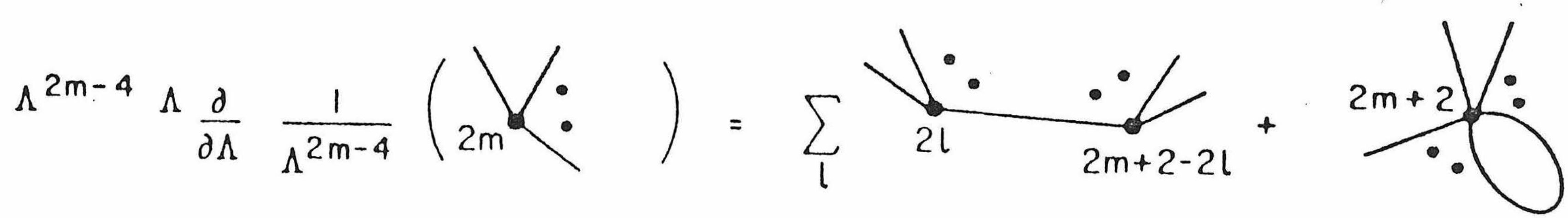

$\Lambda^{2m-4}\ \Lambda \frac{\partial}{\partial\Lambda}\ \frac{1}{\Lambda^{2m-4}}$
2m
=
$\sum_l$
2l
2m+2-2l
+
2m+2

Fig 5

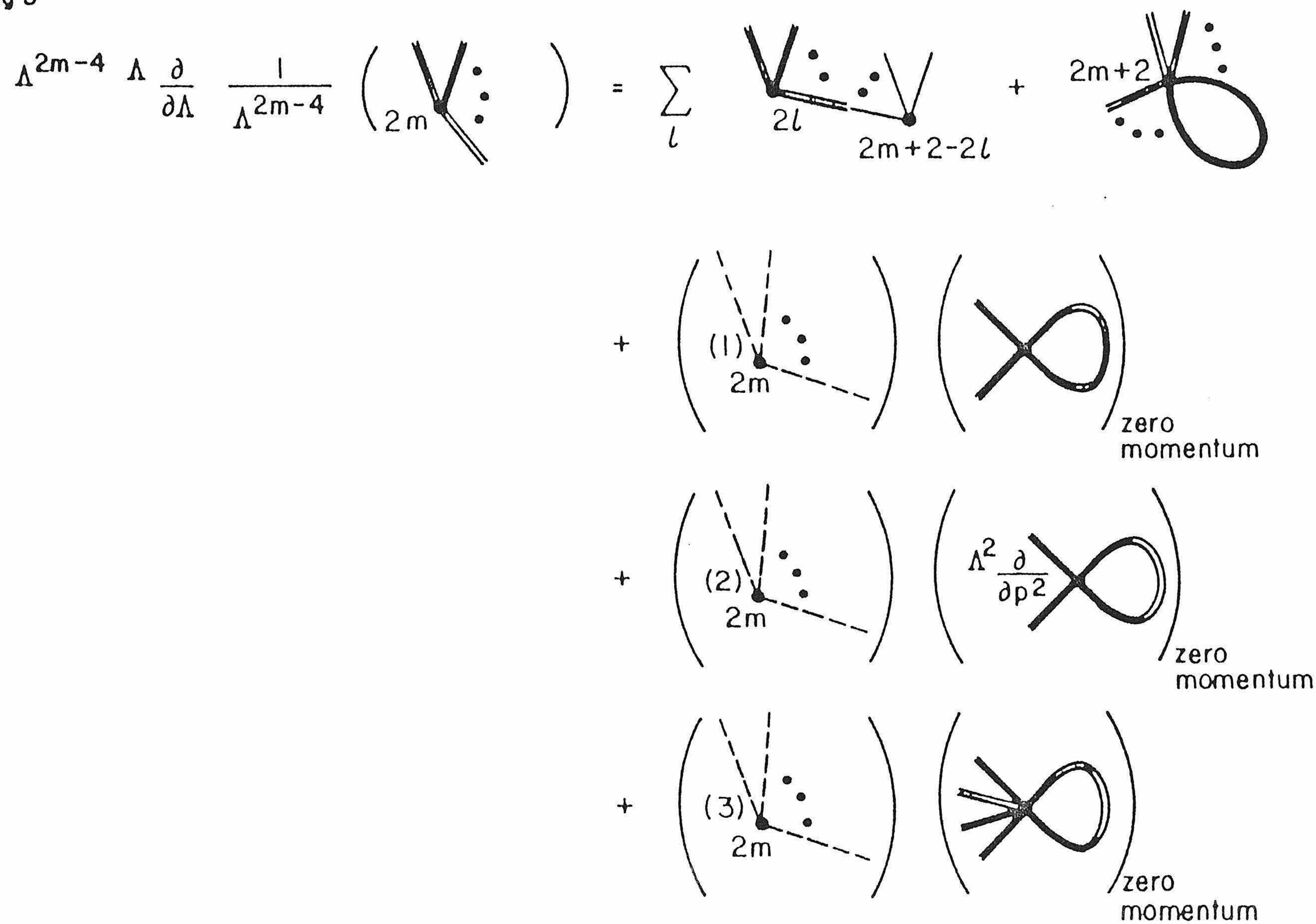

$\Lambda^{2m-4}\ \Lambda \frac{\partial}{\partial\Lambda}\ \frac{1}{\Lambda^{2m-4}}$
2m
$= \sum_{l}$
$2l$
$2m+2-2l$
+
2m+2
+
(1)
2m
zero
momentum
+
(2)
2m
$\Lambda^2 \frac{\partial}{\partial p^2}$
zero
momentum
+
(3)
2m
zero
momentum

Fig 6

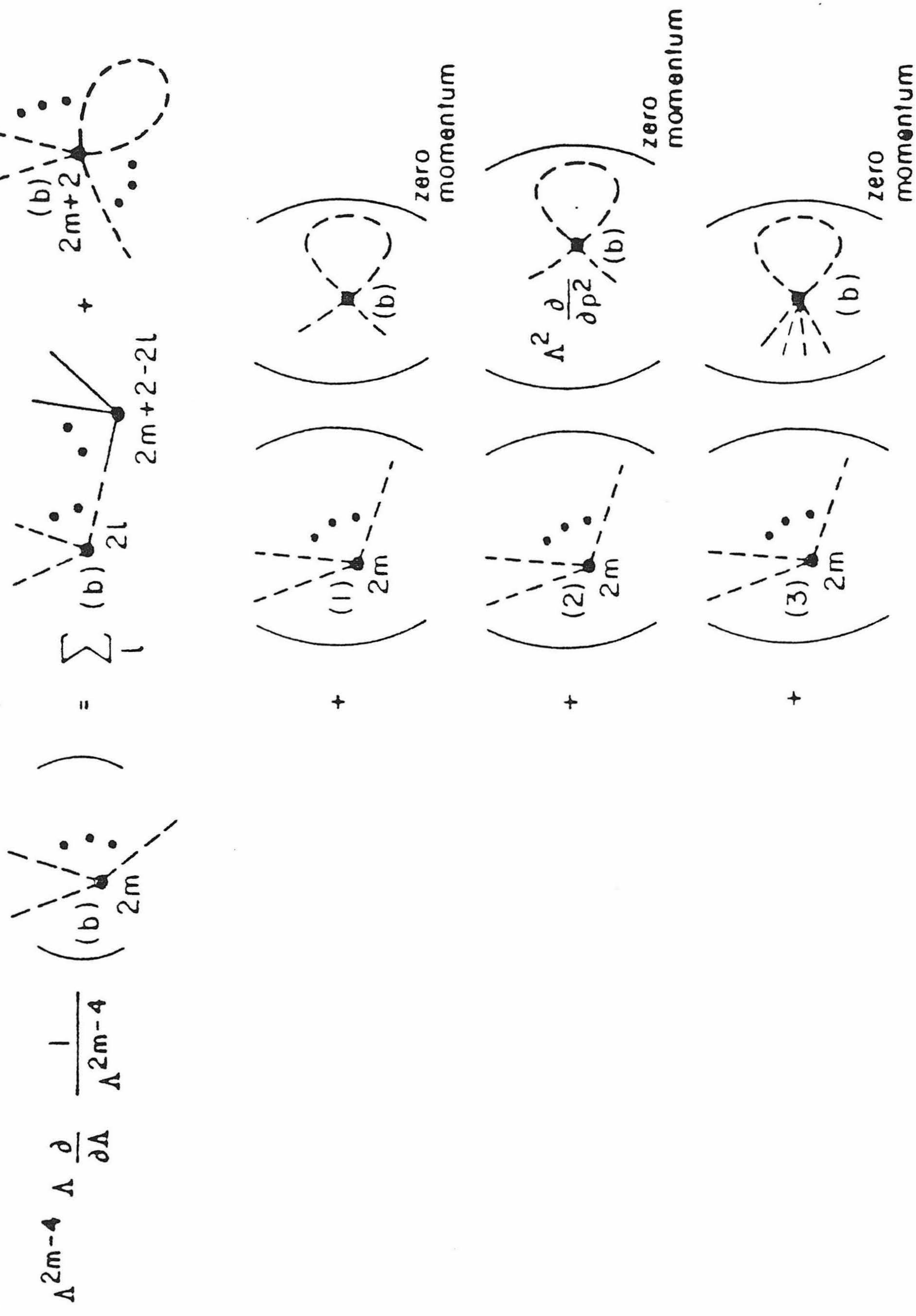


Fig 7

(a) (b)

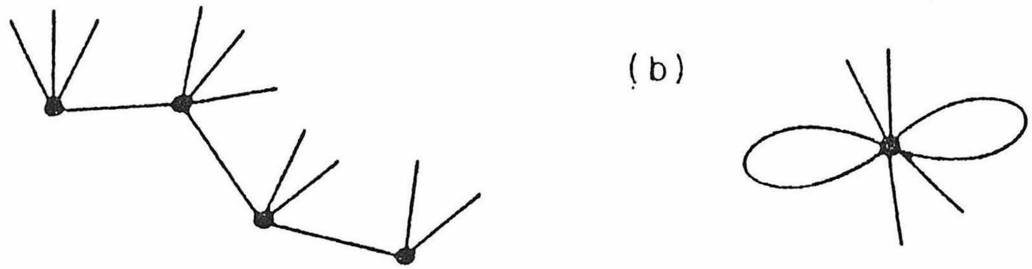


Fig 8

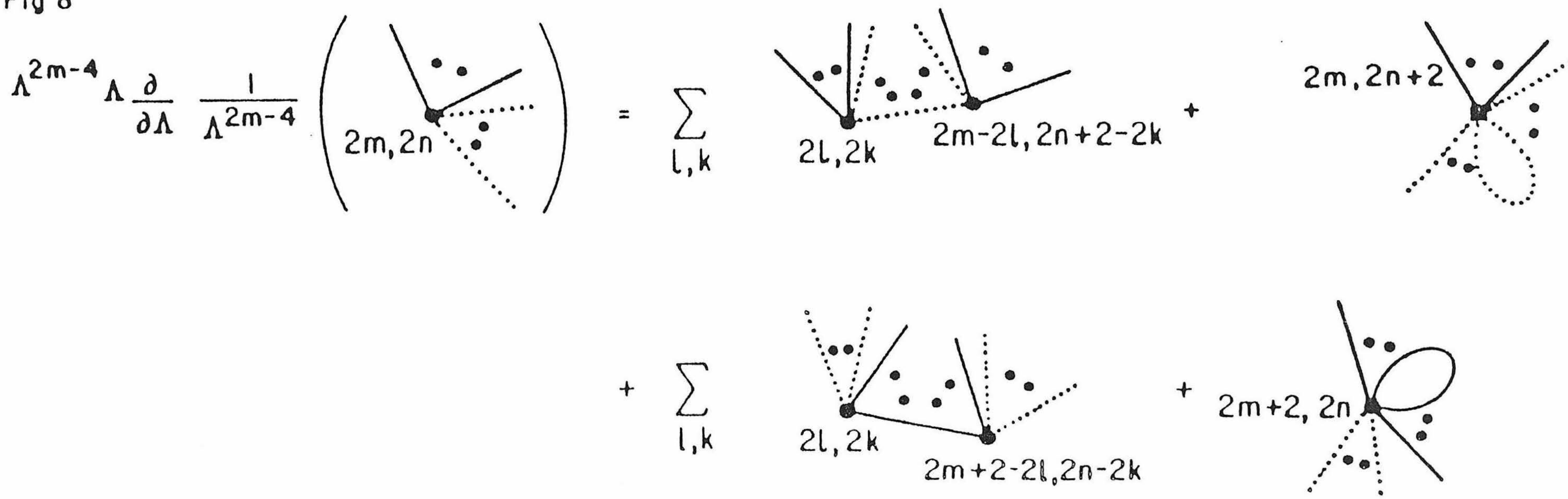

$\Lambda^{2m-4}\Lambda\frac{\partial}{\partial\Lambda}\frac{1}{\Lambda^{2m-4}}$
2m, 2n
=
$\sum_{l,k}$
2l, 2k
2m-2l, 2n+2-2k
+
2m, 2n+2
+
$\sum_{l,k}$
2l, 2k
2m+2-2l, 2n-2k
+
2m+2, 2n

Fig 9

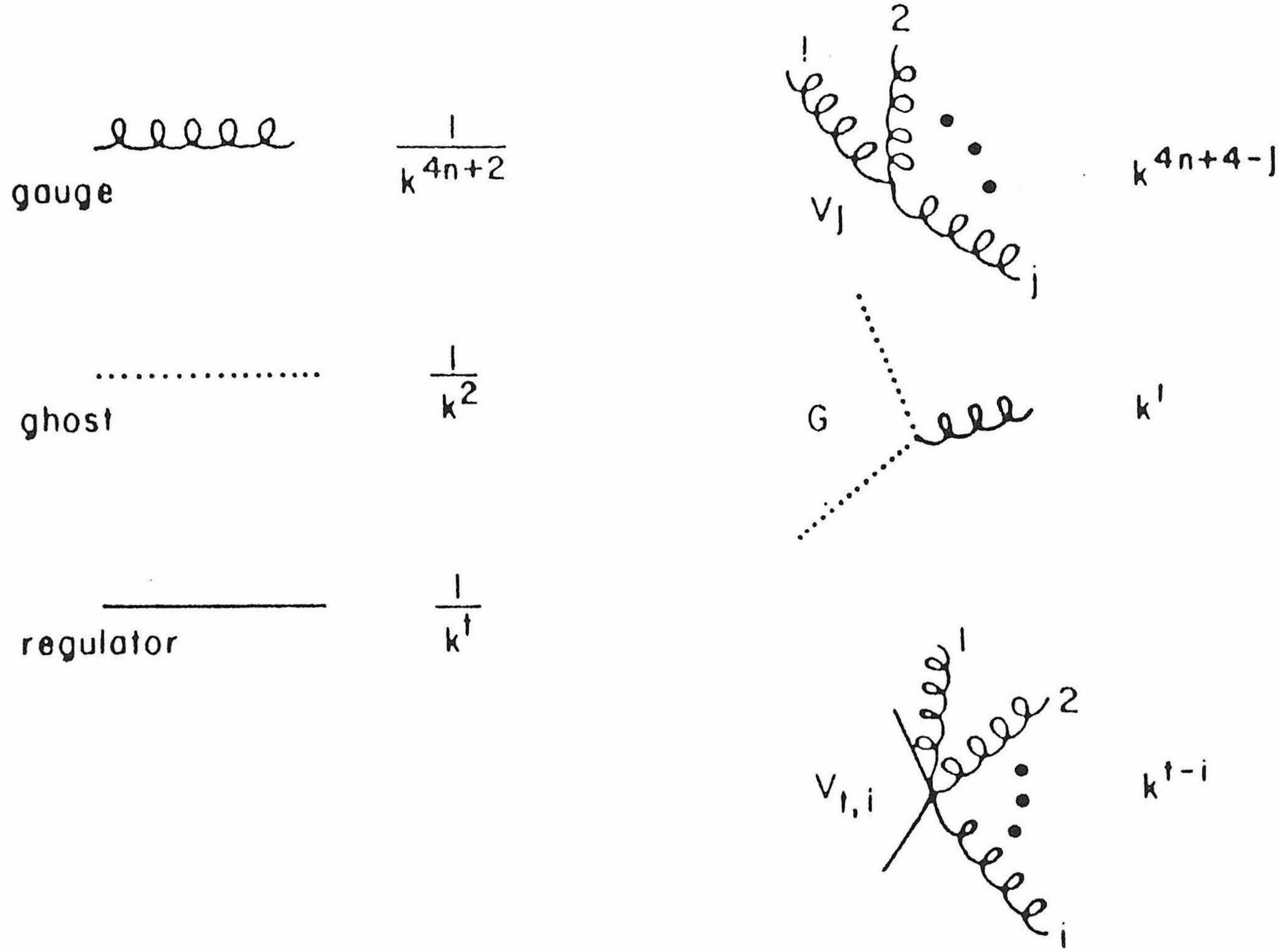

gauge
$\frac{1}{k^{4n+2}}$
ghost
$\frac{1}{k^2}$
regulator
$\frac{1}{k^t}$
1
2
j
V_j
k^{4n+4-j}
G
k^1
1
2
i
$V_{t,i}$
k^{t-i}

Figure 10

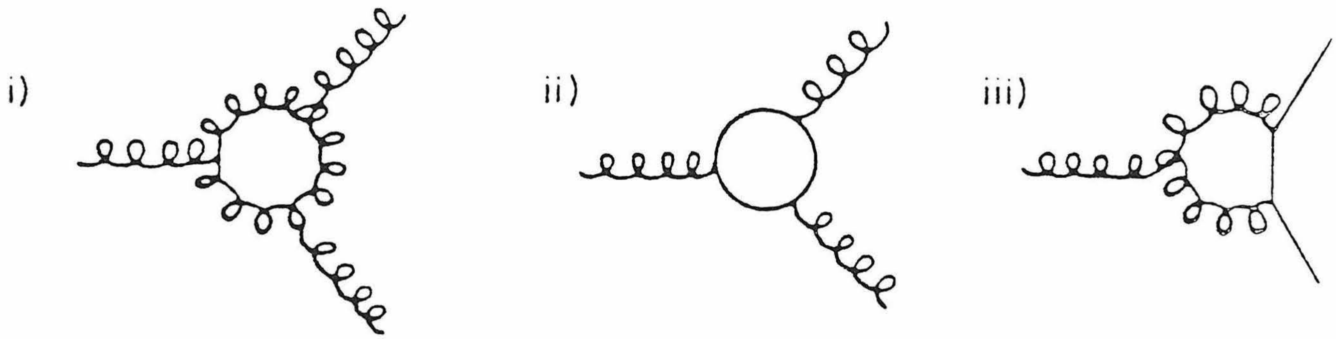


Figure 11

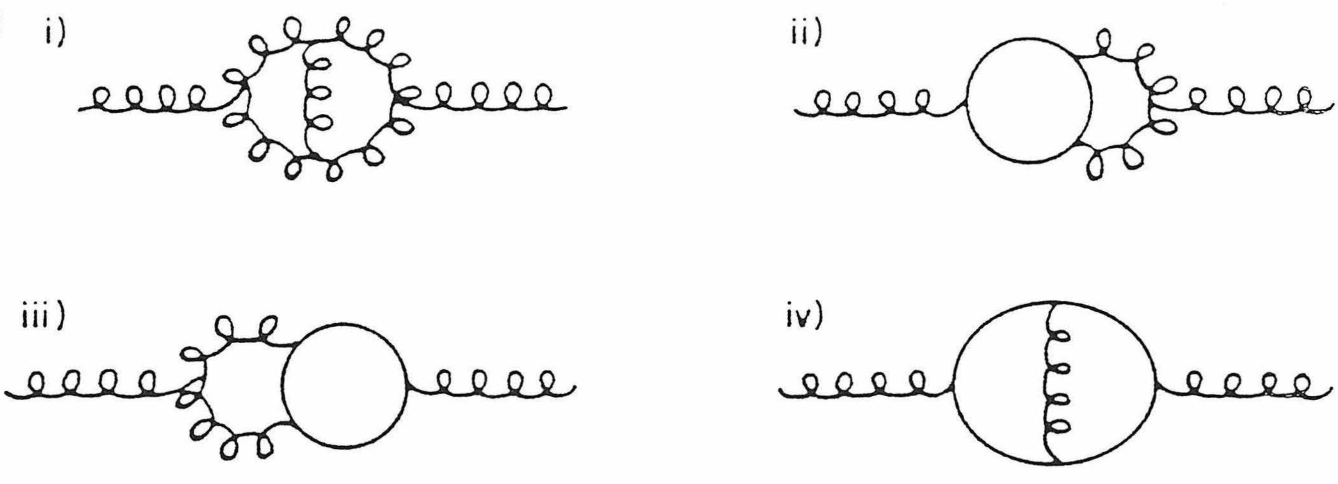
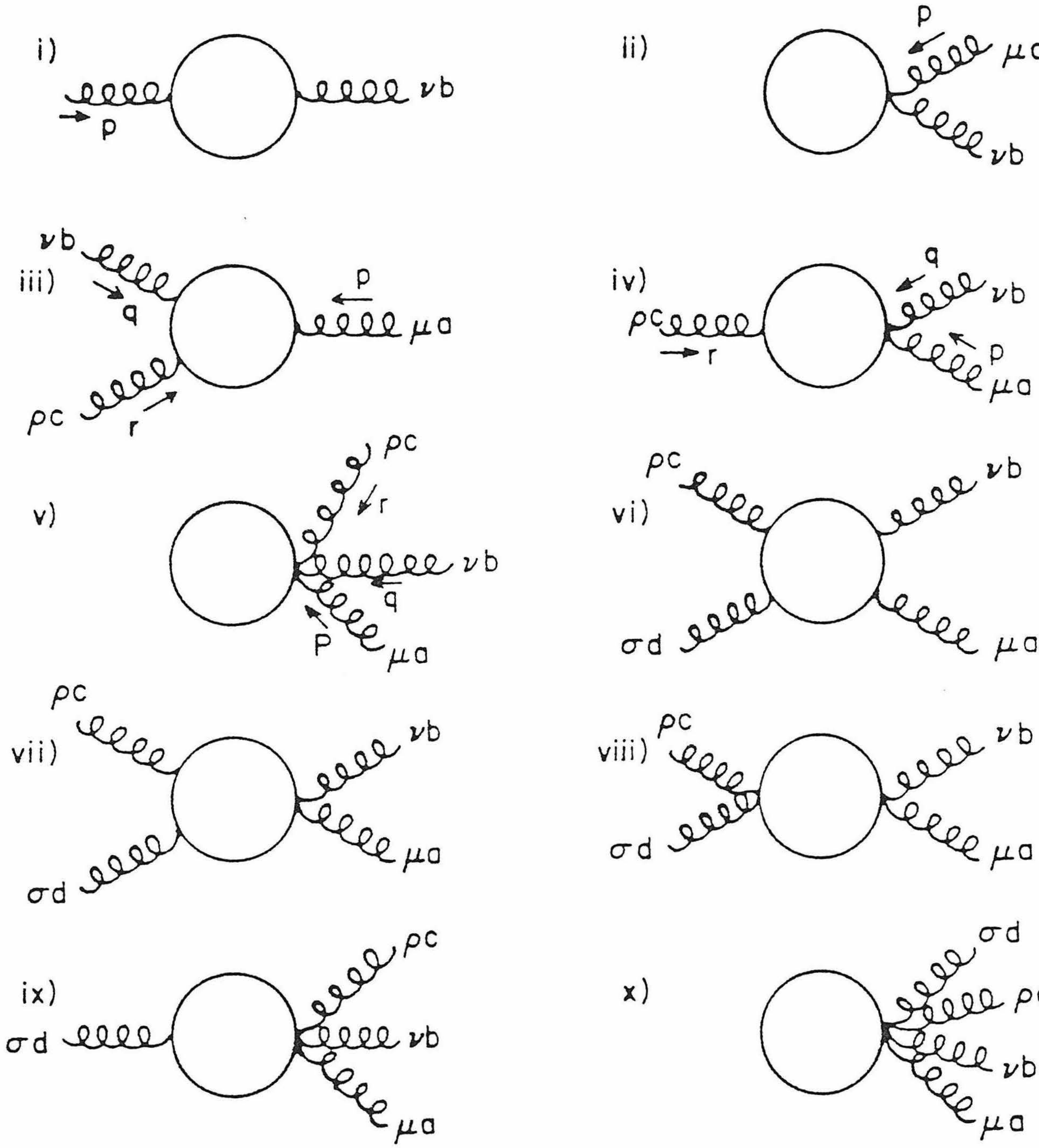
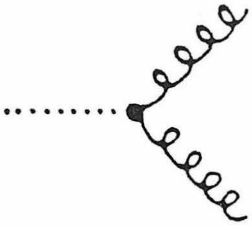
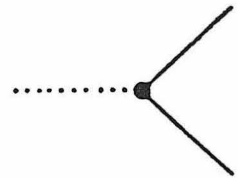
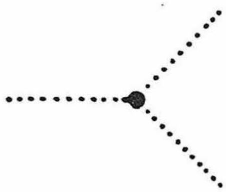


Figure 12

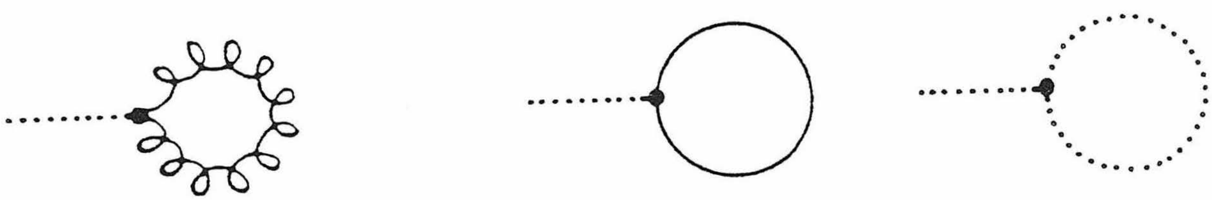

i)
p
νb
ii)
p
μa
νb
iii)
νb
q
p
μa
ρc
r
iv)
q
νb
ρc
r
p
μa
v)
ρc
r
νb
q
P
μa
vi)
ρc
νb
σd
μa
vii)
ρc
νb
σd
μa
viii)
ρc
νb
σd
μa
ix)
ρc
σd
νb
μa
x)
σd
ρc
νb
μa

i)  $\frac{k^{4n+2+2P}}{\Gamma^{2P}}$

ii)  $\frac{k^{1+2P}}{\Gamma^{2P}}$

iii)  $\frac{k^{2+2P}}{\Gamma^{2P}}$

iv)  $\frac{k^{4n+3+2P+4p}}{\Gamma^{2P}m^{4n+4p}}$



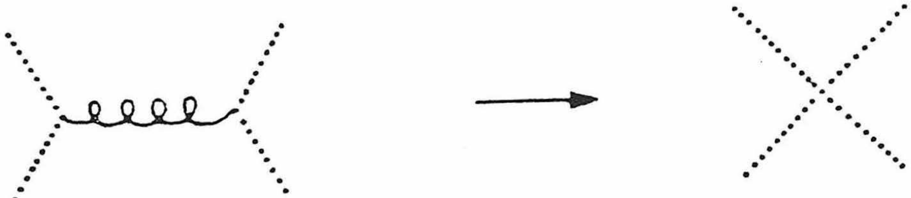


Fig 16

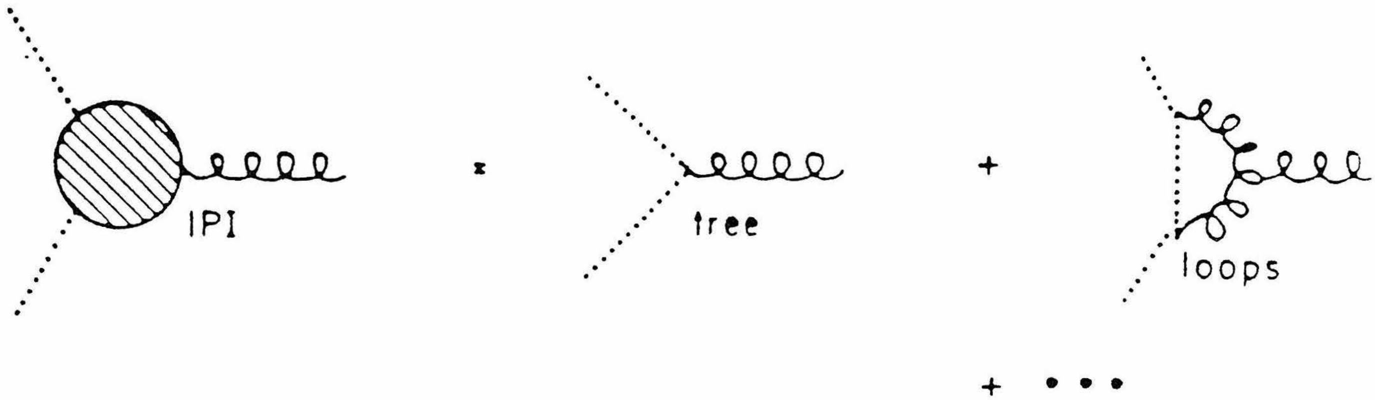


Fig 17

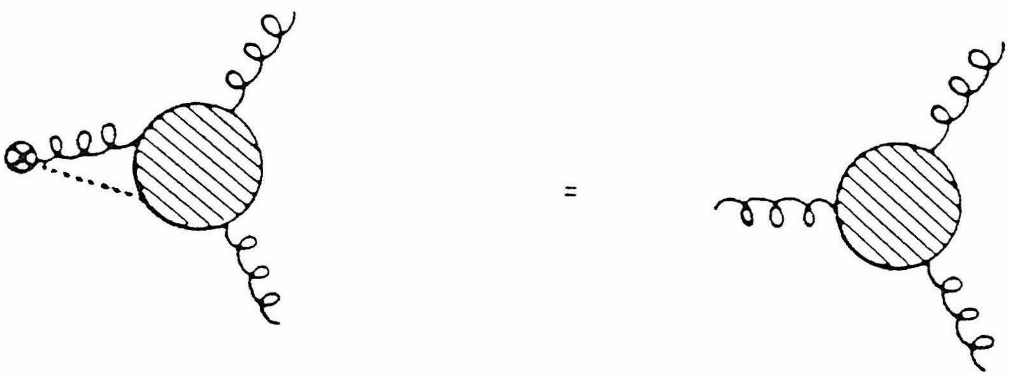


Fig 18

i) ii)

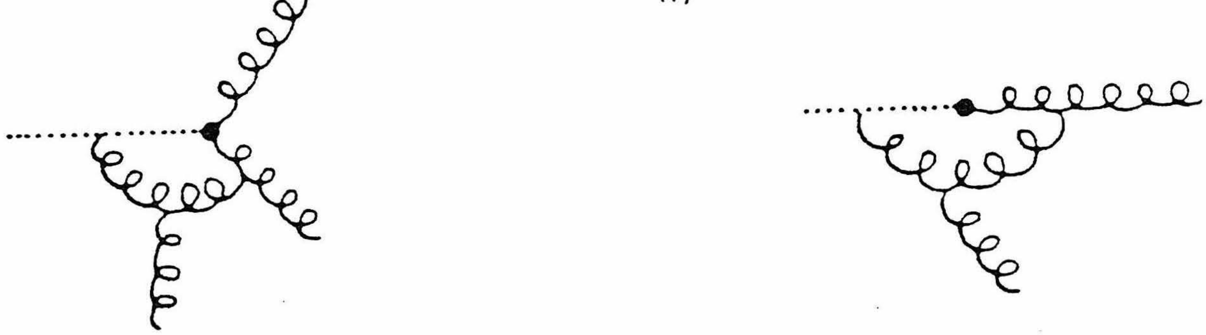
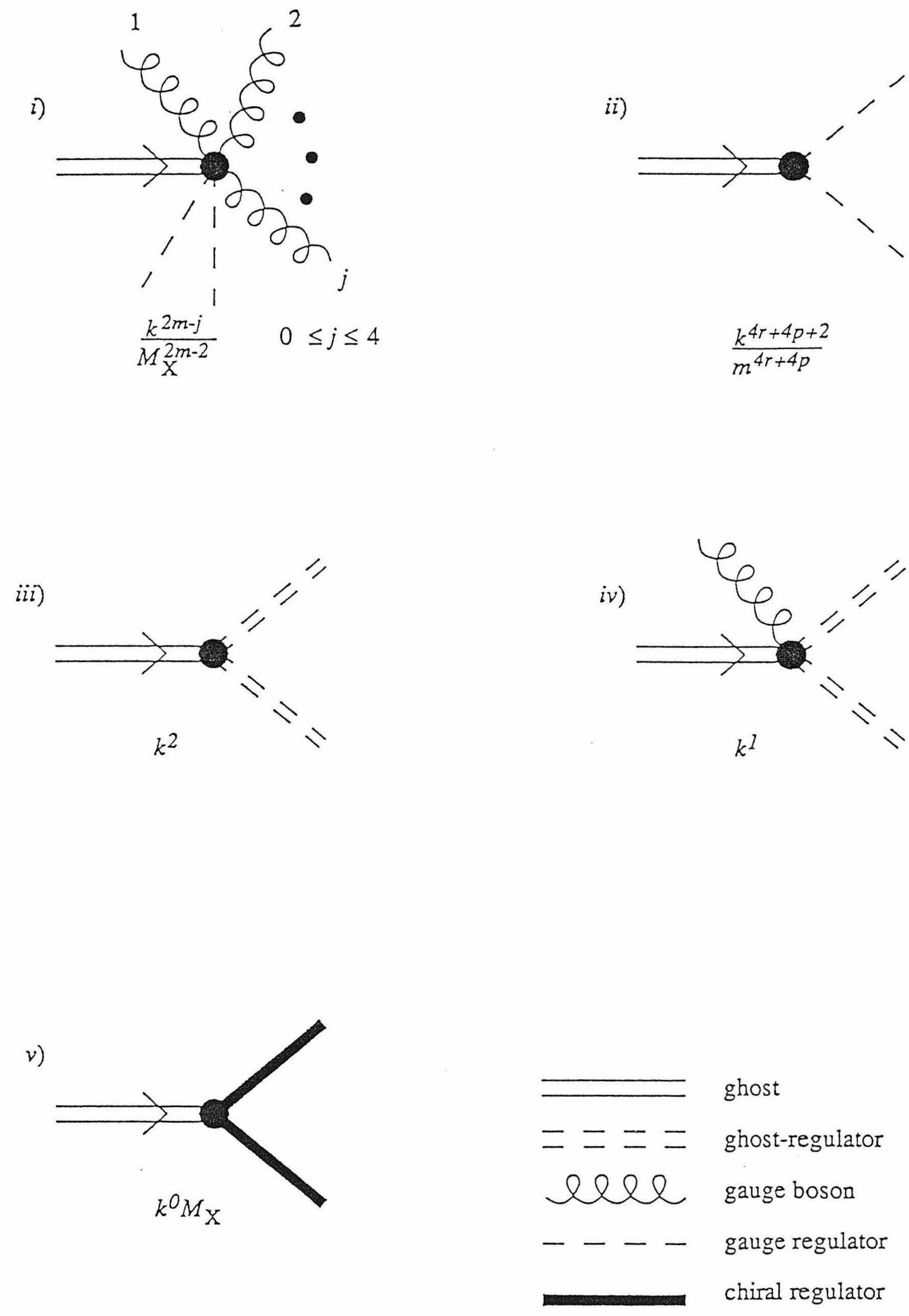
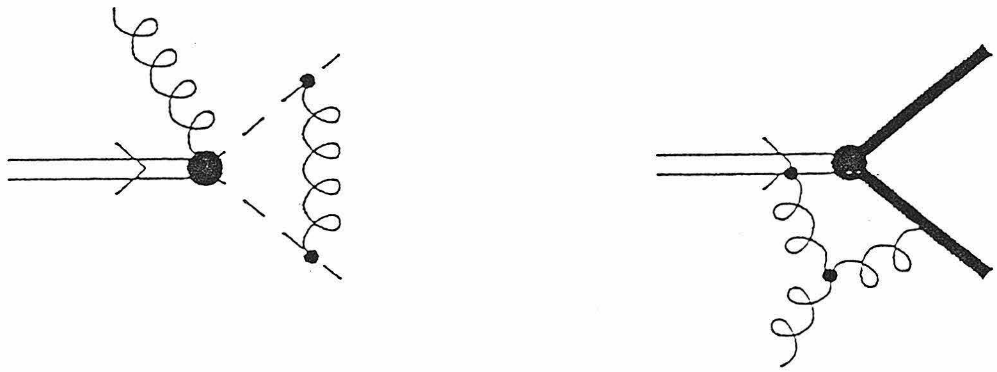


Figure 1



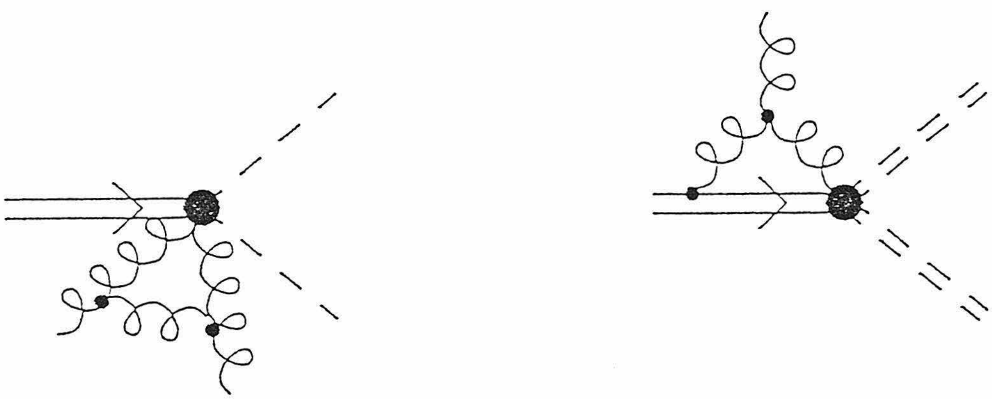
Gauge breaking operator insertions

Figure 2



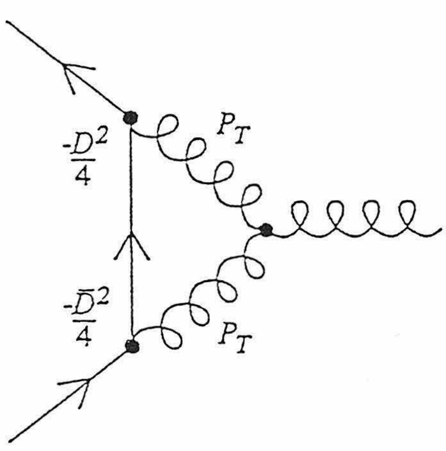
Graphs with internal regulator lines

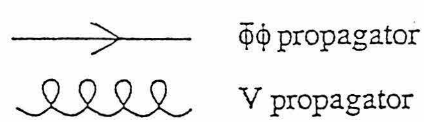
Figure 3



Graphs with no internal regulator lines

Figure 4





This one-loop graph is a correction to the gauge coupling g, and is finite because the V-propagator is proportional to P_T at large momentum.