2 Renormalization

The example considered above tells us that if we calculate an $n-$point Green function defined as

$$(2\pi)^4\delta^4(p_1+\cdots p_n)G^{(n)}(p_1\cdots p_{n-1},\lambda_0,m_0,\Lambda)=\int d^4x_1\cdots d^4x_n e^{i(p_1\cdot x_1\cdots p_n\cdot x_n)}\langle 0|T\phi(x_1)\cdots\phi(x_n)|0\rangle,$$

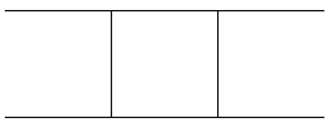
in terms of the bare coupling λ_0 and bare mass m_0, the result will depend explicitly on the cutoff Λ.

This dependence of Λ, however, is such that when expressed in terms of the renormalized quantities λ_R and m_R the "renormalized Green function" defined as

$$G_R^{(n)}(p_1\cdots p_{n-1},\lambda_R,m_R) \;=\; Z^{-(n/2)}G^{(n)}(p_1\cdots p_{n-1},\lambda_0,m_0,\Lambda), \tag{2.1}$$

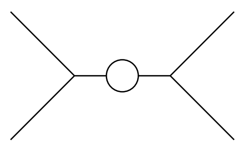
is finite (cut-off independent). It is these renormalized Green functions which are used to construct the S-matrix elements.

It is useful to work in terms of "truncated" or "one-particle irreducible" Green functions. These are Green functions calculated from graphs which cannot be separated into two or more graphs by cutting though one internal line. or a four-point Green function, the box graphs, such as



are one-particle irreducible, since we need to cut through two internal lines to separate them into two graphs.

whereas vertex or self-energy graphs such as



can be cut into two by cutting a single internal line (in several ways), and are therefore not one-particle irreducible.

We use the symbol Γ to refer to one-particle irreducible graphs, and the relation between the renormalized and the bare one-particle irreducible Green functions is

$$\Gamma_R^{(n)}(p_1\cdots p_{n-1},\lambda_R,m_R) \;=\; Z^{n/2}\Gamma^{(n)}(p_1\cdots p_{n-1},\lambda_0,m_0,\Lambda). \tag{2.2}$$

(The self-energy Σ is the same as $\Gamma^{(2)}$.)

2.1 Counterterms

We should think of renormalization as adjusting the masses and coupling constants (by a cut-off dependent amount if necessary), such that the S-matrix elements calculated to higher orders are cut-off independent and expressed in terms of physically measurable masses and couplings. In order to perform these higher order calculations it is convenient to view renormalization as the process of subtracting counterterms, in each order of perturbation theory, for some one-particle-irreducible Green function.

We do this by writing the Lagrangian in terms of bare parameters as a sum of two terms, one being the renormalized Lagrangian in terms of renormalized parameters and renormalized fields and the other being a set of counterterms. Thus for the ϕ^3 theory we have

$$\mathcal{L} \;=\; \frac{1}{2}\left(\partial_\mu\phi\,\partial^\mu\phi - m_0^2\phi^2\right) - \frac{\lambda_0}{3!}\phi^3$$

The fields are the “bare” fields and are related to the renormalized fields by

$$\phi \;=\; \sqrt{Z}\phi_R$$

so we may write the Lagrangian as

$$\mathcal{L} \;=\; \mathcal{L}_R + \mathcal{L}_{CT},$$

where, in terms of renormalized fields, masses and couplings

$$\mathcal{L}_R \;=\; \frac{1}{2}\left(\partial_\mu\phi_R\,\partial^\mu\phi_R - m_R^2\phi_R^2\right) - \frac{\lambda_R}{3!}\phi_R^3$$

and

$$\mathcal{L}_{CT} \;=\; \frac{1}{2}(Z-1)\partial_\mu\phi_R\,\partial^\mu\phi_R - \frac{1}{2}(Z_m-1)m_R^2\phi_R^2 - (Z_1-1)\frac{\lambda_R}{3!}\phi_R^3,$$

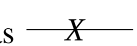
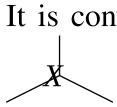
with

$$m_0^2 = \frac{Z_m}{Z}m_R^2$$

and

$$\lambda_0 = \frac{Z_1}{Z^{3/2}}\lambda_R.$$

We now calculate a Green function to any order using $\mathcal{L}_R$, i.e. in terms of renormalized masses and couplings. We will sometimes obtain UV divergences, which are cancelled when we consider a graph of lower order with a counterterm insertion.

It is convenient to view these counterterms as extra Feynman diagrams such as  or , where X represents the counterterm. The subtraction of the counterterm graphs from the unrenormalized Green functions renders them finite up to a factor of $\sqrt{Z}$ for each external line.

For example, for the two-point function (self-energy) we calculate the quantity

$$\Sigma(p^2, \lambda_R, m_R, \Lambda),$$

which has an explicit Λ dependence, but the subtracted quantity

$$\Sigma_R(p^2,\lambda_R,m_R) \;=\; \left(\Sigma(p^2,\lambda_R,m_R,\Lambda) - (Z-1)p^2 - (Z_m-1)m_R^2\right), \tag{2.3}$$

is *finite* and equal (by eq.(2.2)) to $Z\Sigma(p^2,\lambda_0,m_0,\Lambda)$. The contribution from these counterterms is therefore equivalent to replacing the renormalized masses by bare masses and multiplying by Z.

For the three-point function, we have

$$\Gamma^{(3)}(p_1,p_2,\lambda_R,m_R,\Lambda)$$

and the renormalized (finite) quantity is obtained as

$$\Gamma_R^{(3)}(p_1,p_2,\lambda_R,m_R) \;=\; \Gamma^{(3)}(p_1,p_2,\lambda_R,m_R,\Lambda)+(Z_1-1)\lambda_R,$$

which, by eq.(2.2) is equal to $Z^{3/2}\,\Gamma^{(3)}(p_1,p_2,\lambda_0,m_0,\Lambda)$

How many counterterms are needed to make all renormalized Green functions finite? The superficial degree, $\omega(G)$, of divergence for some Feynman graph, G, is

$$\omega(G) \;=\; 4L+\sum_{vertices}\delta_v-I_F-2I_B \;=\; \sum_{vertices}(\delta_V-4)+3I_F+2I_B+4, \tag{2.4}$$

where L is the number of loops, δ_V is the number of derivatives in the Feynman rule for the vertex (each introduces a power of momentum), and $I_{F(B)}$ is the number of internal fermion (boson) lines. Internal fermion lines carry a power of momentum in the numerator of their propagators.

$$L=I_B+I_F+1-V,$$

where V is the number of vertices.

Define

$$\omega_V \;=\; \delta_V+\frac{3}{2}f_V+b_V,$$

where f_V, (b_V) are the number of fermions (bosons) emerging from a vertex. Using the fact that one end of each internal line must end on a vertex we have

$$\sum_{vertices}\omega_V=\sum_{vertices}\delta_V+3I_F+2I_B+\frac{3}{2}E_F+E_B,$$

where $E_{F(B)}$ are the number of external fermions (bosons).

Thus we end up with

$$\omega(G) \;=\; 4-\frac{3}{2}E_F-E_B+\sum_{vertices}(\omega_V-4) \tag{2.5}$$

If $\omega_V>4$ then as we go to higher orders more and more graphs will have a non-negative degree of divergence (E_F and E_B increase for a given degree of divergence). These are non-renormalizable theories, since we need more and more counterterms as we go to higher orders.
Examples of such non-renormalizable theories are theories with interaction terms of the form $\lambda\phi^5$, or $g\bar{\Psi}\gamma^\mu\Psi\partial_\mu\phi$, for which $\omega_V=5$

If $\omega_V = 4$ we have a renormalizable theory. We require counterterms for all one-particle-irreducible Green functions for which $\frac{3}{2}E_F + E_B \leq 4$, but once these counterterms appear at the one-loop level, no further counterterms are required in higher order.
Examples of such renormalizable theories are those with interaction terms of the form $\lambda\phi^4$, $g\bar{\Psi}\gamma^\mu\Psi A_\mu$, $g\bar{\Psi}\phi\Psi$, $g\phi\partial_\mu\phi A^\mu$. These have $\omega_V = 4$.

There is one exception:
The above analysis of the degree of divergence assumes that propagators for bosons of momentum p always behave as $1/p^2$ as $p \to \infty$. For massive vector particles the propagator is

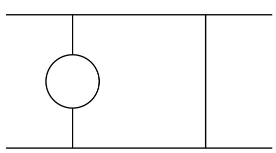
$$-i\left(\frac{g_{\mu\nu} - p_\mu p_\nu/m^2}{p^2 - m^2}\right)$$

and some of the components are constant as $p \to \infty$. Theories involving massive vector particles are in general NOT renormalizable. The exception is the case in which the mass of the vector boson is generated by Spontaneous Symmetry Breaking of a gauge theory. In that case it is possible to choose a gauge in which the propagator of the vector boson does indeed vanish like $1/p^2$ as $p \to \infty$.

If $\omega_V < 4$ we have a super-renormalizable theory in which the number of counterterms needed to render the Green functions finite decreases as the number of loops increases. The interaction $\lambda\phi^3$ is an example of such a theory. The only cut-off dependent counterterm is the mass renormalization and this is only cut-off dependent at one loop. What this actually means is that beyond one loop the counterterms which we introduce in order to express S-matrix elements in terms of physical quantities are cut-off independent.

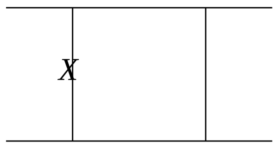
For the rest of this section we assume that we are dealing with a renormalizable, rather than a super-renormalizable, theory with generic coupling constant λ so that we assume that vertex correction graphs are ultraviolet divergent and that the divergences do persist in higher orders.

The real degree of divergence of a Feynman graph is the largest superficial divergence of any subgraph



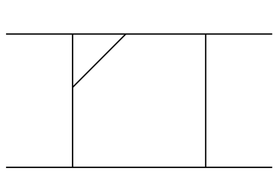
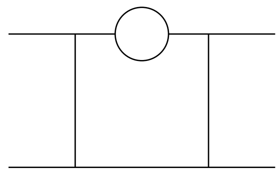
For such graphs, in a renormalizable theory such as QED, the superficial divergence of the entire graph is negative (box-graphs are UV finite †), but the real degree of divergence is the divergence of the self-energy insertion on one of the internal lines. This means that in association with the above graph we require a counterterm graph

†In a non-abelian gauge theory box diagrams with four external gauge-bosons also require renormalization. This is because there is a four-point coupling between four gauge-bosons at the tree level proportional to g^2 and the renormalization of the coupling constant, g gives rise to a counterterm for the four-point gauge-boson graph.

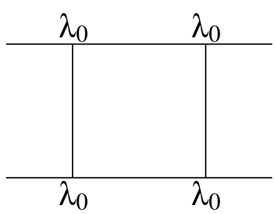


This counterterm renormalizes the mass of the internal line to which it is attached and also contributes to the renormalization of the couplings at either end of that line.

The other graphs which contribute to the renormalization of the couplings are all the remaining vertex corrections at all four vertices and all other self-energy insertions on internal and external lines, e.g.

 ,  etc.

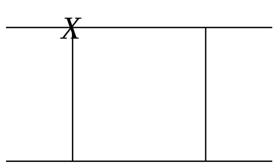
If we look at the one-loop graph with the coupling constants taken to be the bare couplings

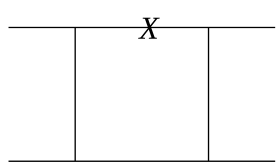


and write

$$\lambda_0 = \frac{Z_1}{Z^{3/2}} \lambda_R,$$

expand Z and Z_1 to order λ_R^2 we get the above diagram with λ_0 replaced by λ_R everywhere plus all the possible counterterm graphs such as



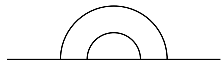
, 

etc.

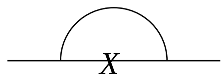
Together with the counterterms associated with mass renormalization, these counterterms render all the divergent subgraphs finite up to a factor of $\sqrt{Z}$ for each external line. The renormalized Green function expressed as a power series in the renormalized coupling and using renormalized masses is therefore cut-off independent, provided the counterterms associated with all the superficially divergent subgraphs have been accounted for.

This technique can also be used for the higher order computation of a Green function which itself has a non-negative degree of superficial divergence.

For example, at two loop level there is a graph contributing to the self-energy which is

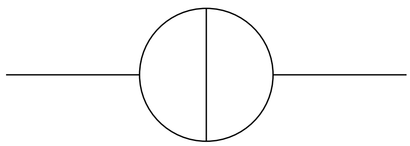


To this we must add a counterterm which renormalizes the mass of the lower propagator, and a counterterm $(Z-1)$ which contributes to the renormalization of the couplings of the one-loop graph.

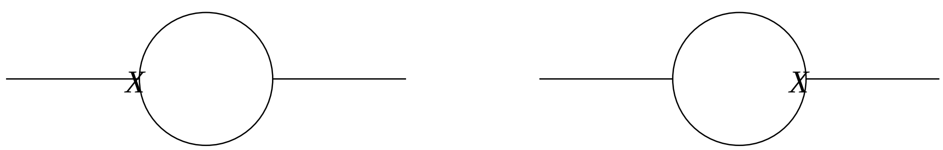


The remaining contribution (which for a general renormalizable theory will still be cut-off dependent), contributes to the λ_R^4 term in the expansions of Z and δm.

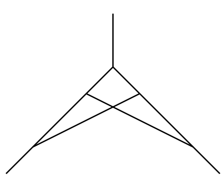
Sometimes we will have overlapping divergent sub-diagrams such as



Associated with this we have two counterterm graphs, corresponding to renormalizations of the vertex on the left- and right- of the graph.



There are also some graphs which have no divergent subgraphs, but which are nevertheless divergent and contribute to a counterterm at order λ_R^4



The subgraphs are all four-point, which we assume to be finite in the theory we are considering, but there is an overall divergence which contributes to Z_1 at order λ_R^4.

The central theorem of renormalization (proved by Bogoliobov, Parasiuk, Hepp and Zimmerman (BPHZ)) states that this procedure can be used to render cut-off independent all renormalized Green functions provided their superficial degree of divergence is negative. In other words, for a renormalizable theory we have a finite number of counterterms, which, in general, have cut-off dependent contributions in all orders in perturbation theory. Provided these counterterms are used in association with all divergent subgraphs, then as we go to higher orders we do *not* have to introduce further counterterms to cancel off infinities that occur in subgraphs.

2.2 Regularization:

A regulator is a process which renders finite a momentum integral which is superficially divergent. Ideally, we would like the regulator to preserve the symmetries of the theory, so that the counterterms calculated using that regulator automatically preserve the symmetries.

The simple cut-off procedure used previously does not in general do this.

Pauli-Villars Regulator
Before the relevance of gauge-theories was recognized, the most popular method of regulating ultraviolet divergent integrals was to replace a propagator

$$\frac{1}{k^2 - m^2}$$

by the regulated propagator

$$\sum_{i=0}^{\infty} a_i \frac{1}{k^2 - m_i^2},$$

where $a_0 = 1$ and $m_0 = m$.

If we expand each term of this sum as a power series in k^2 we get

$$\sum_{i=0}^{\infty} \frac{a_i}{k^2} + \sum_{i=0}^{\infty} \frac{a_i m_i^2}{k^4} + O\left(\frac{1}{k^6}\right).$$

For a renormalizable theory the maximum superficial power of divergence of any integral is quadratic, so that the $O(1/k^6)$ terms are ultraviolet finite. The finiteness of the regulated integral is then guaranteed by requiring that

$$\sum_{i=0}^{\infty} a_i = 0,$$

$$\sum_{i=0}^{\infty} a_i m_i^2 = 0.$$

Dimensional Regularization

The above method of regularization is unsuitable for gauge-theories, because gauge invariance requires that the gauge-bosons should be massless, so that the Pauli-Villars regulated propagator, which introduces masses, would break this gauge invariance.

A more useful method is the method of "dimensional regularization", which relies on the fact that most symmetries (excluding supersymmetry, which will be discussed briefly later) do not depend on the number of dimensions of the space in which we are working.

The integral that we wish to regulate is performed not in four dimensions, but in a number of dimensions, d, for which the integral is finite. An analytic continuation is made in the variable d. This analytic function can be expanded as a Laurent series about $d = 4$ and the fact that the symmetry is preserved in all dimensions means that each term in the series will respect the symmetry. The divergences appear as poles at $d = 4$ and the regularization is effected by removing these poles.

In d dimensions the typical integrals that we obtain after going through the steps of Feynman parametrization, shifting the variable of integration, and rotating to Eucidean space is of the form

$$I_0(\alpha) \equiv i\frac{(-1)^\alpha}{(2\pi)^d}\int \frac{d^dk}{(k^2+A^2)^\alpha} = i\frac{(-1)^\alpha}{(2\pi)^d}\int \frac{d\Omega_d\, k^{d-1}\, dk}{(k^2+A^2)^\alpha}.$$

$$\int d\Omega_d = \frac{2\pi^{d/2}}{\Gamma(d/2)}$$

is the area of the surface of a sphere in d dimensions, and

$$\int_0^\infty \frac{x^{d/2-1}\, dx}{(x+1)^\alpha} = \frac{\Gamma(d/2)\,\Gamma(\alpha-d/2)}{\Gamma(\alpha)},$$

so that the integral we are left with is

$$I_0(\alpha) = i\frac{(-1)^\alpha}{(4\pi)^{d/2}}\frac{\Gamma(\alpha-d/2)}{\Gamma(\alpha)}A^{d-2\alpha} \tag{2.6}$$

If $\alpha > (d/2)$, the integral in finite. As an analytic function of d, it has poles if $(\alpha - d/2)$ is zero or a negative integer.

We perform a Laurent expansion about $d = 4$, defining the quantity ε † by

$$d = 4-2\varepsilon,$$

giving rise to a pole term at $\varepsilon = 0$ for any integral which is superficially divergent in four dimensions, plus terms which are finite as $\varepsilon \to 0$. Each of these terms preserves the (dimension independent) symmetries of the theory.

† Many authors use a definition of $d = 4-\varepsilon$. We choose this definition because $d/2$ appears very often in the formulae for the integrals.

For example of $\alpha = 2$ we get

$$i\frac{\Gamma(\varepsilon)}{16\pi^2}\left(\frac{4\pi}{A^2}\right)^{\varepsilon} = \frac{i}{16\pi^2}\left(\frac{1}{\varepsilon}+\ln\left(\frac{4\pi}{A^2}\right)-\gamma_E + O(\varepsilon)\right),$$

where we have used

$$\Gamma(\varepsilon) = \frac{1}{\varepsilon} - \gamma_E + O(\varepsilon).$$

In the so-called "minimal subtraction (MS)" renormalization scheme the pole part

$$\frac{i}{16\pi^2\varepsilon}$$

is associated with the counterterm and the regularized integral is the remaining part

$$\frac{i}{16\pi^2}\left(\ln\left(\frac{4\pi}{A^2}\right)-\gamma_E\right).$$

However, the $\ln(4\pi)$ and the Euler constant γ_E. always appear in the finite part of the integral. They have no physical significance and are merely an artifact of the subtraction scheme. A more convenient scheme is the so-called "$\overline{MS}$" scheme, in which the $\ln(4\pi)$ and the γ_E are subtracted off along with the pole part, so that the counterterm is

$$\frac{i}{16\pi^2}\left(\frac{1}{\varepsilon}+\ln(4\pi)-\gamma_E\right)$$

and the regulated integral is

$$-\frac{i}{16\pi^2}\ln\left(A^2\right).$$

We note that choosing the counterterms to be the pole part as in the MS scheme or using the $\overline{MS}$ scheme are perfectly valid renormalization prescriptions. However, in these schemes the renormalized coupling constant is not directly related to any physical measurement. Furthermore the renormalized mass is not the physically measured mass. Nevertheless the renormalized couplings and masses obtained in these schemes can be used as parameters, and *all* physical observables - including masses - can be calculated in terms of these renormalized parameters.

A further integral, which will occur often in higher order calculations is of the form

$$i\frac{(-1)^{\alpha}}{(2\pi)^d}\int\frac{d^dk\,k^{\mu}k^{\nu}}{(k^2+A^2)^{\alpha}}$$

Now by symmetry this will only be non-zero if μ and ν are in the same direction, so we may write

$$i\frac{(-1)^{\alpha}}{(2\pi)^d}\int\frac{d^dk\,k^{\mu}k^{\nu}}{(k^2+A^2)^{\alpha}} = -I_2(\alpha)g^{\mu\nu}, \tag{2.7}$$

(the minus coming from the fact that we have rotated to Euclidian space). Contracting both sides with $g_{\mu\nu}$ are recalling that $g^{\mu\nu}g_{\mu\nu} = d$, we have

$$I_2(\alpha) = -i\frac{(-1)^\alpha}{d(2\pi)^d}\int \frac{d^dk\,k^2}{(k^2+A^2)^\alpha}$$

Writing $k^2 = (k^2+A^2) - A^2$, we have

$$I_2(\alpha) = \frac{1}{d}\left(I_0(\alpha-1) - A^2 I_0(\alpha)\right)$$

Using Eq.(2.6), this gives

$$I_2(\alpha) = -i\frac{(-1)^\alpha}{(4\pi)^{d/2}}A^{d+2-2\alpha}\left(\frac{\Gamma(\alpha-1-d/2)}{d\,\Gamma(\alpha-1)} - \frac{\Gamma(\alpha-d/2)}{d\,\Gamma(\alpha)}\right).$$

Manipulating the Γ functions this reduces to

$$I_2(\alpha) = -i\frac{(-1)^\alpha}{(4\pi)^{d/2}}A^{d+2-2\alpha}\frac{\Gamma(\alpha-1-d/2)}{2\,\Gamma(\alpha)} \tag{2.8}$$

Another useful integral is obtained by contracting both sides of eq.(2.7) and using eq.(2.8)

$$i\frac{(-1)^\alpha}{(2\pi)^d}\int \frac{d^dk\,k^2}{(k^2+A^2)^\alpha} = d\,I_2(\alpha) = -i\frac{(-1)^\alpha}{(4\pi)^{d/2}}\frac{d}{2}A^{d+2-2\alpha}\frac{\Gamma(\alpha-1-d/2)}{\Gamma(\alpha)} \tag{2.9}$$

Other consequences of dimensional regularization are:

1. The action, which must be dimensionless, is now

$$S = \int d^{4-2\varepsilon}x\mathcal{L}.$$

From the quadratic part of the Lagrangian density, we conclude that a fermion field, which enters as $\bar{\Psi}\gamma\cdot\partial\Psi$ must have dimension $\frac{3}{2}-\varepsilon$, whereas a bosonic field, which enters as $(\partial_\mu\phi)^2$ or $F_{\mu\nu}F^{\mu\nu}$, must have dimension $1-\varepsilon$.

When we consider the interaction terms, this in turn implies that the couplings acquire a dimension which differs from its dimensionality in four dimensions. For a renormalizable theory, the couplings are dimensionless in four dimensions. However, in $4-2\varepsilon$ dimensions this will not be the case. For example, the bare electromagnetic coupling in QED, defined by the interaction term in the action

$$\int d^{4-2\varepsilon}x e_0\bar{\Psi}\gamma^\mu A_\mu\Psi,$$

must be replaced by

$$\int d^{4-2\varepsilon}x\tilde{e}_0(\mu)\mu^\varepsilon\bar{\Psi}\gamma^\mu A_\mu\Psi,$$

where μ is some mass scale and $\tilde{e}_0(\mu)$ is a dimensionless quantity. In other words the bare coupling e_0 has dimension ε.

When this is expanded as a power series in ε, we find that there is always a term $\ln(\mu^2)$ accompanying the pole at $\varepsilon = 0$. This scale μ serves as the subtraction point in the renormalization procedure.

2. In dimensional regularization, the Dirac algebra must also be carried out in $4-2\varepsilon$ dimensions.

$$\{\gamma^\mu\gamma^\nu\} \;=\; 2g^{\mu\nu},$$

but

$$\begin{aligned} g_{\mu\nu}g^{\mu\nu} &= 4-2\varepsilon \\ \gamma^\mu\gamma_\mu &= (4-2\varepsilon)I \end{aligned}$$

Thus, for example

$$\begin{aligned} \gamma^\mu\gamma^\nu\gamma_\mu &= -2(1-\varepsilon)\gamma^\nu, \\ \gamma^\mu\gamma^\nu\gamma^\rho\gamma_\mu &= 4g^{\nu\rho}-2\varepsilon\gamma^\mu\gamma^\rho, \\ \gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma\gamma_\mu &= -2\gamma^\sigma\gamma^\rho\gamma^\nu+2\varepsilon\gamma^\nu\gamma^\rho\gamma^\sigma. \end{aligned}$$

Strictly, we should also have

$$\mathrm{Tr}I \;=\; 2^{d/2},$$

but it turns out that one can always use

$$\mathrm{Tr}I \;=\; 4,$$

and absorb the factor of $2^{d/2-2}$ into the renormalization of the coupling.

Dimensional Reduction:

This method of regularization is not suitable for dealing with supersymmetry which only holds in a given number of dimensions. For example, in a four dimensional supersymmetric theory a Majorana fermion has two degrees of freedom and is accompanied by two scalar superpartners. A vector field has four degrees of freedom and is accompanied by a Dirac fermion, which also has four degrees of freedom.

The solution to this problem is to introduce extra scalar particles called "$\varepsilon-$scalars" which compensate for the "lost" bosonic degrees of freedom as the number of dimensions is reduced below four. Thus, for example, in three dimensions a vector field is replaced by a vector field with three degrees of freedom plus a new scalar field which interacts with the fermions with the same coupling. In this way, the fermion, the vector field and the extra scalar can be combined into a supermultiplet.

The upshot of this scheme is that the Dirac algebra is once again carried out in four dimensions, using the rules of the four dimensional, whereas the integrals over the loop momentum is carried out in $4-2\varepsilon$ dimensions.