that, in principle, a way to overcome these problems was to introduce a UV and, if needed, an IR cutoff into the theory. However, we soon dismissed this option because it seemed evident that it would lead to spurious non-universal cutoff dependences of physical results.

How do these observations relate to what we are doing presently? Obviously, the present[10] version of the RG procedure also relies on the introduction of a cutoff regularizing the logarithmic UV divergences mentioned above; apparently, the RG procedure shares a lot of structures with the perturbative approach. But, somehow, we managed to extract the information in which we were interested – the dependence of the potential strength on long-range system parameters – in a manner independent of the cutoff.[11] The key to obtaining this information was to introduce not one, but an entire hierarchy of cutoffs and to integrate over each of these domains recursively.

Now, a subtle and important point is that this procedure does not imply that the cutoff or, more generally, short-scale fluctuations of the model have silently made their way out of the theory. After all, the UV divergences mentioned before are manifestations of a large "phase volume" of field fluctuations that are likely to somehow affect the behavior of the system. To understand the "implicit" way through which these fluctuations enter our results, let us return to a remark made on page 425. There, we had noted that, upon scaling frequency/time, each operator changes according to its physical dimension. An operator carrying the dimension $[\text{time}]^d$ would acquire a scaling factor b^d. The scaling dimension of an operator predicted by its "physical" dimension is called the **naive scaling dimension**, the **canonical scaling dimension**, or, for obscure reasons, the **engineering dimension**.

The designations indicate, however, that these dimensions are not the last word on the actual scaling behavior of an operator. Indeed, the net result of the RG analysis was that our operator of interest, $\int d\tau \cos\theta$, an object of engineering dimension 1, changes according to b^{1-g}. The correction to the naive scaling dimension (presently, g) is called the **anomalous dimension** of an operator. Its origin lies in the (cutoff-dependent) phase volume of fluctuations co-determining the change of an operator during each RG step. Put differently, we can say that the cutoff Λ, by itself a quantity of dimension $[\text{time}]^{-1}$, acts as a "gray eminence" implicitly affecting the scaling behavior of an operator. The anomalous scaling dimensions of the theory effectively determine its long-range observable behavior and, therefore, represent quantities of prime interest.

8.3 Renormalization group: general theory

Having discussed two extended examples, we are in a position to attempt a reasonably general outline of the RG strategy. Suppose we are given a field theory defined through the

[10] Below we become acquainted with UV regularization procedures that are not based on introduction of a cutoff.

[11] One may object that the solutions of the β-functions given above actually *do* contain the bare cutoff, through an initial condition; they also depend on the bare coupling strength and, possibly, other "non-universal" parameters. However, that need not worry us: in most applications (both experimental and theoretical) one is interested not so much in the "absolute value" of physical observables (as these usually depend on unknown material parameters anyway) but rather in the way these observables change as a relevant control parameter is varied. The important feature found above is that the rate at which the effective potential strength varies with temperature, say, is largely universal and cutoff-independent.

action

$$S[\phi] \equiv \sum_{a=1}^{N} g_a \mathcal{O}_a[\phi],$$

where ϕ is some (generally multi-component) field, g_a are coupling constants and $\mathcal{O}_a[\phi]$ a certain set of operators. For concreteness, one may think of these operators as $\mathcal{O}_a = \int d^dx\,(\nabla\phi)^n\phi^m$, i.e. as space–time local operators involving powers of the field and its derivatives – although more general structures are conceivable.[12] By "renormalization of the theory," we refer to a scheme to derive a set of **Gell-Mann–Low equations** describing the change of the coupling constants $\{g_a\}$ as fast fluctuations of the theory are successively integrated out.

8.3.1 Gell-Mann–Low equations

There are a number of methodologically different procedures whereby the set of flow equations can be obtained from the microscopic theory. Here, we formulate this step in a language adjusted to applications in statistical field theory (as opposed to, say, particle physics). While there is considerable freedom in the actual implementation of the RG procedure, all methods share the feature that they proceed in a sequence of three more or less canonical steps.

I: Subdivision of the field manifold

In the first step, one may decompose the integration manifold $\{\phi\}$ into a sector to be integrated out, $\{\phi_\mathrm{f}\}$, and a complementary set, $\{\phi_\mathrm{s}\}$. For example:

▷ We may proceed according to a generalized **block spin scheme** and integrate over all degrees of freedom located within a certain structural unit in the base manifold $\{\mathbf{x}\}$. (This scheme is adjusted to lattice problems where $\{\mathbf{x}\} = \{\mathbf{x}_i\}$ is a discrete set of points. However, as pointed out above, even then it is difficult to implement analytically.)

▷ We could decide to integrate over a certain sector in momentum space. When this sector is defined to be a shell $\Lambda/b \le |\mathbf{p}| < \Lambda$, one speaks of a **momentum shell integration**. Naturally, within this scheme, the theory will be explicitly cutoff-dependent at intermediate stages.

▷ Alternatively, we may decide to integrate over all high-lying degrees of freedom $\lambda^{-1} \le |\mathbf{p}|$. In this case, we will of course encounter divergent integrals. An elegant way to handle these divergences is to apply **dimensional regularization**. Within this approach one formally generalizes from integer dimensions d to fractional values $d \pm \epsilon$. One motivation for doing so (for another, see below) is that, miraculously, the formal extension of the characteristic integrals appearing during the RG step to non-integer dimensions are finite. As long as one stays clear of the dangerous values d = integer one can then safely monitor the

[12] In our previous example of the Luttinger liquid, there appeared an operator $\int(d\omega/2\pi)\theta(\omega)|\omega|\theta(-\omega)$. When represented in space-time, this operator is highly non-local.

dependence of the integrals on the IR cutoff λ^{-1}. For a good introduction to dimensional generalization we refer to the textbook by Ryder.[13]

▷ For a discussion of alternative schemes, such as the introduction of short-distance real space cutoffs underlying the so-called **operator product expansion**, we refer to the literature (see, e.g., the excellent text by Cardy.[1])

II: RG step

The second, and central, part of the program is to actually integrate over short–range fluctuations. As exemplified above, this step usually involves approximations. In most cases, one will proceed by a so-called **loop expansion**, i.e. one organizes the integration over the fast field ϕ_f according to the number of independent momentum integrals – loops[14] – that occur after the appropriate contractions. Of course, this strategy makes sense only if we can guarantee that the contribution of loops of higher orders is in some sense small, a precondition that is, alas, often difficult to meet. At any rate, to engage loop numbers as an expansion parameter, we first need to understand the key role played by space-dimensionality in the present context. We return to this point in Section 8.4.

Following the procedure, an expansion over the fast degrees of freedom gives an action

$$S'[\phi_s] \equiv \sum_a g'_a \mathcal{O}'_a[\phi_s],$$

in which coupling constants of the remaining slow fields are altered. Notice that the integration over fast field fluctuations may (and usually does) lead to the generation of "new" operators, i.e. operators that have not been present in the bare action. In such cases one has to investigate whether the newly generated operators are "relevant" (see below) in their scaling behavior. If so, the appropriate way to proceed is to include these operators in the action from the very beginning (with an *a priori* undetermined coupling constant). One then verifies whether the augmented action represents a complete system, i.e. one that does not lead to the generation of operators beyond those that are already present. If necessary, one has to repeat this step until a closed system is obtained.

III: Rescaling

One next rescales frequency/momentum so that the rescaled field amplitude ϕ' fluctuates on the same scales as the original field ϕ, i.e. one sets

$$q \to bq, \quad \omega \to b^z \omega.$$

Here, the **frequency renormalization exponent** or **dynamical exponent** z may be unity, two, or sometimes a non-integer value, depending on the effective dispersion relating frequency and momentum. We finally note that the field ϕ, as an integration variable, may be rescaled arbitrarily. Using this freedom, we select a term in the action which we believe governs the behavior of the "free" theory – in a theory with elastic coupling this might,

[13] L. H. Ryder, *Quantum Field Theory* (Cambridge University Press, 1996).
[14] For the definition of loops, see Section 5.1.

for example, be the leading-order gradient operator $\sim \int d^d r (\nabla\phi)^2$ – and require that it be strictly invariant under the RG step. To this end we designate a dimension L^{d_ϕ} for the field, chosen so as to compensate for the factor b^x arising after the renormalization of the operator. The rescaling

$$\phi \to b^{d_\phi}\phi,$$

is known as **field renormalization**. It renders the "leading" operator in the action scale invariant.

As a result of all these manipulations, we obtain a renormalized action

$$S[\phi] = \sum_a g'_a \mathcal{O}_a[\phi],$$

which is entirely described by the set of changed coupling constants, i.e. the effect of the RG step is fully encapsulated in the mapping

$$\mathbf{g}' = \tilde{R}(\mathbf{g}),$$

relating the old value of the vector of coupling constants, $\mathbf{g} = \{g_a\}$, to the renormalized one, $\mathbf{g}' = \{g'_a\}$. By letting the control parameter, $\ell \equiv \ln b$, of the RG step assume infinitesimal values, one can make the difference between bare and renormalized coupling constants arbitrarily small. It is then natural to express the difference $\mathbf{g}' - \mathbf{g} = \tilde{R}(\mathbf{g}) - \mathbf{g}$ in the form of a **generalized β-function** or **Gell-Mann–Low equation**

$$\boxed{\frac{d\mathbf{g}}{d\ell} = R(\mathbf{g}),} \tag{8.17}$$

where the right-hand side is defined through the relation $R(\mathbf{g}) = \lim_{\ell\to 0} \ell^{-1}(\tilde{R}(\mathbf{g}) - \mathbf{g})$.

INFO As mentioned at the beginning of the section, the formulation of the RG step above is actually not the only one possible. For instance, in high-energy physics, **other renormalization schemes** appear to be more natural. In this area of physics, there is actually no reason to believe in the existence of a well-defined "bare" action with finite coupling constants. (Contrary to the situation in condensed matter physics, the bare action of quantum electrodynamics, say, is in principle inaccessible.) However, one may legitimately require that, after an integration over UV-divergent fluctuations, the "renormalized" coupling constants of the theory (which, in turn, determine observables such as the physical electron mass) are finite. One may then postulate that the bare coupling constants of the theory are actually infinite. The value of these infinities is fine-tuned so as to combine with the fluctuation-induced "infinities" to realize finite renormalized coupling constants. Alternatively, one may deliberately add extra operators, **counter-terms**, to the action which are designed so as to cancel divergences due to fluctuations. However, the net result of all these RG schemes (which are by and large equivalent) is a mapping describing the flow of the coupling constants upon variation of a control parameter.

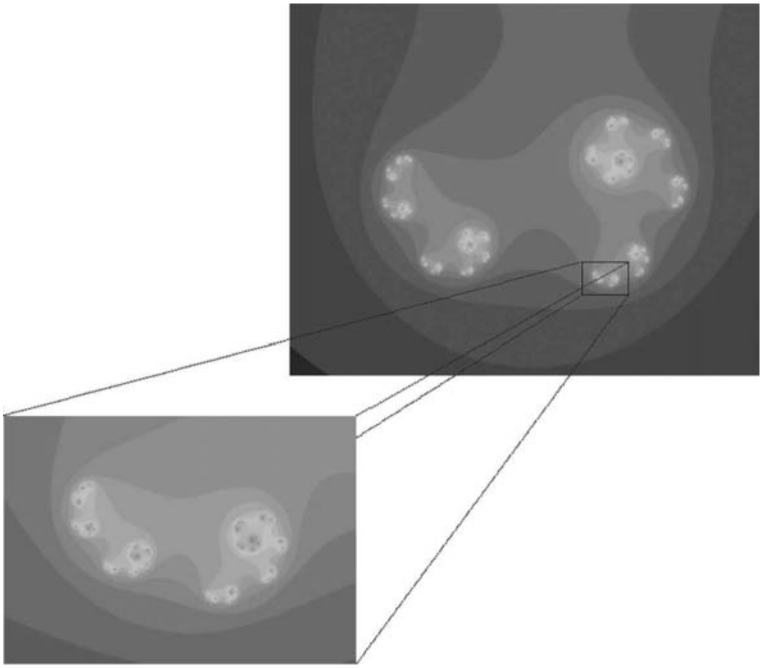


Figure 8.4 The fractal Julia set is self-similar in the sense that any sub-region of it contains the full information of the original set.

8.3.2 Analysis of the Gell-Mann–Low equation

The Gell-Mann–Low equation (8.17) represents the principal result of an RG analysis. Thinking of the control parameter ℓ as a kind of "flow parameter," one may identify this equation as a generalized dynamical system, namely the system describing the evolution of the effective coupling constants of a model upon changing length or time scales. As with any dynamical system, the prime structural characteristic of the set of equations (8.17) is the set of **fixed points**, i.e. the submanifold $\{\mathbf{g}^*\}$ of points in coupling constant space which are stationary under the flow: $R(\mathbf{g}^*) = 0$. Once the coupling constants are fine-tuned to a fixed point, the system no longer changes under subsequent RG transformations. In particular it remains invariant under the change of space/time scale associated with the transformation. Alluding to the fact that they look the same no matter how large a magnifying glass is used, systems with this property are referred to as **self-similar**. (For example, **fractals** such as the Julia set shown in Fig. 8.4 are paradigmatic examples of self-similar systems; the magnification of any sub-region of the fractal looks identical to the full system.)

Now, to each system, one can attribute at least one intrinsic length scale, namely the length ξ determining the exponential decay of field correlations. However, the existence of a finite, and pre-determined, intrinsic length scale clearly does not go together with invariance under scale transformations. We thus conclude that, at a fixed point, either $\xi = 0$ (not so interesting), or $\xi = \infty$. However, a diverging correlation length $\xi \to \infty$ is a hallmark of

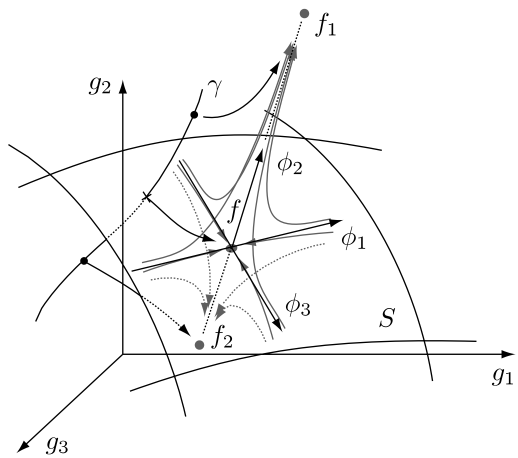


Figure 8.5 Showing the RG flow in the vicinity of a fixed point with two irrelevant (ϕ_1, ϕ_3) and one relevant (ϕ_2) scaling fields. The manifold S defined through the vanishing of the relevant field, $\phi_2 = 0$, is called a **critical surface**. On this submanifold, the RG flow is directed towards the fixed point f. Deviations off criticality make the system approach one of the stable fixed points f_1 and f_2.

a second-order phase transition. We thus tentatively identify fixed points of the RG flow as candidates for "transition points" of the physical system. (For a more comprehensive review of phase transitions and the critical phenomena accompanying them, see the Info block starting on page 436 below.) This being so, it is natural to pay special attention to the behavior of the flow in the immediate vicinity of the fixed-point manifolds. If the set of coupling constants, $\mathbf{g}$, is only close enough to a fixed point, $\mathbf{g}^*$, it will be sufficient to consider the linearized mapping

$$R(\mathbf{g}) \equiv R((\mathbf{g} - \mathbf{g}^*) + \mathbf{g}^*) \simeq W(\mathbf{g} - \mathbf{g}^*), \quad W_{ab} = \left. \frac{\partial R_a}{\partial g_b} \right|_{\mathbf{g}=\mathbf{g}^*}.$$

To explore the properties of flow, let us assume that we had managed to diagonalize the matrix W. Denoting the eigenvalues by $\lambda_\alpha, \alpha = 1, \ldots, N$, and the *left*-eigenvectors[15] by ϕ_α, we have

$$\phi_\alpha^T W = \phi_\alpha^T \lambda_\alpha.$$

The advantage of proceeding via the unconventional set of left-eigenvectors is that it allows us to conveniently express the flow of the physical coupling constants under renormalization. To this end, let v_α be the αth component of the vector $\mathbf{g} - \mathbf{g}^*$ when represented in the basis $\{\phi_\alpha\}$:

$$v_\alpha = \phi_\alpha^T (\mathbf{g} - \mathbf{g}^*).$$

[15] Since there is no reason for W being symmetric, the left- and right-eigenvectors may be different.

These components display a particularly simple behavior under renormalization:

$$\frac{dv_\alpha}{d\ell} = \phi_\alpha^T \frac{d}{d\ell}(\mathbf{g} - \mathbf{g}^*) = \phi_\alpha^T W(\mathbf{g} - \mathbf{g}^*) = \lambda_\alpha \phi_\alpha^T (\mathbf{g} - \mathbf{g}^*) = \lambda_\alpha v_\alpha.$$

Under renormalization, the coefficients v_α change by a mere scaling factor λ_α, wherefore they are called **scaling fields** – a somewhat unfortunate nomenclature. (The coefficients v_α are actually not fields but simply a set of ℓ-dependent coefficients, the vector of coupling constants when expressed in the basis of eigenvectors ϕ_α.) These equations are trivially integrated to obtain

$$\boxed{v_\alpha(\ell) \sim \exp(\ell\lambda_\alpha).}$$

This result suggests a discrimination between at least three different types of scaling fields:

- ▷ For $\lambda_\alpha > 0$ the flow is directed away from the critical point. The associated scaling field is said to be **relevant** (in the sense that it forcefully drives the system away from the critical region). In Fig. 8.5, v_2 is a relevant scaling field.
- ▷ In the complementary case, $\lambda_\alpha < 0$, the flow is attracted by the fixed point. Scaling fields with this property (v_1, v_3) are said to be **irrelevant**.[16]
- ▷ Finally, scaling fields which are invariant under the flow, $\lambda_\alpha = 0$, are termed **marginal**.[17]

The distinction of relevant/irrelevant/marginal scaling fields in turn implies a classification of different types of fixed points:

- ▷ Firstly, there are **stable fixed points**, i.e. fixed points whose scaling fields are all irrelevant or, at worst, marginal. These points define what we might call "stable phases of matter": when you release a system somewhere in the parameter space surrounding any of these attractors, it will scale towards the fixed point and eventually sit there. Or, expressed in more physical terms, looking at the problem at larger and larger scales will make it more and more resemble the infinitely correlated self-similar fixed-point configuration. (Recall the example of the high-temperature fixed line of the one-dimensional Ising model encountered earlier.) By construction, the fixed point is impervious to moderate variations in the microscopic morphology of the system, i.e. it genuinely represents what one might call a "state of matter."
- ▷ Complementary to stable fixed points, there are **unstable fixed points**. Here, all scaling fields are relevant (cf. the $T = 0$ fixed point of the 1-D Ising model). These fixed points represent the concept of a Platonic ideal: you can never get there and, even if you managed

[16] The terminology "irrelevant" indicates that a scaling field of negative dimension usually does not play much of a physical role. There are, however, exceptions to this rule. For instance, it may happen that the free energy of the system depends in a singular manner on an irrelevant scaling variable – in which case the variable is called **dangerously irrelevant**. Dangerously irrelevant scaling variables not only strongly affect the outcome of the theory, but also invalidate the applicability of a number of established concepts of RG theory (such as the scaling laws to be discussed below).

[17] A marginal scaling field corresponds to a direction in coupling constant space with vanishing partial derivative, $\partial_{\phi_\alpha} R|_{\mathbf{g}^*=0} = 0$. In this case, to obtain a refined picture, one sometimes considers the second-order derivative, $\partial^2_{\phi_\alpha} R|_{\mathbf{g}^*=0} \equiv 2x$. In the vicinity of the fixed point, the scaling field then behaves as $d\ v_\alpha = x v_\alpha^2$. For $x > 0$ $(x < 0)$ the field has the status of a **marginally relevant (irrelevant)** scaling field. It is relevant (irrelevant) on account of the non-vanishing direction of the flow. However, it is also "marginal" because the speed of the flow decreases upon approaching the critical regime.

to approach it closely, the harsh conditions of reality will make you flow away from it. Although unstable fixed points do not correspond to realizable forms of matter, they are of importance inasmuch as they "orient" the global RG flow of the system.

▷ Finally, there is the **generic class of fixed points** with both relevant and irrelevant scaling fields. These points are of particular interest inasmuch as they can be associated with **phase transitions**. To understand this point, we first note that the r eigenvectors Φ_α associated with irrelevant scaling fields span the tangent space of an (r)-dimensional manifold known as the **critical surface**. (A schematic illustration for the case $r = 2$ is shown in Fig. 8.5.) This critical manifold forms the **basin of attraction** of the fixed point, i.e. whenever a set of physical coupling constants $\mathbf{g}$ is fine-tuned so that $\mathbf{g} \in S$, the expansion in terms of scaling fields contains only irrelevant contributions and the system will feel attracted to the fixed point as if it were a stable one.

However, the smallest deviation from the critical surface introduces a relevant component driving the system exponentially away from the fixed point. A sketch of the resulting flow is shown in Fig. 8.5 for the case of just one relevant scaling field. For example, in the case of the ferromagnetic phase transition – discussed in more detail in the next section – deviations from the critical temperature T_c are relevant. If we consider a system only slightly above or below T_c, it may initially (on intermediate length scales) appear to be critical. However, upon further increasing the scale, the relevant deviation will grow and drive the system away from criticality, either towards the stable high-temperature fixed point of the paramagnetic phase ($T > T_c$) or towards the ferromagnetic low-temperature phase ($T < T_c$).

This picture actually suggests that systems with generic fixed points typically possess complementary stable fixed points, i.e. fixed points towards which the flow is directed after it has left the critical region. We also notice that a scaling direction that is relevant at one fixed point (e.g. Φ_2 at the critical fixed point) may be irrelevant at others (Φ_2 at the high- and low-temperature fixed points).

INFO The discussion above suggests that the concept of renormalization is intimately linked to the **theory of phase transitions and critical phenomena**, the traditional platform for the development of the subject in the literature. In view of the existing wealth of literature (and acknowledging the fact that we are approaching the field from a more operational perspective), we shall not endeavor to present another "introduction to the theory of critical phenomena." Rather, we will summarize in a concise, but hopefully self-contained, manner, those few tenets and principles that are necessary to place the concept of renormalization into a larger physical context.

The most fundamental[18] signature of a phase transition is its **order parameter**, M, i.e., a quantity whose value unambiguously identifies the phase of the system. Examples from classical statistical mechanics include the magnetization for the ferromagnetic–paramagnetic transition, the density for the liquid–vapor transition, the order parameter amplitude for the BCS transition, etc. (However, to keep the terminology concrete, we shall mostly use the language of the ferromagnetic transition in the following.)

[18] Notice that there are transitions whose order parameter is actually unknown. A famous example is the **quantum Hall transition** discussed in more detail in Section 9.3.4.

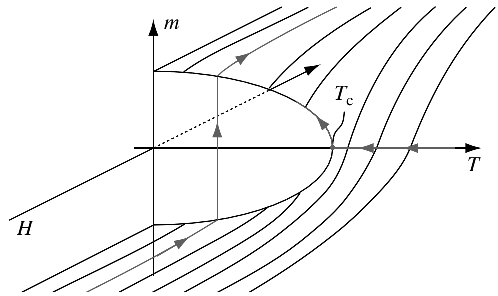


Figure 8.6 Phase diagram of the ferromagnetic transition. Tuning a magnetic field at fixed temperature $T < T_c$ through zero causes the magnetization to jump discontinuously: $[0, T_c]$ is a line of first-order transitions. This line terminates in the unique second-order transition point of the system, $(T = T_c, H = 0)$. Lowering the temperature at $H = 0$ causes the non-analytic, but continuous, development of a finite magnetization at $T < T_c$.

Transitions between different phases of matter fall into two large categories.[19] In **first-order** phase transitions the order parameter exhibits a discontinuous jump across the transition line while, in the complementary class of **second-order** transitions, the order parameter changes in a non-analytic but continuous manner. (The two cases are exemplified in Fig. 8.6 by the classical ferromagnet.)

The phenomenology of second-order transitions is generally richer than that of first-order transitions. As a thermodynamic state variable, the order parameter is coupled to a **conjugate field**, $H : M = -\partial_H F$, where F is the free energy. At a second-order transition, M changes non-analytically, which means that the second-order derivative, a **thermodynamic susceptibility**, $\chi = -\partial_H^2 F$, develops a singularity. Now, you may recall from the discussion of the fluctuation dissipation theorem that the susceptibility is intimately linked to the field fluctuation behavior of the system. More precisely, χ is proportional to the integral over the correlation function C determining the fluctuation behavior of the fields (cf. Eq. (8.3)). A divergence of the susceptibility implies the accumulation of infinitely long-range field fluctuations.

The divergence of the susceptibility goes hand in hand with non-analytic and/or singular behavior of all sorts of other physical quantities. In fact, an even stronger statement can be made. We have seen that, right at the transition/fixed point, the system is self-similar. This implies that the behavior of its various characteristics must be described by **power laws**. Referring for a more substantial discussion to Section 8.3.3 below, we here merely support this statement by a heuristic argument. Consider a function $f(t)$, where f is representative of an observable of interest and t is a control parameter (a scaling field) determining the distance to the transition point. In the immediate vicinity of the transition point, f is expected to "scale," i.e. under a change of the length scale $x \to x/b$, $t \to tb^{-D_t}$, the function f must, at most, change by a factor reflecting its own scaling dimension. $f(t) = b^{D_f} f(tb^{-D_t})$. (A more serious, structural change of the function would be in conflict with asymptotic self-similarity.) Mathematically speaking, this equation amounts to homogeneity of the function f, equivalently expressed by $f \sim t^{D_f/D_t}$.

19 Readers absolutely unfamiliar with the thermodynamics of phase transitions may wish to consult the corresponding section of a textbook on statistical mechanics.

The set of different exponents characterizing the relevant power laws occurring in the vicinity of the transition are known as **critical exponents**. For at least four different reasons, the set of critical exponents represents the most important structural fingerprint of a transition:

1. They carry universal significance, i.e. we do not have to invent a set of critical exponents for each transition anew. (For example, the divergence of the correlation length, $\xi \sim |t|^{-}$, is characterized by a critical exponent commonly, and irrespective of the particular transition under consideration, denoted by ν.)
2. The set of critical exponents carries the same information as the set of exponents of the scaling fields, i.e. knowledge of the critical exponents is equivalent to the knowledge of the linear dynamical system characterizing the flow in the transition region. (In fact, the set of critical exponents overdetermines the scaling field exponents, i.e. it contains redundancy. For example, of the six critical exponents characterizing the magnetic transition, only two are independent. The others are interrelated by[20] **scaling laws** or **exponent identities** to be discussed below.
3. Critical exponents are fully universal; they are numbers depending, at most, on dimensionless characteristics such as the space-time dimensionality or number of components of the order parameter.
4. Perhaps most importantly, the critical exponents represent quantities that can be measured. In fact, their universality and structural importance make them quantities of prime experimental interest.

In the following, let us briefly enumerate the list of the most relevant exponents, α, β, γ, δ, η, ν, and z.[21] Although we shall again make use of the language of the magnetic transition, it is clear that (and, indeed, how) the definitions of most exponents generalize to other systems.

α: In the vicinity of the critical temperature, the **specific heat** $C = -T\partial_T^2 F$ scales as $C \sim |t|^{-\alpha}$, where $t = (T - T_c)/T_c$ measures the distance to the critical point. Note that, by virtue of this definition, a non-trivial statement has been made: although the phases above and below the transition are essentially different, the scaling exponents controlling the behavior of C are identical. The same applies to most other exponents listed below.

β: Approaching the transition temperature from below, the **magnetization** vanishes as $M \equiv -\partial_H F\big|_{H\ \ 0} \sim (-t)^\beta$.

γ: The **magnetic susceptibility** behaves as $\chi \equiv \partial_h M|_{h\ \ 0} \sim |t|^{-\gamma}$.

δ: At the critical temperature, $t = 0$, the **field dependence of the magnetization** is given by $M \sim |h|^{1/\delta}$.

ν: Upon approaching the transition point, the **correlation length** diverges as $\xi \sim |t|^{-}$.

η: This implies that the correlation function,

$$C(\mathbf{r}) \sim \begin{cases} \frac{1}{|\mathbf{r}|^{d-2+\ }}, & |\mathbf{r}| \ll \xi, \\ \exp[-|\mathbf{r}|/\xi], & |\mathbf{r}| \gg \xi, \end{cases}$$

crosses over from exponential to a power law scaling behaviour at the length scale ξ. To motivate the power, one may notice that $C \sim \langle\phi\phi\rangle$ carries twice the dimension of the field

[20] Unfortunately, the language used in the field of critical phenomena makes excessive use of the prefix "scaling."
[21] Historically, the exponents are drawn from the first six letters of the Greek alphabet. The exceptional designation of the last exponent, z, betrays the fact that quantum dynamical fluctuations were considered only later.

ϕ. The engineering dimension of the latter follows from the requirement that the gradient operator $\sim \int d^dr\,(\quad \phi)^2$ be dimensionless: $[\phi] = L^{(2-d)/2}$, according to which $C(\mathbf{r})$ has canonical dimension L^{2-d}. The exponent η, commonly referred to as the **anomalous dimension** of the correlation function, measures the mismatch between the observed and the canonical dimension.

z: A quantum theory can, to a large extent, be viewed as a kind of classical theory in $d+1$ dimensions. The theory is "quantum critical" if the effective classical theory contains a critical point. In the vicinity of that point large fluctuations are observed in both the d spatial directions and the temporal "direction." However, the different physical origin of these dimensions manifests itself in the scaling being anisotropic. Denoting the correlation length in the temporal direction by τ, we define $\tau \sim \xi^z$, where deviations $z = 1$ in the **dynamical exponent** measure the degree of anisotropy.

Now, a moment's thought shows that, of the six classical exponents, only a few can be truly independent. Previously we have noted that, modulo irrelevant perturbations, the flow in the vicinity of a transition point is controlled by the relevant scaling fields. Referring for a more quantitative discussion to Section 8.4 below, we anticipate that, for the magnetic transition, the magnetic field will certainly represent a relevant perturbation (a fact readily expressed by the positivity of the exponent δ). Moreover, deviations from the critical temperature, $t = 0$, are also relevant.[22] However, for the magnetic transition, that exhausts the list; in the asymptotic vicinity of the transition, the flow is controlled by a two-dimensional dynamical system. This suggests that four constraining equations should reduce the set of six classical exponents to only two independent ones. Historically, these **scaling laws** were discovered one by one (at a time when the underlying connections to the system of "scaling fields" had not been fully appreciated). For the sake of reference, these constraint equations (along with the names of the people who discovered them) are listed below. In Section 8.3.3 below, we exemplify how the scaling laws can be transparently derived from the intrinsic structure of the theory.

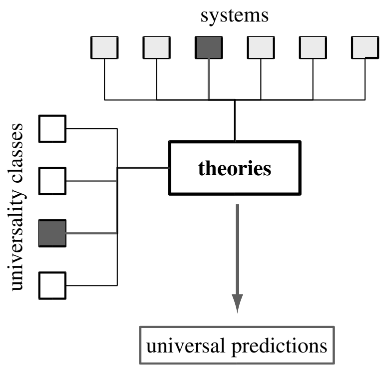
Fisher	$\nu(2-\eta) = \gamma$
Rushbrooke	$\alpha + 2\beta + \gamma = 2$
Widom	$\beta(\delta - 1) = \gamma$
Josephson	$2 - \alpha = \nu d$

For practical purposes, we need only compute/measure two exponents – no matter which – to fully specify the scaling structure of the theory.

In the next section we discover that the dynamical system of scaling fields encapsulates practically all information about "critical" fluctuation phenomena accompanying a phase transition. However, for the moment, we shall restrict ourselves to the discussion of one more aspect of conceptual importance, namely **universality**. In fact, the majority of critical systems can be classified into a relatively small number of **universality classes**. Crudely speaking, leaving apart more esoteric classes of phase transitions there are $\mathcal{O}(10^1)$ fundamentally different types of flow recurrently appearing in practical applications. This has to

[22] If you find it difficult to think of temperature as a "coupling constant," remember that, in our derivation of the ϕ^4-model as the relevant theory of the magnetic transition, the coupling constant of the "mass operator" $r\int d^dr\,\phi^2$ turned out to be proportional to the reduced temperature $t = |T - T_c|/T_c$.

be compared with the near infinity of different physical systems that display critical phenomena. Why, then, is it that the plethora of all these transitions can be grouped into a very limited set of different universality classes? Remarkably, the origin of this universality can readily be understood from the concept of critical surfaces.



Imagine, then, an experimentalist exploring a system that is known to exhibit a phase transition. Motivated by the critical phenomena that accompany phase transitions, the available control parameters X_i (temperature, pressure, magnetic field, etc.) will be varied until the system begins to exhibit large fluctuations. On a theoretical level, the variation of the control parameters determines the initial values of the coupling constants of the model (as they functionally depend on the X_is through their connection to the microscopic Hamiltonian). In Fig. 8.5 the curve in coupling constant space defined in this way is indicated by γ. For microscopic parameters corresponding to a point above or below the critical manifold, the system asymptotically (i.e. when looked at at sufficiently large scales) falls into either the "high-" or the "low-temperature" regime (as indicated by the curves branching out from γ in Fig. 8.5). However, eventually the trajectory through parameter space will intersect the critical surface. For this particular set of coupling constants, the system is critical. As we look at it on larger and larger length scales, it will be attracted by the fixed point at S, i.e. it will display the universal behavior characteristic of this particular point. This is the origin of universality: variation of the system parameters in a different manner (or for that matter considering a second system with different material constants) will generate a different trajectory $g_\alpha(\{X_i\}) = \gamma'$. However, as long as this trajectory intersects with S, it is guaranteed that the critical behavior will exhibit the same universal characteristics (controlled by the unique fixed point).

In fact a more far-reaching statement can be made. Given that there is an infinity of systems exhibiting transition behavior (symbolically indicated by the row of boxes in the upper part of the figure above) while there is only a very limited set of universality classes (the set of boxes on the left), many systems of very different microscopic morphology must have the same universal behavior. More formally, different microscopic systems must map onto the same critical low-energy theory. Examples of these coincidences include (to mention but a few entries of an endless list) the equivalence of the disordered Luttinger liquid to a Josephson junction (cf. Problem 6.7), the equivalence of models of planar magnets (see Section 8.6 below) to two-dimensional classical Coulomb plasmas, and the equivalence of the liquid–gas transition to the ferromagnetic transition. (In all cases, "equivalence" means that the systems exhibit identical scaling behavior and, therefore, fall into the same universality class.) Further coincidences of this type will be encountered below.

8.3.3 Scaling theory

Previously, we have seen that the dynamical system of scaling fields encodes a wealth of information on the large-scale structure and on the phases of a physical system. However, we have not yet established a connection between the concept of renormalization and concrete (i.e. experimentally accessible) data. This is the subject of the present section. Imagine, then, that we had represented some observable of experimental interest, X, in the language of the functional integral. According to the discussion of the previous chapter this means that we have managed to express

$$X = \sum_{\mathbf{p}} C(p_i, g_\alpha),$$

as the sum over an n-point correlation function $C(p_i, g_\alpha) = \langle (\cdots)\phi\phi\cdots\phi\rangle_\phi$, where the (symbolic) notation indicates that C may depend both on the momentum scale at which it is evaluated (e.g. through the explicit momentum dependence of current operators, etc.) and on the coupling constants. The ellipsis $(\cdots)$ stand for optional algebraic elements entering the definition of the correlation function.

We next build on our assumption of renormalizability of the theory, i.e. we make use of the fact that we can evaluate C before or after an RG step; the result must be the same. On the other hand, the RG transformation will, of course, not leave the individual constituents entering the definition of C invariant; it will change coupling constants, g_α, the momenta p_i, and the field amplitudes ϕ according to the prescriptions formulated in the previous section. Expressed in a single formula,

$$\boxed{C(p_i, g_\alpha) = b^{nd_\phi} C(p_i b, g_\alpha b^{\lambda_\alpha}),} \tag{8.18}$$

where we have simplified the notation by assuming that the coupling constants themselves scale (for, otherwise, the matrix elements of a linear transformation mediating between the coupling constants and the scaling fields would appear). For notational convenience, let us also assume that the fixed point values of the coupling constants are specified in such a way that $\mathbf{g}^* = 0$. The factor b^{nd_ϕ} accounts for the explicit rescaling of the n fields entering the definition of C.

Notice that Eq. (8.18) presents a remarkable statement. Although the three different elements (ϕ, p_i, g_α) contributing to the correlation function change under the transformation in seemingly unrelated manners, the net result of the concerted rescaling is nil. Indeed, Eq. (8.18) serves as a starting point for the derivation of various relations of immediate practical relevance.

Scaling functions

Let us return to a principle already employed in connection with the one-dimensional Ising model. For concreteness, imagine that we are working under conditions where there is just a single relevant scaling field g_1, while all $g_{\alpha>1}$ are irrelevant (or, for that matter, marginal). We can then write

$$\begin{aligned} C(p_i, g_1, g_\alpha) &= b^{nd_\phi} C(p_i b, g_1 b^{\lambda_1}, g_\alpha b^{\lambda_\alpha}) = g_1^{-nd_\phi/\lambda_1} C(p_i g_1^{-1/\lambda_1}, 1, g_\alpha g_1^{-\lambda_\alpha/\lambda_1}) \\ &\overset{g_1 \ll 1}{\approx} g_1^{-nd_\phi/\lambda_1} C(p_i g_1^{-1/\lambda_1}, 1, 0) \equiv g_1^{-nd_\phi/\lambda_1} F(p_i g_1^{-1/\lambda_1}). \end{aligned}$$

Here, we have used the freedom of arbitrarily choosing the parameter b to set $g_1 b^{\lambda_1} = 1$ while, in the third equality, we have assumed that we are sufficiently close to the transition that the dependence of C on irrelevant scaling fields is inessential. The function F defined through the relation

$$C(p_i, g_1) = g_1^{-nd_\phi/\lambda_1} F(p_i g_1^{-1/\lambda_1}), \tag{8.19}$$

is an example of a **scaling function**. Alternatively (for example, if C represents a thermodynamic observable or a global transport coefficient) we might be interested in the correlation function $C(g_1, g_\alpha) \equiv C(p_i = 0, g_1, g_\alpha)$ at zero external momentum $p_i = 0$. In this case, a typical question to ask would be the dependence of C on the most relevant *and* the second most relevant control parameter g_2 (where we leave unspecified whether g_2 is relevant, marginal, or irrelevant). Following the same logic as above, one obtains

$$C(g_1, g_2) = g_1^{-nd_\phi/\lambda_1} \tilde{F}(g_2 g_1^{-\lambda_2/\lambda_1}),$$

with some different scaling function $\tilde{F}$.

INFO As an example particularly relevant to the comparison between analytical theory and numerics, we note the concept of **finite-size scaling**. While analytical theories are most conveniently formulated in the thermodynamic limit, numerical simulations are carried out for systems of still very limited size. The need to compare theory and numerical simulations motivates the need to explicitly keep track of the system size under renormalization. Indeed, the system size L has dimension [length] and, therefore, gets rescaled as $L \to L/b$. Setting $L/b = 1$, we obtain a scaling function

$$G(g_\alpha, L) = L^{nd_\phi} F_{\text{fs}}(g_\alpha L^{\quad\alpha}),$$

with explicit system size dependence.

While the construction of any particular scaling function may be context-dependent, the principle behind the derivation is general: once the scaling behavior of a correlation function is known, the arbitrariness of the scaling parameter b can be used to reduce the number of independent variables by one. The reduced correlation function is called a scaling function. As with the response functions discussed in the previous chapter, scaling functions also represent a prime **interface between theory and experiment**. Experimentally, the measurement of an observable X in its dependence on a number of relevant system parameters, t and h say, results in a multi-parameter function $X(t, h)$. In fact, a better way to think about this object is as a set of one-dimensional functions $X_h(t)$ depending on a parameter h. (This is because, in experiment, one typically varies only a single control parameter, e.g. temperature at fixed magnetic field.) Scaling implies that all these functions collapse onto a generic one-dimensional[23] profile, if only the data are plotted as a function of the relevant scaling parameter th^x.

This mechanism can be exploited in several different ways. For example, if there is not yet a theory of the transition phenomenon in question, an experimentalist may empirically identify the relevant scaling parameters and pose the explanation of the observed scaling exponent x – by construction a fully universal number – as a challenge to theorists.[24] Conversely, theorists may suggest a scaling exponent that can be put to the test by checking whether the experimental data collapse onto this exponent. Summarizing, one of the great virtues of the concept of scaling is that it condenses the information exchange between experiment and theory (and analytical theory and numerics for that matter) into a small set of universal numbers.

INFO For the sake of completeness we mention that, especially in the field theoretical community, the information encapsulated in the scale-dependent correlation functions is often represented in a different manner. Starting out from the relation

$$C(p_i, g_\alpha) = e^{n \ \ d_\phi} C(p_i e \ \ , g_\alpha(\ell)),$$

where we have set $b = e$, we can use the ℓ-independence of the left-hand side to write $0 = \frac{d}{d} e^{n \ \ d_\phi} C(p_i e \ \ , g_\alpha(\ell))$. (Notice that, here, we do not need to be in the asymptotic scaling regime, i.e., for the sake of the present construction, the ℓ-dependence of the coupling constants need not be explicitly exponential.) We next carry out the ℓ-differentiation to obtain

$$\boxed{n \ \ d_{\mathrm{e},\phi} + \frac{\eta}{2} \ \ + \partial \ \ + \beta_\alpha(g_\alpha)\partial_{g_\alpha} \ \ C(p_i e \ \ , g_\alpha(\ell)) = 0.} \tag{8.20}$$

Here, $d_{\mathrm{e},\phi}$ is the engineering dimension of the field ϕ and $\eta/2 = d_\phi - d_{\mathrm{e},\phi}$ its anomalous dimension (see the definition of η in the Info block starting on page 436). Further, the partial derivative ∂ acts on the explicit scale dependence of the momentum arguments (or any other explicitly scale-dependent argument for that matter). Finally, $\beta_\alpha(g_\alpha)$ is the β-function defined above. Equation (8.20) is known as a **renormalization group equation**. Both the RG equation and the scaling form that we used to derive it equivalently express the scaling behavior of the correlation function.

Scaling functions and critical exponents

Another important aspect of scaling theory is that it can be used to disclose relations between the seemingly independent[25] critical exponents of the theory. For the sake of concreteness, let us consider the case of the ferromagnetic transition, i.e. a transition we have previously characterized in terms of six critical exponents $\alpha, \ldots, \eta$ (see page 438). However, the flow in the vicinity of the magnetic fixed point is controlled by only two relevant scaling fields, the (reduced) temperature t and the reduced magnetic field $h \equiv H/T$. Neglecting

[23] For an n-dimensional data set, the collapse is to an $(n-1)$-dimensional functional set.

[24] Parenthetically, one may note that the empirical collapse of experimental data onto scaling functions requires a lot of skill. For example, if the data set consists of a number of functional "patches" of only limited overlap, it is quite "easy" to construct a scaling function of, in fact, almost any desired power law dependence. Data of this type tend to contain a lot of statistical uncertainty, which can easily lead to erroneous conclusions.

[25] After all, the critical exponents describe the behavior of quite different physical observables in the transition region.

irrelevant perturbations, we thus conclude that, under a renormalization group transformation, the reduced free energy $f = F/TL^d$ will behave as[26] $f(t,h) = b^{-d} f(tb^{y_t}, hb^{y_h})$. We next fix $tb^{y_t} = 1$ to reduce the number of independent variables to one:

$$\boxed{f(t,h) = t^{d/y_t} \tilde{f}(h/t^{y_h/y_t}).} \tag{8.21}$$

Containing the complete thermodynamic information, Eq. (8.21) is all that we need to compute the critical exponents. Indeed, comparing with the definitions summarized on page 438, it is straightforward to show that

$$\left.\begin{array}{lll} \alpha = 2 - \dfrac{d}{y_t}, & \beta = \dfrac{d - y_h}{y_t}, & \gamma = \dfrac{2y_h - d}{y_t}, \\ \delta = \dfrac{y_h}{d - y_h}, & \nu = \dfrac{1}{y_t}, & \eta = 2 + d - 2y_h, \end{array}\right\} \tag{8.22}$$

from where follow the cross-relations summarized in the table on page 439 by direct comparison. These relations illustrate our previous assertion that, conceptually, the dimensions of the relevant scaling fields have a more fundamental status than the critical exponents.

EXERCISE Verify these statements. To obtain the fifth relation, the **hyperscaling relation**, notice that, under a change of scale, $\xi \to b\xi$. On the other hand, we know that $t \sim \xi^{-1/}$. The sixth relation is obtained from Eq. (8.3) by a substitution of the definition of the spatial profile of the correlation function in terms of the critical exponent η into the integral to obtain a relation between the critical exponents γ and η (Fisher's scaling law).

8.4 RG analysis of the ferromagnetic transition

In the previous section, we became acquainted with some fundamental elements of the structure of RG analyses, and their connection to the theory of critical phenomena. Being kept at a general and conceptual level, the discussion may have seemed somewhat abstract. Therefore, to elucidate the concepts introduced above, and to introduce some more elements of the RG, we turn now to a concrete application of the approach to the classical theory of the (uniaxial) ferromagnetic (or liquid–gas) transition. In Section 5.1, the ϕ^4-theory was identified as an effective low-energy model of the ferromagnetic system. However, beyond the mean-field, we have not yet applied the model to explore the universal characteristics of the transition. In the following, we shall see that RG methods, and only RG methods, can be applied to successfully understand much of the intriguing behavior displayed by the $(d > 2)$-dimensional Ising model in the vicinity of its phase transition.

[26] Here we have made use of the fact that the reduced free energy does not carry an anomalous dimension. By definition, the free energy $F = -T \ln \mathcal{Z}$ does not change under renormalization (which after all, merely amounts to representing the number $\mathcal{Z}$ through functional integrals of different space-time resolution). Thus, the scaling of the reduced free energy is entirely carried by the prefactor L^{-d}.

8.4.1 Preliminary dimensional analysis

The first question that we wish to address has a somewhat technical status: with what justification was the Ising model represented in terms of the model action[27]

$$S[\phi] = \int d^d r \left[\frac{r}{2}\phi^2 + \frac{1}{2}(\nabla\phi)^2 + \frac{\lambda}{4!}\phi^4 - h\phi \right], \tag{8.23}$$

i.e. why was it possible to neglect both higher powers and gradients of the field ϕ that are surely present in the exact reformulation of the Ising problem in terms of ϕ-variables? To rationalize the neglect of these terms, we proceed by dimensional analysis. Anticipating that the "real" dimensions carried by the operators in the action will be not too far from their engineering dimensions (see below), we begin by exploring the latter. We proceed along the lines of the general scheme outlined in the previous chapter and attribute a dimension of unity to the leading gradient term $\int (\nabla\phi)^2$ in the action. This entails the choice $[\phi] = L^{(2-d)/2}$, from where it is straightforward to attribute engineering dimensions to all other operators:

$$\left[\int \phi^2\right] = L^2, \quad \left[\int \phi^4\right] = L^{-d+4}, \quad \left[\int \phi^n\right] = L^{d+(2-d)n/2}, \quad \left[\int (\nabla^m \phi)^2\right] = L^{2(1-m)}.$$

These relations convey much about the potential significance of all structurally allowed operators:

- ▷ The engineering dimension of the non-gradient operator $\sim \phi^2$ is positive in all dimensions, indicating general relevance.
- ▷ The ϕ^4 operator is relevant (irrelevant) in dimensions $d < 4$ ($d > 4$). This suggests that for $d > 4$ a harmonic approximation ($\lambda = 0$) of the model should be reasonable. It also gives us a preliminary clue as to how we might want to approach the ϕ^4-model on a technical level: while for dimensions "much" smaller than $d = 4$ the interaction operator $\sim \phi^4$ is strongly relevant, the dimension $d = 4$ itself is borderline. This suggests that we analyze the model at $d = 4$, or maybe "close"[28] to $d = 4$ where the ϕ^4 operator is not yet that virulent, and then try to extrapolate to infer what happens at the "physical dimensions" of $d = 2$ and 3.
- ▷ Operators $\phi^{n>4}$ become relevant only in dimensions $d < (-1/n + 1/2)^{-1} < 4$. However, even below these threshold dimensions, operators of high powers in the field variable are much less relevant than the dominant non-harmonic operator $\int \phi^4$. This is the *a posteriori* justification for the neglect of $\phi^{n>4}$ operators in the derivation of the model.
- ▷ Similarly, operators with more than two gradients are generally irrelevant and can be neglected in all dimensions.
- ▷ In contrast, the operator $\int \phi$ coupling to the magnetic field carries dimension $1 + d/2$ and is therefore always strongly relevant.

[27] Generalizing our discussion from Section 5.1, we have incorporated a coupling to an external field. (Exercise: Recapitulate the construction of Section 5.1 to convince yourself that, to lowest order in an expansion in terms of ϕ, coupling the system to a magnetic field leads to the fourth term of Eq. (8.23). In case you are too impatient to do this: justify the structure of the term on physical grounds.)

[28] As we see shortly, the analysis of the problem is readily generalized to non-integer dimensions.

Dimensional analysis provides us with some valuable hints as to the importance of various operators appearing in the theory. It also indicates that, in the present context, dimension $d = 4$ might play a special role. Guided by this information, we now proceed to analyze the model in a sequence of steps of increasing sophistication.

8.4.2 Landau mean-field theory

Given an action of the form (8.23), the first thing one might try is a mean-field analysis. That is, assuming that our coupling constants r and λ are sufficiently large we might assume that the functional integral over ϕ is centered around solutions of the equation $\frac{\delta S[\phi]}{\delta \bar{\phi}} = 0$, or

$$r\bar{\phi} + \frac{\lambda}{6}\bar{\phi}^3 - h = 0, \tag{8.24}$$

where we have used the fact that the low-energy mean-field configuration will be spatially constant. Just by inspecting the potential part of the field-free Lagrangian, $\frac{r}{2}\phi^2 + \frac{\lambda}{4!}\phi^4$, it is clear that, depending on the sign of r, the mean-field equation possesses two fundamentally different types of solution. For $r > 0$, the action has a global minimum at $\phi = 0$, implying that $\bar{\phi} = 0$ is the unique mean-field (see Fig. 8.7). Noticing that the amplitude of ϕ represents a measure of the magnetization of the system (which is clear from the way the ϕ^4-action was derived from the Ising model on page 196), we identify $r > 0$ as a phase of zero net magnetism, the **paramagnetic phase**.

In contrast, for $r < 0$, the action has two degenerate minima at non-zero values, $\bar{\phi} = \pm\phi_0 \equiv \pm(6|r|/\lambda)^{1/2}$ (see Fig. 8.7). The system then has to make a choice as to whether it wants to sit in the ground state configuration $\bar{\phi} = \phi_0$ or $\bar{\phi} = -\phi_0$. This is the state of spontaneous symmetry breaking indicative of the low-temperature **ferromagnetic phase**. (Notice that, upon the switching on of a small magnetic field, the degeneracy between the two ground states is lifted and the system will populate a state of predetermined magnetization, $\bar{\phi} = \pm\phi_0$, depending on the sign of h.)

The preliminary analysis above indicates that r has the status of a fundamental parameter tuning the system through the ferromagnetic transition. Indeed, the microscopic analysis in Section 5.1 had indicated that $r \sim T - T_c$ was a function of temperature that changed sign at some critical temperature T_c, the mean-field critical temperature of the transition. However, even if we did not know the microscopics, it would be clear that $r(T)$ is (i) *some*

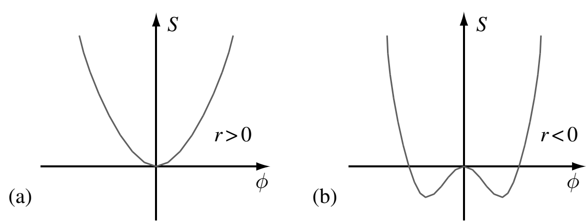


Figure 8.7 Action of the ϕ^4-theory evaluated on a constant field configuration above (a) and below (b) the critical point.

Table 8.1 *Critical exponents of the ferromagnetic transition obtained through di erent methods. Experimental exponents represent cumulative data from various three-dimensional ferromagnetic materials.*

Exponent	Experiment	Mean-field	Gaussian	ϵ^1	ϵ^5
α	0–0.14	0	1/2	1/6	0.109
β	0.32–0.39	1/2	1/4	1/3	0.327
γ	1.3–1.4	1	1	7/6	1.238
δ	4–5	3	5	4	4.786
ν	0.6–0.7		1/2	7/12	0.631
η	0.05		0	0	0.037

Source: Data taken from K. Huang, *Statistical Mechanics* (Wiley, 1987).

function of temperature which (ii) must have a zero at some temperature $T = T_\mathrm{c}$ (otherwise there would be no transition to begin with). Therefore, in the vicinity of $T = T_\mathrm{c}$, we can set $r \sim T - T_\mathrm{c}$ as our prime measure of the distance to the critical point. (This observation is, in fact, in perfect agreement with our earlier observation that the operator $\int \phi^2$ coupled to r is relevant – see the discussion in Section 8.3.2.)

What can mean-field theory say about the prime descriptors of the transition, the **critical exponents**? Identifying the field amplitude ϕ (alias the magnetization) with the order parameter of the transition, and referring back to our list of exponents on page 438, the low-temperature profile is given by $|\bar{\phi}| = (12|r|/\lambda)^{1/2} \sim |t|^{1/2}$, implying that $\beta = 1/2$. The exponent γ is obtained by differentiating Eq. (8.24) with respect to h. With $\chi \sim \partial_h \phi$, it is then straightforward to verify that, on approaching the critical point from either side of the transition, $\chi \sim |t|^{-1}$, implying an exponent $\gamma = 1$. The action evaluated on the mean-field-configuration takes the form

$$\frac{S[\bar{\phi}]}{L^d} = \frac{r}{2}\bar{\phi}^2 + \frac{\lambda}{4!}\bar{\phi}^4 \sim \begin{cases} \lambda^{-1}t^2, & t < 0, \\ 0, & t > 0. \end{cases} \tag{8.25}$$

With the mean-field free energy $F = TS[\bar{\phi}]$ we find that the specific heat $C = -T^2\partial_T^2 F \sim \partial_t^2 S$ behaves as a step function at the transition point, implying $\alpha = 0$. Right at the critical temperature, $r = 0$, the mean-field magnetization depends on h as $\bar{\phi} \sim h^{1/3}$, implying that $\delta = 3$. Finally, the correlation length exponents ν, η cannot directly be computed from plain mean-field theory as they are tied to the spatial profile of fluctuating field configurations.

For the sake of later comparison, the mean-field critical exponents are summarized in Table 8.1. At first sight the differences between the experimentally observed exponents (second column) and the mean-field exponents (third column) do not look too dramatic – apparently the primitive mean-field approach pursued here fares reasonably favorably – which, in view of the accumulation of pronounced fluctuations at the critical point, should come as something of a surprise. On the other hand we must keep in mind that the exponents describe singular power laws in the transition region. In view of that, the difference between

1.3 and 1 does look quite significant. At any rate, we should try to refine our theoretical understanding of the transition and search for the source of the discrepancy with experiment.

8.4.3 Gaussian model

As a first improvement on the mean-field approximation, let us explore the effect of quadratic fluctuations around the constant field configuration $\bar{\phi}$. Approaching the transition point from above, we set $\bar{\phi} = 0$ and approximate the action through its quadratic expansion[29]

$$S[\phi] \approx \int d^d r \left[\frac{r}{2}\phi^2 + \frac{1}{2}(\nabla\phi)^2 - h\phi \right]. \tag{8.26}$$

In this form, one may effect the Gaussian integral over field fluctuations and evaluate the dependence of the free energy on the external parameters h and r. However, in anticipation of our analysis of the full problem below, we here pursue a slighly different, renormalization-group-oriented approach, i.e., pretending that we did not know how to do the Gaussian integral, we subject the quadratic action to a momentum shell RG analysis.

Proceeding along the lines of the canonical scheme, we split our field into fast and slow degrees of freedom $\phi = \phi_{\rm s} + \phi_{\rm f}$ resulting in the, now familiar, fragmentation of the action $S[\phi_{\rm s}, \phi_{\rm f}] = S_{\rm s}[\phi_{\rm s}] + S_{\rm f}[\phi_{\rm f}] + S_{\rm c}[\phi_{\rm s}, \phi_{\rm f}]$. However, the crucial simplification, characteristic of a Gaussian theory, is that the action $S_{\rm c}$ coupling fast and slow components vanishes (exercise), implying that the integration over the fast field merely leads to an inessential constant. The effect of the RG step on the action is then entirely contained in the rescaling of the slow action. According to our previous discussion, the scaling factors thus appearing are determined by the engineering dimensions of the operators appearing in the action, i.e. $r \to b^2 r$ and $h \to b^{d/2+1}h$. Using the fact that $r \sim t$ we can then readily write down the two relevant scaling dimensions of the problem, $y_t = 2$ and $y_h = d/2 + 1$. Comparison with Eq. (8.22) finally leads to the list of exponents,

$$\alpha = 2 - \frac{d}{2}, \quad \beta = \frac{d}{4} - \frac{1}{2}, \quad \gamma = 1, \quad \delta = \frac{d+2}{d-2}, \quad \nu = \frac{1}{2}, \quad \eta = 0.$$

Notice that the exponents now explicitly depend on the dimensionality of the system, a natural consequence of the fact that they describe the effect of spatial fluctuations. Table 8.1 contains the values of the exponents for a three-dimensional system. We cannot really say that the results are any better than those obtained by the mean-field analysis. Some exponents (e.g. δ) agree better with the experimental data, while others (e.g. α) are decidedly worse.

As a corollary to this section, we note that the Gaussian model possesses only one fixed point, namely $r = h = 0$, which in the context of ϕ^4-theory is called the **Gaussian fixed point**.

[29] The appearance of a linear term indicates that we are expanding not around the "true" mean-field, i.e. the exact solution of (8.24), but rather around the solution $\bar{\phi} = 0$ of the field-free system. However, in view of the fact that h has the status of an external perturbation, this choice of the reference configuration is quite natural.

8.4.4 Renormalization group analysis

In the present analysis of the model, we have not really touched upon its principal source of complexity, namely the effect of the "interaction operator" ϕ^4 on the fluctuation behavior of the field. It seems likely that the neglect of this term is responsible for the comparatively poor predictive power both of the straightforward mean-field analysis, and of the Gaussian model. Indeed, the dimensional analysis of Section 8.4.1 indicated that the ϕ^4 addition to the action becomes relevant below four dimensions. A more physical argument, to the same effect, is given in the Info block below.

Although the solution of the general problem posed by the action (8.23) still appears to be hopelessly difficult, there is one aspect we can turn to our advantage. While physical systems exist in integer dimensions $d = 1, \ldots, 4, \ldots$, there is actually no reason why we should not be allowed to evaluate our theory, i.e. the functional integral with action (8.23), in fractional dimensions. In the present context, this seemingly academic freedom turns out to be of concrete practical relevance. The point is that the nonlinear ϕ^4 operator was found to be marginal at $d = 4$ and relevant below. One may thus expect that, in dimensions $d = 4 - \epsilon$, $\epsilon \ll 1$, the operator is relevant but not that relevant, i.e. one may expect that, for sufficiently small deviations off the threshold dimension, four, the theory knows of an expansion parameter, somehow related to ϵ, which will enable us to control the interaction operator. Of course, at the end of the day, we will have to "analytically continue" to dimensions of interest, $\epsilon = 1$ or even $\epsilon = 2$, but, for the present, we will see what we can learn from a $d = 4 - \epsilon$ representation of the theory.

INFO Our previous analysis relied on the assumption that the field integration is tightly bound to the vicinity of the extrema of the action. But let us now ask under what conditions this assumption is actually justified. We should develop some intuition as to the relative importance of the mean-field content of the theory and of the **fluctuations around the mean-field**. While there are several ways to proceed with this program, we will focus on the analysis of the magnetic susceptibility. (At this point, we should warn the reader that the arguments formulated below, while technically straightforward, are conceptually involved. The critical contemplation of the logical steps of the construction is time well invested.) Firstly, let us recall the definition of the susceptibility,

$$\chi = -\partial_H^2 F \sim \int d^d r \, \langle \phi(\mathbf{r})\phi(0)\rangle_c \sim G(\mathbf{k} = 0),$$

where we have used the fact that $\langle \phi(\mathbf{r})\phi(\mathbf{r}')\rangle = G(\mathbf{r}-\mathbf{r}')$ is the Green function of the model. Given this identification, we note that a formal criterion of the transition – divergent susceptibility! – is synonymous with a singularity of the zero-momentum Green function.

On the level of the Gaussian theory (see Eq. (8.26)) $G(\mathbf{k}) = (r + k^2)^{-1}$, i.e. $\chi \sim r^{-1}$. Anticipating troubling observations to come, we reiterate that the mean-field transition temperature is identified by the condition $r \sim t = 0$. Now, let us move on to explore corrections to the mean-field susceptibility on the level of a perturbative one-loop calculation. To this end, we recall that (if necessary, recapitulate the discussion of Section 5.1), due to the presence of the ϕ^4 operator, the Green function acquires a self-energy which, at the one-loop level, is given by $\Sigma = -\frac{1}{2} \sum_{\mathbf{k}'} \frac{1}{r+k'^2}$.

As a consequence, one can identify the susceptibility as

$$\chi^{-1} \sim (G(\mathbf{k}=0))^{-1} = r - \Sigma = r + \frac{\lambda}{2} \sum_{\mathbf{k}'} \frac{1}{r + k'^2}.$$

A first observation to be made is that non-Gaussian fluctuations (physically: interactions between harmonic fluctuations around the mean-field amplitude) lower the transition amplitude, i.e., setting $r \sim T - T_c$, it now takes a smaller temperature T to reach the critical point; in accord with the intuitive expectation that fluctuations tend to "disorder" the system. Frustratingly, one may also observe that the cutoff Λ is needed to prevent the "correction",

$$-\frac{\lambda}{2} \sum_{\mathbf{k}'} \frac{1}{r + k'^2} \sim \lambda \int^{\Lambda} d^d k' \frac{1}{r + k'^2},$$

from diverging in dimensions $d \geq 2$. To deal with this singularity, we have to realize that the effect of fluctuations is actually two-fold: the transition temperature gets shifted and the temperature dependence of the inverse susceptibility is apparently no longer simply linear (by virtue of the r-dependence of the integrand). The two effects can be disentangled by writing

$$\chi^{-1} = r + \frac{\lambda}{2} \left(\frac{L}{2\pi} \right)^d \int^{\Lambda} d^d k' \left(\frac{1}{r + k'^2} - \frac{1}{k'^2} \right) \approx r - \frac{\lambda r}{2} \left(\frac{L}{2\pi} \right)^d \int^{\Lambda} \frac{d^d k'}{(r + k'^2) k'^2}, \tag{8.27}$$

where

$$r \equiv r + \frac{\lambda}{2} \left(\frac{L}{2\pi} \right)^d \int^{\Lambda} \frac{d^d k'}{k'^2},$$

represents the shifted transition temperature while the integral describing the deviation from the linear temperature dependence of the susceptibility is now UV-convergent in dimensions $d < 4$. Notice that in the second equality of Eq. (8.27) we have replaced the parameter r in the integrand by the modified parameter r. To the accuracy of a one-loop calculation, this manipulation is permissible.

Naively, it looks as if this sequence of manipulations has led to a catastrophe: the fluctuation-renormalized transition temperature appears to diverge as one sends the cutoff to infinity, clearly a nonsensical prediction! However, one may note that there was actually no justification for identifying the *physical* transition temperature through the parameter r in the first place. This identification was based on mean-field theory alone, i.e. an approach to the problem which neglected altogether the key effect of fluctuations. However, the bare parameter r appearing in the action carries as little "universal" meaning as the cutoff Λ, or any other microscopic system parameter for that matter!

Once we have acknowledged this interpretation, we should then identify the transition temperature through the singularity of the macroscopically observable properties (e.g. divergence of the susceptibility leads to the vanishing of the modified parameter r at the one-loop level) while the microscopic parameters carry no significance by themselves.

EXERCISE This interpretation closely parallels the philosophy of **renormalization in high-energy physics**. There, the bare parameters of the action are fundamentally undetermined, while the inverse of the Green function at zero external momentum represents a physical observable, e.g. the mass of the electron. Since the loop corrections to this physical quantity appear to be infinite (and the theory does not enjoy the luxury of the presence of a physically motivated cutoff), one postulates that the bare parameters of the action have been infinite by themselves. These singularities are deliberately adjusted so as to cancel the divergence of the fluctuation corrections and to produce finite "physical" quantities. It is instructive to consult a textbook on renormalization in high-energy physics (such as Ryder[30]) to become acquainted with the functioning of this strategy, and with the enormous success it has had in the context of QED and other sub-branches of particle physics.

We now turn to the second effect of the fluctuation correction, namely the deviation from the linear temperature dependence, as described by the integral contribution to Eq. (8.27). On dimensional grounds, the integral depends on the parameter r as $\sim \lambda L^d r^{(d-4)/2}$. The (mean-field + quadratic fluctuations) approach to the problem breaks down when this contribution becomes more important than the leading-order contribution to the susceptibility, i.e. for dimensions $d \quad 4$. This observation is the essence of the so-called **Ginzburg criterion**. The criterion states that mean-field theory becomes inapplicable below the so-called **upper critical dimension** $d_c = 4$. While we have derived this statement for the particular case of the ϕ^4-model, it is clear that similar estimates can be performed for every nonlinear field theory, i.e. as with the lower critical dimension, the upper critical dimension also represents an important threshold separating the mean-field dominated $d > d_c$ from the fluctuation dominated $d < d_c$ behavior. Also notice that the analysis above conforms with our previous observation that the nonlinear ϕ^4 operator is relevant in dimensions $d < 4$. (Convince yourself that the two lines of argument reflect the same principle, namely the dependence of fluctuations on the accessible phase volume, as determined by the dimensionality of the system.)

Before proceeding to the details of the RG program, let us try to predict a number of general elements of the ϕ^4 phase diagram on dimensional grounds. We saw that in dimensions $d > 4$ the ϕ^4 operator is irrelevant and that the Gaussian model essentially dictates the behavior of the system. Specifically, for $d > 4$, the Gaussian fixed point $r = \lambda = h = 0$ is the only fixed point of the system. Below four dimensions, the ϕ^4 operator becomes relevant and the emergence of a richer fixed point structure may be expected. However, for $\epsilon = 4 - d$ sufficiently small, we also expect that, whatever new fixed points appear, they should be close to the Gaussian point. This means that we can conduct our search for new fixed points within a double expansion in ϵ, and the small deviation of the coupling constants r, λ, h around the Gaussian fixed point. (In fact, we will momentarily identify a third expansion parameter, namely the number of momentum loops appearing in fast-field integration.)

Step I

We next proceed to formulate the steps of the RG in detail. To keep things simple, the RG transformation will be carried out to lowest order in a triple expansion in ϵ, the coupling constants, and the number of momentum loops. The rationale behind the loop expansion can be best understood if we assume that the entire action[31] is multiplied by a large parameter (which, in the case of a quantum theory, might be $\hbar^{-1}$). The expansion in the number of loops is then equivalent to an expansion in the inverse of that parameter (for a quantum theory, an expansion away from the classical limit).

[30] L. H. Ryder, *Quantum Field Theory*, (Cambridge University Press, 1996).
[31] Before we rescaled the fields so as to make the leading-order coefficient equal to 1/2.

EXERCISE Verify this statement; to this end, notice that a diagram of nth order in perturbation theory in the ϕ^4 vertex contains a prefactor a^n involving the large parameter. On the other hand, each of the I internal lines, or propagators, contained by the diagrams contributes a factor a^{-1}, so that the overall power is a^{n-I}. Next relate the number of internal lines to the number L of loops. Notice that each line corresponds to a momentum summation. However, the number of independent summations is constrained by the n δ-functions carried by the vertices. Use this information to show that the overall power of the graph is a^{-L+1}, i.e. an expansion in L is equivalent to an expansion in the inverse of a.

Let us now decompose the action in the standard manner, setting $S[\phi_{\rm s},\phi_{\rm f}] = S_{\rm f}[\phi_{\rm f}] + S_{\rm s}[\phi_{\rm s}] + S_{\rm c}[\phi_{\rm s},\phi_{\rm f}]$, where

$$\begin{aligned} S_{\rm f}[\phi_{\rm f}] &= \int d^d r \left[\frac{r}{2}\phi_{\rm f}^2 + \frac{1}{2}(\nabla\phi_{\rm f})^2\right], \\ S_{\rm s}[\phi_{\rm s}] &= \int d^d r \left[\frac{r}{2}\phi_{\rm s}^2 + \frac{1}{2}(\nabla\phi_{\rm s})^2 + \frac{\lambda}{4!}\phi_{\rm s}^4 - h\phi_{\rm s}\right], \\ S_{\rm c}[\phi_{\rm s},\phi_{\rm f}] &= \frac{\lambda}{4}\int d^d r \phi_{\rm s}^2\phi_{\rm f}^2 + \cdots \end{aligned}$$

Several approximations related to the loop order of the expansion are already imposed at this level. We have neglected terms of $\mathcal{O}(\phi_{\rm f}^4)$ because their contraction leads to two loop diagrams. The same applies to terms of $\mathcal{O}(\phi_{\rm s}\phi_{\rm f}^3)$ (exercise). Terms of $\mathcal{O}(\phi_{\rm s}^3\phi_{\rm f})$ do not arise because the addition of a fast momentum and three slow momenta is incompatible with momentum conservation.

Steps II and III

To simplify the notation, let us rescale the momentum according to $\mathbf{q} \to \mathbf{q}/\Lambda$, implying that coordinates are measured in units of the inverse cutoff $\mathbf{r} \to \mathbf{r}\Lambda$. With the coupling constants rescaled according to their engineering dimensions, $r \to r\Lambda^2$, $\lambda \to \lambda\Lambda^{4-d}$, the action remains unchanged, while the fast and slow momenta are now integrated over the dimensionless intervals $|\mathbf{q}_{\rm s}| \in [0, b^{-1}]$ and $|\mathbf{q}_{\rm f}| \in [b^{-1}, 1]$, respectively. We next construct an effective action by integration over the fast field: $e^{-S_{\rm eff}[\phi_{\rm s}]} = e^{-S_{\rm s}[\phi_{\rm s}]}\left\langle e^{-S_{\rm c}[\phi_{\rm s},\phi_{\rm f}]}\right\rangle_{\rm f}$. In performing the average over fast fluctuations, $\langle\cdots\rangle_{\rm f}$, we shall (a) retain only contributions of one-loop order while (b) neglecting terms that lead to the appearance of $\phi_{\rm s}^{n>4}$ contributions in the action. (For example, the contraction $\langle\left(\int \phi_{\rm s}^2\phi_{\rm f}^2\right)^3\rangle$ would lead to such a term.) To this level of approximation, one obtains

$$e^{-S_{\rm eff}[\phi_{\rm s}]} = e^{-S_{\rm s}[\phi_{\rm s}]}\exp\left[-\langle S_{\rm c}[\phi_{\rm s},\phi_{\rm f}]\rangle_{\rm f} + \frac{1}{2}\left\langle S_{\rm c}[\phi_{\rm s},\phi_{\rm f}]^2\right\rangle_{\rm f}^{\rm c}\right],$$

where the superscript c denotes a connected average. (Exercise: It is instructive to check the consistency of this expansion for yourself.) The two diagrams corresponding to the contractions $\langle S_{\rm c}[\phi_{\rm s},\phi_{\rm f}]\rangle_{\rm f}$ and $\langle S_{\rm c}[\phi_{\rm s},\phi_{\rm f}]^2\rangle_{\rm f}^{\rm c}$ are shown in parts (a) and (b) of the figure below, respectively, where the external line segments indicate the passive $\phi_{\rm s}$ amplitudes.

According to the standard rules of perturbation theory, the first of the two diagrams, (a), evaluates to

$$\langle S_{\rm c}[\phi_{\rm s},\phi_{\rm f}]\rangle_{\rm f} = \frac{\lambda}{4}\int_{\rm f}\frac{d^dq'}{(2\pi)^d}\frac{1}{r+q'^2}\int_{\rm s}\frac{d^dq}{(2\pi)^d}\phi_{\rm s}(\mathbf{q})\phi_{\rm s}(-\mathbf{q}).$$

We now consider the summation over fast momenta appearing in this expression. Using the fact that we are in the near vicinity of the critical point and anticipating that we are interested in no more than the expansion of the β-function for small values of the coupling, we now expand the integrand to first order in r, $\int_{\rm f}\frac{d^dq}{(2\pi)^d}\frac{1}{r+q^2} = I_1 - rI_2$, where we have introduced the shorthand notation,

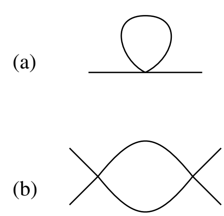
$$I_\alpha \equiv \int_{\rm f}\frac{d^dq}{(2\pi)^d}\frac{1}{q^{2\alpha}}. \tag{8.28}$$

These integrals are straightforwardly computed by switching to polar coordinates,

$$I_\alpha = \Omega_d\int_{b^{-1}}^1 dq\, q^{d-2\alpha-1} = \frac{\Omega_d}{d-2\alpha}(1-b^{2\alpha-d}),$$

where $\Omega_d = (2\pi^{d/2}/\Gamma(d/2))/(2\pi)^d$ denotes the volume of the d-dimensional unit sphere (measured in units of 2π). We thus find that, after the integration over fast modes, and the standard rescaling operation, $\mathbf{q}\to b\mathbf{q}$, $\phi\to b^{(d-2)/2}\phi$, the quadratic part of the action takes the form

$$S^{(2)}[\phi] = \frac{b^2}{2}\left[r + \frac{\lambda\Omega_d}{2(d-2)}(1-b^{2-d}) - \frac{r\lambda\Omega_d}{2(d-4)}(1-b^{4-d})\right]\int d^dr\,\phi^2. \tag{8.29}$$



Turning to the second diagram (b) in the figure, we notice that, owing to the presence of four external legs, its contribution will be proportional to $\phi_{\rm s}^4$. Further, momentum conservation implies that the momenta carried by the internal lines of the diagram will depend on both the fast "internal" momentum and the external momenta carried by the fields $\phi_{\rm s}$. However, we can simplify the analysis by neglecting the dependence on the latter from the outset. The reason is that the integration over the internal momentum followed by Taylor expansion in the slow momenta would generate expressions of the structure $F(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3)\phi(\mathbf{q}_1)\phi(\mathbf{q}_2)\phi(\mathbf{q}_3)\phi(-\mathbf{q}_1-\mathbf{q}_2-\mathbf{q}_3)$, where $\mathbf{q}_{1,2,3}$ represent slow momenta and F is some polynomial. Taking account of the small momenta would thus generate derivatives acting on an operator of fourth order in ϕ, a combination that we saw above is irrelevant.

Neglecting the external momenta, diagram (b) leads to the result

$$\frac{1}{2}\langle S_{\rm c}[\phi_{\rm s},\phi_{\rm f}]^2\rangle_{\rm f} \simeq \frac{\lambda^2}{16}\int d^dr\,\phi_{\rm s}^4\int_{\rm f}\frac{d^dq}{(2\pi)^d}\frac{1}{(r+q^2)^2} = \frac{\lambda^2 I_2}{16}\int d^dr\,\phi_{\rm s}^4 + \mathcal{O}(\lambda^2 r).$$

Evaluating the integral and rescaling, we find that the quartic contribution to the renormalized action reads

$$S^{(4)}[\phi] = b^{4-d}\left(\frac{\lambda}{4!} - \frac{\lambda^2\Omega_d}{16}\frac{1-b^{4-d}}{d-4}\right)\int d^dr\,\phi^4.$$

Finally, there are no one-loop diagrams affecting the linear part of the action, i.e.

$$S^{(1)}[\phi] = hb^{d/2+1} \int d^d r\, \phi,$$

rescales according to its engineering dimension.

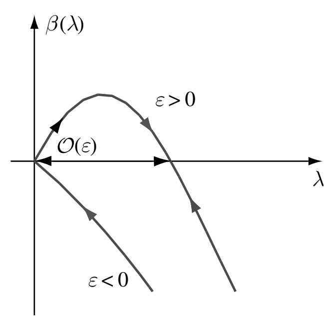
Combining everything, we find that, to one-loop order, the coupling constants scale according to the relations $r \to b^2(r + \frac{\lambda\Omega_d}{2(d-2)}(1 - b^{2-d}) - \frac{r\lambda\Omega_d}{2(d-4)}(1 - b^{4-d}))$, $\lambda \to b^{4-d}(\lambda - \frac{3}{2}\lambda^2\Omega_d \frac{1-b^{4-d}}{d-4})$, and $h \to hb^{d/2+1}$. We next set $d = 4 - \epsilon$ and evaluate the right-hand sides of these expressions to leading order in ϵ. With $\Omega_{4-\epsilon} \approx \Omega_4 = \frac{1}{8\pi^2}$, we thus obtain

$$\begin{aligned} r &\to b^2 \left(r + \frac{\lambda}{32\pi^2}(1 - b^{-2}) - \frac{r\lambda}{16\pi^2} \ln b \right), \\ \lambda &\to (1 + \epsilon \ln b) \left(\lambda - \frac{3\lambda^2}{16\pi^2} \ln b \right), \\ h &\to hb^{3-\epsilon/2}, \end{aligned}$$

which, setting $b = e^\ell$, lead to the **Gell-Mann–Low equations**:

$$\boxed{\begin{aligned} \frac{dr}{d\ell} &= 2r + \frac{\lambda}{16\pi^2} - \frac{r\lambda}{16\pi^2}, \\ \frac{d\lambda}{d\ell} &= \epsilon\lambda - \frac{3\lambda^2}{16\pi^2}, \\ \frac{dh}{d\ell} &= \frac{6-\epsilon}{2} h. \end{aligned}} \tag{8.30}$$

These equations clearly illustrate the meaning of the ϵ-expansion. According to the second equation, a perturbation away from the Gaussian fixed point will initially grow at a rate set by the engineering dimension ϵ. While, on the level of the classical, zero-loop theory, λ would grow indefinitely, the one-loop contribution $\sim \lambda^2$ stops the flow at a value $\lambda \sim \epsilon$.



Equating the right-hand sides of Eq. (8.30) to zero (and temporarily ignoring the magnetic field), we indeed find that besides the Gaussian fixed point $(r_1^*, \lambda_1^*) = (0, 0)$ a **non-trivial fixed point** $(r_2^*, \lambda_2^*) = (-\frac{1}{6}\epsilon, \frac{16\pi^2}{3}\epsilon)$ has appeared. Notice that, in accord with the schematic considerations made at the beginning of the section, the second fixed point is $\mathcal{O}(\epsilon)$ and coalesces with the Gaussian fixed point as ϵ is sent to zero. Plotting the β-function for the coupling constant λ (see figure), we further find that, for $\epsilon > 0$, λ is relevant around the Gaussian fixed point but irrelevant at the non-trivial fixed point.

To understand the full flow diagram of the system, one may linearize the β-function around both the Gaussian and the non-trivial fixed point. Denoting the linearized mappings by $W_{1,2}$, we find

$$W_1 = \begin{pmatrix} 2 & \frac{1}{16\pi^2} \\ 0 & \epsilon \end{pmatrix}, \quad W_2 = \begin{pmatrix} 2 - \frac{1}{3}\epsilon & \frac{1+\epsilon/6}{16\pi^2} \\ 0 & -\epsilon \end{pmatrix}.$$

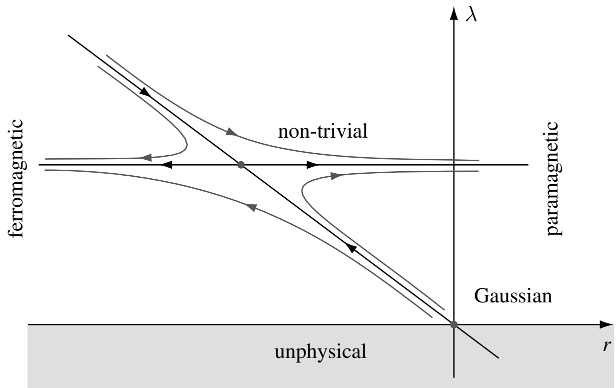


Figure 8.8 Phase diagram of the ϕ^4-model as obtained from the ϵ-expansion.

Figure 8.8 shows the flow in the vicinity of the two fixed points, as described by the matrices $W_{1,2}$ as well as the extrapolation to a global flow chart. Notice that the critical surface of the system – the straight line interpolating between the two fixed points – is tilted with respect to the $r \sim$ temperature axis of the phase diagram. This implies that it is not the physical temperature alone that decides whether the system will eventually wind up in the paramagnetic ($r \gg 0$) or ferromagnetic ($r \ll 0$) sector of the phase diagram. Rather one has to relate temperature ($\sim r$) to the strength of the nonlinearity ($\sim \lambda$) to decide on which side of the critical surface we are. For example, for strong enough λ, even a system with r initially negative may eventually flow towards the disordered phase. This type of behavior cannot be predicted from the mean-field analysis of the model (which would generally predict a ferromagnetic state for $r < 0$). Rather it represents a non-trivial effect of fluctuations. Finally notice that, while we can formally extend the flow into the lower portion of the diagram, $\lambda < 0$, this region is actually unphysical. The reason is that, for $\lambda < 0$, the action is fundamentally unstable and, in the absence of a sixth-order contribution, does not describe a physical system.

What are the critical exponents associated with the one-loop approximation? Of the two eigenvalues of W_2, $2 - \epsilon/3$ and $-\epsilon$, only the former is relevant. As with the Gaussian fixed point, it is tied to the scaling of the coupling constant, $r \sim t$, i.e. we have $y_t = 2 - \epsilon/3$ and, as before, $y_h = (d+2)/2 = (6-\epsilon)/2$. An expansion of the exponents summarized in Eq. (8.22) to first order in ϵ then yields the list

$$\alpha = \frac{\epsilon}{6}, \quad \beta = \frac{1}{2} - \frac{\epsilon}{6}, \quad \gamma = 1 + \frac{\epsilon}{6}, \quad \delta = 3 + \epsilon, \quad \nu = \frac{1}{2} + \frac{\epsilon}{12}, \quad \eta = 0.$$

If we are now reckless enough to extend the radius of the expansion to $\epsilon = 1$, i.e. $d = 3$, we obtain the fifth column of Table 8.1. Apparently the agreement with the experimental results has improved – even in spite of the fact that we have driven the ϵ-expansion well beyond its range of applicability! (For $\epsilon = 1$, terms of $\mathcal{O}(\epsilon^2)$ can, of course, no longer be neglected!)

How can one rationalize the **success of the ϵ-expansion**? Trusting in the principle that good theories tend to work well beyond their regime of applicability, we might simply speculate that nature seems to be sympathetic to the concept of renormalization and the loop expansion. Of course, a more qualified approach to the question is to explore what happens at higher order in the ϵ-expansion. Needless to say, the price to be paid for this ambition is that, at orders $\mathcal{O}(\epsilon^{n>1})$, the analysis indeed becomes laborious. Nonetheless, the success of the first-order expansion prompted researchers to drive the ϵ-expansion up to fifth order! The results of this analysis are summarized in the last column of Table 8.1. In view of the fact that we are still extending a series beyond its radius of convergence,[32] the level of agreement with the experimental data is striking. In fact, the exponents obtained by the ϵ-expansion even agree – to an accuracy better than one percent – with the exponents of the *two*-dimensional model,[33] i.e. for a situation where the "small" parameter ϵ has to be set to two.

However, it is important to stress that the ϵ-expansion is not just a computational tool for the calculation of exponents. On a more conceptual level, its merit is that it enables one to explore the phase diagram of nonlinear theories in a more or less controlled manner. In fact, the ϵ-expansion not only is useful in the study of field theories close to the upper critical dimension (i.e. close to the mean-field threshold) but can equally well be applied to the analysis of systems in the vicinity of the lower critical dimension. In the following section, we consider a problem of this type, i.e. we will apply an ϵ-expansion around $d = 2$ to detect the onset of global thermal disorder in models with continuous symmetries.

8.5 RG analysis of the nonlinear σ-model

The scalar field theory encapsulates a wide class of systems encompassing a single-component order parameter. However, throughout the text, we have encountered problems where the order parameter involves more than one component, e.g. the complex field associated with condensation phenomena, the matrix field associated with the quantum disordered metallic system, or the field theories involving spin. In such cases, one very often finds that the low-energy content of the theory involves a projection which imparts a constraint to the field integral. In the context of condensation phenomena, we saw that, at low temperatures, one can neglect the massive amplitude fluctuations of the order parameter, while the collective fluctuations of the phase mode impacted significantly on the low-energy properties of the system. In this case, the phase degree of freedom is constrained by its topology to lie on the unit circle. Similarly, if we neglect the "high-energy" physics of local moment formation, classical and quantum spin theories are constrained by the normalization of the local spin. When subjected to an auxiliary constraint, theories that are otherwise free are known as nonlinear σ-models. The aim of the present section is to apply methods of the RG to explore the critical properties of a general class of nonlinear

[32] Indeed, it is believed that we are dealing with a series that is only asymptotically convergent. That is, beyond a certain order of the expansion, the agreement with the "true" exponents will presumably become worse.

[33] The latter are known from the exact solution of the two-dimensional model, see L. Omsager, Crystal statistics I. A two-dimensional model with an order–disorder transition. *Phys. Rev.* **65** (1944), 117–49.