IISc-CTS-?/93

RENORMALISATION GROUP[†]

Apoorva Patel[*]

CTS and SERC, Indian Institute of Science
Bangalore-560012, India

ABSTRACT

These lectures give an introduction to the subject of the Renormalisation Group, and some examples of its diverse applications.

1. Problems with Multiple Scales and Renormalisation Group

In the study of natural phenomena, one encounters a great diversity of length scales. In contrast, most theoretical methods work only when there are a few, quite often only one, degrees of freedom. Thus to make the problems solvable, one has to condense information contained in many degrees of freedom at many length scales to only a handful of variables and parameters. It is customary to label these as effective degrees of freedom and effective parameters.

Consider an illustrative example. The motion of a fluid on the macroscopic scale, for instance waves with wavelengths of the order of metres, is described by hydrodynamics. On a much smaller atomic scale, one sees the atomic structure of the fluid, and the physics at this scale must be described by the Schrödinger equation for the electrons making up the atoms. On a still smaller scale, one has to deal with Quantum Chromodynamics to understand how the nuclei of the atoms are made. On the other hand, while studying hydrodynamics, one does not deal with the 10^{23} or so degrees of atomic degrees of freedom. All the information regarding what is going on at the atomic scale is condensed into parameters such as viscosity and temperature, and degrees of freedom such as density, velocity and vorticity. Similarly, while studying atomic physics, one does not worry about how the nuclei are made up from quarks and gluons; its sufficient to parametrise the nuclei in terms of their charge, mass and size.

[†] Lectures at the 9th SERC school on High Energy Physics, University of Gauhati, Guwahati, November 15-27, 1993.

[*] E-mail: adpatel@cts.iisc.ernet.in

This example describes a typical situation where the different scales of the problem are essentially separable. The disparity of scales is so great that the communication between them is reduced to a few effective parameters. These effective parameters contain practically all the information about what influence the smaller scales will have on larger scales, and having identified them, the phenomena associated with each scale can be treated independently, without ever having to worry about what is going on at still smaller scales. In a sense, the effective degrees of freedom and the effective parameters represent the content of the theory after averaging over all fluctuations at scales smaller than the one of interest. The simplification is enormous. In fact the success of almost all theories in physics depends on isolating some limited range of scales, variables and parameters.

Let us look at the range of validity of such an effective description. The size of a liquid can be varied over a wide range without qualitatively changing its properties. A liquid of only 1000 atoms, say, has almost the same energy per unit volume and density as the liquid with 10^{23} atoms. The minimum size one can reach without substantially altering the properties of the system is called the correlation length ξ. Under favourable circumstances, it is only about a few atomic spacings for a liquid. This is the underlying reason for the enormous simplification in the problem. It is only the degrees of freedom within a correlation length that are of importance while studying a problem; the intensive or extensive properties of a macroscopic system can be reconstructed given only a microscopic sample of it of size ξ. (Of course, the above example contains still too many degrees of freedom even within a correlation length, but they can be dealt with using a variety of approximations.)

There exist, however, a class of phenomena where the happenings at many scales of length make contributions of equal importance. For example, a liquid near its critical point develops fluctuations in density at all possible scales, in the form of drops of liquid thoroughly interspersed with bubbles of gas. Both drops and bubbles come in all sizes—from single molecules upto the volume of the specimen. In such a case, the correlation length is large near the critical point; it is infinite precisely at the critical point. The number of degrees of freedom within a correlation length are far too many, and any theory that describes such a phenomenon must take into account the entire spectrum of length scales.

It is often the case that such phenomena display self-similarity or scale invariance. The system appears more or less the same no matter what scale it is viewed at (as long as the scale is smaller than ξ). Obviously, a self-similar system would have to be described by the same variables at every scale. Only the parameters of the theory can change from scale to scale. The "renormalisation group" provides the conceptual framework for dealing with such systems.

It is not a descriptive theory of nature, but a general method for constructing theories. The strategy is to tackle the problem in steps, one step for each length scale. At each step in sequence, the renormalisation group approach integrates out the fluctuations at that scale, say starting with an atomic scale and moving on to successively larger scales, until fluctuations on all scales have been integrated out. Such steps make up a group of transformations that take one from one set of input parameters at one scale to another equivalent set of parameters at a larger scale. Strictly speaking, it is a semi-group: One can define a group composition rule for combining different steps, and there is an identity element, but there is no inverse transformation (there is no unique way to construct what may be going on at a smaller scale from observations only at a larger scale). We generically denote a renormalisation group step which changes the scale by a factor s by the symbol R_s.

When fluctuations at small scales have been averaged out, one can talk about taking a continuum limit. This is possible because the basic equations of all physics are local. The classical continuum limit is the limit which enables one to define derivatives. It is easy to use in problems which have a definite scale. There is relatively little that is going on at much shorter scales, and continuum forms can be introduced for various variables. For example, in a hydrogen atom, the characteristic length scale is provided by the Bohr radius. It is determined by simple dimensional analysis, and any length of importance in the hydrogen atom is proportional to it. The continuum limit is taken holding this physical scale fixed, and letting any shorter scale go to zero. The classical continuum limit is also used in formally defining functional integrals, with independent field variables at every point in a continuous space-time. The field theory problems with infinitely many variables, however, can be more difficult to solve.

The scale invariant systems define a different type of continuum limit, which may be called a statistical continuum limit. There may be a basic underlying structure, say an atomic lattice with a fixed and finite number of independent variables at each lattice site. The continuum limit arises when the correlation length becomes macroscopic—much larger than the lattice spacing. It necessarily involves the thermodynamic limit, i.e. the limit of infinite volume. At the scale of the correlation length the lattice structure is almost invisible, making it plausible to get rid of any reference to the lattice altogether. The number of variables within a correlation length is large, and the continuum limit is taken by letting the correlation length diverge in units of the lattice spacing. Such a limit is much harder to achieve while maintaining locality of the description. The reason for that is the absence of any characteristic length scale in the system; dimensional analysis has no meaning.

It is easy to talk about critical phenomena in position space, just as we observe them. On the other hand, field theories are much easier to handle in momentum space. Technically the two are related just by Fourier transform, and one can easily jump from one view-point to the other as the need may be. In quantum field theories the lack of energy scale arises when one has to sum over intermediate states containing arbitrarily large energies. In many cases, such sums reduce to the logarithmically divergent integral $\int_Q^\infty d^4k/k^4$. To make sense out of the theory, a regularisation of the theory becomes necessary. Such a regularisation is nothing but a prescription for cutting off the divergent integrals, by somehow altering the behaviour of the integrand when $|k|$ is much larger than any other physical scales of interest. In such a case, the statistical continuum limit consists of letting the regulator go to infinity while holding the physical quantities fixed, and the renormalisation group embodies the change in parametrisation of the theory at the regulator scale as the regulator is changed.

In critical phenomena, the dominant fluctuations are neither the fluctuations of size ξ nor the fluctuations of the atomic size, rather they are the ones at all scales between these two limits. A logarithmic divergence is a typical symptom of a problem lacking a characteristic scale. The various energy intervals $[Q, 2Q]$, $[2Q, 4Q]$, ... contribute equal and finite amount to the integral. The divergence arises because of the fact there are infinitely many such intervals. The basic physical idea underlying the renormalisation group approach is that even though there are many length scales in the problem, they are only locally coupled to each other. The result of this hypothesis is that there is a cascade effect in the whole system: the atomic fluctuations (1-2 Å) influence the 2-4 Å fluctuations, which in turn influence the 4-8 Å fluctuations, and so on. Such a cascade picture naturally gives rise to scale invariance for fluctuations of intermediate scales. The scaling fails for fluctuations with scales near a length parameter, e.g. it doesn't hold for fluctuations of atomic size or of order ξ.

A noteworthy feature of the cascade picture is the existence of amplification and deamplification as the cascade develops. For example, consider a small change in temperature of a magnet away from its critical temperature. It has little effect on the atomic scale fluctuations. But as the cascade develops, from 1 Å to 2 Å to 4 Å to 8 Å etc., the effect of the temperature change is amplified, leading to macroscopic changes at large scales. In particular, the temperature change leads to a change in the correlation length, which is a macroscopic change for fluctuations with scales larger than ξ.

Deamplification also occurs in the cascade and it underlies the concept of universality. For example, two different magnetic materials can have quite different atomic structure. But the effect of the different atomic structures usually decreases with each step, finally becoming

negligible at macroscopic scales. This turns out to be so because when many degrees of freedom are behaving cooperatively, the behaviour of the whole system is governed primarily by the nature of the degrees of freedom themselves and the fact that there is a cooperative behaviour amongst them. The detailed nature of the action or Hamiltonian, which define the system, only plays a secondary role. The coupling between different scales, from the atomic to the macroscopic, essentially depends on the dimensionality of the system and the nature of scaling variables, and governs the scaling behaviour.

The renormalisation group methodology of stepping from one scale to the next, in the statistical continuum limit, is an analogue of the derivative in the classical continuum limit. The scale dependence of the parameters of the theory can be expressed in terms of differential equations (the variables themselves remain the same at all scales). In practical problems, such equations have to be solved with specific boundary conditions. The renormalisation group thus expresses the symmetry with respect to the change of parametrisation of the solutions to these equations. It is not a symmetry of the equations, but rather the symmetry of solutions themselves which are functions of essential physical variables and appropriate boundary conditions. The concept of universality boils down to the fact that the choice of regulator does not really matter for scale invariant systems, just the same way that there can be many different finite difference approximations to a derivative in the classical continuum limit.

One important limitation of the renormalisation group must be kept in mind. It only exploits the scale invariance symmetry of the problem and deduces some characteristics of physical quantities. Many quantitative properties of physical systems actually depend on the specific values chosen as boundary conditions for the parameters. The renormalisation group does not have much to say about such properties; it can generally make statements only about the characteristics that are independent of the quantitative values of the parameters. This definitely leads to certain observable predictions, nonetheless it is not a substitute for a complete solution of the problem. On the other hand, the physical insight gained into the dynamics of the problem, although qualitative, can be far more informative than a tedious mathematical solution.

Historically, the process that the theory remains the same in going from one scale to another, but the parameters appearing in the theory change with the scale, was labeled renormalisation. It was discovered as a purely mathematical symmetry, without any physical insight, in field theory calculations. The phenomenon of scale invariance was observed in experiments with condensed matter systems. The two were amalgamated together by

Wilson. Since then the modern theory of the renormalisation group has benefited a lot by cross-fertilisation of different branches of physics, and gone on to describe a wide variety of phenomena. It today forms the basis for our of understanding of how scale invariance arises, in natural phenomena as well as in mathematical models, at a deep level.

Exercise 1: Consider a baloon filled with Helium (atomic weight 4), which is released in the atmosphere (essentially Nitrogen having molecular weight 28). Estimate its acceleration at the instance of its release. Neglect the weight of the empty baloon. Note that the viscous forces are proportional to the velocity and vanish at the instance of release. Remember that one cannot use gravity to create a force stronger than gravity itself. This is an exercise to bring out the concept of bare vs. effective parameters and the role of the background.

2. Phase transitions, Critical Phenomena and Critical Exponents

All materials undergo phase change when the external parameters, such as pressure and temperature, applied to them are varied. A solid may melt, a liquid may vaporise, a magnet may lose its magnetisation, a crystal structure may change, and so forth. It is observed that these changes take place rather abruptly (as opposed to gradually) at specific values of the external parameters. The systems undergo a change in symmetry at these values of the external parameters, and boundaries appear between the distinct phases. Moreover, various thermodynamic quantities display a singular behaviour in the vicinity of these values. Such phenomena are therefore called phase transitions or critical phenomena. A particular class of these phase transitions, called continuous or second order phase transitions, are characterised by a divergent correlation length at the transition point. (Conventionally a phase transition is said to be of n^{th} order, if the n^{th} derivative of the free energy is discontinuous across the transition point. The free energy itself is a continuous function of the external parameters in equilibrium statistical mechanics.) These are the problems amenable to treatment by the renormalisation group framework.

The simplest example to study, as well as to introduce notation, is that of a ferromagnet. In the neighbourhood of its critical temperature T_c, called the Curie temperature, it displays non-analytic behaviour. Various macroscopic observables either diverge or go to zero as one

approaches the critical point. In particular, it is observed that

$$\begin{aligned}
\text{Correlation length}: \quad \xi(t \to 0, h = 0) &\propto |t|^{-\nu} \\
\text{Specificheat}: \quad C_v(t \to 0, h = 0) &\propto |t|^{-\alpha} \\
\text{Magnetisation}: \quad M(t \to 0_-, h = 0) &\propto (-t)^{\beta} \\
\text{Susceptibility}: \quad \chi(t \to 0, h = 0) &\propto |t|^{-\gamma} \\
M(t = 0, h \to 0) &\propto h^{1/\delta} \\
\text{Correlation function}: \quad G^{(2)}(r \to \infty, t \neq 0) &\propto \exp(-r/\xi) \\
G^{(2)}(r \to \infty, t = h = 0) &\propto r^{-(d-2+\eta)}
\end{aligned} \tag{2.1}$$

Here the reduced temperature $t = (T - T_c)/T_c$, and h is the applied magnetic field. In the language of field theory, ξ is the inverse mass gap, the $2-$point connected correlation function $G^{(2)}$ is the usual propagator, and one approaches a massless theory in going towards the critical point. We adopt the convention of always absorbing the Boltzmann constant in the definition of T, and factoring out the volume from all extensive quantities so that C_v, M, χ etc. refer to quantities per unit volume.

ν, α, β, γ, δ and η are called the critical exponents. In principle, ν, α and γ can be assigned different values depending on whether $t \to 0_-$ or $t \to 0_+$. (The proportionality constants appearing in the corresponding equations are indeed different for the two cases $t \to 0_-$ or $t \to 0_+$.) In systems displaying scale invariance (which are the ones studied using renormalisation group), however, the two values turn out to be the same; we therefore do not make any distinction between the two. Also, in a particular problem, some exponent may turn out to be zero. In practical terms, it doesn't necessarily mean an absence of singularity. Instead it just means that the singularity is milder; the actual behaviour may be logarithmic or at times discontinuous.

Widom introduced the concept of scaling on empirical grounds. It amounts to writing the asymptotic singular parts of thermodynamic functions as homogeneous functions of their arguments. In the laguage of renormalisation group, in the vicinity of the critical point, ξ is the only relevant length as far as the singular behaviour of thermodynamic quantities is concerned. The singular behaviour of all other quantities can be expressed in terms of that of ξ. This simple notion is a kind of dimensional reduction in the function space. For example, the equation of state describes a relation between the magnetic field, the magnetisation and the temperature. Scaling implies that near the critical point h is not a function of two

independent variables M and t, but only of the single combination $tM^{-1/\beta}$. Similarly, using the definitions of critical exponents, we find that the free energy depends on the combination $th^{-1/\beta\delta}$, and the 2−point correlation function depends on the combination $tr^{1/\nu}$. Explicitly

$$\begin{aligned} \text{Equation of state}: \quad h(M,t) &= M^{\delta}\ f(M^{-1/\beta}t) \ , \\ \text{Free energy}: \quad F_{sing}(t,h) &= t^{2-\alpha}\ \phi(th^{-1/\beta\delta}) \ , \\ \text{Correlation function}: \quad G^{(2)}(r,t,h=0) &= |r|^{-(d-2+\eta)}\ g(r^{1/\nu}t) \ , \end{aligned} \tag{2.2}$$

where f, ϕ and g are some sufficiently general functions of a single argument.

Together with the definitions of thermodynamic quantities

$$C_v = -T\left(\frac{\partial^2 F}{\partial T^2}\right) \ , \quad M = -\left(\frac{\partial F}{\partial h}\right)_T \ , \quad \chi = -\left(\frac{\partial^2 F}{\partial h^2}\right)_T = TG^{(2)}(k=0) \ , \tag{2.3}$$

these homogeneous relations imply relations among the critical exponents:

$$\begin{aligned} \gamma &= \beta(\delta-1) \ , \\ 2-\alpha &= \beta(\delta+1) \ , \\ \gamma &= \nu(2-\eta) \ , \\ 2-\alpha &= \nu d \ . \end{aligned} \tag{2.4}$$

The last of these relations, involving the dimensionality of the system explicitly, is known as the hyperscaling relation. Merely the definitions of thermodynamic functions are not enough to derive it; one has to assume that the singular part of the free energy, which has dimension d, scales as ξ^{-d} as the critical volume diverges.

With four relations amongst themselves, only two out of the six exponents are independent. We thus have a two-exponent scaling theory. The example has a global symmetry, namely the direction of magnetisation can be freely rotated without affecting any physics. So the two independent exponents can be readily identified as even and odd, depending on whether the corresponding quantities remain invariant under the global symmetry or not. Obviously temperature is even and magnetisation is odd, so ν is an example of an even exponent while β is that of an odd one.

It is also possible to derive relations amongst various amplitudes (the omitted proportionality constants in Eq.(2.1)) on the basis of the scaling hypothesis. Again there are only two independent amplitudes, which can be readily identified as the length scales of temperature and magnetisation. The difference between exponents and amplitudes is that the two

independent exponents are universal, while the two remaining amplitudes are not, e.g. all isotropic ferromagnets display the same critical exponents but have a range of values for T_c. **Exercise 2:** Derive the relations between the critical exponents from the definitions of thermodynamic quantities and the scaling forms. This is only an exercise in dimensional analysis.

3. Landau's Mean Field Theory

Landau took the first steps towards understanding critical phenomena. Although his theory is not correct when applied to real physical systems, it has played an important historical role since it identified the relevant degrees of freedom and showed how thermodynamic functions can become singular for specific values of parameters.

To see the non-triviality of these features, consider the formal prescription of statistical mechanics. The partition function Z of a system is the sum of the Boltzmann factor $\exp(-H/T)$ over all possible configurations of the system, and the free energy is proportional to the negative logarithm of Z. The Boltzmann factor is an analytic function of T near T_c, and a sum of analytic functions is also analytic. The true non-analytic behaviour can therefore only occur in the thermodynamic limit, when the number of configurations become infinite.

We continue with the illustrative example of a ferromagnet. It is too difficult to deal with 10^{23} degrees of freedom contained in a macroscopic physical system. One has to pick only a subset of the degrees of freedom that contain the essential dynamics. Landau assumed that fluctuations are present in the system only at atomic scale (e.g. electrons wandering around the nucleus), and they can be somehow averaged out to define an effective spin variable ϕ at each lattice site. This effective spin is a continuous variable which fluctuates in response to external (may be position dependent) stimuli. There is no loss of generality in taking the effective spin to be a variable defined on a continuum instead of only at lattice sites—it is an averaged variable any way. The continuum language is more suitable for applying field theoretical methods to the problem.

Since the physical interactions are local and the Boltzmann factor analytic, the Hamiltonian for a specific configuration can be written down as

$$H[\phi]/T \;=\; \int d^dx\; [c(\nabla\phi(x))^2 + a_0 + a_2\phi^2(x) + a_4\phi^4(x) - h\cdot\phi(x)] \quad , \tag{3.1}$$

Figure 1: The effective potential of the Landau-Ginzburg model for $a_2\{<,=,>\}0$, or equivalently for $T\{<,=,>\}T_c$ respectively.

where the variables c, a_0, a_2, a_4 and h are analytic functions of external parameters. The gradient term is the leading contribution to the energy from misalignments of adjacent effective spins, and it is customary to choose the scale of ϕ so that $c = 1$. a_0 just shifts the overall energy scale and can be ignored for most practical purposes. The self-energy terms, involving a_2 and a_4, are constrained by the global symmetry that the energy cannot depend on the direction of ϕ in the absence of an external field; hence only even powers of ϕ occur. The final term reflects the interaction of the effective spin with a suitably normalised external magnetic field.

This Hamiltonian is said to be in the Landau-Ginzburg form. It is the simplest form one can write down based on the first few terms of the power series expansion of the energy. It is an analytic function of the effective spin as well as the external parameters. As we shall see, the loss of analyticity in Landau's theory arises when an infinite sum is carried out over all allowed configurations of ϕ.

ϕ is also called the order parameter. It is arrived at after averaging out all microscopic fluctuations in the system, and possesses all the features that are essential to the critical behaviour in a given problem. In particular these feature include: (a) the number of dimensions that ϕ lives in, (b) the number of components of ϕ, and (c) the pattern of global symmetry breaking at the critical point. In fact, the first step in studying any critical phenomenon is to identify the appropriate order parameter.

Landau furthermore observed that interesting things happen when a_2 goes through zero. The location of this zero is identified with the critical point, and in the simplest situation a_2 is proportional to t in the neighbourhood of the critical point. a_4 is always positive to ensure that the energy is bounded from below, and it can be treated as a constant in the

vicinity of the critical point. Figure 1 displays the change in the potential energy around T_c; the phenomenon is called "spontaneous symmetry breaking". In this particular example, it is the discrete symmetry $\phi \leftrightarrow -\phi$ that gets broken below T_c.

In the thermodynamic limit, the free energy is dominated by the most probable configuration, i.e. the configuration which minimises $H[\phi]$. Obviously to minimise the gradient term, ϕ has to be spatially uniform: $\phi(x) = M$. As in any phenomenon of spontaneous symmetry breaking, the direction of magnetisation is selected by the direction of the external magnetic field. The minimisation of the energy gives

$$2a_2 M + 4a_4 M^3 - h \;=\; 0 \quad . \tag{3.2}$$

For small h ($\hat{h}$ denotes the unit vector in the direction of h), the solutions to this equation are

$$\begin{aligned} M &= \frac{h}{2a_2} : \; a_2 > 0 \quad , \\ M &= (\frac{h}{4a_4})^{1/3} : \; a_2 = 0 \quad , \\ M &= (\frac{-a_2}{2a_4})^{1/2}\hat{h} - \frac{h}{4a_2} : \; a_2 < 0 \quad . \end{aligned} \tag{3.3}$$

Obviously, a_2 has to be positive above T_c and negative below.

Ignoring the fluctuations altogether, we can immediately read off $\beta = 1/2$ and $\delta = 3$. The susceptibility is

$$\begin{aligned} \chi &= \frac{1}{2a_2} : \; T > T_c \quad , \\ \chi &= -\frac{1}{4a_2} : \; T < T_c \quad , \end{aligned} \tag{3.4}$$

whence $\gamma = 1$. The minimum free energy is

$$\begin{aligned} F(h=0)/T &= a_0 : \; T > T_c \quad , \\ F(h=0)/T &= a_0 - \frac{a_2^2}{4a_4} : \; T < T_c \quad , \end{aligned} \tag{3.5}$$

which shows that the specific heat has a discontinuity at T_c:

$$\Delta C_v \;=\; \frac{a_2^2}{2a_4 t^2} \quad . \tag{3.6}$$

The 2−point correlation function can be extracted by studying the system when h is replaced by a $\delta-$function localised at $x = 0$. The equation of motion then gives

$$-\nabla^2\phi(x) + a_2 M(x) + 2a_4 M^3(x) \;\propto\; \delta^d(x) \quad . \tag{3.7}$$

When the magnetisation is small (which can be assumed for $T > T_c$), the solution is

$$M(r \to \infty) \;\propto\; r^{(1-d)/2} \exp(-\sqrt{a_2} r) \quad , \tag{3.8}$$

implying that $\xi \propto (a_2)^{-1/2}$ and $\nu = 1/2$. At the critical point itself, for $d > 2$,

$$M(r \to \infty) \;\propto\; r^{2-d} \quad , \tag{3.9}$$

implying that $\eta = 0$.

These exponents obey the scaling relations, except that the hyperscaling relation holds for exactly $d = 4$ only. More importantly, they are all independent of the dimensionality of the system, quite unlike the experimental observations. The reason for this failure, it turns out, is the complete neglect of fluctuations around the most probable configuration. In the vicinity of the critical point, these fluctuations can grow to be of macroscopic size, and the framework of renormalisation group shows how to properly handle them. Landau theory then can be considered as the classical or tree-level approximation to the full theory. Correspondingly, the above values for the critical exponents are referred to as their classical values.

Exercise 3: Derive the critical exponents from mean field theory. In the simplest version, the effective field experienced by each spin is h', and the magnetisation follows the Curie law $M \propto h'/T$. The self-consistency condition implies that the effective field is the sum of externally applied field and the effect of neighbouring spins, i.e. $h' = h + M(a - bM^2)$ where a, b are constants. This is sufficient to calculate α, β, γ, δ. What would ν and η be in this case?

4. Block Spin Transformations and the Renormalisation Group

Kadanoff took the next important step in studying critical phenomena. He emphasised the idea that the divergence of the correlation length at the critical point is responsible for singularities in thermodynamic quantities. This feature implies that near the critical point the order parameter behaves co-operatively over macroscopic distances. In such a case, the order parameter can be averaged over a local region, and it would still behave like the effective variable that it is supposed to be. The interaction Hamiltonian of the locally averaged order parameter would still look similar to that of the original system.

For concreteness, consider a 3−dimensional ferromagnet on a cubic lattice. The atomic spins of the original system can be averaged over a $2\times2\times2$ block of atoms, to construct block spins over a block lattice with 8 times fewer number of sites. The block spins now have simple interactions with their neighbours, just as in the original system. The correlation length, in its own lattice spacing units, however, is smaller by a factor of 2 on the block lattice compared to the original lattice. This can be taken care of by observing that the block system may have effective temperature and external magnetic field that are distinct from the original system. The blocking step can be iterated, and in general effective spins can be defined on a lattice of arbitrary spacing L times the original atomic spacing. Kadanoff's idea was that there would be $L-$ dependent parameters T_L and h_L, such that T_{2L} and h_{2L} would be analytic functions of T_L and h_L. At the critical point, where ξ diverges, T_L and h_L would have fixed values independent of L. If the blocking steps are iterated backwards, then one can increase ξ and get closer to the critical point. The critical exponent ν can be calculated provided one knows how T_{2L} depends on T_L. With this picture, Kadanoff was able to derive the scaling laws of critical exponents.

The explicit construction of block variables and their couplings in terms of those for the original system is called a renormalisation group transformation. Kadanoff took the original spins to have only nearest neighbour couplings, and assumed that this would remain true also for the block spins. Moreover, he did not specify how to calculate T_{2L} in terms of T_L. These shortcomings of his analysis were surmounted by Wilson. By combining techniques of field theory and statistical mechanics, Wilson gave an explicit program for calculating the block couplings in terms of original ones, and showed that the block couplings can get quite complicated although retaining the local character of the interaction Hamiltonian (e.g. the block spins may get couplings other than nearest neighbour type even when the original spins have only nearest neighbour couplings).

Let us see how these formal ideas work in a simple situation—the 1−dimensional Ising model. The model is defined by the partition function

$$\begin{aligned} Z_N \equiv \exp(-FN/T) &= 2^{-N}\mathrm{Tr}[\exp(-H/T)] \\ &= 2^{-N}\sum_{s_i=\pm1}\exp[\sum_i\{(J/T)s_is_{i+1}+(B/T)s_i+C\}] \quad . \end{aligned} \tag{4.1}$$

Here the only degree of freedom for an individual spin is its sign. The nearest neighbour coupling between spins, J, is positive for a ferromagnet and negative for an antiferromagnet. B is a uniform external magnetic field, and C is the zero-point energy of a spin. Z can

be easily evaluated by first calculating it for a linear spin chain of finite length N, and then taking the thermodynamic limit $N \to \infty$. The result is independent of the boundary condition. Ignoring the zero-point energy

$$F(T,h) \;=\; -\ln[\cosh(h) + \sqrt{\sinh^2(h) + e^{-4K}}] \;\;,\;\; K \;=\; J/T \;\;,\;\; h \;=\; B/T \;\;. \qquad (4.2)$$

Various thermodynamic quantities can now be extracted from this free energy. It is easy to see that the singularity occurs only at $T = 0$. The standard definitions of the exponents do not apply at this point, but exponents can still be defined by altering the definition of the reduced temperature, or equivalently, by expressing all the thermodynamic singularities in terms of that for the correlation length ξ. (This is a peculiarity of the $1-$dimensional nature of the model.)

Exercise 4: Evaluate the various thermodynamic quantities and the scaling functions for the $1-$dimensional Ising model.

Now we apply the renormalisation group ideas to this model without using the exact solution. Instead of trying to do the sum over all the N spins at once, we have to do the sum over only some of the spins at a time, in such a way as to preserve the system looking as much as possible like it did before the summation, and in such a way that a spatial rescaling of the system is effected. There is a crucial constraint of unitarity that a renormalisation group transformation must obey. The renormalised Hamiltonian is obtained by performing a partial trace over the original degrees of freedom. If one now sums over the remaining degrees of freedom, the result must be the same as if one had performed the entire trace operation in one go. Mathematically, this means that the partition function is preserved under renormalisation:

$$Z_{N'}[H'] \;=\; Z_N[H] \;\;,\;\; F'[H'] \;=\; (N/N')\; F[H] \;\;, \qquad (4.3)$$

which is essential to the derivation of the hyperscaling relation.

Let us perform the partial trace using a scale-2 renormalisation group transformation, in which the block partition function is obtained just by summing over every alternate spin in the linear chain. Using primed symbols to denote block variables, the block spin variables are $s'_i \equiv s_{2i}$, while s_{2i+1} are summed over. A segment of the partition function thus becomes

$$\begin{aligned} Z \;&=\; ...2^{-1} \sum_{s_{2i+1}=\pm 1} \exp[K s_{2i} s_{2i+1} + (h/2)(s_{2i} + s_{2i+1}) + C] \\ &\qquad \times\; \exp[K s_{2i+1} s_{2i+2} + (h/2)(s_{2i+1} + s_{2i+2}) + C]... \\ &=\; ...2^{-1} \{\exp[K(s_{2i} + s_{2i+2}) + (h/2)(s_{2i} + s_{2i+2} + 2) + 2C] \\ &\qquad +\; \exp[-K(s_{2i} + s_{2i+2}) + (h/2)(s_{2i} + s_{2i+2} - 2) + 2C]\}... \;\;, \end{aligned} \qquad (4.4)$$

where for convenience we have introduced a zero-point energy C. The segment evaluated above is a symmetric function of s_{2i} and s_{2i+2}, with three unknowns. With the magnitude of the spins fixed to be ± 1, there are also only three independent and symmetric configurations of these two spins. Thus we can rewrite the segment as

$$Z \;=\; ... \exp[K' s_{2i} s_{2i+2} + (h'/2)(s_{2i} + s_{2i+2}) + C']... \quad , \tag{4.5}$$

where $[K', h', C'] \equiv R_2[K, h, C]$ can be determined in terms of $[K, h, C]$:

$$\begin{aligned} \exp[4K'] \;&=\; \frac{\cosh(2K+h)\cosh(2K-h)}{\cosh^2 h} \quad , \\ \exp[2h'] \;&=\; \exp[2h]\, \frac{\cosh(2K+h)}{\cosh(2K-h)} \quad , \\ \exp[4C'] \;&=\; \exp[8C]\;\; \cosh(2K+h)\cosh(2K-h)\cosh^2 h \quad . \end{aligned} \tag{4.6}$$

We have thus achieved our aim of determining the renormalised parameters after a discrete scale-2 renormalisation group transformation. With $\xi' = R_2(\xi) = \xi/2$, the iteration of the renormalisation group transformation moves one farther and farther away from the critical point with $\xi = \infty$ towards $\xi = 0$. In fact, by following the renormalisation group in the reverse direction, we can locate the parameter values where the system would display critical behaviour.

Exercise 5: Derive the relations in Eq.(4.6).

The particular renormalisation group transformation considered above, where the block degrees of freedom are just a subset of the original degrees of freedom, is known as decimation. It has an intrinsic limitation, i.e. the correlation functions defined on a block lattice are identical to the ones defined on the original lattice, except for a change in lattice spacing. For instance, it constrains the 2$-$point correlation function, $G^{(2)}_{2L}(r/2) \;=\; G^{(2)}_{L}(r)$, implying that $\eta = 2 - d$. This relation holds for the 1$-$dimensional Ising model, but is not true in general. We thus learn that in order to obtain a non-trivial value for η the renormalisation group transformation must rescale the block variables or define them as non-linear functions of the original variables. This rescaling implies that the variables do not scale according to their classical (or engineering) dimensions; they have anomalous dimensions.

There is another drawback of the scale-2 transformation considered above. If we take the original model to be antiferromagnetic, then the couplings become ferromagnetic after one renormalisation step. The reason is that the block spins are not defined so as to preserve the antiferromagnetic nature of the theory. If we consider a scale-3 transformation, however,

then the antiferromagnetic interaction will be preserved by the block spins. The point to note is that the renormalisation group transformation must be designed so as to follow the symmetries of the problem under consideration.

Exercise 6: Carry out a scale-3 block spin transformation for the 1-dim. Ising model. Ignore the magnetic field if you wish, but identify the difference between ferromagnetic and antiferromagnetic behaviour.

Finally, we note there is no unique definition of a renormalisation group transformation for any given class of Hamiltonians. Among the variety of scale transformations that be defined in principle, the one of practical interest would be the one which focuses on the critical phenomena of interest, i.e. the one which is easy to accurately work with while reaching the critical behaviour in the fewest number of steps. It is instructive to look at two specific renormalisation group transformations that have been used for Ising models in different dimensions. The first is a linear one defined by the kernel

$$R_{N',N}(s',s) \;=\; \prod_{i'}^{N'} \tfrac{1}{2}[1+\rho s'_{i'} \sum_{i\in\text{block } i'}^{N/N'} s_i] \quad . \tag{4.7}$$

ρ is a free parameter, which avoids undue restrictions on η, while still obeying the unitarity constraint (i.e. $\sum_{s'_{i'}} R(s',s) \;=\; 1$). Decimation can be looked upon as having $\rho = 1$ for one of the spins in the block and $\rho = 0$ for the rest. The advantage of the free parameter is that it can be fixed using a self-consistency condition, or optimised on the basis of some variational principle. Results for critical exponents show marked improvement after such an optimisation. The second renormalisation group transformation is a non-linear one:

$$R_{N',N}(s',s) \;=\; \prod_{i'}^{N'} \tfrac{1}{2}[1+s'_{i'}\ \text{sign}(\sum_{i\in\text{block } i'}^{N/N'} s_i)] \quad . \tag{4.8}$$

This defines a majority rule, supplemented by a random choice for $s'_{i'}$ in case of an even split. In actual calculations using these position space block transformations, the complexities of couplings grow rapidly. Typically, after only a few iterations, one has to devise approximation methods or numerical algorithms to keep the problem tractable. Field theoretical methods come handy when one changes over to the momentum space instead, and at present that is where the best results come from.

Figure 2: The renormalisation flow pattern in the $[K, h]$ plane for the ferromagnetic $1-$dimensional Ising model.

5. Renormalisation Flows, Fixed Points and Classification of Operators

Let us go back to the ferromagnetic $1-$dimensional Ising model and see what happens to its couplings under repeated block transformations. The recursion relations, Eq.(4.6), can be looked upon as "flow equations" in the "space of couplings". The variable C does not alter the other two couplings (though it is affected by them), and so we can restrict our attention to the two-dimensional space spanned by temperature and magnetic field. Eq.(4.6) implies that $K' \leq K$, and $h' \geq h$ when $h \geq 0$. The problem also has a discrete symmetry under $h \leftrightarrow -h$. We thus obtain the flow diagram shown in Fig. 2.

It is easy to see by inspection that certain points in the coupling space do not change from their initial value under the renormalisation group transformation. Such points are called "fixed points", and they by definition satisfy

$$R_s(T^*, h^*) \;=\; (T^*, h^*) \quad . \tag{5.1}$$

For the $1-$dimensional Ising model, there is an isolated fixed point at $(K = \infty, h = 0)$, and a line of fixed points at $(K = 0, h)$.

Since a physical system, and therefore its correlation length, is completely specified by the set of couplings appearing in its action or Hamiltonian, and since we have $R_s(\xi) = \xi/s$, there are only two possible values for the correlation length at a fixed point: (i) $\xi = \infty$, which evidently characterises a critical point, and (ii) $\xi = 0$, which corresponds to a trivial fixed point. In the case of $1-$dimensional Ising model, the fixed point at $(T = 0, h = 0)$ is a critical point, while the line of fixed points at $(T = \infty, h)$ is made up of trivial fixed points. Moreover, the absolute scale of free energy is determined by the fixed point; scale invariance demands that the free energy vanishes at the fixed point.

The flow pattern in the neighbourhood of a fixed point classifies the operators that appear (or may appear) in the action or Hamiltonian. The directions in which the renormalisation group flow moves away from the fixed point are called "relevant", the directions in which the flow moves towards the fixed point are called "irrelevant", and the directions in which the flow does not move at all are called "marginal". Clearly the fixed point with $\xi = \infty$ must have at least one relevant operator, while the fixed points with $\xi = 0$ have only irrelevant and marginal operators. In the example under consideration, the fixed point at $(K = \infty, h = 0)$ has both temerature and the magnetic field as relevant directions, while the line of fixed points at $(K = 0, h)$ has temperature as an irrelevant direction and magnetic field as a marginal direction.

Exercise 7: Study the renormalisation flow pattern for the scale-3 block spin transformation for the 1-dim. Ising model.

The existence of fixed points with $\xi = \infty$ justifies the "scaling hypothesis". At these points, the theory possesses a symmetry under the renormalisation group transformation R_s; it does not have any inherent scale. All the macroscopic critical behaviour then must be describable in terms of a single overall scale, which can be conveniently chosen as ξ. Actually, we need an additional assumption of smoothness in the neighbourhood of the fixed point, i.e. if $R(H) = H'$ and $R(H+\delta H) = H'+\delta H'$ then as $\delta H \to 0$ we also have $\delta H' \to 0$. Technically, this means that we can define various derivatives characterising the flow of couplings, such as $\partial K'/\partial K$. We never have exact thermodynamic limit in the real world; the growth of the correlation length is cut off by the physical size of the system. We nonetheless observe critical behaviour in the region where ξ is not infinite but is still much larger than any other inherent scale in the system. Given a uniform thermodynamic limit and dominance of a single length scale, the departure from scaling behaviour can be bounded and the limiting scaling behaviour can be extracted.

The fixed points with $\xi = \infty$, together with the subspace spanned by irrelevant couplings around it, define a "critical surface". All the points on the critical surface have $\xi = \infty$ too, since they converge to the fixed point under renormalisation group transformation. If the critical behaviour of a system is truly macroscopic, then it would be dictated only by the fixed point and the relevant directions around it. The domain of attraction (or stability) formed by a fixed point and the irrelevant couplings in its neighbourhood is the basis for the concept of "universality". Different points on the same critical surface describe scale invariant systems, which differ by irrelevant operators. They have different microscopic Hamiltonians, but the same long distance behaviour (which is extracted by repeated application of a renormalisation group transformation). Thus we have a variety of ferromagnetic materials with quite different atomic structures and interactions, but displaying the same critical exponents.

The "critical region" is defined as the neighbourhood of the fixed point where corrections due to irrelevant operators are negligible. It is this region in where we can experimentally test the scaling hypothesis, since the irrelevant operators typically give corrections to the scaling forms. The subspace of the parameter space spanned by relevant operators only is called the "renormalised trajectory" (it is often a single parameter line). If the original Hamiltonian did not have $\xi = \infty$ exactly, then the renormalisation flows do not stop, but converge towards this renormalised trajectory. The renormalised trajectory leads one away from the fixed point with $\xi = \infty$ towards the trivial fixed point at $\xi = 0$. The corrections to scaling behaviour are expected to be small along this renormalised trajectory, and in practice it is desirable to carry out renormalisation group calculations, particularly the numerical ones, close to it. Note that in an interacting theory, the flow even along the renormalised trajectory can become highly nonlinear and violate scaling far away from the fixed point, so one cannot afford to go too far from the fixed point of interest.

There is another class of operators, labeled "redundant" operators, that show up in renormalisation group analysis of a general problem. They are typically generated by a change of variables in the functional integral, and hence do not affect the free energy at all. Though they have no physical implications, in a specific implementation of a renormalisation group transformation, they can show up as either relevant, irrelevant or marginal operators. This can be a nuisance, particularly in numerical implementations of renormalisation group. They can be turned around to an advantage, however, because the freedom to suitably redefine functional variables can be exploited to simplify the form of the action or Hamiltonian and change the location of the fixed point (and the renormalised trajectory) of interest.

6. Linearised Transformations around Fixed Points and Critical Exponents

Our task now is to understand how the critical behaviour is completely specified by the fixed points and its relevant directions. In general R_s is a non-linear transformation, as is evident for example from Eq.(4.6). In the vicinity of the fixed point, however, we can linearise it on the assumption that it behaves sufficiently smoothly. Let the vector μ denote a point in the coupling space. Around the fixed point μ^*, we formally write

$$\mu \;=\; \mu^* + \delta\mu \;\;,\;\; R_s(\mu) \;=\; \mu' \;\;,\;\; \delta\mu' \;=\; R_s^L \delta\mu \;\;. \tag{6.1}$$

The matrix R_s^L represents the linearised flow around the fixed point, with its elements given by

$$(R_s^L)_{mn} \;=\; \left(\frac{\partial \mu'_m}{\partial \mu_n}\right)_{\mu=\mu^*} \;\;. \tag{6.2}$$

If the system has additional symmetries which are respected by the renormalisation group transformation, e.g. the even and odd couplings in case of a ferromagnet, then the matrix R_s^L has a block diagonal form and the different subspaces can be studied separately.

The natural thing to do is to determine the eigenvalues and eigenvectors of this linearised transformation matrix. The eigenvectors do not depend on the scale factor s, but the eigenvalues clearly do. The eigenvectors define the "scaling fields", i.e. linear combinations of couplings of the theory that get multiplicatively renormalised under a scale transformation. (The scaling fields are in general highly nonlinear, we actually have only a linear approximation for them.) These fields provide the basis for performing dimensional analysis for scale invariant systems; we used them implicitly in deriving the scaling relations between various critical exponents.

Since the scale transformations satisfy the group composition rule, $R_s R_{s'} = R_{ss'}$, the eigenvalues of R_s^L must have the form s^{λ_i}, where λ_i are independent of s. According to the operator classification scheme

$$\begin{aligned} \text{Relevant operators}: &\quad \lambda \;\geq\; 0 \;\;, \\ \text{Marginal operators}: &\quad \lambda \;=\; 0 \;\;, \\ \text{Irrelevant operators}: &\quad \lambda \;\leq\; 0 \;\;. \end{aligned} \tag{6.3}$$

Note that an operator which is marginal in the linear approximation may turn out to be relevant or irrelevant when higher order terms are included. Such marginal operators are important in the analysis of theories with logarithmic (instead of power law) singularities.

A variable is defined to have scaling dimension λ, if it scales as $R(A) = s^{\lambda}A$ under a renormalisation group transformation. We denote this behaviour by $D(A) = \lambda$. Obviously, $D(\xi) = -1$, $D(k) = 1$ and $D(F) = d$, which are nothing but their classical dimensions. From the definition of the 2−point correlation function, we find that $D(\phi(k)) = -1 + \eta/2$, $D(\phi(x)) = (d-2+\eta)/2$ and $D(h) = (d+2-\eta)/2$, which differ from their classical dimensions whenever $\eta \neq 0$.

For a ferromagnet, we have two relevant exponents, λ_t and λ_h, corresponding to the directions representing temperature and magnetic field respectively. The definitions of critical exponents (cf. Eq.(2.1)) imply that the largest exponent in the even sector $\lambda_t = 1/\nu$, while the largest exponent in the odd sector $\lambda_h = \beta\delta/\nu$. Let us now see how the scaling relations arise. The flow of the free energy near the fixed point is given by

$$s^d F(t,h) \;=\; F(s^{\lambda_t}t, s^{\lambda_h}h) \quad . \tag{6.4}$$

Choosing a reference point t_0 such that $s = (t/t_0)^{-1/\lambda_t}$,

$$F(t,h) \;=\; \left(\frac{t}{t_0}\right)^{d/\lambda_t} f(t_0, \frac{h}{(t/t_0)^{\lambda_h/\lambda_t}}) \quad . \tag{6.5}$$

On comparing this with the scaling form in Eq.(2.2), we identify $2-\alpha = d/\lambda_t$ and $\beta\delta = \lambda_h/\lambda_t$. Other relations between critical exponents and critical amplitudes can be derived similarly.

We can now work through the 1−dimensional Ising model example and determine all its critical exponents. The non-trivial fixed point is at $(T^* = 0, h^* = 0)$. The scale−2 transformation gives

$$h = 0: \quad \tanh(K') \;=\; \tanh^2(K) \quad , \quad T = 0: \quad h' \;=\; 2h \quad . \tag{6.6}$$

These results can be generalised to renormalisation group transformation by an arbitrary scale factor s, as

$$h = 0: \quad R_s[\tanh(K)] \;=\; (\tanh(K))^s \quad , \quad T = 0: \quad R_s[h] \;=\; sh \quad . \tag{6.7}$$

We can identify the correlation length as $\xi = -1/\ln(\tanh(K))$, in agreement with the exact 2−point correlation function

$$G^{(2)}(r, K, h = 0) \;=\; (\tanh(K))^r \quad . \tag{6.8}$$

Moreover, $\lambda_h = 1$ which implies that $\eta = 1$.

In the linear approximation, h is the magnetic scaling field, while it is convenient to choose $x = \exp(-4K)$ as the temperature scaling field. The recursion relations for the scale-2 transformation become

$$x' = 4x \ , \ h' = 2h \ , \ F(x,h) = F(4x,2h)/2 \ . \tag{6.9}$$

It follows that $\lambda_x = 2$, and $2 - \alpha_x = \gamma_x = \nu_x = 1/2$. The magnetisation is discontinuous at $T = 0$, $M(T = 0, h) = \text{sign(h)}$, which is in agreement with $\delta = \infty$ prediction of the scaling relations. Finally, we note that $M(T \neq 0, h = 0)$ is always zero, still we can take $\beta = 0$ from the scaling relations.

In principle one can define the renormalisation group in terms of only discrete scale transformations, e.g. $R_s = (R_2)^l$ for $s = 2^l$. A general description in terms of continuous s may not exist. This is often the case in numerical implementation of the renormalisation group in position space. A linear approximation then is defined using discrete changes instead of derivatives as done above. The critical dimensions are extracted as logarithms of the eigenvalues of the discrete scale change matrix, e.g. $\lambda_i = \log(e_i)/\log(2)$ for a scale-2 transformation matrix with eigenvalues e_i.

7. The Gaussian Model and its Renormalisation Group Analysis

We now move on to another exactly solvable model, the Gaussian model, to illustrate the renormalisation group in momentum space using field theory methods. The model is obtained by just retaining the quadratic part of the Landau-Ginzburg Hamiltonian, which we write with a momentum cutoff Λ as

$$\begin{aligned} H[\phi]/T &= a_0 L^d + \sum_{k<\Lambda} (a_2 + ck^2)|\phi(k)|^2 \ , \\ \phi(x) &= L^{-d/2} \sum_{k<\Lambda} \phi(k) e^{ik\cdot x} \ . \end{aligned} \tag{7.1}$$

The Brillouin zone sum over k can be replaced by an integral over a sphere of radius Λ in explicit calculations: $\sum_{k<\Lambda} \to L^d \int^\Lambda d^dk/(2\pi)^d$. For this model, the Hamiltonian is just a sum of independent oscillators, and we can easily write down

$$\begin{aligned} G^{(2)}(k) &= <|\phi(k)|^2> = \tfrac{1}{2}(a_2 + ck^2)^{-1} \ , \\ FL^d &= a_0 L^d T - \tfrac{1}{2} T \sum_{k<\Lambda} \ln[\pi/(a_2 + ck^2)] \ . \end{aligned} \tag{7.2}$$

$G^{(2)}(k)$ is a measure of the fluctuations of the spins.

The critical point of the theory is at $a_2 = 0$, and as before we take a_2 proportional to $T - T_c$ in its neighbourhood. For $T > T_c$,

$$G^{(2)}(k, T = T_c) \ \propto \ k^{-2} \ \ , \ \ G^{(2)}(k = 0, T \sim T_c) \ \propto \ (T - T_c)^{-1} \ \ , \tag{7.3}$$

which yield $\eta = 0$ and $\gamma = 1$. The specific heat is

$$C_v \ = \ \tfrac{1}{2} \left(\frac{a_2}{T - T_c} \right)^2 T^2 \int^{\Lambda} \frac{d^d k}{(2\pi)^d} (a_2 + ck^2)^{-2} \ + \ \ldots \ \ , \tag{7.4}$$

where only the leading singular term is kept. It is proportional to ξ^{4-d}, when the parameters are factored out of the integral using $\xi = \sqrt{c/a_2}$. So we have $\nu = 1/2$ and $\alpha = 2 - d/2$.

For $T < T_c$, it is necessary to introduce a small positive quartic term in the Hamiltonian to keep it bounded from below. The system develops a magnetisation, and we now expand the Hamiltonian up to quadratic order around the minimum given by mean field theory:

$$H[\phi]/T \ = \ H[\overline{\phi}]/T + \sum_{k<\Lambda, k\neq 0} (-2a_2 \overline{\phi}^2 + ck^2) |\phi(k)|^2 \ \ , \tag{7.5}$$

where $\overline{\phi} = \sqrt{-a_2/2a_4}$ is the zero mode solution. We can again calculate $G^{(2)}$, F and C_v, the only change being replacement of a_2 by $-2a_2$. The exponents remain unaffected.

By construction, the Gaussian model has the same β and δ as in mean field theory, and we see that γ, η and ν also agree with mean field theory. In the case of α, however, we see the role played by fluctuations: α is different from mean field theory for $d \neq 4$, and satisfies the hyperscaling relation. We also observe from Eq.(7.4) that the specific heat does not diverge in the infrared at T_c when $d > 4$; the integral is finite. The actual behaviour of the specific heat then has to be obtained by keeping all the non-leading terms that were ignored in Eq.(7.4). The result turns out to be the same as in mean field theory, i.e. a discontinuity in C_v at T_c and effectively $\alpha = 0$.

All this accumulated evidence indicates that the mean field theory is exact for $d > 4$. The fluctuations modify the results in a non-trivial way for $d < 4$, while $d = 4$ is a special borderline case. In fact we see that the effect of fluctuations becomes more and more pronounced as d is decreased. It is not difficult to understand this feature. The equilibrium configuration of a system is the result of a balance between its entropy and energy. With increasing d, the number of neighbours (say on a lattice) of a single variable increase, but its own degrees of freedom remain the same. Thus the potential energy always wins for large

enough d; mean field theory is obviously exact for $d \to \infty$. For small d, the entropy can take over, and the interesting point is to find out the value of d where a cross-over occurs. For Landau-Ginzburg type Hamiltonians, this upper critical dimension turns out to be $d = 4$.

Exercise 8: Propagators of point particles in free field theories can be represented as random walks. Noting that the fractal (or Hausdorff) dimension of a random walk is 2, argue that the upper critical dimension for ϕ^4 type theories should be $d = 4$. What would be the upper critical dimension for a ϕ^6 theory ?

Exercise 9: Let us go back to the $1-$dimensional ferromagnetic Ising model. Find the energy of a kink solution, i.e. the configuration where the magnetisation is of one sign on one side and of opposite sign on the other. A general configuaration is just a superposition of kinks. Show that the entropy of the kinks is always large enough to disorder the system at any non-zero temperature, i.e. $M(T \neq 0, h = 0) = 0$.

The Gaussian model includes the leading effect of fluctuations on top of the tree-level mean field theory. So we can use the size of the correction to estimate the range of validity of the mean field theory. This estimate is known as the "Ginzburg criterion". This criterion is of practical importance, in measuring critical exponents driven by fluctuations. Unless the measurements are made sufficiently close to T_c, the results would turn out to be consistent with the mean field theory expectations. A well-known example illustrating this point is the superconducting transition. The traditional BCS theory ignores fluctuations completely, and still works amazingly well. Inserting typical values for the parameters, we find that the violations of BCS theory are not expected more than $\sim 10^{-15}\ ^\circ K$ away from the transition temperature.

Exercise 10: Compare the Landau theory discontinuity in specific heat, Eq.(3.6), with the singular contribution to specific heat in the Gaussian model, Eq.(7.4). Find the criterion to limit the range of temperature around T_c where the Landau theory breaks down. Express the result in terms of field theory diagrams.

Now we can analyse the same model following the renormalisation group. We ignore the additive constant in the Hamiltonian, and define the parameter space of the model is defined by $\mu = (a_2, a_4, c)$. The first step of the renormalisation process is to integrate out $\phi(k)$ in the range $\Lambda < k < \Lambda/s$. We are focusing our attention on the case $u = 0$, for which the integration is trivial. The second step then is to rescale the variables $\phi(k) \to s^{1-\eta/2}\phi(sk)$.

This rescaling restores the integration range back to its original size and allows the possibility of a non-trivial anomalous dimension:

$$\begin{aligned}(H[\phi]/T)' &= \sum_{k<\Lambda/s} s^{2-\eta}(a_2 + ck^2)|\phi(sk)|^2 \quad , \\ &= \sum_{k'<\Lambda} (a_2 s^{2-\eta} + cs^{-\eta}k'^2)|\phi(k')|^2 \quad .\end{aligned} \tag{7.6}$$

The renormalised Hamiltonian now has the same form as the original one, and we can read off

$$\mu' \;=\; R_s(\mu) \;=\; (a_2 s^{2-\eta}, a_4 s^{4-2\eta}, cs^{-\eta}) \quad . \tag{7.7}$$

When we choose $\eta = 0$, we find the fixed point $\mu_0^* = (0, 0, c)$, It is called the Gaussian fixed point, and corresponds to a massless free field theory. By setting $\eta = 2$, we obtain two other fixed points: $\mu_\infty^* = (a_2 > 0, 0, 0)$ and $\mu_{-\infty}^* = (a_2 < 0, 0_+, 0)$. These two describe the cases where each spin is independent of all the others, i.e. the correlation length is zero. μ_∞^* describes the situation at very high temperatures, while $\mu_{-\infty}^*$ describes a magnetised system far below T_c. In the latter case, the magnetisation is uniform for $c = 0_+$, while $c = 0_-$ can give rise to antiferromagnetic behaviour.

The physically interesting case is the Gaussian fixed point, and we next study its neighbourhood in the linearised approximation. We take the various terms in the power series expansion of the Hamiltonian as forming the basis for the scaling fields. The scaling dimensions of these terms are all classical:

$$\begin{aligned}D(A) \;=\; d \;+\; &(1 - d/2) \times (\text{power of } \phi(x) \text{ in } A \\ &+\; (-1) \times (\text{power of } \nabla \text{ in } A \quad .\end{aligned} \tag{7.8}$$

First we look at the terms involving even powers of ϕ. We see that c is always a marginal parameter, independent of the dimensionality of the system. Moreover, terms involving higher derivatives of ϕ are always irrelevant. When $d > 4$, ϕ^2 is the only relevant operator; increasing powers of $\phi(x)$ leads to more and more irrelevant operators. For $4 > d > 3$, ϕ^4 also becomes relevant. For $3 > d > 2$, ϕ^6 turns relevant too. For $2 > d$, all powers of ϕ are relevant.

Of the terms in the odd sector, the leading magnetic interaction has scaling dimension $1 + d/2$, and is always relevant. At this stage, we do not worry about other terms which represent anisotropic perturbations. We thus find that the Gaussian fixed point is stable for $d > 4$, with two relevant directions: $\lambda_t = 2$ and $\lambda_h = 1 + d/2$.

The scaling relations give $\gamma = 1$ and $\nu = 1/2$ for the Gaussian fixed point. They, however, also give $\alpha = 2 - d/2$, $\beta = (d-2)/4$ and $\delta = (d+2)/(d-2)$, which are in contradiction with the actual solution of the model, for $d > 4$. The reason why the scaling relations fail for α, β and δ has to do with scaling in the presence of "dangerous irrelevant operators". The coupling a_4 of the Landau-Ginzburg model, though it vanishes at the fixed point, is essential to keep the Hamiltonian bounded from below for $T < T_c$. The scaling arguments only give, e.g. for the magnetisation,

$$s^{-1+d/2} M(t, h, a_4) \;=\; M(ts^2, hs^{1+d/2}, a_4 s^{4-d}) \quad . \tag{7.9}$$

The third argument goes to zero at the fixed point for $d > 4$, but we can only drop it if the scaling function remains finite at the fixed point. In this particular case, $M(h = 0)$ diverges like $a_4^{-1/2}$ and $M(t = 0)$ diverges like $a_4^{-1/3}$ near the fixed point. Upon inserting this information, we recover the correct values $\beta = 1/2$ and $\delta = 3$. Similar considerations show that the free energy behaves like a_4^{-1} near the fixed point, yielding the correct value $\alpha = 0$.

The moral of the story is that the scaling assumptions tacitly depend on non-singular behaviour of various scaling functions and their arguments. In many cases, these assumptions hold, but in certain cases they fail and the exponent relations change their form. It is usually the hyperscaling relation, which assumes regularity of the free energy as a function of its arguments near the fixed point, that fails, while the other exponent relations survive. The renormalisation group, unfortunately, does not have anything to say about these caveats. One has to be careful and justify the implicit assumptions, either with explicit solutions or with educated guesses (such as the Gaussian approximation above). We learn that the irrelevant operators are not irrelevant in the literal sense: they of course control the corrections to scaling, but they can even modify the scaling relations in some situations.

8. Landau-Ginzburg Model and $\epsilon-$Expansion

For $4 > d > 3$, the Gaussian fixed point becomes "tricritical" with three relevant directions, and we have to search whether there is any other fixed point with only two relevant directions. If so, the new fixed point will describe the generic critical phenomena for these dimensionalities. We anticipate that the new fixed point would coincide with the Gaussian

Figure 3: The Feynman diagrams representing the leading corrections to the Landau-Ginzburg model parameters a_2 and a_4 in the $\epsilon-$expansion.

fixed point for $d = 4$, and would move away from it for $d < 4$. Dimensional regularisation is the most convenient framework to study such a behaviour.

Following Wilson, we perform a simultaneous expansion in a_4 and $\epsilon = 4 - d$. (We know that $a_4^* \to 0$ as $\epsilon \to 0$, and the results justify the expansion a posteriori.) The leading perturbative corrections to a_2 and a_4 come from the $1-$loop diagrams of Fig. 3. As per the philosophy of the renormalisation group, we carry out momentum integrals only over a limited range $\pi/s < k < \pi$ for the internal loop variables. All the external lines in the diagrams correspond to $k < \pi/s$.

First we note that the only $1-$loop propagator renormalisation diagram is a tadpole. The external momentum does not enter the loop, and so there is no wavefunction renormalisation (i.e. no correction to $(\nabla\phi)^2$ term) at $1-$loop. Thus $d - 2D(\phi(x)) = 2$, and it follows that $\eta = O(\epsilon^2)$. At $O(\epsilon)$, the counting rules for determining scaling dimensions of various operators are the same as in the Gaussian model (cf. Eq.(7.8)). The tadpole does give rise to mass renormalisation:

$$s^{2D(\phi(x))-d} a_2' \;=\; a_2 \;+\; 3a_4 \int_{\pi/s}^{\pi} \frac{d^d k}{(2\pi)^d} \frac{1}{a_2 + ck^2} + O(a_4^2) \quad . \tag{8.1}$$

Similarly the vertex renormalisation gives:

$$s^{4D(\phi(x))-d} a_4' \;=\; a_4 \;-\; 9a_4^2 \int_{\pi/s}^{\pi} \frac{d^d k}{(2\pi)^d} \frac{1}{(a_2 + ck^2)^2} + O(a_4^3) \quad . \tag{8.2}$$

Here we have taken the external momenta to be zero for simplicity and without loss of generality. An expansion in powers of external momenta can be carried out. It produces terms like $(\nabla\phi)^2\phi^2$ in the renormalised Hamiltonian. They represent irrelevant operators at

Figure 4: The renormalisation flow pattern around the Gaussian fixed point μ_0^* and the stable fixed point μ^* for $d = 4 - \epsilon$, $\epsilon > 0$.

the fixed point and hence are ignored. For the same reasons, we also ignore diagrams giving terms like ϕ^6.

The renormalisation integrals can be easily carried out, after expanding the denominators in powers of a_2. With

$$x \;=\; 3\int_{\pi/s}^{\pi} \frac{d^dk}{(2\pi)^d}\frac{1}{ck^2} \quad , \quad y \;=\; 3\int_{\pi/s}^{\pi} \frac{d^dk}{(2\pi)^d}\frac{1}{(ck^2)^2} \quad , \tag{8.3}$$

the renormalisation equations for the parameters become

$$\begin{aligned} a_2' &= s^2(a_2 + xa_4 - ya_2a_4) \;+\; O(a_2^2a_4, a_4^2) \quad , \\ a_4' &= s^\epsilon(a_4 - 3ya_4^2) \;+\; O(a_2a_4^2, a_4^3) \quad . \end{aligned} \tag{8.4}$$

In addition to the Gaussian fixed point ($a_2^* = 0, a_4^* = 0$), we obtain a new fixed point as

$$a_2^* \;=\; -\frac{\epsilon x}{3y}\frac{\ln(s)}{1 - s^{-2}} \;+\; O(\epsilon^2) \quad , \quad a_4^* \;=\; \frac{\epsilon}{3y}\ln(s) \;+\; O(\epsilon^2) \quad . \tag{8.5}$$

The linearised transformations in the neighbourhood of this new fixed point give

$$\begin{pmatrix} a_2' - a_2^* \\ a_4' - a_4^* \end{pmatrix} \;=\; \begin{pmatrix} s^2(1 - ya_4^*) & s^2xa_4^* \\ 0 & s^\epsilon(1 - 6ya_4^*) \end{pmatrix} \begin{pmatrix} a_2 - a_2^* \\ a_4 - a_4^* \end{pmatrix} \tag{8.6}$$

We can read off the eigenvalues $D(\phi^2) = 2 - \epsilon/3 = 1/\nu$ and $D(\phi^4) = -\epsilon$. These eigenvalues show that, for $d < 4$, the new fixed point has one less relevant operator than the Gaussian fixed point, i.e. a_2 is a relevant variable but a_4 is not. The new fixed point is only $O(\epsilon)$ away

from the Gaussian fixed point, and it exchanges stability with the Gaussian fixed point when d crosses 4. An illustration of the renormalisation flow pattern is given in Fig. 4.

Exercise 11: Evaluate the constants x and y in the limit $s \to 1$, and explicitly determine the linearised transformation matrix in Eq.(8.6).

The other critical exponents for the new fixed point can be determined from η and ν using the scaling relations. There are no caveats due to dangerous irrelevant operators, since a_4 is already positive at the new fixed point. We obtain $\alpha = \epsilon/6$, $\beta = 1/2 - \epsilon/6$, $\gamma = 1 + \epsilon/6$, $\delta = 3 + \epsilon$. Obviously the calculation can be extended by including higher order graphs and extending the power series expansion in terms of ϵ. The series, however, turns out to be only asymptotic, and application to physical systems involves setting $\epsilon = 1$. Nonetheless, the results, upto $O(\epsilon^2)$ accuracy, are in good agreement with experimental observations.

The concept of universality provides a basis for the wide range of applicability for the Landau-Ginzburg model. A field theory defined using continuous variables ϕ looks very different from the one defined using Ising spins $s = \pm 1$, yet they show the same critical behaviour. Let us understand how using some algebraic manipulations. Ignoring zero-point energy, the partition function for a generalised Ising model is

$$Z \;=\; \sum_{s_i=\pm 1} \exp[\sum_{ij} s_i K_{ij} s_j + \sum_i h_i s_i] \quad . \tag{8.7}$$

Using identities of Gaussian integrals, it can be rewritten as

$$\begin{aligned} Z \;&=\; \sum_{s_i=\pm 1} \int_{-\infty}^{\infty} \prod_i d\phi_i \exp[-\tfrac{1}{4} \sum_{ij} \phi_i K_{ij}^{-1} \phi_j + \sum_i (\phi_i + h_i) s_i] \\ &=\; \int_{-\infty}^{\infty} \prod_i d\phi_i \exp[-\tfrac{1}{4} \sum_{ij} (\phi_i - h_i) K_{ij}^{-1} (\phi_j - h_j) + \ln(\cosh(\phi_i))] \quad . \end{aligned} \tag{8.8}$$

An expansion,

$$\ln(\cosh(\phi_i)) \;=\; \tfrac{1}{2}\phi_i^2 \;-\; \tfrac{1}{12}\phi_i^4 \;+\; \ldots \;, \tag{8.9}$$

then gives the mapping between Ising and continuous variable theories. We have a meaningful correspondence, as long as the interaction K_{ij} has a local inverse and the higher order terms in Eq.(8.9) correspond to irrelevant operators.

The Landau-Ginzburg model can be easily generalised by making the order parameter an $n-$component vector field $\vec{\phi}$. The Ising model corresponds to $n = 1$ and the spontaneous breakdown of Z_2 symmetry. For larger values of n, the $O(n)$ symmetry can break in different

ways, leading to different critical phenomena. (A simple magnetisation as in the Landau-Ginzburg model corresponds to $O(n)$ breaking to 1.) Typically Goldstone bosons arise and give new features to the critical behaviour of the system. The critical exponents for these models have been worked out in the $\epsilon-$expansion.

As we have noted earlier, the critical exponents are essentially governed by the symmetry aspects of the order parameter: dimensionality, number of components and nature of symmetry breaking. A number of physical systems exist corresponding to different values for these. For example, bulk materials have $d = 3$, layered systems and thin films have $d = 2$, while long polymers can be described essentially by $d = 1$. Self-avoiding random walks are described by the limit $n = 0$, binary alloys and liquid-gas phase transitions by $n = 1$, superfluidity in liquid He^4 by $n = 2$ (the order parameter is the phase of the complex wavefunction), isotropic magnets by $n = 3$, and superfluidity in liquid He^3 by $n = 18$. For $d = 2$, Kosterlitz-Thouless transition provides a different type of scenario where the phase transition occurs not by symmetry breaking but by vortex condensation. The renormalisation group framework has been successfully applied to all these systems.

9. Field Theories and the Renormalisation Group Equation

Now let us move on from condensed matter systems to field theories. Here the physical world is in $d = 4$, which is the borderline case. Zero coupling fixed points, logarithmic singularities and marginal operators abound, instead of finite transition temperatures and power law singularities typically seen in condensed matter systems.

There is a basic difference between statistical mechanics and continuum field theories. Both can be described using functional integrals, but the formal involves the weight function $\exp[-H/T]$ while the latter involves the weight function $\exp[iS/\hbar]$. The two formulations can be related by analytic continuation between Euclidean and Minkowski spaces, and there is a well-established prescription to convert one type of results into other. It is assumed that there are no singularities in the complex energy plane that come in the way of rotating integration contours, and that the amplitudes are well-behaved at infinity. In such a case, the prescription is easily implemented by inserting appropriate $i\epsilon$ terms in the energy denominators. We shall not worry about these technicalities, and instead directly deal with field theories in Euclidean space.

The other major difference arises from the way the continuum limit is taken, holding a physical scale fixed while letting the ultraviolet regulator go to zero. (Contrast this with critical phenomena, where the ultraviolet regulator is fixed but the correlation length diverges.) The irrelevant operators generally become negligible as the ultraviolet regulator disappears from the theory; their scaling dimensions typically have only logarithmic corrections to their classical values. The easiest choice is to just ignore them, and we hardly see field theory actions involving operators with dimensionality larger than that of the system (i.e. operators which can appear in the action only with suppression factors of the cutoff). Instead of the infinite dimensional space of all possible couplings, we then have only a finite number of relevant operators to deal with, and the physics is dictated by the flow of the couplings along the renormalised trajectory. The fixed point behaviour of the continuum theory is thus specified by two aspects: (a) The (anomalous) dimensions of various field operators, and (b) The rate of flow in the parameter space (along the renormalised trajectory) with respect to a logarithmic change in scale. Conventionally, these are called the $\gamma-$ and $\beta-$functions of the theory, while the final statement embodying the scale invariance symmetry of the theory is known as the renormalisation group equation. (Quite understandably, historical development of the subject was in the opposite direction. The simplified renormalisation prescription of continuum field theories led to the powerful framework applicable to critical phenomena.)

The continuum integrals are also performed in a different manner. Instead of integrating only over the momentum shell $\Lambda/s < k < \Lambda$, the integrals are evaluated over the complete range of momenta $0 \leq k < \Lambda$. Complications often arise in evaluating such integrals due to infrared divergences (the $k = 0$ mode) when there are massless fields in the theory. The desired information is extracted at the end by differentiating the results with Λ, which is tantamount to performing the momentum shell integrals without having to worry about the infrared divergences. All this correspondence is straightforward, when the ultraviolet regulator is employed in position or momentum space. The more popular choice, in weak coupling perturbation theory, however, is dimensional regularisation (due to its mass independent nature) where the regulator scale somewhat mysteriously appears from the change in dimensionalities of the couplings. The final result of renormalisation group analysis, of course, has to be the same for any renormalisation prescription, when expressed in terms of appropriate variables. The change in the nature of the regulator can be absorbed into a change in the boundary conditions for the couplings at that scale, which of course is the essence of renormalisation.

The dynamics of the continuum field theory is determined by its $n-$point correlation functions, $G^{(n)}(k_1, k_2, ..., k_n)$. The field theory action is typically written in terms of "bare" couplings and fields. The bare Green's functions, when evaluated in terms of the bare variables, are then found to be divergent. It becomes necessary to introduce a regulator to make them finite. Counter-terms, which themselves are divergent, are added to the theory. The fields and couplings are then rescaled, e.g. $\phi_0 = Z^{1/2}\phi(\mu)$, and the results are expressed in terms of the new "renormalised" variables. The merit of renormalisation lies in the fact that the renormalisation $Z-$factors can be defined in such a manner that (a) the renormalised variables are finite, and (b) they depend on the physical observation scale (also called the renormalisation point) μ, but not on the arbitrary regulator of the theory.

It is worth contrasting this field theory version of renormalisation group, with the physical approach we have studied earlier. The field theory version is rather algorithmic, cancellation of infinity with infinity results in a finite contribution. The bare variables have no physical interpretation; only the renormalised variables correspond to physical observables. On the other hand, in the physical approach, only renormalised variables appear. They are always specified at some physical scale. There are no divergences and no cancellation of infinities is needed at any stage.

Let us go back to the field theory version and obtain predictions which do not involve the variables at all. These predictions will then contain the consequences of the scale invariance of the theory. The "bare" Green's functions do not depend on the renormalisation point, but the renormalised Green's functions, $G_R^{(n)} = Z^{-n/2}G_B^{(n)}$, do. This condition gives us the renormalisation group equation:

$$\frac{dG_B^{(n)}}{d\mu} \;=\; 0 \quad . \tag{9.1}$$

Noting that the Green's functions depend on the couplings of the theory which in turn depend on the renormalisation point, this result can be rewritten as

$$\left(\mu\frac{\partial}{\partial\mu} \;+\; \beta(g)\frac{\partial}{\partial g} \;+\; n\gamma(g)\right) G_R^{(n)}(k_i, g, \mu) \;=\; 0 \quad , \tag{9.2}$$

where we have defined

$$\mu\frac{\partial}{\partial\mu}g(\mu) \;=\; \beta(g(\mu)) \quad , \quad \mu\frac{\partial}{\partial\mu}\ln Z(\mu) \;=\; 2\gamma(g(\mu)) \quad . \tag{9.3}$$

$g(\mu)$ is known as the running or renormalised coupling constant (as opposed to the bare coupling constant which appears in the action).

Eq.(9.2) is also called the Callan-Symanzik equation. It is less general than its analogue in critical phenomena; the r.h.s. of the equation has contributions coming from irrelevant operators which contain suppression factors of the cutoff and are usually ignored. It can be formally solved as

$$G^{(n)}(k_i, g, \mu) \;=\; \exp\left(-n\int^g dg' \frac{\gamma(g')}{\beta(g)}\right)\; \Gamma^{(n)}\left[\frac{k_i}{\mu}\exp\left(\int^g \frac{dg'}{\beta(g)}\right)\right] \quad . \tag{9.4}$$

The scale invariance is reflected in the fact that $G^{(n)}$ is not a function of two independent variables k/μ and g, but is just a function of the single combination $\left[\frac{k_i}{\mu}\exp\left(\int^g \frac{dg'}{\beta(g)}\right)\right]$. The explicit dependence of $\Gamma^{(n)}$ on μ cancels with its implicit dependence on μ through g, making $\Gamma^{(n)}$ truly scale independent. The factor in front of $\Gamma^{(n)}$ in Eq.(9.4) takes care of the anomalous dimensions of the operators involved.

Exercise 12: Check by explicit substitution that Eq.(9.4) indeed satisfies the renormalisation group equation, Eq.(9.2).

Traditionally the various renormalisation constants are called wavefunction (or field) renormalisation, mass renormalisation, coupling (or vertex) renormalisation, and so on. There is one more such label, operator renormalisation, that appears when the calculations involve composite operators. Another multiplicative renormalisation constant is introduced to define the scaling behaviour:

$$(G^{(n)}_{\mathcal{O}})_R \;=\; Z^{-n/2} Z_{\mathcal{O}} (G^{(n)}_{\mathcal{O}})_B \quad . \tag{9.5}$$

The departure of the scaling dimension of the operator from that of the fields appearing in it is then given by

$$\mu\frac{\partial}{\partial\mu}\ln Z_{\mathcal{O}} \;=\; \gamma_{\mathcal{O}} \quad . \tag{9.6}$$

$\gamma_{\mathcal{O}}$ is called the anomalous dimension of the operator $\mathcal{O}$. Note that composite operators typically appear in effective field theories (to be discussed later), when heavy fields are eliminated from the theory using operator product expansion.

With the couplings restricted to only a finite dimensional subspace of relevant and marginal operators, it is straightforward to carry to extend the renormalisation group analysis to higher orders. It is no longer necessary to stick only to the linear region around the fixed point, and field theory calculations are typically carried out to several loop orders. $\beta-$ and $\gamma-$functions then look like polynomials in the couplings. These perturbative expansions

Figure 5: The $\beta-$functions for hypothetical field theories: (a) The ultraviolet stable fixed point is at g^*, while the infrared stable fixed point is at $g = 0$; (b) The infrared stable fixed point is at g^*, while the ultraviolet stable fixed point is at $g = 0$.

are typically asymptotic, and cannot be self-consistently used in regions where the couplings become large.

The $\beta-$function represents the rate of flow of the couplings with the logarithm of the regulator scale, and the fixed points of the theory correspond to the zeroes of the $\beta-$function. Let us look at two illustrative examples, assuming that we know the exact $\beta-$functions of these theories. Figure 5a displays a situation where $\beta(g)$ starts out positive for small g and then turns over and becomes negative, crossing the axis at $g = g^*$. We can analyse the behaviour of g near g^* by expanding $\beta(g)$ about g^*. The crucial feature is the sign of $d\beta/dg$. For the case shown in Fig.5a, the coupling g is driven towards g^* with increasing μ, both from above and from below. Such a fixed point is called ultraviolet stable, and a continuum field theory can be defined at $g = g^*$. The point $g = 0$, on the other hand, is infrared stable. It represents a trivial non-interacting theory at long distances.

Figure 5b depicts a situation where $\beta(g)$ starts out negative for small g and then turns over and becomes positive, crossing the axis at $g = g^*$. With increasing μ, the coupling is driven towards zero, and the perturbation expansion becomes better at shorter distances. Such theories are called "asymptotically free", and the $g = 0$ fixed point is ultraviolet stable. The fixed point at $g = g^*$ is an infrared fixed point representing a non-trivial interacting theory at long distances.

A point to note is that field theorists (with their preoccupation to get rid of the regulator) are more interested in the ultraviolet fixed points; while the infrared fixed points representing the long distance behaviour (i.e. diverging correlation length) are of higher interest to condensed matter physicists. An important issue arises regarding the meaning of the theory when the only infrared fixed point happens to be at $g = 0$. Such theories are called "trivial", since they become non-interacting at long distances. Two well-known examples in this category are QED and the ϕ^4 theory at $d = 4$. If we want to keep the coupling finite at some fixed physical distance, then it starts growing with increasing μ and there is no consistent way to take a continuum limit. The only way out is to replace the theory by another one at some high energy scale. We shall see later how the concept of effective field theories helps in making sense out of such situations.

Exercise 13: Consider the ϕ^4 theory, whose $\beta-$function looks like $\beta(\lambda) = b_0\lambda^2+O(\lambda^3)$. Show that the coupling blows up at a finite value for the cutoff, when the higher order corrections are neglected. Find the location of this Landau singularity. What are the possibilities for the theory when higher order terms are included?

10. Lattice Gauge Theories and QCD

Applications of renormalisation group in perturbation theory are useful for studying problems having small couplings. Perturbation theory, more often than not, is found to be asymptotic; still one can hope to make some sense out of it if the first few terms seem to converge reasonably well. Some problems, however, due to their intrinsic complexities, demand a non-perturbative application. A typical situation is that of Quantum Chromodynamics, the theory of strong interactions. In this case, the coupling is asymptotically free at short distances (hence the ultraviolet regulator can be removed smoothly), but it grows at long distances. To deal with such situations, one has to fall back to applying renormalisation group with position or momentum space regularisation.

A position space lattice regulator provides a convenient formulation for a rigorous study of non-perturbative gauge theories. Derivatives for matter fields appearing in the action can be discretised by a variety of schemes. More care is needed to maintain exact gauge invariance of the theory. The vector potential $A_\mu(x)$ are not the suitable variables for this purpose; the appropriate choice is the "gauge connection",

$$U_{x,\mu} \;=\; \exp[ig\int_x^{x+\hat{\mu}} A_\mu(x')dx' \quad . \tag{10.1}$$

For example, a simple choice for the Euclidean functional integral defining QCD is

$$\begin{aligned} Z &= \int [dU]\ [d\overline{\psi}\ d\psi]\ \exp[-S_G - S_F] \quad , \\ S_G &= (6/g^2) \sum_{x,\mu<\nu} [1 - \mathrm{Tr}(U_{x,\mu} U_{x+\mu,\nu} U^\dagger_{x+\nu,\mu} U^\dagger_{x,\nu})] \quad , \\ S_F &= \sum_x \overline{\psi}_x \psi_x - \kappa \sum_{x,\mu} [\overline{\psi}_x (1-\gamma_\mu) U_{x,\mu} \psi_{x+\mu} + \overline{\psi}_{x+\mu} (1+\gamma_\mu) U^\dagger_{x,\mu} \psi_x] \quad . \end{aligned} \tag{10.2}$$

Here the integration variables are defined on a hypercubic space-time grid, the group matrices $U \in SU(3)$, and ψ are the quark fields. Under a gauge transformation, $\{\xi(x) \in SU(3)\}$,

$$\overline{\psi}_x \ \rightarrow\ \overline{\psi}_x \xi(x)^{-1} \quad , \quad \psi_x \ \rightarrow\ \xi(x)\psi_x \quad , \quad U_{x,\mu} \ \rightarrow\ \xi(x) U(x,\mu) \xi(x+\mu)^{-1} \quad . \tag{10.3}$$

The lattice theory does not have the same symmetry properties as the continuum field theory. It is anticipated that the explicitly broken symmetries (e.g. rotational and chiral) would be recovered in the continuum limit (as the lattice spacing a is taken to zero holding the physical scale fixed). On the other hand, demonstration of confinement and spontaneous chiral symmetry breaking is easy in lattice theory at strong coupling. The connection between the desired weak coupling fixed point and trivial strong coupling fixed point has to be left to numerical simulations. Monte Carlo calculations have indeed shown that a smooth connection is possible, and nowadays many QCD results evaluated from first principles come from such calculations.

Exercise 14: Check the gauge invariance of the lattice action, Eq.(10.2), and show that it reproduces the usual continuum action in the limit $a \rightarrow 0$.

The lattice theory does not have an inherent scale; all lattice results come out in units of the spacing a. In other words, only dimensionless quantities, such as mass ratios, are uniquely predicted in lattice calculations. The absolute value of the lattice cutoff a has to be fixed by assigning some dimensionful physical quantity its experimental value, and afterwords the results can be expressed in physical units (say in GeV). The renormalisation group tells us how the regulator scale a is related to the coupling of the theory. For QCD, the fixed point at $g = 0$ is asymptotically free with marginally relevant coupling. In leading order perturbation theory,

$$a \frac{\partial}{\partial a} (\frac{1}{g^2}) \ = \ -\frac{(33 - 2N_f)}{24\pi^2} \ + \ O(g^2) \quad , \tag{10.4}$$

which yields (N_f is the number of quark flavours)

$$\Lambda_{QCD} a \ \underset{a \rightarrow 0}{\approx}\ \exp[-\frac{24\pi^2}{(33-2N_f)g^2}] \quad . \tag{10.5}$$

Implementation of renormalisation group in the lattice formulation allows this mapping to be extended (at least numerically) to non-perturbative but scaling region. Note that the corrections to the $\Lambda-$ parameter in Eq.(10.5) are of the form $(1+O(g^2))$, but the deviations of dimensionless ratios from their fixed point values are of the form $(1+O(a/\xi)) \sim (1+O(e^{-cg^{-2}}))$. Another use of renormalisation group in the problem is to find an "improved" discretisation of the continuum action such that the corrections to scaling (due to irrelevant operators suppressed by factors of a), which are lattice artifacts, are small.

The future objective of lattice QCD is to first verify and refine what we already know about QCD from indirect methods (quark models, perturbation theory, spectral sum rules, large$-N_c$ expansions etc.), and then proceed on to predict unknown parameters and new phenomena. It is the combination of both analytical and numerical techniques that has brought the subject of lattice QCD to the stage it has reached today, and developments on both these fronts are needed to make the results still more subtantiative in future.

There are also non-perturbative problems other than QCD, where lattice technology has made important contributions. The Kondo problem was first "solved" using numerical renormalisation group methods. Many other problems are being studied using lattice formulations these days. These include the electro-weak sector of the standard model (Higgs and Yukawa theories), random surfaces (quantum gravity), correlated electron systems (high T_c superconductivity). The interested reader is refered to the proceedings of the annual lattice field theory meetings for more details.

11. Effective Field Theories

Up to now, we have discussed field theories for which a continuum limit can be defined. The renormalisation group then puts theories with asymptotically free couplings at fixed points on a sound basis. On the other hand, there are situations where no fixed points exist or the couplings are not asymptotically free at the fixed points. It is still useful to talk about these theories in the renormalisation group framework, provided there exists a range of scales where the correlation length is much larger (though not infinitely larger) than other scales in the problem. Such cases are called "effective field theories" and they provide a new point of view about the meaning of renormalisation.

It is clear that we cannot look at such theories only in the subspace of relevant and marginal operators; we must include the irrelevant operators from the beginning. When the cutoff of the theory is in a suitable range, the contribution of irrelevant operators to the

action would be small, and one can hope to treat them in a perturbation theory framework. Given some boundary conditions, i.e. the normalisations of these extra terms at some scale, the renormalisation group then quantifies the range of scales where the effective description can be treated self-consistently using a truncated perturbative expansion.

The general prescription for constructing such theories is to split the fields in to high and low energy components using some cutoff, then perform the functional integral over the high energy components, arriving at an effective action containing only the low energy components. Schematically

$$\phi \;=\; \phi_H(k>\Lambda) + \phi_L(k<\Lambda) \quad , \quad \exp[-S_\Lambda(\phi_L)] \;=\; \int [d\phi_H]\; \exp[-S(\phi_H,\phi_L)] \quad . \qquad (11.1)$$

The cutoff Λ is finite here, and the effective action S_Λ contains all possible operators (including the irrelevant ones) allowed by the symmetries of the problem. How to carry out this procedure in practice is a technical detail, called the operator product expansion, but we won't discuss that here. The leading terms in the expansion of the effective action may be different for different boundary conditions, and one may have to use different perturbative expansions in different circumstances. Such calculations with the effective theory can be organised such that they are accurate upto a known power of the cutoff.

The irrelevant operators are also called "non-renormalisable". Their coefficients in the action grow in size, if one actually attempts to remove the cutoff from the theory (i.e. move in the direction opposite to the renormalisation flow). In the usual renormalisation theories, they are strictly forbidden. But if they are observed to have non-zero coefficients in practice, they actually tell us where the cutoff of the theory must appear. This is important information, the point where the coefficients of irrelevant operators become $O(1)$ is the point where the effective description breaks down. For still higher energies, a new theory must take over the physics, from the low energy effective theory. As we have noted earlier, there exist trivial theories with marginally irrelevant operators. They are considered renormalisable in the textbook sense, but physically their status is not very different from non-renormalisable theories, i.e. they are effective theories and will be replaced by some new theories at high energies. Note that the lattice formulation is useful for studying both fixed-point field theories (where the lattice regulator is ultimately removed by taking the scaling limit), and effective field theories (where the lattice regulator acts as a fixed cutoff beyond which the theory loses its meaning).

Let us look at some examples. The electro-weak sector of the standard model is a $SU(2)_L \times U(1)$ gauge theory, coupled to quarks, leptons and Higgs fields. The theory has an

inherent scale, i.e. the point where the theory undergoes a spontaneous symmetry breakdown. Numerically, it can be taken as $M_W \sim M_Z \sim 100$ GeV. At energies higher than this scale, the theory can be studied with all the fields included. At lower energy scales, however, W, Z and Higgs fields can be integrated out of the theory. What we are left with is the $4-$Fermi weak interaction theory and QED. It is not the case that the full electro-weak theory ceases to be valid at lower energy scales, rather the effective theory is much more convenient to handle at lower energies without any loss of predictive power. There are no marginal or relevant weak interactions in the effective theory, the leading $4-$Fermi term is an irrelevant operator. The consequences of this irrelevant operator are directly observable, since it breaks the flavour symmetries respected by the free field theory. The coefficient of the irrelevant operator, the Fermi constant G_F, is determined by comparing the effective theory results with the full theory ones at an appropriate energy scale. In practice, it is convenient to use dimensional regularisation for performing calculations with both the full and effective theories, and the matching boundary condition is imposed at the energy scale where the heavy fields are integrated out of the theory. This prevents logarithms of large ratios of scales from appearing in the effective description, keeping a perturbative expansion in effective couplings under control. Note that, historically, the observed value of G_F determined the energy scale where one would find W and Z, before their direct experimental observation.

Another interesting case is that of QCD. The theory becomes strongly coupled at lower energies, making it impossible to obtain a simple description of the theory using quark and gluon fields. An effective theory can be obtained, however, if one abandons these fundamental fields in favour of new composite fields, the mesons and the baryons. The resulting theory then is less complicated and still has certain predictive power. The degrees of freedom surviving below the QCD scale (which can be taken as $M_p \sim 4\pi f_\pi \sim 1$ GeV), are those guaranteed by Goldstone's theorem. Their dynamics, which is called current algebra, is governed by the spontanous breakdown of the chiral symmetry, $SU(N_f)_L \times SU(N_f)_R \to SU(N_f)_V$. Since the original theory is strongly coupled, we cannot impose the boundary conditions directly; the couplings of the effective fields are determined empirically. In addition, the ratio of the energies of the effective fields to the QCD scale is not very small, often as large as 0.5, making it necessary to include more irrelevant operators in the effective theory to increase the accuracy of the results. Even with these limitations, one learns many things about the strong interactions from the effective description—that is its merit.

These two examples illustrate crucial points about the way we use field theories to describe nature. May be no theory we have encountered is complete and no fields we have used

are fundamental up to arbitrarily high energies. They may be all effective descriptions. The meaning of renormalisation is different; it is not the artificial "cancellation of infinities" as appears in many textbooks, but much more physical. Renormalisation tells us that the low energy physics depends on the theory at high energies through relevant and marginal couplings, and possibly through some leading irrelevant couplings if one measures small enough effects. Many attempts at grand unification, indeed involve replacing the Standard Model and gravity by another theory, which on aesthetic grounds should be general, symmetrical and unique. The renormalisation principle puts a severe restriction on what we can say about the high energy theory, since it can be probed only through a small number of parameters. Physics often becomes philosophy at this stage, because noone can say with certainty where the effective description would end. One can only conjecture whether nature actually has a fundamental structure, and if so what are the degrees of freedom involved.

There are some hurdles to be overcome, however. Slow evolution of couplings of marginal operators is useful in creating large ratios of scales between various interactions observed in nature. The coefficients of truly relevant operators, however, cannot be easily controlled in the low energy theory. Since they rapidly grow with the renormalisation flow, the only natural scale for them is $O(1)$ in units of the cutoff. Such terms are also called "super-renormalisable". A particularly notorious example is the mass term for the fields. It may be protected by an exact symmetry of the theory, gauge or chiral, and can be exactly zero. But without a protecting symmetry, its natural scale is $M \sim \Lambda$; to produce a much smaller coefficient in the effective low energy theory requires an artificial fine-tuning of the parameters in the original theory. Without fine tuning, the fields are so massive that they cannot be self-consistently kept in the effective description. So we must have a criterion of naturalness, based on some exact or dynamical symmetry in the infrared limit of the theory, to keep the super-renormalisable under control. The Standard Model is not natural in this sense, since a large mass for the Higgs scalar is not forbidden by any symmetry. What happens to it above the weak scale is an open question, which will have to be answered by future accelerators.

Another relevant operator, even more troublesome than the mass term, is the identity operator. We have ignored it so far, on the grounds that the zero-point energy really does not matter in the evaluation of the correlation functions. We can no longer do so, if we want to include the coupling to gravity in the theory. Obviously $\mathbf{1}$ is rather symmetric and hard to forbid, and it is the cosmological constant in the theory of gravity. To suppress it to the degree that has been observed in nature is a difficult problem.

12. Some Examples in Quantum Mechanics

Singular potentials.
2-dim.: $\delta-$function, $1/r^2$

13. Application to Superconductivity

Renormalisation of Coulomb repulsion
Competition between phonon and photon exchanges
Contrast with the exact BCS solution

14. Miscellaneous Topics

Multicritical points
Limit cycles and strange attractors
Fractals and chaos
Turbulence
Kondo model

Acknowledgements: There is no claim for originality in these notes. I have generously borrowed material from many lectures, review articles and books on the subject. I list below the most prominent of them, and apologise for not individually crediting many other contributions to the field.

[1] [2] [3] [4] [5] [6] [7] [8]
[9] [10] [11] [12] [13] [14] [15] [16] [17] [18] [19] [20] [21] [22] [23] [24] [25] [26]

References

[1] K.G. Wilson and J.B. Kogut, *Phys. Rep.* **12C** (1974) 77.

[2] S.-K. Ma, *Modern Theory of Critical Phenomena* (Benjamin, New York 1976).

[3] F. Wegner, in *Phase Transitions and Critical Phenomena*, Vol. VI, eds. C. Domb and M.S. Green (Academic Press, New York, 1976).

[4] E. Brézin, J.C. Le Guillou and J. Zinn-Justin, in *Phase Transitions and Critical Phenomena*, Vol. VI, eds. C. Domb and M.S. Green (Academic Press, New York, 1976).

[5] D.J. Amit, *Field Theory, the Renormalization Group and Critical Phenomena* (World Scientific, Singapore, 19??).

[6] P. Ramond, *Field Theory: A Modern Primer* (Benjamin/Cummings, Reading, Massachusetts, 19??).

[7] J. Collins, *Renormalization* (Cambridge University Press, Cambridge, 1984).

[8] M.E. Fisher, *Scaling, Universality and Renormalization Group Theory*, Lecture Notes in Physics, Vol. 186 (Springer-Verlag, Berlin, 1983).

[9] E. Stückelberg and A. Peterman, *Helv. Phys. Acta* **26** (1953) 499.

[10] M. Gell-Mann and F.E. Low, *Phys. Rev.* **95** (1954) 1300.

[11] N.N. Bogoliubov and D.V. Shirkov, *Dockl. Akad. Nauk. SSSR* **103** (1955) 203; 391.

[12] L.P. Kadanoff, *Physics* **2** (1966) 263.

[13] K.G. Wilson, *Phys. Rev.* **B4** (1971) 3174; 3184.

[14] L.P. Kadanoff *et al.*, *Rev. Mod. Phys.* **39** (1967) 395.

[15] P.G. de Gennes, *Phys. Lett.* **A38** (1972) 339.

[16] K. Wilson, *Phys. Rev.* **D10** (1974) 2445 ;
K. Wilson, in *New phenomena in subnuclear physics*, Erice lectures 1975, ed. A. Zichichi, (Plenum, New York, 1977).

[17] K.G. Wilson, *Rev. Mod. Phys.* **47** (1975) 773.

[18] D.R. Nelson and M.E. Fisher, *Ann. Phys.* **91** (1975) 226.

[19] G. Jona-Lasinio, *Il Nuovo Cimento* **26** (1975) 99.

[20] J. Polchinski, Harvard University preprint HUTP-83/A018 (1983).

[21] J. Polchinski, TASI Lectures (1992), hep-th 9210046.

[22] S.R. White, J.W. Wilkins and K.G. Wilson, *Phys. Rev. Lett.* (1985) ???.

[23] D.V. Shirkov, KEK Report 91-13 (1992).

[24] P. Gosdzinksy and R. Tarrach, University of Barcelona preprint UB-ECM-PF 12/89 (1989).

[25] K.S. Gupta and S.G. Rajeev, University of Rochester preprint UR-1310 (1993).

[26] C. Manuel and R. Tarrach, University of Barcelona preprint UB-ECM-PF 19/93 (1993).