UNIVERSITY OF SOUTHAMPTON
FACULTY OF SCIENCE
DEPARTMENT OF PHYSICS

Statistical Conformal Field Theory & Information Loss in the Renormalisation Group

by

Gari D. Clifford

October 1995

A dissertation submitted in partial fulfillment of the requirements
for the degree of Master of Science (M.Sc.)
in Applied Mathematics and Theoretical Physics
by instructional course.

A great truth is a truth whose opposite is also a great truth.

Neils Bohr.

Acknowledgments

I am especially indebted my supervisor and friend, Dr Nouri Mohammedi. I am grateful to all my lecturers for all their help and advice and I would particularly like to thank Prof. Ken Barnes for creating such an amicable atmosphere in which to pursue our follies in the Theory Group. Many thanks also go to Dr D. Chillingworth, and Dr T.R. Morris, for their helpful suggestions and performing the necessary task of proof reading.

I would like to thank my family, particularly my parents for their support and understanding throughout my studies and Neil, my brother for endless computer advice. Thanks go to Dave and Di for playing host to me during the summer over which this was written.

Finally I must thank all my friends who have filled me with the joy of life and encouraged me in all that I do. In naming people I will inevitably offend those who I forget (though not through lack of care, rather through the undefined manner in which I store and retrieve information) and for this I apologise in advance. For the sake of brevity I will restrict the list to those without whose help the struggle may have been too great - Vicki, Marcus, Steve, Sacha, Nicoletta, Ollie, Treeve, Nouri, Steve, Lee, Ben,...

Gari Clifford

Southampton,

October 1995

Contents

Preface

1 **Introduction** 2

2 **Critical phase transitions** 3

2.1 'Spontaneous' magnetization 4

2.1.1 Critical exponents and correlation length 6

2.1.2 Dimensional dependence and the Ising model 7

3 **Renormalization** 7

3.1 The Renormalization Group 9

3.2 Critical behaviour in the RG 10

4 **Universality and scale invariance** 17

5 **The Conformal Group** 19

5.1 The algebra in two dimensions 19

6 **The generalized quantum case - the conformal anomaly** 22

6.1 The stress tensor as the generator of scaling 25

6.2 The trace anomaly 28

6.3 The operator product expansion 29

6.4 Virasoro algebra and Fourier mode expansion 32

6.5 The free boson example 34

6.5.1 The free fermion 36

7 The continuum limit **37**

8 Zomolodchikov's C-theorem **40**

9 Introduction to Information Theory **43**

10 Signal generation and transmission **43**

10.1 The discrete noiseless channel 44

10.1.1 The information source 45

11 Choice, capacity, uncertainty and entropy **46**

11.1 Relative entropy and redundancy 48

11.2 The discrete channel with noise 49

11.3 Continuous Information 53

11.4 The entropy of a continuous distribution 53

11.4.1 The capacity of a continuous channel 55

11.5 A coordinate transformation example 57

12 Conclusions & Discussion **58**

Appendix60 References61

List of Figures

1 Dependence of magnetization M, on temperature T, in an ideal ferromagnet domain. 5

2 Graphs of QCD β-function and coupling constant behaviour at different energy scales. 12

3 Renormalized QED interaction strength behaviour under energy scaling. 13

4 The critical surface in (renormalized) coupling space. 15

5 General RG transformations. (a) corresponds to a scale transformation, (b) to a conformal transformation and (c) to a general coordinate transformation 17

6 Evaluation of the commutator on the conformal plane. 33

7 The equivocation possible for a given input entropy to a channel. 52

Preface

At critical phase transition points a systems correlation lengths become infinite. Since these tell us how the system behaves under scaling, we can infer that the system is unaffected by small scale changes and the only relevant perturbations are global or large scale ones. At criticality the values of observables fluctuate on all length scales and thus any accurate model must be invariant under scaling. It is therefore consistent that conformally invariant quantum field theories turn out to be useful descriptors of critical behaviour and in fact such critical theories are invariant under the full conformal group. However, it is only in restricting the number of dimensions to two that we gain more information other than simple scale invariance, since the conformal algebra becomes infinitely dimensional, leading to increased restrictions on the theory.

Statistical mechanical models of phase transitions at critical points are also scale (conformally) invariant at large distances with corresponding infinite correlation lengths. This, coupled with its deep connections to many different areas of mathematics and physics (including string theory), has led to increased attention of two dimensional conformal field theory in recent years.

This dissertation is intended to give an overview of these subjects and their connections, particular in the light of entropy and information theory.

Critical phase transitions ... *or do you not know that between two opposites there exists a mean.*

Plato.

1 Introduction

In this chapter I shall introduce some of the connections of QFT with statistical physics. In particular, since both theories can be formulated in the context of a finite spaced lattice, one would expect them to share the quality of information loss that is inherent in this *coarse graining*. However, this important concept which occurs in many areas of physics will not be adequately explored until the end of chapter 2 since we must highlight the considerable background reading needed in order to place such concepts into context. The notation will follow the standard texts as much as possible and I shall assume the reader is familiar with the basics of quantum field theory (QFT), tensor calculus, group theory and thermodynamics (see [6, 8, 14, 23, 27]).

One phenomenon of physics that is increasingly being recognized as a useful modeling tool is that of *scale invariance*. Many physical problems are made tractable by theories that concentrate on a small restricted range of parameter (such as length, time, and mass) scales around some characteristic point. For example, a theoretical model of galactic motions ignores the interactions of the constituent planets since there are far too many to make a calculation feasible.

This is essentially the method of statistical mechanics where the only important degrees of freedom are those that characterize the system as a whole such as free energy, temperature, entropy and these tend to be present at a certain scale - thermodynamics is not relevant for each particle. The 10^{20} or more internal degrees of freedom resulting from the properties of the constituents are effectively ignored. In these types of systems an effective Hamiltonian is the descriptor and it

changes as attention is moved from one characteristic scale to another.

There are many systems that do not exhibit this scale dependence and their Hamiltonians remain invariant under scale changes. Examples of this range from economic predictions and works of art to data storage, certain QFT's and phase transitions [24]. In this dissertation we are exploring the latter two since both theories often utilize the powerful techniques of renormalization and lattice formulation. It is therefore not surprising that connections abound. In particular, the mean field theories break down near the critical points of these models and the best description turns out to be via scale invariant power laws.

2 Critical phase transitions

In studying macroscopic systems such as crystals using statistical mechanics, we find that observables have critical values that characterize the discrete transition of the system between different configurations or phases with an accompanying change in internal symmetry. An example of this is the transition of water from a solid to a liquid state at a particular temperature and pressure - its melting point. These critical values define the critical points of the relevant theory.

There is a (statistical) physical explanation for the succession of different temperature dependent phase changes a system goes through, apparent in the free energy $F = U - TS$, which is minimized by the state of thermodynamic equilibrium. U is the internal energy and S is the *entropy* (a measurement of disorder since it is defined to be proportional to the logarithm of the number of ways of arranging the system: $S = k_B \log W$. There is trade off between U, which favours order, and S which favours disorder; a rise in temperature shifts the balance towards in-

creased entropy and hence disorder. It is therefore likely that at a certain temperature the entropy will 'avalanche' and a sudden change of order or phase will become apparent in the system.

The mean field theory describing behaviour away from the critical point assumes that the internal field acting at each point in the domain is the same, and equal to the mean internal field for the whole domain. However, in regions near the critical point the fluctuations in order are no longer small and the interactions producing the internal field become long range, averaging over the states of many sites. At critical phase transition points there is participation from all the degrees of freedom and thus fluctuations on all scales are equally relevant - the correlation length becomes infinite. The assumption that the internal field is related to the mean degree of order is no longer valid and a new mathematical description is needed.

The phase transition between different crystal configurations involves a sudden rearrangement of the crystal lattice and the state of the body changes in a discontinuous manner leading to a change of symmetry. We therefore expect a sudden change in gradient of any function describing an observable[1] near a critical point. The powers that the parameters in the theory are raised to very near this point are known as the *critical exponents*. They express the functional behaviour of quantities as fractional power laws.

2.1 'Spontaneous' magnetization

For example, magnetism in metallic materials is explicable in terms of a model involving cooperative alignment of the microscopic magnetic moments (spins) of electrons with the mediating

[1] Usually described by a second order differential equation, hence the term *second order* phase transition.

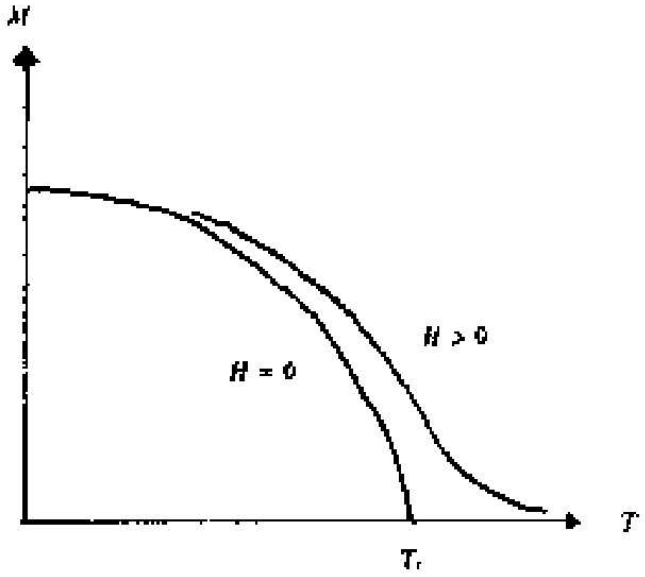

Figure 1: Dependence of magnetization M, on temperature T, in an ideal ferromagnet domain.

force involving a quantum mechanical description[2]. The strength of this coupling is temperature dependent and the magnetization drops to zero in the absence of an applied magnetic field, H, as we approach a critical temperature, the Curie temperature T_c, from below (as in fig. 1).

The loss of magnetization is due to the reduction in the alignment of the spins from thermal noise and not from the electrons' microscopic magnetic moments disappearing. As we lower the system through the critical point ($T = T_c,\ H = 0$) a (second order) phase transition occurs between the paramagnetic ($T > T_c$) and ferromagnetic ($T < T_c$) states. It is characterized by the absence of latent heat and the continuous growth from zero of an order parameter (the magnetization M in this case) and this corresponds to a breaking of a symmetry [3]. In the high T phase the system

[2]The alignment actually only occurs within the domains, formed to minimize the dipole interaction between the spins. An external magnetic field H must be applied to align the domains and magnetize the material.

[3]Which is that of rotations in this example.

is invariant under rotations about all three axes, whilst in the low T phase it remains invariant only under rotations around the axis of spontaneous magnetization.

2.1.1 Critical exponents and correlation length

Consider a paramagnetic material at very high temperatures where the electron spins essentially act independently of each other. As the temperature is lowered the correlation between neighbouring spins increases as the coupling begins to dominate the 'noise'. The range of this *cooperation* is thus know as the correlation length[4], ξ, (although in the absence of an applied field there is no magnetization) and as $T \to T_c$ it becomes infinite[5].

The magnetization curve in fig 1 can be described close to T_c by

$$M \propto (T_c - T)^p \qquad (H = 0,\ T < T_c) \tag{1}$$

where p is an example of a critical exponent. These parameters are extremely useful for describing a systems behaviour near critical points. In particular the correlation length's approach to infinity is described by the exponent ρ in

$$\xi \propto (T - T_c)^{-\rho} \qquad (H = 0,\ T > T_c) \tag{2}$$

as $T \to T_c$. It is important to note that the non-integer nature of the exponents implies that ξ and M (and in fact the thermodynamic quantities as it turns out) do not depend analytically upon their variables. The resulting singular behavior means that perturbation theory is ineffective.

[4]Which therefore increases as the temperature drops.

[5]It is possible to draw an analogy here with the phenomenon of critical opalescence in liquid-gas systems which arise close to the critical point with large correlated functions.

2.1.2 Dimensional dependence and the Ising model

A typical model of this magnetic behaviour is to use a lattice of binary (up or down) spin sites with a nearest neighbour interaction coupling. In a one dimensional system such as a line of spins there is never true long range order and thus no phase transition. This is because the change from a spin up to a spin down has a finite probability anywhere on the line. In two or more dimensions the dominant region tends to surround the minority region producing a shrinking effect upon them. It has been shown that in four or more dimensions the pressure from the dominant phase would be so strong that even near the critical point the fluctuations would be negligible and the mean field theory would still be valid.

Two and three dimensional systems lie between these two limits but little progress has been made on exact solutions for the latter. However, the two dimensional (Ising) model was solved by Onsager in 1940 [29] and this provided both qualitative and quantitative descriptions of phase transitions. The exponents in (1) and (2) turn out to be $p = 1/8$ and $\rho = 7/4$, in close agreement to many experimental values. The mean field predictions, on the other hand, are $1/2$ and 1 respectively.

3 Renormalization

In the phase transition from a paramagnetic state to a ferromagnetic one a difficult theoretical problem arises, in as much as all fluctuations within a magnetic system with a range between the lattice spacing and the correlation length become important. We have already seen that as the system approaches its critical point, the correlation length becomes macroscopic. So, although

each spin only interacts directly with locally positioned spins, the phase transition occurs through the cooperative overall group effect of the interactions between a large number of spin sites. We cannot therefore directly use approximations which only take into account the interactions between a few spins. One approach particularly suited to tackling this kind of problem is through utilizing the renormalization group (RG), first formulated for use in relativistic quantum mechanics by Gell-Mann and Low (1954) and Stueckelberg and Petermann (1953) ([10]) to remove divergent integrals in the Green function calculations for processes beyond tree level.

These divergences result in infinite S-matrix elements and thus infinite probability amplitudes for the corresponding process. If the path integral is regularized in some way, the divergent parts of the path integral may be separated and then removed by a renormalization procedure. One method of regularizing is to place the QFT on a lattice possessing a microscopic spacing a, which acts as a low energy cut-off. The renormalized operators[6] $\phi(r)$ are related to the lattice operators $\phi_{lat}(r)$, by a function of the lattice spacing and the scaling dimension Δ (the exponential factor by which the function changes under scale transformations) such that

$$\phi(r) \sim a^{\Delta}\phi_{lat}(r). \tag{3}$$

The lattice correlation function therefore asymptotically tends towards the renormalized one for distances much greater than a. In the continuum limit these functions contain no dependence on lattice spacing. It is therefore possible to infer that the two-point Green function of a renormalized operator which is scalar under rotations has a power law behaviour of the form

$$< \phi(r)\phi(0) >= |r|^{-2\Delta} \tag{4}$$

[6]Although the local renormalized fields are referred to as scaling operators, they only become operators when we leave Euclidean space in favour of a Hilbert space formalism.

and this correlation function's scaling is only relevant on large distance scales. This corresponds to the required behaviour for analyzing critical phase transitions.

3.1 The Renormalization Group

The RG approach to many-body problems such as the critical magnetization phenomenon is to set up a transformation which is applied many times successively to remove the effect of fluctuations of increasingly larger sizes at the expense of renormalizing (redefining) the effective interactions[7]. There are many methods for RG formulation, but all contain the main concept of reducing the number of spins[8] in a region the size of the correlation length. At each successive step several neighbouring spins are eliminated and replaced by a new single (meta-) spin which reproduces the effect of the old *community* of spins. The process of decimation thus removes a fraction of the degrees of freedom and consequently creates a new effective Hamiltonian for the remaining degrees of freedom at each successive step. This coarse graining effect is the crucial concept in this discussions since it is the source of the effects and connections that we will be considering.

After a certain number of steps the system possesses a correlation length to spin separation ratio ξ/a, that is not large. The system is therefore far away from its critical point and any suitable perturbation theory is permissible. Since the new meta-spin must simulate the role of all the old spins, the new Hamiltonian must be specially chosen to include interaction coupling parameters which simulate the old ones. These coupling constants, λ_i, therefore depend on the ones which

[7] i.e. the strength of the coupling, λ, between the lattice sites.

[8] As in the case of the magnetic spin (Ising) model, although the observable may be a different quantity, such as the density of spins as in the QFT case.

parameterized the old system. They are redefined l times by the recursion relation

$$\lambda_i^{(l)} = f_i(\lambda^{(l-1)})$$

which upon iteration thins the number of spins and generates a new Hamiltonian at each step. As l increases, the systems described by $\{\lambda_i^{(l)}\}$ become increasingly further from the critical point since a region the size of the correlation length contains fewer and fewer spin sites.

In practice it is not possible to calculate the functions f_i since the original Hamiltonian cannot be exactly solved and some approximation scheme must be used. It is important to note that these functions are regular in the critical region, in contrast to the singular nature of thermodynamic quantities and thus approximation schemes that fail in thermodynamics are potentially useful here.

3.2 Critical behaviour in the RG

The renormalization group trajectories are created by considering the influence of a RG transformation upon a point. It is assumed that there exists a single-parameter group of motions in the space Q of coupling constants λ, s.t. $B_t Q \to Q$, of such a nature that a field theory describable by an action $S'(B_t\lambda, e^t a)$ is equivalent to the original action $S(\lambda, a)$ in the sense that all the correlation functions of the two theories agree at scales $x \gg e^t a$ for $(t > 0)$. The components of the vector field which generate the renormalization group are called β-functions where $\beta^i(\lambda) = \frac{d\lambda^i}{dt}$.

When dimensional regularization is used to absorb the divergent integrals present in loop expansions the redefinition of the coupling constants involves introducing a dimensionful mass parameter μ in order to keep the coupling constants dimensionless and preserve renormalizability. The renormalized couplings, denoted by a prime, become $\lambda' = \mu^{d-\frac{(d-2)n}{2}}\lambda$ where n is the number

of fields at a vertex which become irrelevent when the number of dimensions ,d, are reduced to two. The β-function can then be written as $\beta = \mu \frac{d\lambda}{d\mu}$. The appearance of this mass parameter is due to the breaking of scale invariance by the regulator[9]. The freedom to choose μ at intermediate steps allows us to explore the scaling properties of Green functions; since they scale as $\Gamma^n(tp_i, tm, t\mu) = t^{4-n}\Gamma^n(p_i; m)$ in four dimensions, with momenta p_i and mass m, taking derivatives with respect to t and removing the μ-derivative leads to the renormalization group equation [14]

$$\left(-t\frac{\partial}{\partial t} + \beta(\lambda)\frac{\partial}{\partial \lambda} + [\gamma_m(\lambda) - 1]m\frac{\partial}{\partial m} - n\gamma_d(\lambda) + 4 - n\right)\Gamma^n(tp_i, \lambda, m) = 0 \tag{5}$$

where we have defined

$$\begin{aligned}
\beta(\lambda, \epsilon) &= \mu \left.\frac{d\lambda}{d\mu}\right|_{m_0, \lambda_0, \epsilon} \text{....the } \beta\text{−function,} \\
\gamma_d(\lambda) &= \left.\frac{\mu}{2Z_\phi}\frac{dZ_\phi}{d\mu}\right|_{m_0, \lambda_0, \epsilon} \text{....the anomalous dimension of field,} \\
\gamma_m(\lambda) &= \left.\frac{\mu}{m}\frac{dm}{d\mu}\right|_{m_0, \lambda_0, \epsilon} \text{....the anomalous dimension of mass}
\end{aligned}$$

and Z_ϕ is the wavefunction renormalization constant. The RG flow equation (5) describes the evolution of the Green functions as their external momenta are scaled, the solution of which is

$$\Gamma^n(tp_i, \lambda, m; \mu) = \left[t^{4-n}e^{-n\int_1^t \frac{ds}{s}\gamma_d(\lambda'(s))}\right]\Gamma^n(tp_i, \lambda'(t), m'(t); \mu). \tag{6}$$

Thus the action of scaling the momentum by t is equivalent to changing the coupling (from λ to λ'). In general the anomalous dimensions will not be constants except near a zero of the β-function where $\frac{\partial \lambda'}{\partial t} = \beta(\lambda') = 0$, i.e. $\beta^i(\lambda_*) = 0$ and thus the zeros are the fixed (critical) points of the renormalization group. A fixed point is defined to be ultra-violet (UV) stable if a non-zero first derivative of its generating β-function is negative since you are always driven to it as the

[9]Or the finite cut-off.

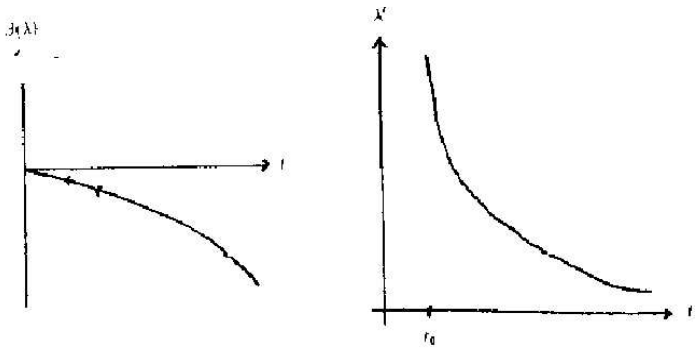

Figure 2: Graphs of QCD β-function and coupling constant behaviour at different energy scales.

momentum scaling increases ($t \to \infty$) such as for QCD (fig. 2). Conversely, it is infra-red (IR) stable for a positive first derivative since you are always driven to this point as $t \to 0$.

Furthermore, since the renormalized coupling may increase or decrease with t (i.e. as we scale to higher momentum and energy), the validity of perturbation theory may depend on the momentum range we are probing. For example, in quantum chromodynamics (QCD), $\lambda'(t) \to 0$ as $t \to \infty$ (fig. 2) and perturbative expansions in the coupling are not reliable for relatively low momenta. Consequently, QCD is a high energy effective theory and perturbation theory is valid only at short distances (small λ) giving it the quality of asymptotic freedom (quark confinement). In contrast to this, if we study fig. 3 we see quantum electro-dynamics (QED) is a low energy model since $\lambda'(t) \to \infty$ as $t \to t_0$, where t_0 is the Landau pole signifying the breakdown point of QED.

However, all physical quantities such as the Green functions are independent of the mass parameter μ and are therefore RG invariant. It does not matter how you renormalize the theory, the final quantifiable answers should always be the same. Thus a transformation in coupling space

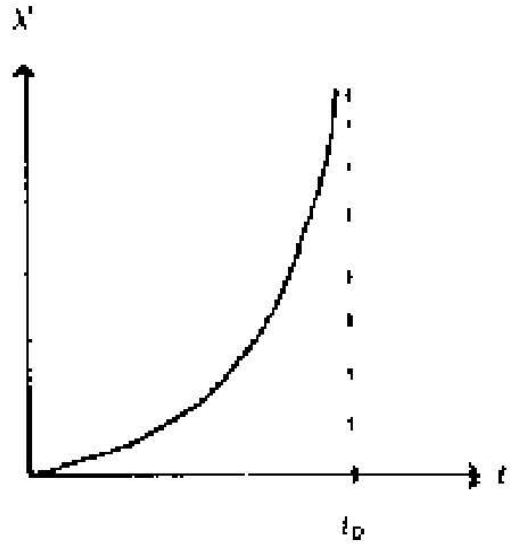

Figure 3: Renormalized QED interaction strength behaviour under energy scaling.

(a renormalization) should not cause any difference in the results, it simply causes a movement along a RG trajectory. Consider fig. 4, a plot of the numerically computed solutions[10] to the exact (Wilson) renormalization group flow equations[11],

$$\begin{aligned} \frac{\partial\sigma}{\partial t} - 2\sigma &= \frac{\gamma_4}{1+\sigma}, \\ \frac{\partial\gamma_4}{\partial t} &= \frac{\gamma_6}{1+\sigma} - 3\frac{\gamma_4^2}{(1+\sigma)^2}, \\ \frac{\partial\gamma_6}{\partial t} + 2\gamma_6 &= -15\frac{\gamma_4\gamma_6}{(1+\sigma)^2} + 30\frac{\gamma_4^3}{(1+\sigma)^3}, \end{aligned} \tag{7}$$

calculated to two loops for a single component scalar field. The one particle irriducible (1PI) self energy is

$$\sigma(t) = \frac{m^2}{\Lambda^2} = -\frac{\alpha}{2} + \frac{\alpha^2}{2},$$

and is thus scaled by Λ (the low energy (IR) cut-off scale). γ_4 is the four point function and is equal to the coupling α, whilst the six point function γ_6, turns out to be

$$g_6\Lambda^2 = \gamma_6 = 15\alpha^3 - \frac{45}{2}\alpha^4$$

giving a β-function of

$$\beta(\alpha) = \frac{\partial\alpha}{\partial t} = -\Lambda\frac{\partial\alpha}{\partial\Lambda} = -3\alpha^2 + 12\alpha^3.$$

At the critical point ($t = 0$) $\Lambda = \Lambda_0$, the UV cut-off, (although it is usually far less than this high energy bound) and we have a free theory. Furthermore, $\gamma_6 = 0$, $\gamma_4 = \alpha = \alpha_0 = \lambda_0/(4\pi)^2$ (the bare coupling) and $\sigma = \sigma_0 = m_0^2/\Lambda_0^2$, the scaled bare mass. $\Lambda \to 0$ corresponds to increased energy scaling ($t \to \infty$), so as we move out to higher energy scales we move away from the gaussian fixed

[10] Supplied by T. R. Morris.

[11] Which were derived from the truncated Wilson action $S = \frac{1}{2}(\partial_\mu\phi)^2 + \frac{1}{2}m^2\phi^2 + \lambda\phi^4/4! + g_6\phi^6/6!$. Usually there are infinitely many couplings, but this renders the computation impossible.

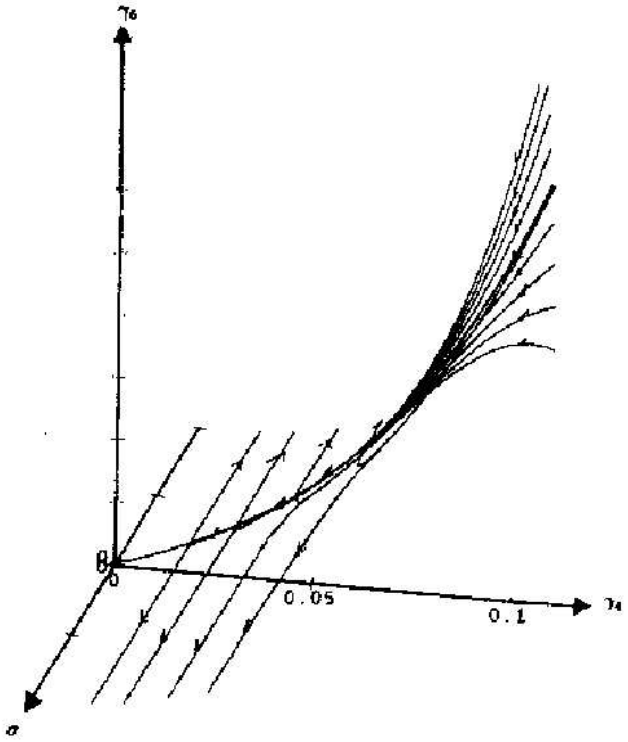

Figure 4: The critical surface in (renormalized) coupling space.
σ is the mass, γ_4 positive dimensional (relevant) couplings and γ_6 negative dimensional (irrelevant) couplings.

point and a symmetry breaking is induced. It is in fact the aquisition of mass since all but one of the trajectories[12] lie in the $\sigma \neq 0$ region.

The RG trajectories in the coupling space have the physical interpretation of being the change in the Green functions as their cut-off is lowered. Furthermore they can be regarded as the flows of the Wilsonian effective action[13] as the effective UV energy cut-off is lowered, integrating out more energy scales in such a way as to leave the physics below this cut-off unaltered. So if we reset the origin to some finite point on the renormalization trajectory it is equivalent to defining a *perfect action* - a special form of the Wilsonian effective action with no cut-off effects (even though μ is finite) because the flow through this Wilsonian action just reproduces the continuum limit of the renormalized Green functions. At the end of chapter 2, when we consider a general result for RG flows, we will see that the flows induced by renormalization necesarily end up at critical points of the RG.

The critical behaviour of a QFT is governed by the fixed point(s) of the RG transformation and it can be used to predict the phase transition temperature, scaling behaviour and the values of the critical exponents. These ideas may be applicable in any circumstance where the interactions are formulated at one length scale and cooperative fluctuations are important over a whole range of length scales.

[12] Which can thus be considered that of a theory describing a photon-like object.
[13] Which gives the same physics as the bare action below some finite energy cut-off.

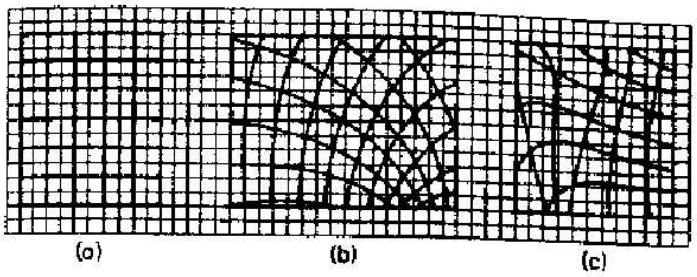

Figure 5: General RG transformations. (a) corresponds to a scale transformation, (b) to a conformal transformation and (c) to a general coordinate transformation

4 Universality and scale invariance

Remarkably critical exponents are not simply arbitrary. There are groups of systems whose exponents are (almost) identical and systems undergoing varied transition processes can be assigned to a small number of classes each characterized by a certain set of values for the exponents. Often the relations are as simple as the addition of a few exponents being equal to a small integer and are called scaling laws. In the case of the magnetic phase transition the relations (1) and (2) exhibit scaling behaviour since, for example, the scale change $(T - T_c) \to b(T - T_c),\ M \to b^p M$ leaves (1) unchanged.

The existence of a microscopic length a to facilitate regularization of a QFT produces a non-trivial scaling behaviour. Under a dilation (transformation) $r \to r' = b^{-1}r$, the power law behaviour (4) of the correlation functions corresponds to the transformation law

$$\phi(r) \to b^{\Delta}\phi(r').$$

If we imagine the lattice imposing a coordinate grid upon the plane. An RG transformation

then leads to an increase in the coordinate grid or lattice spacing from a to ba and the short-wavelength degrees of freedom are integrated out of the partition function so as to leave the large distance physics unchanged. Once again we are utilizing a coarse graining procedure to deal with infinities that would otherwise occur. The action, S, is modified to $S + dS$ unless it lies at a RG fixed point S^*. (At a RG fixed point the action remains invariant and we may renormalize using scale transformations without changing the Green functions). Thus a coordinate change is not just a simple relabeling of the reference points in the plane but actually has a physical effect. Consequently the coarse graining also has a physical implication, which turns out to be a restriction of the direction of flow of the RG trajectories. Fig. 5 demonstrates how the new (transformed) lattice gives a grid with respect to which the coordinates r' of a given point are defined.

The concept that there are universal sets of critical exponents which apply to certain critical points in systems involving different physics is known as universality and there exists a whole hierarchy of degrees of this. For example, the critical temperature depends upon every minor detail, whilst the critical exponents are pure numbers which depend only on the most general properties of the system such as the spatial and order parameter dimensionality, the range of the interactions and the underlying symmetry - they are independent of the nature of the transition and the detailed features of the interaction.

Conformal field theory and its connection to statistical physics *It may be said of me that I have lost my sense of scale, but never that I have lost my sense of proportion.*

Scale, Will Self.

5 The Conformal Group

It has been shown [20] that local scale (i.e. conformal) invariance can be used to construct theories near a critical point and so it is my aim in this chapter is to provide an insight into the important connections between statistical mechanics and conformal field theory, However, the highly theoretical nature of these subjects necessitates a thorough grounding in the relevant techniques and I will begin by defining the key concepts, mostly with reference to two-dimensional cases as these will prove the most useful and applicable. I refer the reader to appendix A for a list of formal definitions.

5.1 The algebra in two dimensions

Consider a d-dimensional Riemann manifold upon which a metric $g_{\mu\nu}$ is defined. If the metric is flat it is equal to $= \eta^{\mu\nu}$, a matrix with the only non-zero elements being ± 1 on the diagonal[14]. The invariant interval between two points on the manifold (line element) is given by

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu. \tag{8}$$

A coordinate transformation (a change of coordinates $x \to x'$) generates a change in the metric tensor

$$g_{\mu\nu} \to g'_{\mu\nu}(x') = \frac{\partial x^\alpha}{\partial x'^\mu} \frac{\partial x^\beta}{\partial x'^\nu} g_{\alpha\beta}(x). \tag{9}$$

[14] In Euclidean geometry $\eta_{\mu\nu} = \delta_{\mu\nu}$.

The *conformal group* is defined to be the subgroup of transformations that leaves the metric unchanged apart from a non-zero differentiable scaling function $\Omega(x)$ such that

$$g_{\mu\nu} \to g'_{\mu\nu}(x') = \Omega(x) g_{\mu\nu}. \tag{10}$$

These metrics are said to be conformally related; the angle between vectors and the ratios of magnitudes of vectors is the same on both manifolds, and the geodesics of both metrics coincide. However, since the lengths of vectors differ, this transformation corresponds to a rescaling. The Poincaré group[15] leaves the metric invariant and is thus a subgroup of the conformal group with $\Omega = 1$. In $d = 2$ the Reimann-Christoffel or curvature[16] tensor, $R^{\mu}_{\nu\vartheta\varphi}$ has one independent component, the Ricci scalar $R = g^{\nu\varphi} g^{\vartheta\mu} R_{\mu\nu\vartheta\varphi}$. The infinitesimal generators of the conformal group can be determined by analysis of an infinitesimal coordinate transformation $x^{\mu} \to x^{\mu} + \epsilon^{\mu}$, which causes a line element change

$$ds^2 \to ds^2 + (\partial_{\mu}\epsilon_{\nu} + \partial_{\nu}\epsilon_{\mu}) dx^{\mu} dx^{\nu}. \tag{11}$$

It can therefore be seen that (8) and the conformal transformation condition (10) necessitate

$$\eta_{\mu\nu}(\Omega(x) - 1) dx^{\mu} dx^{\nu} = (\partial_{\mu}\epsilon_{\nu} + \partial_{\nu}\epsilon_{\mu}) dx^{\mu} dx^{\nu}, \tag{12}$$

so $(\partial_{\mu}\epsilon_{\nu} + \partial_{\nu}\epsilon_{\mu})$ must be proportional to $\eta_{\mu\nu}$. Tracing with the flat metric gives $d \propto 2\partial \cdot \epsilon$ and hence the constant of proportionality as $\frac{2}{d}(\partial \cdot \epsilon)$. The scaling function is therefore $\Omega(x) = 1 + (\frac{2}{d})(\partial \cdot \epsilon)$ which leads to

$$(\eta_{\mu\nu} \Box + (d-2)\partial_{\mu}\partial_{\nu}) \partial \cdot \epsilon = 0. \tag{13}$$

Dimensional restriction to $d = 2$ leads to increased information about the transformation properties of the theory and so it is worth noting that in two dimensions (12) reduces to the the Cauchy-

[15] The semidirect product of translations and the Lorentz transformations of flat space.
[16] $R^{\mu}_{\nu\vartheta\varphi}$ tells us in what manner the manifold curves.

Riemann equation

$$\partial_1\epsilon_1 = \partial_2\epsilon_2, \qquad \partial_1\epsilon_2 = -\partial_2\epsilon_1. \tag{14}$$

If we then use the complex coordinates $z, \bar{z} = x^1 \pm ix^2, z = x + iy$ and write the infinitesimal change as $\epsilon(z) = \epsilon^1 + i\epsilon^2$ and $\bar{\epsilon}(\bar{z}) = \epsilon^1 - i\epsilon^2$ then the conformal transformations coincide with the coordinate transformations[17]

$$z \to f(z), \qquad \bar{z} \to \bar{f}(\bar{z}). \tag{15}$$

These are the generators of the local[18] conformal algebra with Euclidean line element

$$ds^2 = dx^2 + dy^2 = dzd\bar{z} \to \left|\frac{\partial f}{\partial z}\right|^2 dzd\bar{z} = \Omega dzd\bar{z}. \tag{16}$$

If we take the transformed (primed) coordinates

$$z \to z' = z + \epsilon_n(z), \qquad \bar{z} \to \bar{z}' = \bar{z} + \bar{\epsilon}_n(\bar{z}) \qquad (n \in \mathbf{Z})$$

as a basis, with z small such that

$$\epsilon_n(z) = -z^{n+1} \quad \text{and} \quad \bar{\epsilon}_n(\bar{z}) = -\bar{z}^{n+1},$$

then infinitesimal generators are

$$L_n = -z^{n+1}\partial_z \quad \text{and} \quad \bar{L}_n = -\bar{z}^{n+1}\partial_{\bar{z}}. \tag{17}$$

These satisfy the commutation relations

$$\begin{aligned} [L_m, L_n] &= (m-n)L_{m+n}, \\ [\bar{L}_m, \bar{L}_n] &= (m-n)\bar{L}_{m+n}, \\ [L_m, \bar{L}_n] &= 0 \end{aligned} \tag{18}$$

[17] These have an infinitely dimensional local algebra.

[18] The distinction between local and global conformal algebras only occurs in two dimensions. In $d > 2$ the local one doesn't exist.

and the Jacobi Identity,

$$[L_l,[L_m,L_n]]+[L_n,[L_l,L_m]]+[L_m,[L_n,L_l]]=0. \tag{19}$$

In two dimensions the group of conformal transformations is infinitely dimensional since any function mapping the complex plane to itself is conformal. (18) therefore constitutes an infinite Lie algebra[19]. The latter condition that the ordinary and barred generators commute shows that we can consider the local conformal algebra as a direct sum of two independent isomorphic[20] subalgebras A and $\bar{A}$. The action of the $d=2$ conformal group factorizes into independent actions on z and $\bar{z}$ and the Green functions of the field theory are such that z and $\bar{z}$ can be treated as independent variables.

6 The generalized quantum case - the conformal anomaly

When we consider the involvement of QFT we must look at the action of this algebra on a Hilbert space. In general, the fields in a conformal field theory can be grouped into families $\{\Phi_n\}$ containing one primary field Φ defined by the two dimensional generalization of the transformation law (9)

$$\Phi(z,\bar{z})\rightarrow\left(\frac{\partial f}{\partial z}\right)^h\left(\frac{\partial f}{\partial \bar{z}}\right)^{\bar{h}}\Phi(f(z),f(\bar{z})), \tag{20}$$

and an infinite set of secondary fields (including the derivative of the primary) called descendants (of the primary field). Φ is a covariant tensor of rank $h+\bar{h}$ with conformal (highest) weights h and $\bar{h}$ [21]. These are the irreducible representations of the conformal group where the primary is the highest weight of the representation.

[19] A good text on this subject can be found in volume 2 of [6]

[20] With a one to one mapping between their elements.

[21] Independent real eigenvalues, so $\bar{h}$ is not the complex conjugate of h. It is a separate weight born out of the process of using complex coordinates.

In Dirac notation we have $|h>=\Phi(0)|0>$, a highest weight state $|h>$ satisfying

$$L_0|h>=h|h>$$

created by acting with a primary field Φ on the SL(2,R) invariant vacuum $|0>$ satisfying

$$L_n|0>=0$$

for $n \geq -1$. Descendent states are created by successively acting on $|h>$ with a series of L_{-n_i}'s where $n_i \geq 0$. These states can be regarded as the result of a descendent field acting on the vacuum:

$$L_{-n}|h>=L_{-n}(\phi(0)|0>)=(L_{-n}\phi)(0)|0>=\phi^{(-n)}(0)|0>\ .$$

Every consistent representation of the this algebra is characterized by such a highest weight state. For example, consider operators of a fixed conformal weight[22] acting on an arbitrary eigenstate $|\Psi>$ such that $L_0|\Psi>=h_\Psi|\Psi>$. Since $[L_0, L_n]=-nL_n$, then

$$L_0L_n|\Psi>=(h_\Psi-n)L_n|\Psi>$$

and L_n lowers the eigenvalue of the state upon which L_0 acted (for $n>0$). Dilation on the complex plane implies that L_0 and $\bar{L}_0$ should be (separately) bounded from below in any consistent QFT. By acting with L_n's we will ultimately reach a state annihilated by L_n (for $n>0$) called the highest weight or primary state, $|h>$. We can therefore regard the L_n's as an infinite number of harmonic oscillator annihilation operators and the $L_n^\dagger = L_{-n}$'s as creation operators. If we draw an analogy between L_0 and the angular momentum operator J_3 and also between the infinite number of $L_{\pm n}$'s and (the finite number of) lowering and raising operators, $J_\mp$'s, we see that the above representation resembles that of SU(2).

[22]Since these are scaling operators which we shall be most concerned with.

Since we have bounded states we must introduce normal ordering into our operators to prevent contradictions such as annihilation on the vacuum. This is a type of coarse graining regularization which causes anomalous additions to the group algebra (18). The modifications turn out to be

$$\begin{aligned} [L_n, L_m] &= (n-m)L_{n+m} + \frac{c}{12}(n^3 - n)\delta_{n+m,0}, \\ [\bar{L}_n, \bar{L}_m] &= (n-m)\bar{L}_{n+m} + \frac{\bar{c}}{12}(n^3 - n)\delta_{n+m,0}, \\ [L_n, \bar{L}_m] &= 0. \end{aligned} \tag{21}$$

where c is known as the *conformal anomaly*, or the *central charge* and the resulting quantum algebra is two copies of an infinitely dimensional group known as a *Virasoro* algebra[23]. Every conformally invariant QFT has a particular representation of this algebra with some value for c and $\bar{c}$.

The global conformal algebra (SL(2,C) group) is generated by $\{L_{-1}, L_0, L_1\} \cup \{\bar{L}_{-1}, \bar{L}_0, \bar{L}_1\}$ and is often useful for characterizing properties of physical states. For example, since the correlation functions are invariant under this group we have a scaling and rotation of a field such that

$$< \phi(re^{i\theta})\phi(0) >= r^{-2(h+\bar{h})} e^{-2i\theta(h-\bar{h})}$$

and thus ϕ has a scaling dimension

$$\Delta = h + \bar{h}$$

and a rotational scaling phase given by the spin $s = h - \bar{h}$. The two-point Green function of a renormalized operator (4) in two dimensions and complex coordinates is therefore

$$< \phi(z, \bar{z})\phi(0, 0) >= |z|^{-2h} |\bar{z}|^{-2\bar{h}} \tag{22}$$

[23]Note that the Virasoro algebra reduces to the classical case (18) for $c = \bar{c} = 0$ which therefore has no anomaly.

and h and $\bar{h}$ are thus referred to as the complex scaling dimensions of ϕ (although they are real numbers).

Although the form of the Virasoro algebra changes for this as the L_n's are shifted by constants, (21) is restricted by the requirement that this subalgebra satisfy

$$[L_{\mp 1}, L_0] = \mp L_{\mp 1} \qquad [L_1, L_{-1}] = 2L_0,$$

with no anomaly term. Therefore the SL(2,C) group remains an exact symmetry group even though there is a central charge present in the Virasoro generalization.

6.1 The stress tensor as the generator of scaling

Consider a general coordinate transformation $r^\mu \to r^\mu + \alpha^\mu(r)$ where the variation in the action is assumed to depend locally upon the function $\alpha^\mu(r)$ such that

$$dS = -\frac{1}{2\pi} \int T_{\mu\nu}(r) \partial^\mu \alpha^\nu(r) d^2 r, \tag{23}$$

where the most relevant term depends on the *strain tensor*, $\partial^\mu \alpha^\nu$. (23) therfore defines the *stress tensor*[24], $T_{\mu\nu}$ and has the simple physical interpretation that the interactions of the statistical mechanics model give rise to an effective elasticity of the medium. The contributions to the Hamiltonian or action is simply the integral of the stress tensor contracted with the strain tensor.

In Hilbert space, the quantum Hamiltonian $\hat{H}$ is proportional to the integral of the time-time component of $T_{\mu\nu}$ over the space-like curve on which we choose to quantize the theory so that

$$\hat{H} = \frac{1}{2\pi} \int \hat{T}_{00}(x) dx. \tag{24}$$

[24]The Euclidean version of the energy-momentum tensor.

This relation is the Euclidean version of the familiar QFT statement that the Hamiltonian is the space integral of the time-time component of the energy-momentum tensor.

$T_{\mu\nu}$ may be decomposed into an antisymmetric, diagonal and traceless symmetric pieces. The antisymmetric piece always vanishes if the system is rotationally invariant as we are assuming here, even though statistical mechanical systems are defined on a lattice. In most cases rotational invariance is regained at the large distance limit since the operators which break this symmetry are irrelevant in the context of the RG.

Now consider a dilation $x^\mu \to x^\mu + \Omega^{-\frac{1}{2}} x^\mu$ with an associated dilation current $j_\mu = T_{\mu\nu} x^\nu$. This current must be divergence-free at a critical point since we have scale invariance and the invariant action ($dS = 0$) means that the trace of the stress tensor[25] ,Θ, must vanish. $T_{\mu\nu}$ is therefore traceless:

$$\Theta \equiv T^\mu_\mu = \partial \cdot j = 0.$$

We therefore have the implication that local scale invariance is equivalent to the conservation of the stress energy tensor, although this is violated when we introduce QFT because of its quantized nature. Furthermore, since $T_{\mu\nu}$ is a local field and the traceless symmetric part of $\partial^\mu x^\nu$ vanishes, $dS = 0$ under all transformations. These transformations are local rotations plus dilations - the set of conformal transformations.

In complex coordinates we have $\partial_{\bar{z}} x^z = \partial_z x^{\bar{z}} = 0$ and we must refer the metric to the complex coordinate frames such that

$$g_{zz} = g_{\bar{z}\bar{z}} = 0, \qquad g_{z\bar{z}} = g_{\bar{z}z} = \frac{1}{2}.$$

[25] and its expectation value $<\Theta>$.

The stress-energy tensor can then be written

$$T_{zz} = \frac{1}{4}(T_{00} - 2iT_{10} - T_{11}),$$

$$T_{\bar{z}\bar{z}} = \frac{1}{4}(T_{00} + 2iT_{10} - T_{11}),$$

$$T_{z\bar{z}} = T_{\bar{z}z} = \frac{1}{4}(T_{00} + T_{11}) = \frac{1}{4}T^{\mu}_{\nu}.$$

$T_{\mu\nu}$ conservation, together with the traceless condition $T_{z\bar{z}} = T_{\bar{z}z} = 0$, implies the conservations $\partial_{\bar{z}}T_{zz} = 0$ and $\partial_z T_{\bar{z}\bar{z}} = 0$. The stress-energy tensor therefore factorizes into independent *left and right* (unbarred and barred) pieces since we are left with the two non-vanishing components $T(z) \equiv T_{zz}(z)$ and $\bar{T}(\bar{z}) \equiv T_{\bar{z}\bar{z}}(\bar{z})$ which have holomorphic and anti-holomorphic dependencies[26] respectively. T and $\bar{T}$ generate local conformal transformations on the complex plane.

In two dimensions $T_{\mu\nu}$ has three independent components $T \equiv T_{zz}, \bar{T} \equiv T_{\bar{z}\bar{z}}$ and $\Theta = 4T_{z\bar{z}} = 4T_{\bar{z}z}$. The conservation $\partial_\mu T_{\mu\nu} = 0$ means

$$\begin{aligned} \partial_{\bar{z}}T + \frac{1}{4}\partial_z\Theta &= 0, \\ \partial_z\bar{T} + \frac{1}{4}\partial_{\bar{z}}\Theta &= 0. \end{aligned} \tag{25}$$

Since $\Theta = 0$ at a critical point T depends only on z, and $\bar{T}$ only on $\bar{z}$. In the next section we will see that taking the operator product expansion of the stress tensor $T(z)$ with any scaling operator gives its two point function in the form

$$< T(z)T(0) >= \frac{c/2}{z^4} \tag{26}$$

and this serves to define the central charge, c. $T(z)$ transforms under a finite conformal transformation as a scaling operator with scaling dimensions (2,0). In this way the stress tensor serves as

[26] i.e. functions of only one coordinate.

the generator of scale and conformal transformations.

6.2 The trace anomaly

In this section we shall also see how the conformal anomaly can be thought of as generated by the coupling of QFT to $d = 2$ gravity. The argument that scale invariance implies $<\Theta>= 0$ at a critical point is only valid for flat space. If we consider a fixed curved geometry then it is consistent that $<\Theta>$ is no longer zero, since the curvature provides a scale; local geometry fluctuations are different in magnitude to large scale curvature. In fact, we must assume that the expectation value of Θ must only depend on the local geometry. Dimensional analysis and general coordinate invariance imply

$$<\Theta> \propto R,$$

where R is the (non-zero) scalar curvature. The left hand side of this relation is known as the *trace anomaly*[27] in the sense that it may appear to vanish in classical conformal field theory, but in QFT it does not. This is because regularization introduces an infra-red cut-off scale (the lattice spacing) into the theory which explicitly breaks conformal invariance in the plane and thus the theory itself is not at a RG fixed point. So although the theory is clasically scale invariant, quantization destroys this and pushes it away from the fixed point massless regime and coupling to gravity causes a mass aquisition. In this case the RG goes through a phase transition and the order parameter that aquires a non-zero expectation value is the trace of the stress tensor, $<\Theta>$. The constant of proportionality turns out ([11] Vol. 1 p.144-7) to be the central charge such that

$$<\Theta>= -\frac{c}{12}R, \tag{27}$$

[27] Or conformal anomaly.

since the orogin of the trace anomaly is the same as that of the conformal anomaly - the non-conservation of $T_{\mu\nu}$. This is not surprising as the order of operations in quantum mechanics is important and thus normal ordering of operators on the manifold imposes additional constraints which prevent this conservation.

It is not as yet clear that this can be linked with the mass generation process of the Higgs mechanism, although certain similarities do exist such as symmetry breaking and non-zero order parameter expectation values.

6.3 The operator product expansion

The (descendent) fields in the representation, $\{\Phi_n\}$, are found by taking successive operator products with the stress-energy tensor. If we then let the stress-energy tensor act upon itself we can determine the form of the central charge as we shall now see.

One consequence of the symmetry properties of a QFT is that it possesses a conserved Noether current j^μ ($\mu = 1...d$). If we then construct a charge $Q = \int d^{d-1}x j_0(x)$, the source of the field, it is simple to see that this too is conserved. It generates an infinitesimal symmetry variation in the field Φ,

$$\delta_\epsilon \Phi(x) = \epsilon[Q(x), \Phi(x)]. \tag{28}$$

according to the equal-time[28] Euclidean space commutator. The stress-energy tensor[29] $T_{\mu\nu}$, is the conserved current that corresponds to translation invariance and is therefore divergenceless[30] and symmetric.

[28] Or rather *constant radius* (when mapping onto a cylinder) since we are not using Minkowski space and there is no time coordinate.

[29] Or energy-momentum tensor.

[30] Although in this case radial ordering prevents this conservation.

Placing the conserved charges on the conformal z-plane gives

$$Q = \frac{1}{2\pi i}\oint \left(dzT(z)\epsilon(z) + d\bar{z}\bar{T}(\bar{z})\bar{\epsilon}(\bar{z})\right) \tag{29}$$

and the line integral is performed anti-clockwise over some fixed radius circle, centered about a point w and thus leads to the variation (28)

$$\begin{aligned}\delta_{\epsilon,\bar{\epsilon}}\Phi(w,\bar{w}) &= \frac{1}{2\pi i}\int\left[dzT(z)\epsilon(z),\Phi(w,\bar{w})\right] + \left[d\bar{z}\bar{T}(\bar{z})\bar{\epsilon}(\bar{z}),\Phi(w,\bar{w})\right] \\ &= h\partial\epsilon(w)\Phi(w,\bar{w}) + \epsilon(w)\partial\Phi(w,\bar{w}) + \bar{h}\bar{\partial}\bar{\epsilon}(\bar{w})\Phi(w,\bar{w}) \\ &\quad +\bar{\epsilon}(\bar{w})\bar{\partial}\Phi(w,\bar{w}).\end{aligned} \tag{30}$$

The latter form is derived by radial mapping to generate radial ordering and using the primary field transformation law (20) with $f(z) = z + \epsilon(z)$. A more detailed derivation of this result can be found at [12], (p17-22).

So that the field source (29) induces the correct infinitesimal conformal transformations the short distance singularities of T and $\bar{T}$ with Φ should be

$$\begin{aligned}T(z)\Phi(w,\bar{w}) &= \frac{h}{(z-w)^2}\Phi(w,\bar{w}) + \frac{1}{z-w}\partial_w\Phi(w,\bar{w}) + ... \\ \bar{T}(\bar{z})\Phi(w,\bar{w}) &= \frac{\bar{h}}{(\bar{z}-\bar{w})^2}\Phi(w,\bar{w}) + \frac{1}{\bar{z}-\bar{w}}\partial_{\bar{w}}\Phi(w,\bar{w}) + ...\end{aligned} \tag{31}$$

This is the short distance *operator product expansion*[31] for the holomorphic and anti-holomorphic stress-energy tensors T and $\bar{T}$ with a primary field, Φ. These short distance properties therefore effectively define the quantum stress-energy tensor and the notion of a primary field[32]. Furthermore they usually result in coordinate transformation generators and thus encode the conformal transformation properties of Φ.

[31] The radial ordering is implicit.

[32] Secondary fields contain terms higher than this double pole singularity.

The descedent fields create descendent states when they act upon the vacuum and the conformal Ward identities produce differential equations that determine the correlation functions of these fields in terms of the primaries. In fact, the most singular part of the operator product expansion defines a closed, finite-dimensional algebra of fields in a theory. The complete information needed to specify a two dimensional conformal field theory is given by the conformal weights $(h, \bar{h})$ of the Virasoro highest weight states, and the operator product coefficients between the primary fields that create them, although the values of these parameters cannot be determined by only the conformal symmetry condition.

Since we can think of the central charge as arising due to the non-conservation of the stress energy tensor, then certain operations on T should also generate it. This field does not obey (20) or (31) and its transformation properties are usually more complicated than those of primary fields[33]. Two successive conformal transformations allow us to determine the operator product with itself and reveals a higher than double pole singularity[34];

$$T(z)T(w) = \frac{c/2}{(z-w)^4} + \frac{2}{(z-w)^2}T(w) + \frac{1}{z-w}\partial T(w) + \cdots, \tag{32}$$

involving the central charge, c. This constant[35] is therefore generated by determining the nature of the stress-energy tensor through letting it operate upon itself. However, it cannot be calculated from the requirement that the T's generate conformal transformations, since that involves the commutator of T with other operators, although since $< T(z)T(0) >= \frac{c/2}{z^4}$, it is expected that $c \geq 0$. If the constant was not present (32) would simply be a statement that $T(z)$ is a conformal field of weight (2,0) - the classical case. The anti-holomorphic operator product gives

[33] As is usually true of derivative fields.
[34] and hence, by definition, its secondary nature.
[35] Which depends upon the theory under consideration.

the same (anti-holomorphic) result as (32) with central charge $\bar{c}$ and modular invariance imposes the constraint[36] $c - \bar{c} = 0$. This central charge turns out to be a useful tool for classifying and analyzing theories as we shall see in section 6.5, for the free boson and fermion cases.

6.4 Virasoro algebra and Fourier mode expansion

To generate the conformal anomaly it is useful to define a Laurent expansion of the stress-energy tensor in terms of operator modes L_n;

$$T(z) = \sum_{n \in \mathbf{Z}} z^{-n-2} L_n, \qquad \bar{T}(\bar{z}) = \sum_{n \in \mathbf{Z}} \bar{z}^{-n-2} \bar{L}_n \tag{33}$$

which clasically are simply fourier modes but become operators themselves in the context of QFT. The exponent is chosen to give the operators a scaling dimension of n under a scale change $z \to z/b$, since $T(z) \to b^2 T(z/b)$ and hence $L_n \to b^n L_{-n}$. To derive the algebra satisfying the above modes we must commute the operators[37] L_n and hence we need to invert (33) to give

$$L_n = \oint \frac{dz}{2\pi i} z^{n+1} T(z), \qquad \bar{L}_n = \oint \frac{d\bar{z}}{2\pi i} \bar{z}^{n+1} \bar{T}(\bar{z}). \tag{34}$$

This gives the commutator of two contour integrations,

$$[L_n, L_m] = \left(\oint \frac{dz}{2\pi i} \oint \frac{dw}{2\pi i} - \oint \frac{dw}{2\pi i} \oint \frac{dz}{2\pi i} \right) z^{n+1} T(z) w^{m+1} T(w). \tag{35}$$

If we then use the result of the stress-energy tensor self product (32), the commutator becomes

$$\oint \frac{dz}{2\pi i} \oint \frac{dw}{2\pi i} z^{n+1} w^{m+1} \left(\frac{c/2}{(z-w)^4} + \frac{2}{(z-w)^2} T(w) \frac{1}{z-w} \partial T(w) + \cdots \right).$$

The z contour integration can then be performed by fixing w and deforming the difference between

[36] This equivalent to local gravitational anomalies canceling, and allows consistent coupling to two dimensional gravitational theories. For example, ghosts are added to heterotic string theory to make $c = \bar{c} = 0$.

[37] and the anti-holomorphic operators, $\bar{L}_n$, to give a separate algebra.

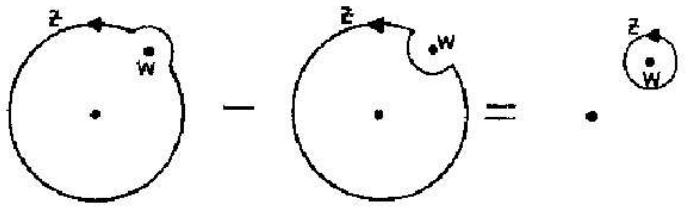

Figure 6: Evaluation of the commutator on the conformal plane.

the two z integrations into a single z contour that winds tightly around w (fig. 6) as was done for the equal time commutator (30). In the limit as $z \to w$ we have, to leading order

$$\begin{aligned}[L_n, L_m] &= \oint \frac{dw}{2\pi i} \Big(\frac{c}{12}(n+1)n(n-1)w^{n-2}w^{m+1} \\ &\quad + 2(n+1)w^n w^{m+1}T(w) + w^{n+1}w^{m+1}\partial T(w) \Big) .\end{aligned}$$

We can then evaluate the w integral. If we integrate the last term by parts and combine it with the second term[38] the result is $(n-m)w^{n+m+1}T(w)$, giving the commutation relation (21) for the Virasoro algebra at the start of section 6.

An identical calculation for $\bar{T}$ results in the anti-holomorphic commutation relation. The fact that the holomorphic and anti-holomorphic operators commute stems from the absence of power law singularities in the operator product of $T(z)$ and $\bar{T}(\bar{z})$. We therefore have two copies of an infinitely dimensional group known as the Virasoro algebra.

[38] Noting also that $\frac{1}{3!}\partial_z^3 z^{n+1}|_{z=w} = \frac{1}{6}(n+1)n(n-1)w^{n-2}$.

6.5 The free boson example

We will now consider a single massless free boson[39] using the above notation on the complex plane. The motivation for this is that the mass parameter in QFT corresponds to the inverse correlation length in statistical mechanics which therefore tends to zero at a critical point.

Consider the normalization[40] for the action

$$S = \int d^2 z \mathcal{L} = \frac{1}{2\pi} \int dz d\bar{z} \partial X \bar{\partial} X, \tag{36}$$

with $X(z, \bar{z})$ having the propagator $< X(z,\bar{z})X(w,\bar{w}) >= -\frac{1}{2}\log|z-w|$. If we then use the Euler-Lagrange equations of motion and try find the solutions, it can be shown that $X(z,\bar{z}) = \frac{1}{2}\left(x(z) + \bar{x}(\bar{z})\right)$ and thus splits into two independent holomorphic and anti-holomorphic pieces, known respectively as massless left-movers and right-movers. The two corresponding propagators are

$$\begin{aligned} & & < x(z)x(w) > &= -\log(z-w) \\ \text{and} & & < \bar{x}(\bar{z})\bar{x}(\bar{w}) > &= -\log(\bar{z}-\bar{w}). \end{aligned}$$

Taking two derivatives of this we see that the derivative field $\partial x(z)$ has a leading short distance expansion of

$$\partial x(z) \partial x(w) = -\frac{1}{(z-w)^2} + \cdots. \tag{37}$$

Although the field $x(z)$ is not actually conformal, (37) has the right scaling properties to be a $(1,0)$ conformal field.

[39] Also known as the Gaussian model.
[40] Used in string theory.

If we define the holomorphic stress-energy tensor to be

$$\begin{aligned} T(w) &= -\frac{1}{2} : \partial x(z) \partial x(w) : \\ &\equiv -\frac{1}{2} \lim_{z \to w} \left[\partial x(z) \partial x(w) + \frac{1}{(z-w)^2} \right] \end{aligned} \tag{38}$$

then we can compute the singular part of the operator product expansion using Wick rules and then Taylor expanding:

$$\begin{aligned} T(w)\partial x(w) &= -1/2 : \partial x(z) \partial x(z) : \partial x(w) \\ &= -\frac{1}{2} \partial x(z) < \partial x(z) \partial x(w) > \cdot 2 + \cdots \\ &= \partial x(z) \frac{1}{(z-w)^2} + \cdots \\ &= \left(\partial(w) + (z-w)\partial^2 x(w) \right) \frac{1}{(z-w)^2} + \cdots, \end{aligned} \tag{39}$$

which becomes

$$T(z)\partial x(w) \sim \frac{\partial x(w)}{(z-w)^2} + \frac{1}{z-w} \partial^2 x(w) + \cdots \tag{40}$$

in the $z \to w$ limit. This agrees with (31) where the derivative field transforms as a (1,0) primary field. A further check we can perform involves substituting this field back into (30) to give

$$\begin{aligned} \left[\oint \frac{dz}{2\pi i} T(z) \epsilon(z) \partial x(w) \right] &= \oint \frac{dz}{2\pi i} \epsilon(z), \left(\frac{\partial x(w)}{(z-w)^2} + \frac{\partial^2 x(w)}{z-w} + \cdots \right) \\ & \qquad \partial\epsilon(w) \partial x(w) + \epsilon(w) \partial^2 x(w). \end{aligned}$$

The above variation accords with the transformation $z \to z + \epsilon$, since $x(z) \to x(z) + \epsilon \partial x(z)$ and hence

$$\partial x(z) \to \partial x(z) + \partial\epsilon \partial x(z) + \epsilon \partial^2 x(z).$$

This result states that the derivative field ∂x transforms as a tensor of conformal weight (or mass dimension) $h = 1$.

We can use the expansion (32) and the free boson stress-energy tensor (38) forward to calculate the operator self product;

$$T(z)T(w) = \frac{1/2}{(z-w)^4} + \frac{2}{(z-w)^2}\left(-\frac{1}{2}(\partial x(w))^2\right) + \frac{1}{z-w}\partial\left(-\frac{1}{2}(\partial x(w))^2\right). \tag{41}$$

By comparing this with (32) it can be seen that the leading term is normalized such that a single free boson has a central charge of $c = 1$.

6.5.1 The free fermion

As a contrast to the previous example, consider a free massless fermion with an action

$$S = \frac{1}{8\pi}\int dz d\bar{z}\left(\psi\bar{\partial}\psi + \bar{\psi}\partial\bar{\psi}\right), \tag{42}$$

that involves the left- and right-moving chiralities $\psi(z)$ and $\bar{\psi}(\bar{z})$. In $d = 2$ the Dirac operator can be represented by

$$\not{\partial} = \sigma_x\partial_x + \sigma_y\partial_y = \begin{pmatrix} & \partial_x + i\partial_y \\ \partial_x + i\partial_y & \end{pmatrix} \sim \begin{pmatrix} & \partial \\ \bar{\partial} & \end{pmatrix}, \tag{43}$$

so that the operators ∂ and $\bar{\partial}$ are projected out by $\frac{1}{2}(1 \pm \sigma_z)$. The normalization of (42) is chosen so that the leading short distance singularities are

$$\psi(z)\psi(w) = -\frac{1}{z-w} \quad \text{and} \quad \bar{\psi}(\bar{z})\bar{\psi}(\bar{w}) = -\frac{1}{\bar{z}-\bar{w}}. \tag{44}$$

The stress-energy tensors are

$$T(z) = \frac{1}{2} : \psi(z)\partial\psi(z) : \quad \text{and} \quad \bar{T}(\bar{z}) = \frac{1}{2} : \bar{\psi}(\bar{z})\bar{\partial}\bar{\psi}(\bar{z}) : \tag{45}$$

and satisfy (32) with $c = 1/2$. Operator product expansions of $T(z)\psi(w)$ and $\bar{T}(\bar{z})\bar{\psi}(\bar{w})$ show that ψ and $\bar{\psi}$ are primary fields of weight $(\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$.

So far we have seen that the representation theory of the Virasoro algebra allows us to describe the possible scaling dimensions of fields of $d = 2$ conformal field theories. Determination of the possible critical indices of two dimensional systems at a phase transition is therefore facilitated. It is in fact possible to give a complete classification of all possible two dimensional critical behaviour in the case of unitary systems with $c \leq 1$. Although for $c < 1$ we have a discrete series corresponding to a set of Virasoro algebra representations, it is not totally obvious that constructing statistical models at their critical points should realize this algebra. However, various authors ([9], [3]) have demonstrated the correspondence between particular models.

The unitary[41] representation theory of the Virasoro algebra plays the same role in studying two dimensional critical phenomena as representation theory of finite and Lie groups play in other branches of physics. After a particular symmetry group of a physical system is identified it is fairly straight forward to analyze its spectrum and interactions. Given a critical statistical model, the scaling weights of the operators can be identified using the two-point correlation functions. Often it suffices to merely identify the relevant Virasoro algebra representation.

7 The continuum limit

If we now turn our attention towards system behaviour in the continuum limit away from criticality, it is important to realize that we are dealing with *renormalized* quantum field theories. Therefore any operators, ϕ_i, will be finite in the limit of a continuous manifold and the coupling constants, λ_i, will be renormalized to give finite Green functions. In general, the Lagrangian counter terms set by the relation between the renormalized and bare field operators will depend upon the interaction

[41] With $|c|, |h| \geq 0$.

coupling λ whose positive dimensional quality implies a finite set of counter terms.

Consider a perturbation of the fixed point action S^* by a relevant[42] scalar operator ϕ with scaling dimensions (h,h) such that the new action is

$$S = S^* - \lambda \int \phi(z,\bar{z}) d^2 z. \tag{46}$$

The coupling constant λ has scaling dimensions $(1-h, 1-h)$ and an overall dimension $[\lambda] > 0$ and hence we have a super-renormalizable theory.

Consider S^* to be at the Gaussian fixed point so that it then corresponds to the free field theory action. ϕ in (46) is defined to be renormalized at $\lambda = 0$ and performing a perturbative expansion in λ by calculating some connected correlation function of ϕ gives

$$<\phi(z,\bar{z}) \cdots> = <\phi(z,\bar{z}) \cdots>_{S^*} + \lambda \int <\phi(z,\bar{z}) \phi(w,\bar{w}) \cdots>_{S^*} d^2 w + \cdots. \tag{47}$$

The integrand may be estimated to leading order by using the operator product expansion method of section 6.3. For convenience we shall assume that no counter terms[43] arise either because the operators ϕ_i coupling to the ϕ^n fields at order λ^n do not exist or because the scaling dimensions of the fields $h_i \leq 0$. However, there is at least one counter term for the stress energy tensor. Consider performing the same type of procedure as in (47) for the zz component of the stress tensor, T, giving

$$<T(z,\bar{z}) \cdots> = <T(z) \cdots>_{S^*} + \lambda \int <T(z) \phi(w,\bar{w}) \cdots>_{S^*} d^2 w + \cdots. \tag{48}$$

If we study the operator product expansion (32) it can be seen that the integral in (48) is ultraviolet

[42] So that $h < 1$, since the RG eigenvalue of λ is $y = 2 - 2h$ and the trajectory stays on the critical surface under such a perturbation.

[43] Of which there are a finite number needed to renormalize ϕ in the $\lambda \neq 0$ theory.

divergent and should be regulated[44] with a cut-off. This leads to $\partial_{\bar{z}}T$ no longer vanishing;

$$\partial_{\bar{z}}T = \lambda \int \frac{(1-h)}{(z-w)}(z-w)\partial_z\phi(z,\bar{z})d(|z-w|^2-a^2)d^2w + \cdots, \tag{49}$$

although the most singular term does. If we compare the conservation condition (25) of $\partial_{\bar{z}}T + \frac{1}{4}\partial_z\Theta = 0$ imposed upon the stress tensor it can be seen that

$$\Theta(z,\bar{z}) = -4\pi\lambda(1-h)\phi(z,\bar{z}) + \cdots. \tag{50}$$

We ignore the higher order terms because of the earlier assumptions made about the lack of counter terms. Now Θ is the response of the action to a scale transformation $z \to b^{-1}z$ and since this rescaling means $d^2z \to b^{-2}d^2z$ and $\phi \to b^{2h}\phi$ then the perturbed action (46) becomes

$$S = S^* - \lambda b^{2(h-1)} \int \phi(z,\bar{z})d^2z. \tag{51}$$

If $(b-1)$ is infinitesimal, this Taylor expands to be the first term in (50) and is indeed the scaling induced action change. This is valid all the way along a renormalization group trajectory leaving the fixed point, until it reaches another IR stable fixed point. This latter fixed point determines the IR behaviour at any point on the trajectory. Furthermore this indicates a change in the UV behaviour. This requires a renormalization of ϕ such that $\phi_{ren} \sim a^{-const}\phi$. Thus in the renormalized theory, ϕ, and hence $\Theta(z,\bar{z})$ given by (50) are zero as we would expect at a fixed point.

Thus, if the action is perturbed away from some critical point by a relevant operator it results in a movement in coupling space on the critical surface formed by the RG trajectories (see fig. 4). The movement is always confined to the surface for small enough perturbations and the direction of movement is towards the critical point(s). The crucial result is that this movement only occurs

[44] One way to achieve this would be to put a step function cut-off $H((z-w)(\bar{z}-\bar{w})-a^2)$ into the integral.

in one direction and we shall explore the reason for this in the next section.

8 Zomolodchikov's C-theorem

It was demonstrated by A. Zomolodchikov in 1986 [31] that there exists a function $C(\{\lambda\})$ of the coupling constants in a two-dimensional renormalizable field theory with the following properties:

- $C(\{\lambda\}) \geq 0$ and monotonically decreases along RG trajectories. Hence

$$\frac{dC(\{\lambda\})}{dt} \equiv \beta_i(\lambda)\frac{\partial C(\lambda)}{\partial \lambda_i} \leq 0. \tag{52}$$

- C is stationary (takes constant values) only at critical fixed points where $\lambda = \lambda_*$. Thus $\beta_i(\lambda_*) = 0$.

- At a fixed point with coupling λ_*, C becomes equal to the value of the central charge, c, of a Virasoro algebra of the corresponding conformal field theory- $C(\lambda_*) = c(\lambda_*)$.

The motivation for the existence of such a function is that the renormalization coarse-graining procedure implies a loss of information about the degrees of freedom (and thus the UV behaviour of the field theory whose wavelength is of the same order as the cut-off, since it is not logically consistent to examine correlations at scales below the cut-off). The RG flows are therefore irreversible and the transformations correspond to irreversible processes such as those found in the time evolution of dissapative systems. It is thus reasonable to expect an entropy function that measured this information loss to exist. If we consider figure 4 this rectified flow is seen as any perturbation on the critical surface causing a point to move towards the fixed point at the orogin.

The proof of this C-theorem is based on the conditions of renormalizability, positivity, the translational symmetry of the field theory, rotational invariance (and hence the conservation of the stress tensor) and certain properties of a $d = 2$ conformal field theory:

Consider a theory at some point on the renormalization group trajectory possessing a flow from a UV stable fixed point to a relatively IR stable fixed point with its coupling dependece suppressed. Since T, Θ and $\bar{T}$ are respectively the spin 2, 0, -2 components of the stress tensor, their two point functions must have the form

$$\begin{aligned} <T(z,\bar{z})T(0,0)> &= F(z\bar{z})/z^4, \\ <\Theta(z,\bar{z})T(0,0)> = <T(z,\bar{z})\Theta(0,0)> &= G(z\bar{z})/z^3\bar{z}, \\ <\Theta(z,\bar{z})\Theta(0,0)> &= H(z,\bar{z})/z^2\bar{z}^2 \end{aligned} \tag{53}$$

If we now take the correlation function of the condition implied by the conservation of the stress tensor (25) with $T(0,0)$ and with $\Theta(0,0)$ we have enough information to construct (the defined) C-function

$$C \equiv 2F - G - \frac{3}{8}H \tag{54}$$

and show that $\dot{C} = -\frac{3}{4}H$.

Since Θ is scalar, reflection positivity means $H \geq 0$ and thus $C(B, \{\lambda\})$ must be a non-increasing function of $B \equiv (z\bar{z})^{\frac{1}{2}}$ at a fixed set of couplings $\{\lambda\}$ (the first condition above). This function is dimensionless and independent of the UV cut-off and should therefore satisfy the Callan-Symanzik equation [14],

$$\left(B\frac{\partial}{\partial B} + \sum_i \beta_i(\{\lambda\})\frac{\partial}{\partial \lambda_i} \right) C(B, \{\lambda\}) = 0, \tag{55}$$

with the rate of change of C along the RG trajectory at a fixed B being given by

$$\frac{dC}{dt} = -\sum_i \beta_i(\{\lambda_i\})\frac{\partial C}{\partial \lambda_i}. \tag{56}$$

If we define $C(\{\lambda\}) \equiv C(1, \{\lambda\})$ then this satisfies the first part of Zomolodchikov's theorem. The second condition is also fulfilled since C is stationary if and only if $H = 0$ implying that $\Theta = 0$ and the theory is at a critical fixed point. Furthermore, at such a point $G = H = 0$ and $F = \frac{1}{2}c$, so $C = c$, the third condition of Zomolodchikov's theorem.

The interpretation given to this theorem is that renormalization group flows go 'downhill' towards fixed critical points of the theory and no limit cycles and other strange behaviours exist - figure 4 is a smooth monotomically decreasing surface. C effectively counts the number of degrees of freedom given by the Lagrangian. Hence for a single free (bosonic) field[45] $C = c = 1$. Fermions appear in pairs in the Lagrangian and so it is consistent to regard them as one field, hence $c = 1/2$ for one fermion. Since C is monotomically decreasing, no RG trajectory exists that can increase the degrees of freedom of a system and thus the inherent loss of information, as one would expect from the coarse graining proceedure of renormalization. It is important to note that a simple generalization of this idea fails in more than two dimensions because more invariant amplitudes in the correlation functions arise.

The fact that it is only possible to lose information under transformations is closely linked to Shannon's information theory, which we shall now explore.

[45] As in section 6.5

Information theory *Truth is much too complicated to allow anything but approximations.*

J. von Neumann.

9 Introduction to Information Theory

The fundamental problem of communication (information transmission) lies in the reproduction of a message at one point that was selected at another point. The reproduction may be exact or approximate, although the latter will often involve a degradation in the amount of information that can be conveyed. In 1948 Claude Shannon published two papers [25] on *The mathematical theory of communication*, a product of his thoughts concerning the growing field of the technology of transmitting signals (particularly with respect to telegraphy). A basis for this information theory was provided through the work done by Nyquist [18], [19] and Hartley [13] in the 1920's, which Shannon extended to form a more coherent and general field of research. Some of the new features included the effect of noise in the channel and the possible savings due to the nature of the final destination and the statistical structure of the original message. Initially we shall explore the basic theory and then apply these new tools to the renormalization group.

10 Signal generation and transmission

Shannon ignored the fact that messages carry a semantic meaning in as much as he believed that the correlations made by consciousness to a system possessing certain physical or conceptual entities is irrelevant to the engineering problem. It may initially seem like quite a large leap of faith to ignore the interpreters of the message, but the theory's success should convince us that this 'approximation' is valid. For example, in his first paper on the subject Shannon demonstrates

how transmission probabilities between words in the English language can be used to generate a reasonable copy of this communication method. It is not wholly unreasonable therefore, to suggest that a machine capable of passing the *Turing* test (see [21]) could be constructed, given enough data on the probabilities of the response to various combinations of words[46].

Shannon's theory is centered on the concept that the actual message is selected from a set of possible messages. If the number of messages, s, in the set is finite then s or any monotonic function of s can be regarded as a measure of the information produced when one message is chosen from the set, all choices being equally likely[47]. A logarithmic measure is chosen for its practical usefulness, its mathematical suitability and its intuitive suitability (since many people presume that two channels have twice the capacity for transmitting information than one - they are not aware that the number of ways of arranging information increases exponentially with the number of channels or *states*). The base of the log is often chosen to be 2 since this corresponds to the choice of unit for measuring information (the bit). For example, flip flops store 1 bit of information (2 states) and n relays store N bits. The total number of possible states is 2^n, so $N = \log_2 2^n = n$. In analytical work we shall use the natural logarithm since calculus will be employed. The resulting units of information are called natural units[48].

10.1 The discrete noiseless channel

Consider a discrete transmission channel, a system consisting of a sequence of choices from a finite set of elementary symbols $s_1 ... s_n$ which can be transmitted from one point to another. Each

[46] However, as Penrose points out, this is no vindication of the *strong Artificial Intelligence* claim that consciousness is simply a result of computer-like computational tasks.

[47] Although this definition is considerably modified by the effects of statistics and for the continuous message range case.

[48] We can change from base a to base b by multiplying by $\log_b a$.

symbol has (not necessarily the same) duration t_i and every sequence of s_i's doesn't have to be transmittable. The allowed sequences *are* the possible signals for the channel and the number of these of duration T are denoted N(T). The maximum possible transmission rate of information in this type of channel is defined to be its *capacity*

$$C \equiv \lim_{T \to \infty} \frac{\log N(T)}{T} = \log W \tag{57}$$

where W is the largest real root of

$$\left| \sum_s W^{-b_{ij}^{(t)}} - \delta_{ij} \right| = 0 \tag{58}$$

and $b_{ij}^{(t)}$ is the duration of the s^{th} symbol which is allowable in state i and leads to state j. The log of the number of possible signals in a discrete channel increases linearly with time. The capacity to transmit the information can then be specified by giving this rate of increase, the number of bits per second required to specify the signal used.

10.1.1 The information source

A discrete source can be considered to be a Markoff process, a stochastic iterative process involving the application of a matrix of transition probabilities to a set of finite states. There exits a subset of Markoff processes known as *ergodic* in which every sequence produced by this process shares the same statistical properties. Thus the symbol (e.g. letter) frequencies will approach definite limits independent of the particular sequence as the lengths of the sequences increase - statistical homogeneity. If P_i is the probability of a state i ($|i>$) and $p_i(j)$ the transition probability to $|j>$ then

$$P_j = \sum_i P_i p_i(j) \tag{59}$$

is the stationary equilibrium condition. For an ergodic source it can be shown that with any starting conditions the probabilities $P_j(N)$ of being in $|j>$ after N symbols approach the equilibrium values as $N \to \infty$. It is these sources that we shall be mostly concerned with.

11 Choice, capacity, uncertainty and entropy

Shannon defines a function H that is a measure of how much choice is involved in the selection of an event, or how uncertain we are about the outcome, i.e. how much information we have. It is given by

$$H = -K \sum_{i=1}^{n} p_i \log p_i \tag{60}$$

where K is a positive constant and really amounts to the choice of unit of measurement. H is therefore recognizable as the entropy $H = -\sum p_i \log p_i$ of the set of probabilities $p_1, ..., p_n$ of a system being in cell i of its phase space (or of possible events occurring, in the context information theory). It is also related to the capacity of a binary symmetric channel with error probability p by

$$C(p) = 1 - H(p). \tag{61}$$

It is this choice measure that we shall be comparing to Zomolodichov's C-function since the entropy function has the following properties:

1. H is a continuous function of $\{p_n\}$.

2. H is positive unless all p_i but one are zero and the non-zero probability is 1, then $H = 0$ since no choice exists and we are certain of the outcome.

3. If all p_i are equal and thus equal to $1/n$, H is a monotonically increasing function of n equal to $\log n$ and is at its maximum. This is the most uncertain situation since we have an increasing level of choice (and hence uncertainty) the more closely the probabilities are grouped.

4. If a choice can be broken down into two successive choices, the original $H(x)$ (where x is the label of the chance variable and not the argument of the function) should be the weighted sum of the individual values of H. For example, $H(\frac{1}{2}, \frac{1}{3}, \frac{1}{6}) = H(\frac{1}{2}, \frac{1}{2}) + \frac{1}{2}H(\frac{2}{3}, \frac{1}{3})$.

5. This has the consequence that any averaging operation on the p_i's of the form $\overline{p_i} = \sum_j a_{ij} p_j$ with $\sum_i a_{ij} p_j = \sum_j a_{ji} p_j = 1$ and all $a_{ij} \geq 0$, will cause H to increase (except in the special case where this transformation is simply the permutation of the p_j and H does not change).

6. The uncertainty of a joint event is less than or equal to the sum of all the individual uncertainties - $H(x, y) \leq H(x) + H(y)$. The equality occurs when the events are independent $(p(i, j) = p(i)p(j))$.

7. The uncertainty of a joint event x, y is the uncertainty of x plus the uncertainty of y when x is known - $H(x, y) = H(x) + H_x(y)$ where the *conditional* entropy[49] of y, $H_x(y) = -\sum_{i,j} p(i, j) \log p_i(j)$ measures how uncertain we are, on average, of y when we know x.

8. These last two points lead to the statement $H(y) \geq H_z(y)$, the uncertainty of y is never increased by knowledge of x. It will be decreased unless x and y are independent events, in which case it is not changed.

Property 3 shows that C is a monotonically decreasing function as we would expect for a measurer

[49]The average of the entropy of y for each value of x.

of transmitted (or transformed) information. Thermodynamics tells us that it is reasonable to expect an increase in entropy from any process unless it is completely reversible. Since we know that the coarse graining procedure used in the RG causes a loss of information and consequently the transformation is not totally reversible, there must be an increase in entropy and thus a reduction in capacity, C. This is consistent with the behaviour of Zomolodichov's information measuring C function. Property 5 further highlights this since we can consider the coarse graining to be an averaging operation, thus increasing H.

It is expected that the final property will cause some problems in the context of quantum mechanics since it is inconsistent with the uncertainty principle - the uncertainty of an observable *can* be increased by the knowledge of another.

11.1 Relative entropy and redundancy

The ratio of the entropy source to the maximum value it could have whilst still restricted to the same symbols is called its *relative entropy*[50]. The *redundancy* is given by one minus the relative entropy. Shannon calculates the redundancy of ordinary[51] English to be about 50% inferring that when English is written, half is determined by the structure of the language and half is chosen freely [26]. Thus it is possible to restore a passage of English if 50% or less of the letters are deleted. A typical basic English vocabulary is limited to 850 words and the redundancy is much higher, as one would expect since a translation into basic English results in an expansion. Conversely James Joyce often wrote with an extended vocabulary and is alleged to have achieved a compression of semantic content. This accounts for the general acceptance that English is very easy to learn on a

[50] It turns out to be the maximum compression possible when we encode a sequence into the same alphabet.
[51] By ignoring statistical structure over greater string lengths than about eight letters.

basic level, but extremely hard to become fluent in (since there are so many continuously evolving, culturally influenced exceptions to learn).

In the next section we shall see the relevance of redundancy and how it restricts information transmission.

11.2 The discrete channel with noise

If the signal is perturbed by noise[52] during the transmission or at one of the terminals where the signal is transmitted or received and the transmitted signal always produces the same (modified) signal, then the effect is termed *distortion*. The received signal is a definite function of the transmitted signal and if this function has an inverse, distortion may be (in principle) corrected by performing the inverse function on the received signal.

More interestingly, if the signal does not always undergo the same change in transmission then we consider the noise to be a chance variable which can be represented by some suitable stochastic process. If a noisy channel is fed by a source there are two statistical processes at work - the noise and the source. In the noiseless case the entropy[53] of the output channel (the received signal) $H(y)$, is equal to that of the input channel (and the source if the emitter is non-singular) $H(x)$. If the joint entropy of input and output is denoted $H(x,y)$ and the conditional entropies $H_x(y)$ (of the output when the input is known) and $H_y(x)$ (of the input when the output is known) then the following relation holds:

$$H(x,y) = H(x) + H_x(y) = H(y) + H_y(x). \tag{62}$$

[52] i.e. any unknown function
[53] Measured per second or per symbol.

If the channel signal is noisy then it is not generally possible to operate on the received signal and completely reconstruct the original message or the transmitted signal with certainty. This is what happens when we perform renormalization since the coarse graining procedure does not allow us to restore the exact original Lagrangian. The new one can be produced from a range of bare Lagrangians. We can therefore consider the coarse graining procedure a type of noise.

However, there exist optimal ways of transmitting information so as to reduce the noise. The proper correction to apply to the amount of information transmitted is the amount of missing information in the received signal. It is therefore logical that we use the conditional entropy of the message, knowing the received signal, $H_y(x)$, which is related to the rate of actual transmission R, by

$$H_y(x) = H(x) - R, \tag{63}$$

where the effective rate of transmission of information R, is the amount of information sent less the uncertainty of what was sent [54]:

$$\begin{aligned} R &= H(x) - H_y(x), \\ &= H(y) - H_x(y), \\ &= H(x) + H(y) - H(x,y). \end{aligned} \tag{64}$$

$H_y(x)$ is known as the *equivocation* in this case. The second expression above measures the amount received less the part of this due to noise and the third expression is the sum of the two amounts less the joint entropy and is thus the number of bits per second that is common to the two.

The capacity of a noisy channel should be the maximum possible rate of transmission. This

[54] The difference between the rate of production (the entropy of the source) and the average rate of conditional entropy.

occurs when the source is properly matched to the channel and the channel capacity is therefore defined to be

$$C = \text{Max } (H(x) - H_y(x)) \tag{65}$$

where the maximum is with respect to all possible information sources used as input to the channel. If the channel is noiseless, $H_y(x) = 0$ and the above definition reduces to the statement that the maximum entropy for a noiseless channel is its capacity. So, although we can never send certain information, there still exists a definite capacity which serves as a reference to reduce the probability of the errors by.

This reduction is achieved by sending the message in a redundant form such as repeating the message several times to enable a statistical study of the different versions. Initially, it might seem that to make this probability of errors approach zero, the redundancy of the encoding must increase indefinitely. However, if this were true there would only be a capacity for a given frequency of errors or equivocation, with the capacity falling as the error requirements are made more stringent. In fact, Shannon shows that for a discrete source with an entropy per second of H and a discrete channel with capacity C, if $H \leq C$ it is possible to send information at the rate C through the channel, with as small a frequency of errors or equivocation as desired, if the proper encoding is used. If an attempt is made to transmit at a rate $C + \delta \geq C$ (so $H \geq C$), then there will be an equivocation equal to or greater than the excess δ. It is possible to encode the source so that the equivocation is less than $H - C - \delta$ but there is no method of encoding which gives an equivocation less than $H - C$.

These results are illustrated in fig. 7 and are Shannon's main justification for the definition of C. $H(x)$ is the rate of information into the channel and $H_y(x)$ is the equivocation. Any point

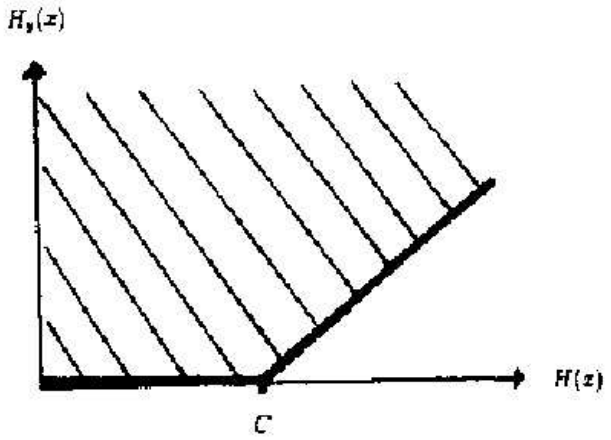

Figure 7: The equivocation possible for a given input entropy to a channel.

above the heavy line in the shaded region can be attained and any point below cannot[55]. The slope of the boundary of these two regions for $H(x) > C$ is one since $H(x) = H_y(x)$ is the limiting condition and any increase in $H(x)$ will be compensated for by an increase in $H_y(x)$.

We can interpret this as 'nature taking payment' by requiring just that much uncertainty so that we are not getting any more information through the channel than the capacity C allows. In terms of Zomolodchikov's theorem we can interpret the attempt to transmit at a rate $C + \delta$ as a transformation that gives us more information. Since this is not possible (the coarse graining will always reduce the amount of information) there must always be a renormalization proceedure to produce a compensatory loss of information.

The modification of the definition of capacity (57) for a noisy channel where the receiver is

[55]The points on the line cannot in general be attained but there are usually two points that can.

constructed to select[56] the most probable cause from a subset of signals, is

$$C = \lim_{T\to\infty} \frac{\log N(T,q)}{T}, \qquad (q \neq 0,1) \tag{66}$$

where $N(T,q)$ is the maximum number of signals we can choose for a subset of signals selected with equal probability, such that the probability of an incorrect interpretation is $\leq q$. So no matter how we set our limits of reliability, we can distinguish enough messages reliably in time T, to correspond to about CT bits when T is sufficiently large.

11.3 Continuous Information

We shall now consider the continuous case since Zomolodichov's theorem is formulated in the continium limit.

Usually the case where the signals or the messages or both are continuously variable, can be obtained through a limiting process from the discrete case. This is done by dividing the continuum of messages and signals into a large but finite number of small regions and calculating the various parameters involved on a discrete basis. As the size of the regions is decreased these parameters generally approach the proper values for the continuous case, although a few new effects do appear.

11.4 The entropy of a continuous distribution

The entropy is defined in an analogous way to the discrete case such that in d dimensions

$$H = -\int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} p(x_1...x_d) \log p(x_1,...,x_d) dx_1...dx_d, \tag{67}$$

[56] As the original signal.

where $p(x_1 ... x_d)$ is a density distribution function. If there are two arguments x and y (which may also be multi-dimensional), the joint and conditional entropies of $p(x,y)$ are given by

$$\begin{aligned} H(x,y) &= -\int\int p(x,y)\log p(x,y)dxdy, \\ H_x(y) &= -\int\int p(x,y)\log \frac{p(x,y)}{p(x)}dxdy, \\ H_y(x) &= -\int\int p(x,y)\log \frac{p(x,y)}{p(y)}dxdy, \end{aligned}$$

where

$$\begin{aligned} p(x) &= \int p(x,y)dy, \\ p(y) &= \int p(x,y)dx. \end{aligned}$$

The entropies of continuous distributions have most of the properties of the discrete case with a few minor modifications. However, in the discrete case the entropy measures in an *absolute* way the randomness of the chance variable. In the continuous system the measurement is *relative to the coordinate system.* If we change coordinates the entropy will generally be changed too. In fact, if we change to the new coordinates $y_1 ... y_n$, the entropy changes from (67) to

$$H = -\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p(x_1 ... x_d) J\left(\frac{x}{y}\right) \log p(x_1 ... x_d) J\left(\frac{x}{y}\right) dy_1 ... dy_d, \tag{68}$$

where $J\,(x/y)$ is the Jacobian of the coordinate transformation. Expanding the logarithm and changing variables to $x_1 ... x_d$ gives

$$H(y) = H(x) - \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p(x_1, ..., x_d) \log J\left(\frac{x}{y}\right) dx_1 ... dx_d. \tag{69}$$

Thus the new entropy is the old entropy degraded by the expected logarithm of the Jacobian. In the continuous case the entropy can be considered a measure of randomness *relative to an assumed*

standard. This standard is the coordinate system chosen with each small volume element $dx_1...dx_d$ given equal weight. When we transform to the new coordinates the entropy in the new system measures the randomness when equal volume elements $dy_1...dy_d$ in the new system are given equal weight.

It is important to note that the derived concepts of information rate and channel capacity depend on the difference of two entropies which itself does not depend on the coordinate frame since each of the two terms change by the same amount and so the concept of entropy is just as important as in the discrete case. This is very important if it is going to be used consistently in *relativistic* quantum mechanics. The scale of measurements sets an arbitrary zero, corresponding to a uniform distribution over a unit volume. A distribution which is more confined than this has less entropy and is therefore negative. However, the rates and capacities will always be non-negative.

11.4.1 The capacity of a continuous channel

If we consider both the transmitted and received signals to be limited to a certain band W. They can thus be specified for a time T, by $2TW$ numbers and their statistical structure by finite dimensional distribution functions. The statistics of the transmitted signal will then be determined by,

$$P(x_1, ..., x_n) = P(x)$$

and those of the noise by the conditional probability distribution

$$P_{x_1,...,x_n}(y_1, ..., y_n) = P(x).$$

The rate of transmission is as for the continuous case[57] and the channel capacity turns out to be

$$C = \lim_{T\to\infty} [\text{Max } P(x)] \frac{1}{T} \int\int P(x,y) \log \frac{P(x,y)}{P(x), P(y)} dxdy. \tag{70}$$

R and C are therefore independent of the coordinate system since the numerator and denominator in the log will be multiplied by the same factors when x and y are transformed in any one-to-one way. This integral expression for C is more general than the continuous case (65) since it will always exist whilst $H(x) - H_y(x)$ may assume an intermediate form of $\infty - \infty$ in some cases. An example of this behaviour is when x is limited to a surface of fewer dimensions than d in its d-dimensional approximation.

In general however, the discrete capacity tends to C above in the limiting case. This can be seen on a 'physical level' by dividing the space of signals into a large number of small cells, sufficiently small so that the probability density $P_x(y)$ of a signal x being perturbed to point y is substantially constant over a cell (either of x or y). If the cells are considered as distinct points then the situation is essentially the same as a discrete channel. However, it is clear that quantizing the volume into individual points cannot alter the final answer significantly in any practical situation (provided the regions are sufficiently small). The continuous capacity will therefore be the limit of the capacities for the discrete subdivisions, as given above. However, when we introduce quantum effects in such small volumes it would not be unreasonable to see modifications perhaps due to the uncertainty principle as we have already mentioned.

[57] See equation (64).

11.5 A coordinate transformation example

Consider a linear coordinate transformation

$$y_i = \sum_i a_{ij} x_i, \tag{71}$$

which has the simple Jacobian $J = |a_{ij}|^{-1}$. (69) gives us a new entropy of

$$H(y) = H(x) + \log |a_{ij}| . \tag{72}$$

The entropy therefore always increases under a transformation and thus the capacity of the channel (61) is reduced. This is as we would expect since Zomolodichov's theorem also implies a loss of information under transformations.

In the case of any measure preserving transformation such as coordinate rotation, $J = 1$ which gives $H(y) = H(x)$ and hence no information loss. In terms of the RG, we are not attempting any decimation by performing a rotation (the scale remains the same) and thus we do not actually expect any information loss.

Conclusion and remarks *As far as the laws of mathematics refer to reality they are not certain; and as far as they are certain they do not refer to reality.*

A theory is more impressive the greater the simplicity of its premise is, the more different kinds of things it relates and the more extended its area of applicability. ... it (thermodynamics) is the only physical theory of universal content concerning which I am convinced that, within the framework of applicability of its basic concepts, it will never be overthrown.

Albert Einstein.

12 Conclusions & Discussion

We have seen how the intimate relationship between QFT, statistical mechanics and information theory arises due to their utilization of similar mathematical techniques. The coarse graining procedure that is so vital for QFT to produce sensible answers seems to lead to an inherent information loss. RG flows must always move to a fixed point since transformations cannot increase the amount of information available. It is therefore possible to draw an analogy between Shannon's 'Capacity' and Zomolodchikov's C-function, since they both 'measure' information and monotonically decrease under the action of the RG.

However, this dissertation has merely highlighted one connected area and it seems likely that many more exist. For example, it was stated that the entropy of a continuous distribution can be negative since the scale upon which we are measuring defines an arbitrary zero. Since we can redefine this origin we might expect phenomena such as scale invariance to occur under certain conditions. Where, is not quite clear as yet, but consistency would imply some sort of critical point, if one could be defined in information theory. It is known that certain functions (e.g. Weierstrass

[24]) have a fractal nature and thus encode information on all scales. It would not be surprising then, for these functions to have a scale invariant entropy descriptor.

It is obvious that the fusion of these fields can only help in the pursuit of scientific knowledge, although serious technical difficulties remain. For example, it is not presently possible to couple gravity with renormalizable conformal field theories in more than two dimensions. However, it is hoped that the introduction of information theory may shed some light on these types of major problems in physics today and help towards a unification of the ever more disparate fields within physics and mathematics. After all, we must understand the limitations of ourselves and our tools as well as those of our environment.

Conformal Field Definitions A conformally invariant theory is defined to satisfy the following properties:

- $\exists$ a set of fields $\{A_i\}$, where the index i labels the particular field which is generally infinite and contains the derivatives of all the fields.

- $\exists$ a subset of fields $\{\phi_j\} \subset \{A_i\}$ that are known as *quasi-primary*[58], that covariantly transform as

$$\phi_j(x) \to \left|\frac{\partial x'}{\partial x}\right|^{\frac{\Delta_j}{d}} \phi_j(x'), \tag{73}$$

  where Δ_j is the dimension of ϕ_j. The correlation functions therefore satisfy

$$< \phi_1(x_1)...\phi_n(x_n) >= \left|\frac{\partial x'}{\partial x}\right|^{\frac{\Delta_1}{d}}_{x=x_1} ... \left|\frac{\partial x'}{\partial x}\right|^{\frac{\Delta_n}{d}}_{x=x_n} < \phi_1(x_1')...\phi_n(x_n') > . \tag{74}$$

- The remaining $\{A_i\}$'s can be given by linear combinations of the quasi-primary fields and their derivatives.

- $\exists$ a vacuum $|0>$ which is invariant under the global conformal group.

[58] Or sometimes SL(2,C) primaries.

References

[1] I. J. R. Aitchison, *An Informal Introduction to Gauge Field Theories*, Cambridge Univ. Press, 1982.

[2] D. J. Amit, *Field theory, the Renormalization Group and Critical Phenomena*, McGraw-Hill, 1978.

[3] G. E. Andrews, R. J . Baxter & J. P. Forrester, *Eight-vertex SOS model and generalized Rogers-Ramanujan-type identities*, J.Stat. Phys. 35 (1984) 193.

[4] E. Brézin & J. Zinn-Justin (Ed.), *Fields Strings and Critical Phenomena*, Les Houches 1988, session XLIX, North-Holland, 1988.

[5] J. L. Cardy, *Conformal Invariance and Statistical Mechanics, Fields Strings and Critical Phenomena*, Les Houches 1988, session XLIX, North-Holland, 1988.

[6] J. F. Cornwell, *Group Theory In Physics Vols. 1 & 2*, Academic Press, 1984.

[7] P. Coveney & R. Highfield, *The Arrow of Time*, Flamingo, 1991.

[8] R. D'Inverno, *Introducing Einstein's Relativity*, Claredon Press, 1992.

[9] D. Friedan, Z. Qiu & S. Shenker, *Conformal invariance, unitarity, and critical exponents in two dimensions*, Phys. Rev. Lett. 52 (1984) 1575.

[10] M. Gell-Mann & F. E. Low, *Quantum electrodynamics at small distances*, Phys. Rev., 95 (1954) 1300 - 12.

[11] M. B. Green, J. H. Schwarz & E. Witten, *Superstring theory*, Cambridge Univ. Press, 1987.

[12] P. Ginsparg, *Applied Conformal Field Theory, Fields Strings and Critical Phenomena*, Les Houches 1988, session XLIX, North-Holland, 1988.

[13] V. L. Hartley, *Transmission of information*, Bell Sys. Tech. J., 7, p535, July 1928.

[14] C. Itzykson & J. Zuber, *Quantum Field Theory*, McGraw-Hill, 1988.

[15] C. Itzykson & J. Drouffe, *Statistical Field Theory*, Cambridge Univ. Press, 1989.

[16] J. L. Kelly, Jr., *A new interpretation of information rate*, Bell Syst. Tech. J., 35, pp. 917-926, July, 1956.

[17] L. D. Landau & E. M. Lifshitz, *Statistical Physics*, Part 1, Pergamon Press, 1976.

[18] H. Nyquist, *Certain factors affecting telegraphy speed*, Bell Sys. Tech. J., 3, p324, April 1924.

[19] H. Nyquist, *Certain topics in telegraph transmission theory*, AIEE Trans., 47, p617, April 1928.

[20] A. M. Polyakov, JETP Lett. 12, p381, 1970.

[21] R. Penrose, *The Emporer's New Mind*, Oxford Univ. Press, 1989.

[22] P. Pfeuty & G. Toulouse, *Introduction to the Renormalization Group and Critical Phenomena*, John Wiley & Sons, 1977.

[23] L. H. Ryder, *Quantum Field Theory*, Camb. Univ. Press, 1991.

[24] M. Schroeder, *Fractals, Chaos, Power Laws*, W. H. Freeman & Company, 1991.

[25] C. E. Shannon, *A mathematical theory of communication*, Bell Syst. Tech. J., 27, pp. 379-423, 623-656, July, October, 1948.

[26] C. E. Shannon, *Prediction and entropy of printed English*, Bell Syst. Tech. J., 30, pp. 50-64, January, 1951.

[27] J. R. Waldram, *The Theory of Thermodynamics*, Cambridge Univ. Press, 1989.

[28] D. J. Wallace & R. K. P. Zia, *Gradient flow and the renormalization group*, Phys. Letts. 48**A**, no.5, (1974) 325.

[29] D. J. Wallace & R. K. P. Zia, *The Renormalization Group Approach to Scaling in Physics*, Rep. Prog. Phys. 41, (1978) 1-85.

[30] W. Weaver, *Recent Contributions to the Mathematical Theory of Communication*, Univ Illinois Press, 1949.

[31] A. B. Zomolodchikov, *'Irreversibility' of the flux of the renormalization group in a 2 d field theory*, JETP Lett. 43 (1986) 730.