Frustration and disorder in discrete lattice models

Jesper Lykke Jacobsen

Frustration And Disorder In Discrete Lattice Models

Jesper Lykke Jacobsen

Institute of Physics and Astronomy
University of Århus

July 1998

Abstract

Part I: Disorder

The effect of quenched impurities on systems which undergo first-order phase transitions is studied within the framework of the q-state Potts model. For large q a mapping to the random field Ising model is introduced which provides a simple physical explanation of the absence of any latent heat in two dimensions, and suggests that in higher dimensions such systems should exhibit a tricritical point with a correlation length exponent related to the exponents of the random field model by $\nu = \nu_{\rm RF}/(2 - \alpha_{\rm RF} - \beta_{\rm RF})$. A phase diagram unifying pure, percolative and non-trivial random behaviour is proposed.

In two dimensions we analyze the model using finite-size scaling and conformal invariance, and find a continuous transition with a magnetic exponent β/ν which varies continuously with q, and a weakly varying correlation length exponent $\nu \approx 1$. For $q > 4$ the first-order transitions of the pure model are softened due to the impurities, and the resulting universality class is different from that of the pure Ising model. We find strong evidence for multiscaling of the correlation functions, as expected for such random systems.

Part II: Frustration

Exact results for conformational statistics of compact polymers are derived from the two-flavour fully packed loop model on the square lattice. This loop model exhibits a two-dimensional manifold of critical fixed points each one characterised by an infinite set of geometrical scaling dimensions. We calculate these dimensions exactly by mapping the loop model to an interface model whose scaling limit is described by a Liouville field theory. The formulae for the central charge and the first few scaling dimensions are compared to numerical transfer matrix results and excellent agreement is found. Compact polymers are identified with a particular point in the phase diagram of the loop model, and the non-mean field value of the conformational exponent $\gamma = 117/112$ is calculated for the first time. Interacting compact polymers are described by a line of fixed points along which γ varies continuously.

Preface

The present volume has been submitted to the Faculty of Natural Sciences at the University of Århus in partial fulfillment of the requirements for the Ph.D. degree in physics.

The work presented here has been done in collaboration with my supervisor, Hans C. Fogedby, during my time as a Ph.D. student at the Institute of Physics and Astronomy, University of Århus, and with John L. Cardy during two visits (September 1996 to August 1997, and January to July 1998) at the Department of Theoretical Physics, University of Oxford. I have also had the pleasure of collaborating with Jané Kondev from Princeton University.

List of Publications

My scientific production during my years as a Ph.D. student has lead to the following publications (ordered according to date of submission):

1. Jesper Lykke Jacobsen and Hans C. Fogedby, *Monte Carlo study of correlations near the ground state of the triangular antiferromagnetic Ising model*, Physica A **246**, 563–575 (1997).

2. Jesper Lykke Jacobsen, *Comment on "Duality relations for Potts correlation functions" by Wu*, Physics Letters A **233**, 489–492 (1997).

3. John L. Cardy and Jesper Lykke Jacobsen, *Critical behavior of random bond Potts models*, Physical Review Letters **79**, 4063–4066 (1997).

4. Jesper Lykke Jacobsen and John L. Cardy, *Critical behavior of random bond Potts models: A transfer matrix study*, Nuclear Physics B **515** [FS], 701–742 (1998).

5. Jesper Lykke Jacobsen and Jané Kondev, *Field theory of compact polymers on the square lattice*, cond-mat/9804048. Submitted to Nuclear Physics B.

6. Jané Kondev and Jesper Lykke Jacobsen, *Conformational entropy of compact polymers*, cond-mat/9805178. Submitted to Physical Review Letters.

Paper 3–4 form the backbone of Part I of the present thesis. The content of Paper 2 is of a rather technical nature, and has been summarised in Section 4.3.2.

The results given in Part II have first been presented in Paper 5–6. Paper 1 pertains to the subject of geometrical frustration, but for reasons of continuity it has not been included in the thesis.

Acknowledgments

This work is mainly a result of the excellent advice that I have received from my two supervisors, Hans C. Fogedby and John L. Cardy. Apart from that the present thesis has benefited from the various discussions that I and my co-authors have had with M. Aizenman, J. Chalker, C. Denniston, E. Domany, Vl. S. Dotsenko, B. Duplantier, D. S. Fisher, C. L. Henley, K. Herrmann, H. Y. Huang, M. Kardar, D. P. Landau, A. Ludwig, M. Mezard, O. Mouritsen, M. Picco, V. N. Plechko, T. Prellberg, S. de Queiroz, Z. Racz, T. Spencer, F. Y. Wu and C. Zeng. J.-S. Caux read through the final draft of the thesis and contributed with many valuable comments.

Warm thanks are also due to my officemates during the years, Chris, Christa, Dirk, Francesco, Jacob, Jean-Sébastien, Jesper, Karen, Kim, Roderich and Sathya, who taught me a lot and made long afternoons seem short, and in particular to Fritz for taking me to lunch at Somerville College almost every day during my time in Oxford.

I am very grateful for the hospitality shown to me by the staff at the Institute of Theoretical Physics during my stay (April to July 1997) at the University of California at Santa Barbara, and to Beverly Potts of the International Office and Norma McManaway, secretary at Somerville College, who arranged a prolongation of my COIMBRA group status in order that I could return to Oxford in 1998 for two further trimesters.

My research during the last four years has been supported through a scholarship from Århus Universitets Forskningsfond. In addition, the research in Oxford was supported in part by the National Science Foundation under Grant PHY94-07194 and that in Santa Barbara was supported in part by the Engineering and Physical Sciences Research Council under Grant GR/J78327.

Resumoj en Esperanto[1]

Post ĉiu ĉapitro la enhavo estas mallonge resumata en la internacia lingvo. Mi ŝuldas multan dankon al Konrad Hinsen pro liaj saĝaj terminaj konsiloj, kaj al Marjorie Boulton pro la uzo de ŝia studocentro. Pri kelkaj sufiĉe specialigitaj terminoj kiuj ankoraŭ ne troviĝas en la vortaroj respondecas kompreneble mi mem. Sugestoj kaj komentoj bonvenas kaj estu sendataj retpoŝte al *lykke@dfi.aau.dk*.

[1] Each chapter is followed by a brief summary in Esperanto.

Contents

Auch kleine Dinge können uns entzücken,
Auch kleine Dinge können teuer sein.
Bedenkt, wie gern wir uns mit Perlen schmücken;
Sie werden schwer bezahlt und sind nur klein.
Bedenkt wie klein ist die Olivenfrucht,
Und wird um ihre Güte doch gesucht.
Denkt an die Rose nur, wie klein sie ist,
Und duftet doch so lieblich, wie ihr wißt.
—Paul Heyse, *Italienisches Liederbuch*

Chapter I

Introduction to scaling and renormalisation

Much effort has been dedicated to the study of phase transitions in two-dimensional lattice models. Although these models can constitute quite accurate realisations of real experimental systems exhibiting magnetic ordering, such as thin magnetic films or layered three-dimensional crystals with a large interlayer separation, we shall take the theoretician's point of view that they are physically (and mathematically) interesting on their own right. By avoiding obscuring the very simple formulation of the models with considerations on their experimental realisations we also attain a certain transparency that is highly appropriate for such versatile models, which have been applied within such different areas as nuclear physics, the social sciences [1] and, of course, condensed matter physics.

In order to give a flavour of the physics controlling the above-mentioned phase transitions we shall briefly consider the canonical example of the class of discrete lattice models under consideration, namely the celebrated Ising model. This model has discrete degrees of freedom $s_i = \pm 1$ defined on each of the vertices i of a lattice $\mathcal{L}$, which for simplicity we shall take to be a d-dimensional hypercube of side L. In accordance with the most common interpretation of the model we shall call the s_i spins; note, however, that they are completely classical, as opposed to quantum mechanical, variables. Each pair of nearest neighbour spins (s_i, s_j) interacts ferromagnetically with an exchange interaction[1] $J_{ij} \equiv J$, and with the symbol $\langle ij \rangle$ we shall designate the totality of such nearest neighbour pairs on the lattice $\mathcal{L}$. Though again a classical object, the energy functional is usually called the Hamiltonian $\mathcal{H}$, and can be written as

$$\mathcal{H} = -J \sum_{\langle ij \rangle} s_i s_j . \tag{1.1}$$

The thermal equilibrium of the system at a temperature T is described by the

[1] This nomenclature, however, reveals the quantum mechanical origin of the interaction. Indeed, for a system of two spins it is a simple matter to show that J is simply the singlet-triplet splitting, which can be conveniently cast as an exchange integral [2].

partition function

$$Z = \sum_{\{s\}} \exp\left(-\frac{\mathcal{H}(\{s\})}{k_B T}\right), \tag{1.2}$$

where $\{s\}$ designates the configuration of the spin degrees of freedom and k_B is Boltzmann's constant.

The free energy F defined by $F = -k_B T \ln Z$ can be cast as

$$F = E - TS, \tag{1.3}$$

where E is the energy and S the entropy of the system. As is well known, the free energy is minimised in thermal equilibrium. At very low temperatures ($T \to 0$) this is accomplished by minimising E, and from Eq. (1.1) we see that the ground state is two-fold degenerate according to whether all spins point up ($s_i \equiv 1$) or down ($s_i \equiv -1$). In both cases the system is in a ferromagnetic phase with all spins aligned, and since fluctuations are frozen out it has to choose *one* of the two ground states, hence breaking the spin reversal symmetry manifest in $\mathcal{H}$. On the other hand, at very high temperatures ($T \to \infty$) we have to maximise S, and therefore the spins will tend to be completely uncorrelated. The system is thus in a paramagnetic phase.

In between these two extremes it then appears that some phase transition must take place[2]. This must happen at a critical temperature T_c determined by $k_B T_c \sim J$ up to some numerical constant of order unity, J being the only energy scale in the problem. Since the transition separates the high-temperature phase with short-range correlations and the low-temperature phase where correlations extend throughout the system, it appears reasonable to suppose that exactly at T_c fluctuations occur on all length scales intermediate between the lattice constant and the size of the system L, and a thorough study of Eq. (1.2) shows that this is indeed the case. In particular, in the thermodynamic limit $L \to \infty$ fluctuations on *all* length scales occur in the critical system, and there is thus no characteristic length scale in the problem. As a consequence we expect thermodynamic quantities to exhibit scale-independent power law relationships. An example of such a relationship is the divergence of the correlation length ξ upon approach of the critical point

$$\xi \propto |t|^{-\nu}, \tag{1.4}$$

where $t \equiv (T - T_c)/T_c$ is a dimensionless measure of the deviation from T_c. Away from the critical point, on the other hand, $\xi < \infty$ serves as a non-trivial length scale of the problem, and thermodynamic quantities are expected to exhibit scale-dependent exponential relationships.

[2] If the dimensionality d is too low it may happen that the ground state is unstable with respect to even the slightest thermal fluctuation. In this case the phase transition moves down to $T = 0$. Let us for the moment assume that d is greater that the lower critical dimension $d_l = 1$.

The number ν introduced in Eq. (1.4) is an example of a so-called critical exponent associated with the phase transition. For the Ising model in $d = 2$ it can be shown that $\nu = 1$ [3]. In general, all relationships between thermodynamic quantities at a critical point are governed by a very small number (two for the Ising model) of such critical exponents. For some deep reasons, that can only be fully understood within the framework of conformal field theory [4] the exponents are typically simple rational numbers. But even then, not all combinations of rational numbers are valid sets of critical exponents. Instead one finds that the exponents are very robust to quantitative changes of the microscopic interactions present in $\mathcal{H}$. For instance, if we augmented Eq. (1.1) with a ferromagnetic interaction between next-nearest neighbours, or assigned different strengths to the vertical and the horizontal coupling constants, we would end up with exactly the same critical exponents. Each possible set of exponents defines what is known as a universality class, of which there is only a very restricted number. It is reassuring to see that the idea of universality to a large extent makes an exact quantitative knowledge of the microscopic interactions immaterial.

A profound consequence of the scale invariance at the critical point is that the system, when viewed on different length scales, looks not only qualitatively, but even quantitatively the same. This observation forms the basis of the so-called renormalisation group approach in which the nature of the large-scale fluctuations of a critical system is inferred by viewing the system on successively larger and larger length scales.

Having introduced these central ideas on phase transitions of lattice models in the context of a simple example, we shall dedicate the rest of this chapter to a more general (and formal) presentation. Since all of this material can be found in any modern textbook on the subject (see Ref. [5] for a delightful presentation) we shall be very brief and focus on those topics needed in later chapters.

1.1 Real space renormalisation

The basic assumption that the critical system looks statistically the same on all length scales gives rise to the so-called block spin construction. Useful both as a conceptual framework bringing out the ideas of scaling and universality and as an approximate calculational scheme, this construction amounts to thinning out the degrees of freedom of the model at hand and reexpressing the interesting physical quantities by means of new block degrees of freedom.

Specifically, one divides the original N-site lattice $\mathcal{L}$ into small identical blocks of b^d vertices and assigns a so-called block spin $s'_{i'} = \pm 1$ to each vertex in this "renormalised" N'-site lattice $\mathcal{L}'$, which is taken to be isomorphic to the original one in the limit $N, N' \to \infty$. The integer b is called the rescaling factor. Projecting the original spins onto the block spins is accomplished by a weight factor $P(s', s)$, which must be positive and satisfy the completeness relation $\sum_{\{s'\}} P(s', s) = 1$. The weight factors used in practice almost invariably implement either decimation or majority projection, the terminology being

self-explanatory.

The Hamiltonian $\mathcal{H}'(s')$ of the block spin system is determined by the relation

$$e^{G+\mathcal{H}'(s')} = \sum_{\{s\}} P(s', s) e^{\mathcal{H}(s)}, \tag{1.5}$$

known as the renormalisation transformation, where the quantity G is chosen so that $\sum_{\{s'\}} \mathcal{H}'(s') = 0$. As the blocking transformation is iterated, of course, the renormalised Hamiltonians will develop interaction terms of longer and longer range. A key assumption then is that the *dominant* interactions will remain short-ranged.

The renormalisation transformation (1.5) can be regarded as the definition of a flow in the space of all possible coupling constants $\{K\}$

$$\{K'\} = \mathcal{R}\{K\}. \tag{1.6}$$

Although this space is in general infinite-dimensional, the above assumption means that we can capture the physics of the problem by retaining only a finite number of terms. The feasibility of actual calculations within this scheme should now be obvious[3]. In general $P(s', s)$ must be chosen so as to render $\mathcal{R}$ an analytic function within the domain of interest.

Similarly, the flow of the specific (*i.e.*, intensive) free energy is obtained as

$$f(\{K\}) = g(\{K\}) + b^{-d} f(\{K'\}), \tag{1.7}$$

where the rescaling factor b^{-d} follows directly from the free energy's being an extensive quantity, and $g(\{K\})$ is a regular contribution arising from the trace over the short-range degrees of freedom within each block.

1.2 Fixed points and scaling fields

A fixed point $\{K^*\}$ for the flow (1.6) is defined by $\mathcal{R}\{K^*\} = \{K^*\}$. Usually there exist two trivial fixed points, since at zero (infinite) temperature ($K^* = \infty$ and $K^* = 0$, respectively) the system should retain its complete (lack of) order upon renormalisation. In the vicinity of $\{K^*\}$ we introduce the linear operator T^* through its matrix elements $T^*_{\alpha\beta} = \left.\frac{\partial K'_\alpha}{\partial K_\beta}\right|_{K^*}$ and obtain the linearised flow equations

$$K'_\alpha - K^*_\alpha = \sum_\beta T^*_{\alpha\beta}(K_\beta - K^*_\beta). \tag{1.8}$$

[3] To proceed, further approximations are necessary, though. One possible approach is to split $\mathcal{H}(s)$ of Eq. (1.5) into a sum of uncoupled cell Hamiltonians $\mathcal{H}_0$ and a remainder, treated as a small perturbation. The quantity $e^{\mathcal{H}(s)}$ is then evaluated through a cumulant expansion [6].

In terms of the left eigenvectors $\{\phi^i_\alpha\}$ of T^*, which are given by $\sum_\alpha \phi^i_\alpha T^*_{\alpha\beta} = \lambda_i \phi^i_\beta$, we define a new set of coordinates, called the scaling fields, by $u_i = \sum_\alpha \phi^i_\alpha (K_\alpha - K^*_\alpha)$. The flow (1.8) can now be rewritten as

$$u_i' = \sum_\alpha \phi^i_\alpha \sum_\beta T^*_{\alpha\beta}(K_\beta - K^*_\beta) = \lambda_i u_i, \tag{1.9}$$

i.e. the scaling fields transform multiplicatively in the vicinity of the fixed point.

Assuming the eigenvalues λ_i to be real, positive quantities we rewrite them as $\lambda_i = b^{y_i}$, where the y_i are the so-called renormalisation group eigenvalues which can be directly related to the critical exponents. A scaling field is called *relevant*, *irrelevant* or *marginal* according to whether the corresponding value eigenvalue y_i is greater than, less than or equal to 0. The rationale of this nomenclature is obvious, since an irrelevant scaling field tends to zero upon repeated renormalisation.

The region of coupling constant space where the linear approximation is valid and where all relevant scaling fields vanish is called the critical surface. By continuity, of course, we expect this surface to persist in some finite region around the fixed point, regardless of the validity of Eq. (1.9). We may say that the fixed point acts as a sink for its adjacent critical surface.

Returning now to our introductory example of the Ising model it is known that two experimental parameters, the reduced temperature t and the external magnetic field h, must be adjusted in order to render the model critical. We therefore expect there to be two relevant scaling fields, u_t and u_h. The redundancy of an next-nearest neighbour ferromagnetic interaction is born out by the fact that it corresponds to an irrelevant scaling field and vanishes under repeated renormalisation. At the same time, however, the reduced nearest neighbour coupling K flows towards its value at the critical fixed point, and we thus expect the presence of irrelevant couplings to change the critical temperature. The Ising model has one non-trivial (critical) fixed point, and in general we expect each universality class to correspond to a distinct critical fixed point.

1.3 Scaling and critical exponents

The presence of a number of relevant scaling fields along with Eq. (1.7) immediately leads to a scaling form for the free energy that will allow us to relate the critical exponents to the renormalisation group eigenvalues. In particular we shall be able to deduce certain scaling laws.

As an example we again consider the Ising model with two relevant scaling fields that by (spin-reversal) symmetry must take the form $u_t \sim t/t_0$ and $u_h \sim h/h_0$, the constants t_0 and h_0 setting the scale of the temperature and the field respectively. On a small portion of the critical surface in the vicinity of the critical fixed point, the singular part of the free energy scales like

$$f_{\rm s}(u_t, u_h) = b^{-n\,d} f_{\rm s}(b^{n y_t} u_t, b^{n y_h} u_h), \tag{1.10}$$

as is seen by iterating Eq. (1.7) n times. Here y_t and y_h are respectively the renormalisation group eigenvalues of the thermal and the magnetic scaling field, as defined above. Since the linear approximation is only valid for small values of u_t and u_h we halt the iteration when $|b^{ny_t}u_t| = u_t^0$, where u_t^0 is small but finite. Absorbing u_t^0 into a redefinition of t_0 we arrive at the fundamental scaling form

$$f_{\rm s}(t,h) = |t/t_0|^{d/y_t}\Phi\left(\frac{h/h_0}{|t/t_0|^{y_h/y_t}}\right), \tag{1.11}$$

in which the only non-universality enters through the scaling factors t_0 and h_0.

Now taking the appropriate derivatives of $f_{\rm s}$, four of the six usual critical exponents immediately follow

$$\left\{\begin{array}{lll} \text{Specific heat} & C(h=0)\propto|t|^{-\alpha} & \alpha = 2-d/y_t \\ \text{Spontaneous magnetisation} & m(h=0)\propto|t|^{\beta} & \beta = (d-y_h)/y_t \\ \text{Susceptibility} & \chi(h=0)\propto|t|^{-\gamma} & \gamma = (2y_h-d)/y_t \\ \text{Equation of state} & h\sim|m|^{\delta} & \delta = y_h/(d-y_h) \end{array}\right. \tag{1.12}$$

along with two scaling laws

$$\left\{\begin{array}{ll} \text{Rushbrooke's law} & \alpha+2\beta+\gamma = 2 \\ \text{Griffith's law} & \alpha+\beta(1+\delta) = 2, \end{array}\right. \tag{1.13}$$

named after those who rigorously established them as inequalities (with $\geq$).

The last two critical exponents pertain to the scaling of the connected two-point correlator

$$G(r_1-r_2) = \langle s(r_1)s(r_2)\rangle - \langle s(r_1)\rangle\langle s(r_2)\rangle. \tag{1.14}$$

Analogously to Eq. (1.10) this renormalises to

$$G(r,t) = b^{-2n(d-y_h)}G(r/b^n, b^{ny_t}t) \tag{1.15}$$

after n iterations, the front factor of b^{2ny_h} originating from the renormalisation of the magnetic field and that of b^{-2nd} being due to the fact that a block spin consists of b^d spins, each of which has almost identical correlation with some spin far away from that block.

Choosing n as before we arrive at

$$G(r,t) = |t/t_0|^{2(d-y_h)/y_t}\Psi\left(r|t/t_0|^{1/y_t}\right), \tag{1.16}$$

and stipulating that $G(r,t)\propto\exp(-r/\xi)$ for $r\gg 1$ the remaining critical exponents follow

$$\left\{\begin{array}{lll} \text{Correlation length} & \xi(h=0)\propto|t|^{-\nu} & \nu = 1/y_t \\ \text{Correlation function} & G(r,0)|_{h=0}\propto r^{2-d-\eta} & \eta = d+2-2y_h \end{array}\right. \tag{1.17}$$

along with two further scaling laws

$$\begin{cases} \text{Fisher's law} & (2-\eta)\nu = \gamma \\ \text{Josephson's law} & \nu d = 2-\alpha. \end{cases} \tag{1.18}$$

The latter scaling law is particularly interesting because it involves the dimensionality of the system. For that reason it is called a hyperscaling relation, and we shall later see that this relation may fail when a dangerous irrelevant variable (such as the temperature at a zero-temperature fixed point) modifies the scaling form of the free energy but not that of the two-point correlator.

Resumo en Esperanto

La Ising-modelo (1.1) estas la klasika ekzemplo de du-dimensia modela sistemo havanta duaordan faztransiron. Je altaj temperaturoj la entropio dominas kaj la spinoj estas en paramagneta fazo, dum je malaltaj la sistemo elektas unu el la du statoj kun kompleta magneta ordo (spontanea rompo de simetrio). Inter tiuj du limoj estas faztransiro kiun karakterizas fluktuoj je ĉiuj longoskaloj kaj nevario je renormigaj transformoj (1.6). Kritaj eksponentoj priskribas la makroskopan konduton de la sistemo; tiuj ĉi ne dependas de *kvantaj* mikroskopaj detaloj kaj plenumas skalumajn leĝojn (1.12 kaj 1.17).

Part I

Disorder

Chapter II

Random fields and random bonds

In real life no material is perfectly homogeneous, and to attain a better description of magnetic systems we must take into account the possibility of having some kind of impurities distributed randomly throughout the material. In particular, for a pure system exhibiting a phase transition, the imposition of such randomness may give rise to a range of different scenarios.

Since impurities will inevitably tend to disorder the system we expect the mildest possible effect to be a lowering of the critical temperature. For a second-order phase transition this effect may or may not be accompanied by a change of the universality class. Whenever such a change occurs the randomness is called *relevant*, and the critical behaviour of the random system must, from a renormalisation group point of view, be governed by a new fixed point, which in the limit of vanishing randomness must somehow merge with the fixed point controlling the critical behaviour of the pure system. This observation can be systematised and used to perturbatively calculate the change in the critical exponents [5].

A more dramatic consequence is encountered when the randomness changes the order of the phase transition. Intuitively it seems reasonable that inclusion of randomness in a system exhibiting a first-order phase transition may under some circumstances smooth out the discontinuities of the thermodynamic quantities across the phase boundary, thus resulting in a second-order phase transition. The investigation of this issue forms a main theme of Part I of the present thesis.

Certain severe types of randomness may drive the transition temperature all the way down to zero, or even eliminate the phase transition altogether. These eventualities can be detected by investigating the stability of the ordered phase upon imposition of the randomness at zero temperature.

Clearly, in order to further discuss the issue of impurities we must decide how to implement the effect of the randomness in the theoretical models describing the pure system. Depending on the particular physical situation at hand this may lead us to consider random bonds (site dependent coupling constants, $J \to J_{ij}$), random fields ($h \to h_i$), and even models defined on graphs with random

coordination number or subject to dynamical triangulations (random gravity graphs).

In all cases a first distinction to be made is that between quenched and annealed randomness. In the latter case we consider the magnetic spins and the impurities to be in thermal equilibrium. This is an appropriate description if we want to study the system on time scales during which the impurities are free to wander away from their original positions, and accordingly the partition function is formed by tracing over both the magnetic degrees of freedom and the positions of the impurities. Random gravity is an example of a problem with annealed randomness.

For most systems that have a straightforward interpretation in terms of real magnetic materials, however, the time scale associated with the motion of the impurities is so huge that their position can essentially be regarded as fixed. This is the case of quenched randomness to which we dedicate the remainder of this chapter.

2.1 Quenched bond randomness

Let us consider a magnetic material with impurities which is initially held at some very high temperature where the impurities are in thermal equilibrium with the magnetic degrees of freedom. Now imagine abruptly lowering the temperature, thus quenching the material to a state where the impurities are positionally fixed whilst the magnetic ions organise themselves so as to attain thermal equilibrium at the new temperature. In this situation the Hamiltonian $\mathcal{H}(\{s\},\mathcal{D})$ is a function of both the ionic spins $\{s\}$ and the fixed disorder configuration $\mathcal{D}$. The appropriate partition function is

$$Z(\mathcal{D}) = \mathrm{Tr}_s \mathrm{e}^{-\mathcal{H}(\{s\},\mathcal{D})} \tag{2.1}$$

and depends explicitly on the disorder.

For illustrative purposes we shall consider the simple example of the random-bond Ising chain [7] with $\mathcal{H} = -\sum_{i=1}^{N} K_i s_i s_{i+1}$. The quenched disorder configuration is here specified by the values of the random exchange constants, $\mathcal{D} = \{K\}$, and the partition function is easily found to be

$$Z = \prod_{i=1}^{N} c_i, \tag{2.2}$$

where $c_i = 2\cosh K_i$.

2.1.1 Self-averaging

At this point we must confront the problem that even though the probability distribution $P(\mathcal{D})$ presumably exhibits translational invariance, any particular realisation of the randomness drawn from it does not. In particular, local

physical quantities computed from a fixed randomness realisation are not translationally invariant. In order to proceed we must somehow average over the disorder, but the partition function Z is not the appropriate place to do so since this would correspond to the case of annealed randomness.

Instead, the average must be performed on the level of the free energy $F(\mathcal{D}) = -\ln Z(\mathcal{D})$. This comes about in the following way. Imagine dividing a huge sample into many macroscopically large domains. For each domain the disorder configuration, which we think of as drawn from some ensemble governed by the probability distribution $P(\mathcal{D})$, differs, but since the free energy of the entire sample is equal to the sum of the free energies for each domain (apart from surface effects which are irrelevant in the limit of sufficiently large domains) the specific free energy in the thermodynamic limit will, according to the central limit theorem, be equal to the free energy averaged over the ensemble. The physics of the disordered system is thus described by suitable derivatives of the quenched average free energy

$$\overline{F} = \mathrm{Tr}_{\mathcal{D}} P(\mathcal{D}) F(\mathcal{D}), \tag{2.3}$$

where the overline is the standard notation for a disorder average. In particular we note that translational invariance has been restored to the problem.

The fact that F is a *self-averaging* quantity in the sense just described is nicely illustrated within the context of our simple example, where the free energy is

$$F = -\sum_{i=1}^{N} \log c_i. \tag{2.4}$$

It is evident that the quenched average of this quantity also describes the behaviour of a typical system with some definite realisation of the randomness, since this average coincides with the most probable value: $\overline{F} = F_{\mathrm{m.p.}} = -N\overline{\log c_i}$. As explained above this property is exactly accounted for by the fact that F is a *sum* of random numbers and hence normally distributed. Relative fluctuations away from this result hence die out like $1/\sqrt{N}$ as $N \to \infty$, N being the (one-dimensional) volume of the system.

In the case of the partition function the situation is different. Since Z is a *product* of random numbers it is *log*-normally distributed, and in particular the quenched average $\overline{Z} = \exp[N \log \overline{c_i}]$ does *not* coincide with the most probable value $Z_{\mathrm{m.p.}} = \exp[N\overline{\log c_i}]$. In other words, Z is a non-self averaging quantity. Now, normally Z is not a physically observable quantity, but correlation functions certainly are. The importance of the lack of self-averaging can therefore be appreciated by noting that the two-point correlator

$$\langle s_1 s_R \rangle = \prod_{i=1}^{R-1} \tanh K_i \tag{2.5}$$

is again a product of random numbers and hence non-self averaging. In particular $\overline{\langle s_1 s_R\rangle^2} \neq \overline{\langle s_1 s_R\rangle}^2$, a remarkable occurrence that leads to *multiscaling*, and to which we shall return in later chapters.

2.1.2 The replica method

Taking the disorder average on the level of the free energy and not in Z constitutes a technical complication which can be circumvented by means of the so-called replica method. This relies on the identity

$$\ln Z = \lim_{n\to 0} \frac{Z^n - 1}{n}, \tag{2.6}$$

which is easily proved by expanding the right-hand side as a power series in n. Thus the disorder average *can* be performed at the level of Z, but a the price of having to raise it to the n-th power before doing the average.

For positive integral n the replica method is tantamount to replicating the system n times, thus generalising the spin degrees of freedom to $s_{a,i}$, where the index $a = 1, 2, \ldots, n$ runs over the replicas. Each replica is subject to the same disorder configuration. Then

$$\overline{Z^n} = \mathrm{Tr}_{\{s_a\}} \mathrm{Tr}_{\mathcal{D}} P(\mathcal{D}) \exp\left(-\sum_{a=1}^{n} \mathcal{H}(\{s_a\}, \mathcal{D}) \right), \tag{2.7}$$

and since it is usually easy to perform the $\mathrm{Tr}_{\mathcal{D}}$ operation (e.g. using a cumulant expansion) one is left with a customary trace over the replicated spins. The price for doing this is of course that the different replicas are now coupled.

Finally, the limit $n \to 0$ is taken by analytic continuation. This is normally done without much ado, but when dealing with spin glasses complications may arise, leading to the eventuality of replica symmetry breaking. For the disordered ferromagnets which are our prime concern no such problem occurs, and the $n \to 0$ limit is trivially taken at the end of the calculations.

2.2 The Harris criterion

The effect of quenched bond randomness on a classical statistical mechanics system whose pure version undergoes a second-order phase transition is well understood. Namely, the so-called Harris criterion states that if the critical exponent α^{pure} governing the divergence of the specific heat at the transition point of the *pure* system is negative, weak bond randomness is irrelevant in the renormalisation group sense and the pure fixed point is stable [8]. On the other hand, if $\alpha^{\mathrm{pure}} > 0$ the randomness is relevant and causes a crossover to critical behaviour governed by a new random fixed point nearby, at least if the crossover exponent α^{pure} is small.

To prove this statement we consider any system in which there is some randomness coupling to the local energy density. Apart from the case of quenched

bond randomness this also addresses site dilution, where a certain fraction of the magnetic spins are replaced by vacancies (at quenched random positions). Letting $\mathcal{H}^*$ designate the Hamiltonian at the fixed point of the pure system, and passing on to the continuum limit for notational beauty, we thus have

$$\mathcal{H} = \mathcal{H}^* + \int \mathrm{d}^d x\, m(\mathbf{x})\epsilon(\mathbf{x}), \tag{2.8}$$

where $m(\mathbf{x})$ is the field due to the random impurities, and $\epsilon(\mathbf{x})$ is the local energy density.

Using the replica method, the trace over the disorder in the interaction part of $\overline{Z^n}$ can be evaluated by means of a cumulant expansion

$$\begin{aligned} &\mathrm{Tr}_{\mathcal{D}}\, P(\mathcal{D}) \exp\left(-\sum_{a=1}^{n} \int \mathrm{d}^d x\, m(\mathbf{x})\epsilon_a(\mathbf{x})\right) \\ &= \exp\left(-\overline{m}\sum_{a=1}^{n}\int \mathrm{d}^d x\, \epsilon(\mathbf{x}) + \frac{1}{2}g \sum_{a,b=1}^{n}\int \mathrm{d}^d x \int \mathrm{d}^d x'\, \epsilon_a(\mathbf{x})\epsilon_b(\mathbf{x}') + \cdots\right), \end{aligned} \tag{2.9}$$

where the second cumulant of $m(\mathbf{x})$, $g \equiv (\overline{m(\mathbf{x})m(\mathbf{x}')} - \overline{m}^2)$, plays the role of a coupling constant determining the interaction between the replicas. The higher-order contributions on the right-hand side can be seen by power counting to generate only less relevant modifications to what has already been written out.

The important new term is the one coupling the energy densities of two *different* replicas, since the term with $a = b$ just contains the energy density itself in its operator product expansion.[1] Together with the term proportional to $\overline{m}$ this diagonal term leads to a shift of the critical temperature, which in this context is trivial. The issue of whether $m(\mathbf{x})$ is a relevant perturbation is thus settled by examining the relevance of $\sum_{a\neq b} \epsilon_a(\mathbf{x})\epsilon_b(\mathbf{x})$. By performing the relevant contractions its two-point function is easily calculated

$$\left\langle \sum_{a\neq b} \epsilon_a(\mathbf{x})\epsilon_b(\mathbf{x}) \sum_{a'\neq b'} \epsilon_{a'}(\mathbf{x}')\epsilon_{b'}(\mathbf{x}') \right\rangle = 2n(n-1)\langle \epsilon_a(\mathbf{x})\epsilon_a(\mathbf{x}')\rangle^2, \tag{2.10}$$

and we conclude that its scaling dimension is twice that of the energy operator.

The renormalisation group eigenvalue of the most relevant term generated by the random impurities is therefore

$$y = d - 2x_\epsilon = d - 2(d - y_t) = 2/\nu - d, \tag{2.11}$$

and randomness is irrelevant if $y < 0$, or, since the pure system can usually be assumed to exhibit hyperscaling, $\alpha < 0$.

[1] In conformal field theory, the product of a number of operators evaluated a nearby points can be replaced by a linear combination of the scaling operators in the theory, at least as far as the evaluation of its correlations with a product of other, distant local operators is concerned.

2.3 Quenched field randomness

The above discussion on quenched bond randomness should be contrasted with the more dramatical effects of randomness in the field conjugate to the local magnetisation. Such randomness can eliminate low-dimensional phase transitions altogether, and at least it always changes the values of the critical exponents [9]. For this reason most early research concentrated on field randomness.

In this context a particularly popular model is the random field Ising model (RFIM) defined by the Hamiltonian

$$\mathcal{H}^{\mathrm{RFIM}} = -J \sum_{\langle ij \rangle} s_i s_j - \sum_i h_i^{\mathrm{RF}} s_i - h \sum_i s_i \,, \tag{2.12}$$

where the $h_i^{\mathrm{RF}} = \pm h^{\mathrm{RF}}$ are quenched random variables with mean $\overline{h} = 0$. The phase transition between the ordered and the disordered phase occurs for vanishing uniform magnetic field, $h = 0$.

Repeating the argument given above for the case of bond randomness we see that the most relevant perturbation due to the randomness is of the form $\sum_{a \neq b} s_a(\mathbf{x}) s_b(\mathbf{x})$, which at the pure fixed point has renormalisation group eigenvalue

$$y = d - 2x_h = 2y_h - d = \gamma y_t \,. \tag{2.13}$$

Since t is a relevant scaling field in the pure Ising model and $\gamma > 0$ for any phase transition, we conclude that random fields are always relevant.

2.3.1 The Imry-Ma argument

A classical argument due to Imry and Ma [10] predicts that the lower critical dimension of the RFIM is $d_{\mathrm{l}} = 2$. This argument, in which a comparison is made between the field fluctuations and the stabilising effect caused by the formation of a domain wall, turns out to be applicable for discussing the effect of bond randomness imposed on a first-order phase transition as well, so we will repeat it in its most basic form here. In a slightly improved version the Imry-Ma argument can also be used to deduce the renormalisation group equations for an RFIM interface, which is going to be an essential ingredient of the next chapter.

To find out the value of d_{l} for the RFIM we study the model at zero temperature in arbitrary dimension d. The question to be settled is whether the ordered ground state (say, $s_i \equiv 1$) is stable to the disordering effect of the fluctuating random field. If not, it will be energetically advantageous to introduce very large domains of the opposite phase ($s_i = -1$), and the ordered state will be destroyed.

Imagine introducing such a region R of linear dimension L. This will of course lead to an energy penalty $2JL^{d-1}$ due to the formation of a domain wall. The random term, on the other hand, changes the energy by an amount

$2\sum_{i\in R} h_i$ which by a judicious choice of R can be made negative[2]. Indeed, since this is a sum of random variables it is normally distributed in the limit of large L. In particular the width of the distribution is $(\overline{h^2}L^d)^{1/2}$, and the proper choice of R means that the energy change will be of the order *minus* this amount. For $d < 2$ and large enough L this outweighs the domain wall term, and we conclude that the ordered state is unstable, whilst for $d > 2$ the opposite is true. Therefore $d_l = 2$.

2.4 Discontinuity fixed points

Consider the RFIM with $h = 0$ at zero temperature. For $h^{\mathrm{RF}} < zJ$, where z is the coordination number of the lattice, the aligning tendency of neighbouring spins is stronger the the disordering effect of the random field. Since we are at $T = 0$ thermal fluctuations are frozen out, and the ground state is the completely ordered configuration $s_i \equiv 1$. For $h^{\mathrm{RF}} > zJ$, on the other hand, the random field is strong enough to break the exchange coupling and since, once again, entropic effects are ruled out the ground state is the disordered configuration in which each spin is aligned with the random field at that site, $s_i = h_i^{\mathrm{RF}}/h^{\mathrm{RF}}$.

It is thus seen that at $h^{\mathrm{RF}} = zJ$ a phase transition must take place. From the above argument it follows that the specific heat diverges like $|h^{\mathrm{RF}} - zJ|^{-1}$. Furthermore, since both the magnetisation and the derivative of the energy with respect to h^{RF} change discontinuously, the transition is of the first order. Interestingly, since there are no fluctuations the correlation length is zero.

Now imagine gradually increasing the temperature. Since thermal fluctuations constitute a further disordering tendency the transition is expected to take place at lower and lower values of h^{RF}. Eventually, above the critical temperature T_{c}, thermal effects are so strong that even the pure system ($h^{\mathrm{RF}} = 0$) is disordered, and the phase transition disappears. These facts are presented in the schematic phase diagram given in Fig. 2.1. Here RF is the fixed point controlling the zero-temperature phase transition, whilst I is the usual critical point of the pure Ising model.

It remains to discuss the topology of the projected RG flows, and to that end we set out discussing the fixed point RF. Let us consider the scaling variable $w \equiv h^{\mathrm{RF}}/J - (h^{\mathrm{RF}}/J)^*$, where $(h^{\mathrm{RF}}/J)^*$ is the value of h^{RF}/J at RF, and define its RG eigenvalue to be y. Analogously to the derivation of the critical exponent α we expect the specific heat to diverge like $|w|^{d/y-2}$ as $w \to 0$, and comparing that to the exact result $C \sim |w|^{-1}$ we deduce that

$$y = d, \tag{2.14}$$

meaning that w is strongly relevant at the transition. This is a very general feature of zero-temperature fixed points. To check this result we remark that

[2] In practice, of course, R will automatically adjust itself so as to minimise the energy. A thorough analysis of exactly how this happens leads to the RG equations for the interface, and will be presented in Chapter 3.

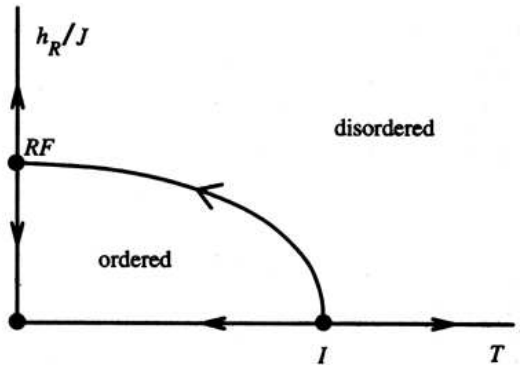

Figure 2.1: Schematic phase diagram and projected renormalisation group flows for the random-field Ising model in dimension $d = 2 + \epsilon$. RF is the zero-temperature fixed point which leads to a violation of hyperscaling, and I is the standard pure Ising fixed point. Taken from Ref. [5].

it implies $x = d - y = 0$, whence the correlation functions are supposed to be trivial. But this is exactly what one would expect for a phase transition where fluctuations are absent.

The fixed point RF is known as a *discontinuity fixed point*, since upon crossing the phase boundary the variable conjugate to w changes discontinuously. A further characteristic of this fixed point is that the temperature T is irrelevant. To see this we note that as $J \sim 1/T \to \infty$ the flow equation giving the change of J upon rescaling the system with a linear factor of b is simply

$$J' \sim b^{d-1} J, \tag{2.15}$$

since near $T = 0$ we expect the spins to be completely aligned. In other words, the thermal exponent $y_T = 1 - d$, and temperature is irrelevant for any $d > 1$.

To summarise, we have found that at RF w is relevant and T is irrelevant. On the other hand, since I is just the ordinary Ising fixed point we know that T is relevant there, and the projected RG flows given in Fig. 2.1 follow.

2.4.1 Modified scaling at a zero-temperature fixed point

For a critical fixed point situated at $T = 0$ the scaling arguments for the free energy given in Sect. 1.3 are no longer appropriate, and as a consequence the usual scaling laws have to be modified [11]. This is due to an additional rescaling of the exchange constant J, which would be a constant at a *thermal* fixed point. At a zero-temperature fixed point, however, it is the ratio h^{RF}/J that is constant.

Under a scale transformation $\mathbf{x} \to \mathbf{x}' = b\mathbf{x}$ the parameters h, J and w become

$$h' = b^{y_h} h, \quad J' = b^{\theta} J, \quad w' = b^{y_t} w, \tag{2.16}$$

defining the quantities y_h, θ and y_t. The notational clash between y_t thus defined and the customary thermal exponent is not coincidental: as we shall soon see w is completely analogous to a temperature variable in the modified scaling relations. Inserting into the expressions for the singular part of the energy density $e(w, hJ^{-1}) = J^{-1}E_{\mathrm{s}}$ and the correlation length $\xi(w, hJ^{-1})$, and using the same tricks as in Sect. 1.3, we arrive at the scaling forms

$$e(w, hJ^{-1}) \propto w^{2-\alpha}\tilde{\Phi}(hJ^{-1}/w^{\Delta}), \tag{2.17}$$

$$\xi(w, hJ^{-1}) \propto t^{-\nu}\Xi(hJ^{-1}/w^{\Delta}), \tag{2.18}$$

where

$$\nu = 1/y_t, \quad \Delta = (y_h - \theta)/y_t, \quad \alpha = 2 - (d - \theta)/y_t. \tag{2.19}$$

Defining the magnetisation and the susceptibility as appropriate derivatives of E_{s}, which replaces the customary f_{s}, we can furthermore deduce the following critical exponents:

$$\beta = (d - y_h)/y_t, \quad \gamma = (2y_h - \theta - d)/y_t, \quad \delta = (y_h - \theta)/(d - y_h). \tag{2.20}$$

Until now everything looks much like in Sect. 1.3, apart from the fact that w takes over the role of the temperature variable, and that the rescaling of J occasionally modifies the expressions for the critical exponents through the presence of θ.

When we turn to the scaling forms for the correlation function it is found that its connected and disconnected parts have to be investigated separately:

$$G(r) = \overline{\langle s_0 s_r\rangle - \langle s_0\rangle\langle s_r\rangle} = r^{2-d-\eta} g(r/\xi), \tag{2.21}$$

$$G_{\mathrm{dis}}(r) = \overline{\langle s_0\rangle\langle s_r\rangle} = r^{4-d-\tilde{\eta}} g_{\mathrm{dis}}(r/\xi). \tag{2.22}$$

However, the appearance of *two* anomalous dimensions

$$\eta = d + 2 + \theta - 2y_h, \quad \tilde{\eta} = d + 4 - 2y_h \tag{2.23}$$

should hardly come as a surprise, since starting out from three scaling fields (h, J and w) will necessarily lead to three independent exponents instead of the usual two. The modified scaling laws can now be written as

$$\begin{aligned} \alpha + 2\beta + \gamma &= 2, & \nu(2 - \eta) &= \gamma, \\ \alpha + \beta(1 + \delta) &= 2, & 2 - \alpha &= (d - \theta)\nu, \end{aligned} \tag{2.24}$$

and the only modification is the *violation of hyperscaling* effected by θ.

2.4.2 T as a dangerous irrelevant variable

A very interesting situation occurs when, as is the case in the RFIM, $\theta < 0$. From the scaling treatment just given we see that even though T is irrelevant in the RG sense it still leads to a violation of hyperscaling. In this situation T acts as a *dangerous irrelevant variable*.

By definition an irrelevant variable u is dangerous if the thermodynamic functions depend on it in a singular manner as $u \to 0$. As a result u may not simply be set equal to zero in calculations, and the way that the free energy, and possibly other quantities, depend on a small but finite u should be thoroughly investigated.

Apart from the temperature in the RFIM another celebrated example of a dangerous irrelevant variable is that of the quartic coupling in the Landau-Ginzburg-Wilson model above its upper critical dimension $d_{\rm u} = 4$ [5]. This situation is quite general. Above $d_{\rm u}$ the critical exponents are known to be constant and assume their mean-field values. On the other hand, since one of the scaling relations depends continuously on the dimensionality there has to be some violation of hyperscaling in order to keep α and ν constant as functions of d.

2.5 Softening of first-order phase transitions

Though the Harris argument provides a simple criterion for the relevance of quenched bond randomness on a second-order phase transition the issue of such randomness imposed on a system that undergoes a thermal *first*-order phase transition has been less studied. An adaptation of the Imry-Ma argument can be established by noting that the bond randomness couples to the local energy density, which differs for the two phases that co-exist at the critical point of the pure system, in exactly the same way that the random field couples to the local magnetisation in the RFIM. Consequently the existence of a non-vanishing latent heat for $d < 2$ can be ruled out. Early work by Imry and Wortis [12] furnished a heuristic argument, reminiscent of that of the Harris criterion, that the bond randomness indeed softens any such phase transition in $d = 2$ to a continuous one. A subsequent phenomenological RG argument by Hui and Berker [13] confirmed that the lower critical dimension for random-bond tricriticality and end-point criticality is $d_{\rm l} = 2$. As the dimensionality increases, tricritical points and critical end points emerge from $T = 0$. Finally, a mathematically rigorous theorem by Aizenman and Wehr [14] stated quite generally that for $d \leq 2$ an *arbitrarily* weak amount of quenched bond randomness leads to the elimination of any discontinuity in the density of the variable conjugate to the fluctuating parameter.

The question then emerges whether this softening of the phase transition can be verified for specific models and, if so, what are the universality classes of these novel second-order phase transitions. An investigation along these lines has recently been initiated by Cardy [15], by considering a system of N two-

dimensional Ising models coupled by their energy operators which, according to mean-field theory (MFT), is supposed to display a second-order phase transition. For $N > 2$, however, the RG flow of the model exhibits a runaway behaviour, which is characteristic of a fluctuation-driven first-order transition [5]. In this sense the transition is only weakly first-order and hence amenable to perturbative calculations. On adding weak bond randomness it was found that the RG trajectories curl back towards the pure decoupled Ising fixed point, and consequently Ising exponents are expected, up to possible logarithmic corrections. This study was extended by Pujol [16] to the case of N coupled q-state random-bond Potts models for $2 \leq q \leq 4$, but here the universality class of the impurity-softened transition was found to depend on the coupling between the models.

2.5.1 The random-bond Potts model

A more interesting model for studying the effect of quenched bond impurities on a first-order transition is the q-state random-bond Potts model (RBPM). This generalisation of the Ising model is defined by the reduced Hamiltonian [17]

$$\mathcal{H} = -\sum_{\langle ij \rangle} K_{ij} \delta_{\sigma_i \sigma_j}, \tag{2.25}$$

where the spins, situated on the vertices of the square lattice, can now take the values $\sigma_i = 1, 2, \ldots, q$ (in the Ising model $q = 2$), and the summation is over all nearest neighbour bonds on the lattice. We shall specialise to the ferromagnetic case, where the reduced couplings $K_{ij} \geq 0$ measure the strength of the aligning tendency of nearest-neighbour spins.

Although the free energy of the pure model ($K_{ij} \equiv K$) is not known in closed form for general q, a wide range of exact results is nevertheless available [18]. In particular it is well-known that the model exhibits a second-order phase transition for $q \leq 4$ [19] and both the critical exponents and the central charge are known exactly from conformal field theory. For $q > 4$ the phase transition is first-order with a latent heat that is an increasing function of q [19]. In fact, since the transition is first-order already in MFT, on the RG level it is controlled by a zero-temperature discontinuity fixed point with the eigenvalue of the relevant scaling operator being $y = d$ [5]. Quenched randomness coupling to the local energy density thus has the eigenvalue $d - 2(d - y) = d$ and is strongly relevant, whence an RG treatment appears to be problematic.

The work undertaken until now has therefore mainly been numerical. Extensive Monte Carlo (MC) simulations have been carried out for $q = 8$ by Chen, Ferrenberg and Landau [20] confirming the transition softening scenario outlined above, and finding critical exponents numerically consistent with those of the *pure* Ising model. Similar conclusions were reached by Wiseman and Domany [21] for $q = 4$ and also for the Ashkin-Teller model. It thus appears that in a variety of situations the universality class of the bond disordered models is that of the Ising model, *irrespective* of the symmetry underlying the original model.

To explain these findings Kardar *et al.* [22] have proposed an interface model for the RBPM which, after several approximations, is amenable to an RG treatment that is exact on the hierarchical lattice. In the pure model the interface exhibits a branching structure with fractal dimension at criticality, but when randomness is present the critical interface is asymptotically linear. Assuming that the vanishing of the interfacial free energy is governed by a zero-temperature fixed point, the Widom exponent μ turns out to be *independent* of q for all sufficiently large q, taking the Ising value $\mu = 1$.

This is in contrast to the perturbative expansion in powers of $(q-2)$ investigated by Ludwig and Cardy [23], Ludwig [24, 25], and Dotsenko *et al.* [26]. Using the RG approach for the perturbation series around the conformal field theories representing the pure models, these authors find the critical behaviour of the RBPM to be controlled by a new random fixed point which merges with the pure fixed point as $q \to 2$. Critical exponents are found to depend continuously on q, at least for $(q-2)$ small, and in the case of the magnetic exponent x_h a calculation to three loop order yields a prediction which is supposed to be very precise even up to $q = 3$ [26]. Unfortunately, extending these results beyond $q = 4$ is impossible, even in principle, since this is the limiting case in the range of minimal conformal theories around which the perturbative calculations take place. Another interesting implication of this line of research is that the local operators exhibit *multiscaling* [25], meaning that correlation functions of different moments of such operators decay with powers that are, in general, independent.

It has been suggested by Kardar *et al.* [22] and Cardy [15] that these contrasting theories describe very different fixed points. Indeed, it can be argued that the interface model pertains to the case of strong non-self dual randomness, whilst the $(q-2)$-expansion is relevant for weak self-dual randomness. Also, even though it may turn out that the critical exponents do not depend on q, the central charge c evidently must, since even when the critical behaviour is controlled by a decoupled Ising fixed point there is generally not just *one* Ising model but several.

The work described in Part I of this thesis was originally motivated by the wish to resolve this controversy. The outcome is given by our numerical results presented in Chapter 6. We shall see that although the critical exponent ν only depends weakly on q and is numerically consistent with its Ising value, the magnetic exponent $x_h = \beta/\nu$ exhibits a pronounced q-dependence. In particular, the results of Ref. [20] for $q = 8$ must now be dismissed, since both our results and two subsequent Monte Carlo studies [27, 28] are consistent with $x_h(q = 8) \simeq 0.15$, far away from the Ising value.

Further evidence for this conclusion is furnished by our proposed phase diagram, to which we turn next.

Resumo en Esperanto

Parto I de la jena tezo celas eltrovi kiel malpuraĵoj influas unuaordan magnetan faztransiron. Ĉefe interesas nin la $q > 4$ stata Potts-modelo (2.25) kun aleatoraj kuplaĵoj. Por $q < 2$ la malpuraĵoj nur malaltigas la kritan temperaturon, aŭ, per aliaj vortoj, la malordo ne modifivas. En la intervalo $2 < q \leq 4$ ŝanĝiĝas ankaŭ la kritaj eksponentoj (modifiva malordo), kaj kiam $q > 4$ la unuaorda transiro de la pura modelo fariĝas duaorda kun novaj eksponentoj kalkulendaj—ĉu ili vere estas Isingaj, kiel asertis Ref-o [20]?

Pli drastas la efiko de aleatora magneta kampo, kiu ĉiam modifivas. Ĝi kaŭzas malkontinuan fikspunkton je temperaturo nula. Pro renormiĝo de la kuplaĵo, la temperaturo rolas kiel danĝera nemodifiva variablo kaj kondukas al malobeo de hiperskalumo.

Chapter III

Phase diagram

After these preliminaries we now propose a phase diagram for the random-bond Potts model, based on a mapping between the interfacial models governing the random cluster model in the limit of an infinite number of states ($q \to \infty$) and the random-*field* Ising model in the zero-temperature limit ($T \to 0$). The portion of the phase diagram near $q = \infty$ can then be inferred from the known renormalisation group behaviour of the Ising interface [11], whereas the properties for small q are furnished by a perturbative expansion around the point where the bond randomness is marginal [24, 25, 26]. Taking into account exact results valid in the limits of zero [19] and infinite [29] randomness the global structure of the renormalisation group flows is then established.

From the explicit relations between the scaling variables in the two models we are also able to establish two exact relations between the respective critical exponents, thus determining the universal behaviour of the random-bond Potts model in terms of that of the random-field Ising model.

The discussion of the numerical evidence supporting the correctness of the proposed phase diagram, at least for $d = 2$, is the subject of Chapter 6.

3.1 Mapping between interface models

As will become clear, many of our results generalise, but let us for definiteness consider a Potts model on the square lattice with degrees of freedom σ_i taking q values, and a reduced Hamiltonian

$$\mathcal{H} = -\sum_{\langle ij \rangle} K_{ij} \delta_{\sigma_i \sigma_j}, \tag{3.1}$$

where the sum is over nearest neighbour pairs. The ferromagnetic couplings $K_{ij} > 0$ are quenched random variables, taking the values K_1 and K_2, each with probability one half

$$P(K_{ij}) = \frac{1}{2}[\delta(K_{ij} - K_1) + \delta(K_{ij} - K_2)]. \tag{3.2}$$

For the special choice

$$\left(\mathrm{e}^{K_1}-1\right)\left(\mathrm{e}^{K_2}-1\right)=q \tag{3.3}$$

this model is on average self-dual, as discussed in more detail in Chapter 4. Assuming that the phase transition is unique the model is therefore at its critical point [30]. It is useful to parametrise

$$\mathrm{e}^{K_{ij}}-1=u_{ij}=q^{\frac{1}{2}+w_{ij}}, \tag{3.4}$$

where $w_{ij}=\pm w$, and $w>0$ measures the strength of the randomness.

The partition function of this model may be mapped onto that of the random cluster model [31], in which each bond of the lattice is either occupied, when it is counted with weight u_{ij}, or empty, in which case it is counted with weight 1. The partition sum is over all such configurations, in which each connected cluster of sites is weighted by a factor q

$$Z=\sum_{G}\left(\prod_{\langle ij\rangle\in G}u_{ij}\right)q^{C}. \tag{3.5}$$

Here C is the total number of such clusters in the graph configuration G. Note in particular that a single isolated site is to be counted as a cluster on its own.

3.1.1 The pure model

Let us first consider the pure model, with $w=0$. In the limit $q\to\infty$, the sum over configurations is dominated by only two: the *empty* lattice, in which no bonds are occupied, which contributes a factor q^N, where N is the total number of sites, and the *full* lattice, with a weight $q(\sqrt{q})^{2N}$, since the number of bonds per site is exactly two in the limit of an infinite system. The extra factor of q, originating from the number of clusters in the full lattice being one, is attributable to the fact that the full lattice really corresponds to q distinct contributions, one for each possible value of the order parameter in the completely ordered state.

All other configurations are down by powers of $q^{\frac{1}{2}}$. At the self-dual point, there are therefore two coexisting states, namely the ordered phase corresponding to the full lattice and the disordered phase corresponding to the empty lattice. These phases have identical bulk free energy, since their weight in the partition function is the same, but different internal energy densities of $-K$ and 0 respectively. This indicates, as expected, that the transition has a non-vanishing reduced latent heat per bond of

$$L=K=\ln(\sqrt{q}+1)\sim\frac{1}{2}\ln q. \tag{3.6}$$

For the pure model, this analysis may be extended to take into account higher order corrections in $1/\sqrt{q}$, with no essential change in the physical picture.

The enumeration of such sub-dominant configurations has recently been carried through to 10th order in this small parameter [32].

Now consider an *interface* between these two phases. For large q, the lowest energy interface is parallel to a lattice direction, say the x-axis, and is such that all the bonds with $y \leq$ some integer are occupied, and those above this are empty (or vice versa). There will also be entropic fluctuations $y = h(x)$ of this interface, described by the usual solid-on-solid interfacial Hamiltonian, proportional to the length of the interface.

In order to determine the interfacial tension it is useful to convert the equality of the weight factors for the dominant contributions into a local balance criterion. Consider therefore the weight factor of the full lattice relative to that of the empty lattice. Each site in the full lattice constitutes the loss of a potential cluster, and thus carries a relative weight of $1/q$. This can be distributed as a factor of $q^{-\frac{1}{4}}$ to each of the four adjacent bonds. On the other hand, these bonds each carry a relative weight of $\sqrt{q}$, which we distribute as a factor of $q^{+\frac{1}{4}}$ to each end of the bond. At every bulk site inside a portion of the full lattice there is therefore an exact balance due to the cancellation of these factors of $q^{-\frac{1}{4}}$ and $q^{+\frac{1}{4}}$.

This balance is not satisfied at the interface. Neglecting for the moment the possibility of overhangs and of 'bits sticking out' (*i.e.*, assuming all the square lattice plaquettes in the full phase to be surrounded by four occupied bonds) it is easily found that the number of excess factors of $q^{-\frac{1}{4}}$ is precisely equal to the length of the interface. The interfacial tension for large q is therefore

$$\sigma \sim \frac{1}{4} \ln q, \tag{3.7}$$

independent of the local shape of the interface. This is to be compared with $\sigma \sim 2J$ between the two *ordered* phases of a low temperature Ising model with reduced exchange coupling J.

A more careful analysis reveals that the result

$$\text{Excess factors of } q^{-\frac{1}{4}} = \text{Interfacial length} \tag{3.8}$$

actually holds true even when overhangs and 'bits sticking out' are present. This relies on a more careful definition of the interfacial length. As described in more detail in Chapter 4 the empty and the full phases are dual in the sense that every empty bond is crossed by an occupied bond on the dual lattice. Thus, between the occupied bonds constituting the full phase and the occupied *dual* bonds forming the empty phase there is a gap of half a lattice spacing. We now define the interface to reside in the middle of this gap, *i.e.*, a quarter of a lattice spacing from the nearest occupied bond, whether original or dual. By construction, the interfacial length is now invariant under a duality transformation. Moreover, by starting from a flat interface and recursively adding any desired number of bonds to the full phase it can be shown that Eq. (3.8) is satisfied as long as the interface stays simply connected.

On the other hand, when bubbles of the wrong phase occur within either the empty or the full phase Eq. (3.8)—and hence the analogy between the Ising and Potts interfaces—breaks down. The correcting factors of q can be related to the Euler characteristic χ of the interface, but whether this observation is useful remains an interesting open question.

3.1.2 Including the bond randomness

Now consider the effect of adding bond randomness to the random cluster model. As is seen from Eq. (3.5) the clusters are weighted as in the pure model, whereas Eq. (3.4) implies assigning an extra weight of $q^{w_{ij}}$ to each bond. Since these are only present below $h(x)$ the interfacial Hamiltonian will now take the form

$$\mathcal{H}_{\rm int} = \sigma A + \sum_x \sum_{y<h(x)} w(x,y) \ln q, \tag{3.9}$$

where (x,y) labels bond positions. The first term corresponds to the surface tension given by Eq. (3.7), where A designates the area of the interface (in two dimensions, of course, this has the dimension of a length). The second may be rewritten, up to a term independent of $h(x)$, as

$$\frac{1}{2}\sum_x \left(\sum_{y<h(x)} - \sum_{y>h(x)} \right) w(x,y) \ln q. \tag{3.10}$$

Now compare this with the energy of an interface between the spin-up and spin-down phases of the random field Ising model (RFIM) with spins $s(x,y) = \pm 1$ coupled to a reduced random field $h(x,y) = \pm h_{\rm RF}$

$$\mathcal{H}^{\rm RFIM} = -J \sum_{\langle ij \rangle} s_i s_j - \sum_i h_i^{\rm RF} s_i - h \sum_i s_i. \tag{3.11}$$

Taking the spin-up phase to reside below an interface parametrised by $y = h(x)$ the interfacial energy is now given by

$$\mathcal{H}_{\rm int}^{\rm RFIM} = 2JA + \sum_x \left(\sum_{y<h(x)} - \sum_{y>h(x)} \right) h(x,y). \tag{3.12}$$

Thus the interfacial models are seen to be identical with the correspondence $J \leftrightarrow \frac{1}{8}\ln q$, and $h_{RF} \leftrightarrow \frac{1}{2} w \ln q$. In addition, the imposition of a uniform reduced magnetic field h on the RFIM, which distinguishes between the two coexisting phases, is seen to be equivalent to a deviation $t \equiv (T - T_c)/T_c$ in the Potts model *temperature* variable away from the critical self-dual point. Since this couples to the energy density we find the correspondence $h \leftrightarrow \frac{1}{4} t \ln q$.

At this point a comment is in order. Until now we have only considered an interface between the ordered and the disordered phase in the Potts model, and

it should hardly be surprising that this two-state picture leads to an equivalence with the Ising model. One may then raise the objection that since the ordered phase can correspond to any of the q different values of the Potts model order parameter, it is in fact (for $q > 2$) possible to have a second type of interface between any two of the ordered phases. Fortunately this is of no consequence to the mapping that we have established. Consider namely an interface of this type, which we for simplicity take to be straight ($h(x) \equiv y_0$). This interface consists of a row of empty vertical bonds separating the two portions of the full lattice, and since each of the two ends of these empty bonds carries an uncompensated factor of $q^{-\frac{1}{4}}$ the (unrenormalised) interfacial tension is *twice* that of Eq. (3.7). In the limit $q \to \infty$ this type of interface is therefore strongly suppressed in comparison with the one studied above.

We conclude that the hitherto unknown behaviour of the Potts model interface can be read off from the better studied RFIM, provided that one translates quantities between the two models using the 'dictionary'

$$\begin{aligned} h^{\mathrm{RF}} &\leftrightarrow \frac{1}{2} w \ln q, \\ J &\leftrightarrow \frac{1}{8} \ln q, \\ h &\leftrightarrow \frac{1}{4} t \ln q. \end{aligned} \tag{3.13}$$

Of course, this is strictly valid only as $q \to \infty$. At finite q the q-dependence of cluster configurations with more complicated topologies is not simply accounted for by the interfacial tension. For the same reason, the mapping is not between *bulk* configurations of the two models. However, it will be argued that certain universal properties are controlled by an RG fixed point at infinite q, and for these the mapping should be asymptotically exact. Although this has been described in terms of a two-dimensional self-dual model, it should be clear that it is more general: lack of self-duality corresponds to a skewness in the distribution of the random fields $h(x, y)$, which may be compensated by adding a suitable uniform field (corresponding to a shift in the T_{c} of the Potts model), and, similarly, higher dimensions may be taken into account by appropriately replacing $\sqrt{q}$ by $q^{1/d}$.

3.2 Renormalisation of the Ising interface

A keystone in the derivation of the RG equations for an RFIM interface is a refinement of the Imry-Ma argument presented in Sect. 2.3.1. The question to be answered is how does the adjustment of the region R take place when viewed on different length scales. Using what is essentially the Imry-Ma argument on each successive length scale, Binder [33] found an explicit expression for the energy gain associated with the optimum adjustment of R. With this information at hand Bray and Moore [11] have established the corresponding infinitesimal RG equations for the quantities h^{RF}, J and h. After reviewing this we shall

demonstrate in the following section how this determines the $q \to \infty$ part of the phase diagram of the random-bond Potts model.

3.2.1 Optimisation of the interfacial shape

Consider a RFIM (with $h = 0$) defined on a portion of the square lattice having the geometry of a strip of width L and length N. We are interested in studying the formation of an interface across the strip in the limit $N \gg L \gg a$, the lattice constant being a. To that end we impose boundary conditions along the sides of the strip so that, for $h^{\rm RF} = 0$ and vanishing temperature, there is a straight interface separating the spin-up domain from the spin-down domain with interfacial energy $E = 2JLa$.

In the presence of a random field the interface may choose to move a distance $w_0(h^{\rm RF}) \equiv w(L, h^{\rm RF})$ without changing its shape, in order to minimise its energy with respect to the field. This is illustrated in Fig. 3.1.a. By doing so there will be an energy penalty associated with the boundary conditions and a further change due to $h^{\rm RF}$

$$\Delta E_0 = 4\frac{w_0}{a}J \mp 2\sum_{i\in R_0} h_i^{\rm RF}, \tag{3.14}$$

where R_0 is the region enclosed by the initial and the present position of the interface. The minus sign correspond to an enlargement of the spin-up domain and *vice versa*.

The direction of the displacement will be automatically chosen so as to render the sum on the right-hand side negative. Repeating the Imry-Ma argument we find that it will typically assume a value of $-h^{\rm RF}(w_0 L/a^2)^{1/2}$, up to an unimportant constant of order unity. Minimising ΔE_0 with respect to w_0 yields

$$w_0(h^{\rm RF}) \sim (h^{\rm RF}/8J)^2 L, \qquad \Delta E_0 \sim -4J(h^{\rm RF}/8J)^2(L/a), \tag{3.15}$$

which for $h^{\rm RF}/J \ll 1$ does not yet compensate the initial energy cost E.

Further adjustments are possible, however, if we allow for displacements on successive length scales L/n^l, $l = 1, 2, \ldots$, where we have stipulated that $L = n^l a$ for $n, l \gg 1$. Although treating the adjustment of each interface segment as independent is going to imply some double counting relative to the previous step (as illustrated by the cross-hatched areas in Fig. 3.1) this correlation can be neglected in the limit of large n. The direction of the displacement within each segment is going to optimise the shape of the interface relative to the random field, and in step k the magnitude of the displacement and the corresponding energy gain are of the order

$$w_k(h^{\rm RF}) \sim (h^{\rm RF}/8J)^2\frac{L}{n^k}, \qquad \Delta E_k \sim -4J(h^{\rm RF}/8J)^2\frac{L/a}{n^k}. \tag{3.16}$$

The minimal length scale on which adjustments can take place is $L/n^{k_{\rm max}}$, where $w_{k_{\rm max}}(h^{\rm RF}) = a$, the lattice constant. Summing up the energy gains

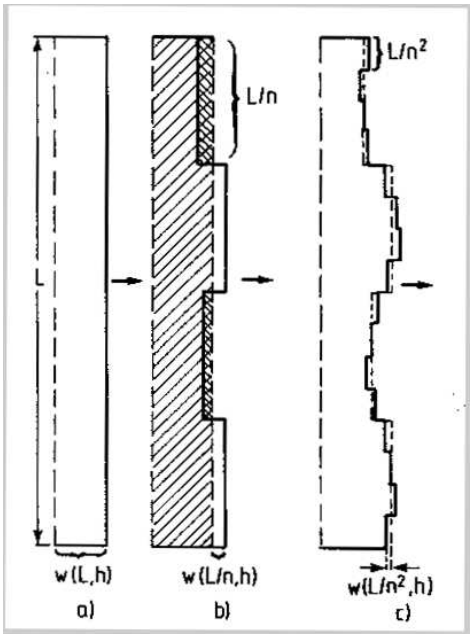

Figure 3.1: Renormalisation of the zero-temperature Ising interface. By successive uniform distortions on length scales L (a), L/n (b), and L/n^2 (c), the interface fits is shape to obtain an energetic optimisation relative to the configuration of the random field. The double-counting, here shown by the cross-hatched areas, is of no consequence as discussed in the text. Taken from Ref. [33].

leaves us with the important result

$$\Delta E \sim \sum_{k=0}^{k_{\max}} n^k \Delta E_k \sim -J(h^{\mathrm{RF}}/J)^2 \frac{L}{a} \frac{\ln(L/a)}{\ln n}. \tag{3.17}$$

There is only a weak dependence on n, and we conclude that the total interfacial energy $E + \Delta E$ becomes negative for L larger than

$$L_0/a \sim \exp[C(J/h^{\mathrm{RF}})^2], \tag{3.18}$$

where C is a constant of order unity. In other words, beyond the length scale set by L_0 the system will spontaneously break up into domains.

3.2.2 Infinitesimal renormalisation group equations

The zero-temperature analysis of the Ising interface just given is particularly relevant in view of the fact that the fixed point governing the critical behaviour

of the low-temperature RFIM is situated exactly at $T = 0$. This was discussed at length in Sect. 2.4 where we also found that since there were *three* scaling fields at play the universal behaviour had to be specified through three independent critical exponents. These are y_h and y_t (the ratio $w \equiv h^{\mathrm{RF}}/J$ playing the role of a temperature variable[1]), and θ governing the violation of hyperscaling.

To find the exponents to order $\mathrm{O}(\epsilon)$ in an expansion around the lower critical dimension it is convenient to write the infinitesimal RG equations corresponding to a length scale factor $b = \mathrm{e}^l$ in the following form

$$\frac{\mathrm{d}h^{\mathrm{RF}}}{\mathrm{d}l} = h^{\mathrm{RF}} f_1(w), \qquad \frac{\mathrm{d}J}{\mathrm{d}l} = J f_2(w), \qquad \frac{\mathrm{d}h}{\mathrm{d}l} = h f_3(w). \tag{3.19}$$

A trivial change of variables yields

$$\frac{\mathrm{d}w}{\mathrm{d}l} = w\big(f_1(w) - f_2(w)\big), \tag{3.20}$$

whence $f_1(w^*) = f_2(w^*)$ at the fixed point.

The exponents y_h and θ are immediately read off from the RG equations

$$y_h = f_3(w^*), \qquad \theta = f_1(w^*) = f_2(w^*), \tag{3.21}$$

whereas to get y_t the functions $f_1(w)$ and $f_2(w)$ have to be expanded to first order in w (*i.e.*, we 'linearise around the fixed point'):

$$y_t = w^*\big(f_1'(w^*) - f_2'(w^*)\big). \tag{3.22}$$

In the pure system ($w = 0$) the RG equations are trivial. Namely, since the fixed point is at zero temperature fluctuations are absent, and the renormalisation originates from simple geometrical considerations. The coupling constant is multiplied by a factor giving the area of the wall separating two neighbouring block spins, $J' = b^{d-1}J$, and the magnetic field scales like the volume of a block spin, $h' = b^d h$. Finally, by the central limit theorem we have $h^{\mathrm{RF}\prime} = b^{d/2} h^{\mathrm{RF}}$. Thus

$$f_1(0) = d/2, \qquad f_2(0) = d - 1, \qquad f_3(0) = d. \tag{3.23}$$

In particular we find that $\mathrm{d}w/\mathrm{d}l = (1 - d/2)w + \mathrm{O}(w^2)$, and since from the phase diagram of Sect. 2.4 we expect the fixed point to move towards $w^* = 0$ when we approach the lower critical dimension from above we infer that $d_{\mathrm{l}} = 2$.

To obtain the critical exponents to first order in $\epsilon \equiv d - 2$ we need the functions $f_i(w)$ to second order in w. Actually we expect the expressions for $f_1(w = 0)$ and $f_3(w = 0)$ just given to be correct to *all* orders in w. Indeed, the renormalised field is related to the energy cost of reversing a block of b^d spins, and the argument given above should hold true whenever all spins within

[1] Although the definitions are different we use the same symbol w for the strength of the randomness in the RFIM and in the RBPM.

a given block can be expected to be predominantly parallel. The only way this could fail is by the random field's creating a domain of flipped spins, but the probability hereof is exponentially small [11].

On the other hand, the coupling constant J has a non-trivial renormalisation to order $\mathrm{O}(w^2)$. This is exactly given by the energy associated with the adjustment of the domain wall to the local shape of the random field. To the order needed it is enough to evaluate this energy at $d = 2$, and using Eq. (3.17) we thus find

$$f_2(w) = d - 1 - Aw^2, \tag{3.24}$$

where $A > 0$ is a non-universal constant. Inserting the $f_i(w)$ in Eq. (3.20) we get

$$\mathrm{d}w/\mathrm{d}l = -(\epsilon/2)w + Aw^3, \tag{3.25}$$

whence the fixed point is situated at $w^* = (\epsilon/2A)^{1/2}$. The critical exponents are then finally evaluated as

$$y_h = 2 + \epsilon, \quad \theta = 1 + \epsilon/2, \quad y_t = \epsilon. \tag{3.26}$$

3.3 The proposed phase diagram

The infinitesimal RG equations for the RBIM can now be inferred from the similar results for the RFIM by translating via Eq. (3.13). Near $d = 2$ they read

$$\begin{aligned} dw/dl &= -(d/2-1)w + Aw^3 + \cdots && (3.27)\\ d(\ln q)^{-1}/dl &= -(\ln q)^{-1}\big((d-1) - Aw^2 + \cdots\big) && (3.28)\\ dt/dl &= t(1 + Aw^2 + \cdots), && (3.29) \end{aligned}$$

where we recall that $A > 0$ is a non-universal constant. Corrections to these equations are supposed to be higher order in w and in $q^{-1/2}$. The reader who is concerned that the number of Potts states q is allowed to flow in these equations is reminded that q is merely a parameter of the random cluster model. Only if magnetic quantities are considered is the permutation symmetry of the model revealed, and this is lost in the mapping to the cluster model. The RG flows for $d > 2$ and the consequent phase diagram are shown in Fig. 3.2.

In the pure models, for $q >$ some $q_2(d)$ (low T in the RFIM), there is phase coexistence with a non-vanishing latent heat (spontaneous magnetisation), controlled by a fixed point at infinite q ($T = 0$). For $d > 2$ this persists into the shaded region, bounded by a line of tricritical points where the latent heat, vanishes. The universal behavior along this line is controlled by the fixed point R at $w = O\big((d-2)^{1/2}\big)$ and infinite q.

Above the line Rq_2, the flows go to large w beyond the validity of Eqs. (3.27-3.29). In addition, the renormalised interfacial tension flows to zero and the

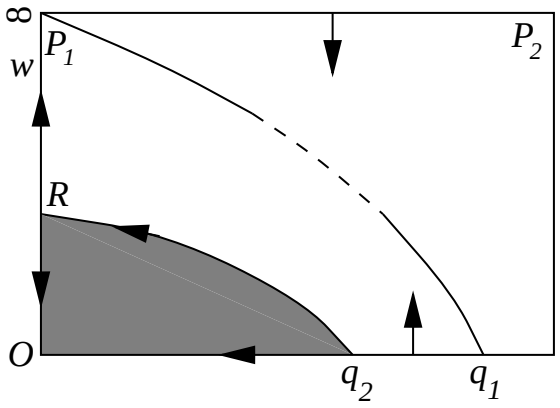

Figure 3.2: Schematic phase diagram in the critical surface for $d > 2$. q increases to the left and w is the disorder strength, with P_1P_2 being the percolation limit. RG flows are indicated. The latent heat is non-vanishing within the shaded region, and elsewhere the transition is continuous, controlled by the line of fixed points P_1q_1. As $d \to 2$ the shaded region collapses to a line q_2O of first-order transitions in the pure system. For $d = 2$ we have $q_1 = 2$ and $q_2 = 4$.

mapping between the models breaks down as domains of different topologies proliferate. However, for infinite q the mapping remains exact and the flows go to infinite w. This cannot happen for finite q since this is the percolation limit $K_1/K_2 = 0$, at which w^{-1} is *relevant* [29]. There must therefore exist another line of stable fixed points emerging from P_1, which control the universal continuous transition for large, but finite, values of w and q. It is tempting to conjecture, as indicated by the dashed line in Fig. 3.2, that this connects on to that found by expansion in powers of $q - q_1$ [24], where q_1 is the point where the exponent α of the pure model changes sign [8]. In Chapter 6 we present numerical evidence that, at least for $d = 2$, this is the case.

In two dimensions (when $q_1 = 2$, $q_2 = 4$ [19]), the shaded region collapses, and for any non-zero w the renormalised interfacial tension, and thus the latent heat, vanishes. The flows should be towards the line P_1q_1, with a crossover length which, from Eq. (3.27), has the form

$$\xi_X \sim \mathrm{e}^{1/2Aw^2} \tag{3.30}$$

and therefore may become very large for weak randomness. (To see this, note that the RG eigenvalue of ξ_X must equal -1, whence $\mathrm{d}\xi_X/\mathrm{d}l = -\xi_X$. Using Eq. (3.25) we then find the differential equation

$$\frac{\mathrm{d}\xi_X}{\mathrm{d}w} = \frac{\mathrm{d}\xi_X}{\mathrm{d}l}\frac{\mathrm{d}l}{\mathrm{d}w} = -\xi_X\left(\frac{\mathrm{d}w}{\mathrm{d}l}\right)^{-1} = -\xi_X\frac{1}{Aw^3}, \tag{3.31}$$

which is readily solved.)

3.4 Exact exponent relations

Using the correspondences between scaling variables in the RFIM and the RBPM given by Eq. (3.13) it is possible to establish exact exponent relations between the two models. Of course, the derivations presented here are only strictly valid close to $d = 2$, but, if the topology of the RG flows does not change, the resulting exponent relations should hold also in three and higher dimensions.

3.4.1 The relation $\beta = \beta_{\mathrm{RF}}$

For the RBPM we define β_{RF} as the exponent governing the vanishing of the latent heat as the line Rq_2 is approached from below

$$L(\ln q, w) \sim (w_c - w)^{\beta_{\mathrm{RF}}}, \tag{3.32}$$

and we aim at showing that β_{RF} is identical to the magnetisation exponent in the RFIM, namely $\beta = (d - y_h)\nu$.

Imagine starting out from some point $(\ln q, w)$ just below the line of tricritical points in the phase diagram of Fig. 3.2 and following an RG trajectory, parametrised by l, that brings us within the vicinity of the fixed point R situated at (∞, w^*). By definition of the RG transformation the partition function stays constant under the renormalisation, whence

$$-\frac{\ln Z}{V} = f(t, \ln q, w) = \mathrm{e}^{-dl} f\big(t(l), \ln q(l), w(l)\big). \tag{3.33}$$

The latent heat is obtained by differentiating with respect to t and taking the limit $t \to 0$. Using the chain rule on the right-hand side this becomes

$$L(\ln q, w) = \mathrm{e}^{-dl} \left.\frac{\partial t(l)}{\partial t}\right|_{t=0} L\big(\ln q(l), w(l)\big). \tag{3.34}$$

The partial derivative can be evaluated by employing the correspondence $t \leftrightarrow \frac{1}{2} h T_{\mathrm{RF}}$, remembering that T_{RF} is irrelevant at R with RG eigenvalue $-\theta$.[2] We find that

$$\frac{\mathrm{d}t}{\mathrm{d}l} = \frac{1}{2}\frac{\mathrm{d}h}{\mathrm{d}l} T_{\mathrm{RF}} + \frac{1}{2} h \frac{\mathrm{d}T_{\mathrm{RF}}}{\mathrm{d}l} = (y_h - \theta) t, \tag{3.35}$$

which upon integration becomes $t(l_2) = \exp[(l_2 - l_1)(y_h - \theta)] t(l_1)$, so that finally

$$\frac{\partial t(l_2)}{\partial t(l_1)} = \mathrm{e}^{(l_2 - l_1)(y_h - \theta)}. \tag{3.36}$$

[2] This is to be contrasted with the situation where we move towards the fixed point 0 along the line of pure models. In this case $T = 1/J$ is trivially related to the surface area via the interfacial model, and the corresponding RG eigenvalue is $d - 1$.

Inserting this in Eq. (3.34) we arrive at

$$L(\ln q, w) = \mathrm{e}^{(y_h-\theta-d)l} L\big(\ln q(l), w(l)\big), \tag{3.37}$$

and since it is expected from Eq. (3.6) that $L\big(\ln q(l), w(l)\big) \propto \frac{1}{2}\ln q(l)$ the next task is to find the l dependence of $\ln q(l)$ in the vicinity of R. This is done using the correspondence $T \leftrightarrow 8(\ln q)^{-1}$, which follows from Eq. (3.13). In complete analogy with Eq. (3.36) we find that

$$\big(\ln q(l_2)\big)^{-1} = \mathrm{e}^{-\theta(l_2-l_1)}\big(\ln q(l_1)\big)^{-1}, \tag{3.38}$$

and in particular $\ln q(l) = \mathrm{e}^{\theta l}\ln q$. This cancels the θ-dependence of Eq. (3.37), which now assumes the form

$$L(\ln q, w) = \mathrm{e}^{(y_h-d)l}\ln q. \tag{3.39}$$

It only remains to relate l to the distance from the transition point. This is easily done by considering the correlation length, which increases according to $\xi(l) = \mathrm{e}^l\xi(0)$, and since by definition of ν we have $\xi \sim (w^*-w)^{-\nu}$ we find that $\mathrm{e}^l \sim (w^*-w)^{-\nu}$. It follows that

$$L(\ln q, w) \sim (w^*-w)^{(d-y_h)\nu}\ln q, \tag{3.40}$$

which is what we wanted to show.

3.4.2 The relation $\nu = \nu_{\mathrm{RF}}/(2-\alpha_{\mathrm{RF}}-\beta_{\mathrm{RF}})$

In contradistinction to what was the case for the exponent β the task of relating ν to quantities in the RFIM is quite elementary, since the latter exponent has identical definitions in the two models. Namely, the correlation length is supposed to diverge like $\xi(t) \sim t^{-\nu}$ upon approach of the critical point, where $\nu = 1/y_t$. From the relation $t \leftrightarrow \frac{1}{2}hT_{\mathrm{RF}}$ we infer that y_t in the RBPM equals $y_h - \theta$ in the RFIM; see also Eq. (3.35).

The values of y_h and θ in the RFIM can be inferred from the scaling relations given in Sect. 2.4. These are

$$\beta_{\mathrm{RF}} = (d-y_h)\nu_{\mathrm{RF}}, \qquad 2-\alpha_{\mathrm{RF}} = (d-\theta)\nu_{\mathrm{RF}}, \tag{3.41}$$

where we recall that hyperscaling is violated through the presence of θ. We thus end up with the relation

$$\nu^{-1} = y_h - \theta = (2-\alpha_{\mathrm{RF}}-\beta_{\mathrm{RF}})/\nu_{\mathrm{RF}}. \tag{3.42}$$

Since there are only two independent critical exponents in the RBPM, the two exponent relations given this far suffice to relate the universal information about the RBPM fixed point R to the corresponding information about the RFIM.

Resumo en Esperanto

Ni proponas fazdiagramon por la Potts-modelo kun aleatoraj kuplaĵoj en dimensio $d \geq 2$ (vidu Fig-on 3.2). Ĉi tie w estas la forto de la malordo kaj la nombro da statoj q kreskas maldekstren. Por $d = 2$, $q_1 = 2$ kaj $q_2 = 4$. La sagoj indikas la renormiĝan fluon. La specialaj kazoj de la pura modelo [19] kaj la perkola limo [29] estas bone konataj. Perturba kalkulo ĉirkaŭ q_1 [24, 25, 26] demonstras la emerĝon de linio da stabilaj fikspunktoj.

Proksime de $q = \infty$ du statoj, la tute ordita kaj la tute malordita, dominas la statan sumon. De tie sekvas ekvivalenton inter la koresponda interfaca modelo kaj tiu de la malalttemperatura Ising-modelo kun aleatora magneta kampo. En la griza regiono la transiro havas pozitivan latentan varmon, kaj la linion Rq_2 da trikritaj punktoj kontrolas la fikspunkto R. La najbaran renormiĝan fluon, Ekvojn (3.27–3.29), ni transprenas de Ref-o [11]. La menciita ekvivalento ankaŭ implicas ekzaktan rilaton inter la kritaj eksponentoj de la du modeloj (vidu sekcion 3.4). El P_1 emerĝas alia linio da stabilaj fikspunktoj; ni konjektas ke ĝi konektiĝas al tiu ĉe q_1 kaj prezentos nombran apogon tiurilate en Ĉapitro 6.

Chapter IV

Potts duality

Many discrete lattice models allow for a duality transformation, and in the case of the Potts model this is a most important ingredient in the analysis of the critical behaviour. Roughly speaking duality amounts to a transformation of the original lattice model under which each plaquette is turned into a spin site and *vice versa*. There is a one-to-one correspondence between interacting pairs of spins in the original and the dual model, but the interaction strengths transform non-trivially under duality.

The duality of the Ising model partition function has been known for almost as long as the model itself [34]. The generalisation to the q-state Potts model is relatively straightforward, but its statement in terms of the equivalent random cluster model is particularly illuminating, and we shall give the details shortly. Arguably the greatest benefit arising from the duality of the partition function is the resulting *exact* knowledge of the critical temperature. The duality arguments can easily be adapted to the case of bond randomness, a feature which shall turn out to be crucial for extracting the thermal scaling dimension in Chapter 6 using the technique of phenomenological renormalisation.

Duality also applies to the correlation functions between n spins situated at the boundary of the system. Quite naturally these results depend on the particular geometry imposed (planar, cylindrical, ...), and consequently the boundary conditions have to be specified carefully. For a planar geometry, such as a long strip with open boundary conditions, a clever diagrammatic method recently invented by Wu and coworkers [35, 36, 37, 38] now allows for the mapping of any n-point function in this way. The case of cylindrical geometry, which is particularly relevant to long strips with periodic boundary conditions along the transverse direction, and hence to the construction of the transfer matrices of Chapter 5, is more evasive. Results for a general n-point function ($n \geq 3$) are presently unknown, but we shall see how the two-point function is mapped onto a disorder operator in the dual model. Eventually this will allow us to devise an ingenious way of extracting magnetic properties from the transfer matrix through the imposition of so-called twisted boundary conditions. This point is elaborated in Chapter 5.

4.1 Duality of the partition function

The notion of duality applies to a lattice model through a transformation of both the model itself and of the underlying lattice. For lattices that can be embedded in a surface of genus g the duality transformation amounts to converting all plaquettes of the lattice into vertices of the dual lattice. Each edge of the original lattice is intersected by exactly one edge of the dual lattice, so that any two plaquettes separated by a common edge is transformed under duality into two vertices connected by a common dual edge. Evidently, a second application of the duality transformation gives us back the original lattice, whence the terminology.

A common example among the infinite regular lattices is the duality between the triangular and the honeycomb lattice. The square lattice is its own dual, that is to say it is *self-dual*. Finite regular lattices are usually not self-dual, since the exterior infinite plaquette is transformed into a vertex with a (large) coordination number equal to the number of edges along the circumference of the original lattice. Clever arrangements of the boundaries are possible, however, so as to obtain a finite self-dual lattice which resembles a square lattice at every bulk edge [39, 40]. Such lattices are important for the investigation of exact partition function zeroes [41].

Quite generally the dual of a lattice model is obtained by replacing the degrees of freedom on the vertices of the lattice with (possibly other) degrees of freedom on the vertices of the dual lattice. These new degrees of freedom may of course interact in an altogether different fashion than the old ones. As a relatively simple example we mention the duality between the SOS and the XY models [5].

The Ising and Potts models both allow for a duality in which the lattice and the interaction strengths are altered whereas the degrees of freedom themselves as well as the type of interaction remain the same. The duality of the Ising model is by now standard textbook material, and we shall not repeat it here. As far as the Potts model goes we begin by examining the duality of the partition function. Since the main purpose is to see how the free energy per site maps under duality in the limit of an infinite system, boundary conditions are not of much consequence. In particular the argument determining the critical temperature essentially applies to any boundary conditions on a lattice that is self-dual (*i.e.*, square) in the bulk.

4.1.1 Connection to the random cluster model

The q-state Potts model with degrees of freedom $\sigma_i = 1, 2, \ldots, q$ defined on the vertices i of a lattice $\mathcal{L}$ can be mapped onto the so-called random cluster model introduced by Kasteleyn and Fortuin [31]. Writing the partition function of the

Potts model without a magnetic field as

$$Z = \sum_{\{\sigma\}} \prod_{\langle ij \rangle} \exp(K_{ij} \delta_{\sigma_i \sigma_j}), \tag{4.1}$$

where $\prod_{\langle ij \rangle}$ is a product over the edges of $\mathcal{L}$, and K_{ij} are the edge-dependent (reduced) coupling constants, we notice that the main obstacle to the evaluation of the sum $\sum_{\{\sigma\}}$ is the fact that the Kronecker deltas $\delta_{\sigma_i \sigma_j}$ appear exponentiated. Fortunately the latter can take only two values and can hence be written

$$\exp(K_{ij} \delta_{\sigma_i \sigma_j}) = 1 + u_{ij} \delta_{\sigma_i \sigma_j}, \tag{4.2}$$

where we have defined the variables $u_{ij} = \exp(K_{ij}) - 1$. Now imagine expanding out the product $\prod_{\langle ij \rangle}$ of these binomials. Each term in the resulting sum is given a graphical representation G in which each choice of the factor u_{ij} (1) is associated with the edge $[ij] \in \mathcal{L}$ being present in (absent from) G. An example of a valid configuration G is shown in Fig. 4.1. The deltas enforce the equality between all spins belonging to the same connected cluster of G, and the sum $\sum_{\{\sigma\}}$ can now be performed with the result that each such cluster (including isolated vertices) is weighted with a factor of q. Thus

$$Z = \sum_{G \subseteq \mathcal{L}} \left(\prod_{\langle ij \rangle \in G} u_{ij} \right) q^{C(G)}, \tag{4.3}$$

where $C(G)$ designates the number of independent clusters on G.

We first stipulate the duality between two very special graphs. Namely, the full graph $G = \mathcal{L}$ with partition function

$$Z_{\text{full}}(\{u_{ij}\}) = q \prod_{\langle ij \rangle \in \mathcal{L}} u_{ij} \tag{4.4}$$

is taken to be dual to the empty graph $G^* = \emptyset$ with

$$Z^*_{\text{empty}}(\{u^*_{ij}\}) = q^{N^*}, \tag{4.5}$$

where the number of dual sites N^* is fixed by the Euler relation $|\langle ij \rangle| = N + N^* - 2$.

Establishing the duality then amounts to ascertaining that all other graphs have the same weight relative to this reference state as is the case in the dual model. In the terminology introduced above, duality means that a graph configuration G on the original lattice $\mathcal{L}$ is dual to a configuration G^* on the dual lattice $\mathcal{L}^*$ in which every bond of strength u_{ij} being 'present' in G corresponds to the dual bond of strength u^*_{ij} being 'absent' from G^* and *vice versa*.

In particular, removing one bond from the full graph (relative weight: $1/u_{ij}$) must correspond to adding the corresponding dual bond to the empty dual graph

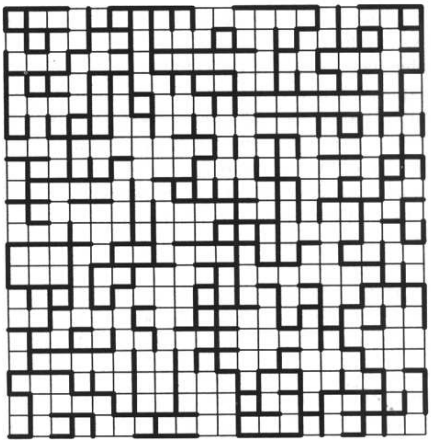

Figure 4.1: Graphical representation of a typical term contribution to the partition function of the random cluster model. Taken from Ref. [42].

(relative weight: u^*_{ij}/q), meaning that the bond strengths and their duals must obey the relation

$$u_{ij}u^*_{ij} = q. \tag{4.6}$$

When removing further bonds from G it may happen that a new cluster is separated from the rest of the graph, yielding an additional factor of q. But such a cluster formation corresponds precisely to a loop closure on the dual lattice, also giving an extra factor of q! Since all graph configurations G can be constructed by successive removals of bonds from the full reference state we have thus proven the fundamental duality relation [43]

$$Z(\{u_{ij}\}) = qCZ^*(\{u^*_{ij}\}), \tag{4.7}$$

where $C = q^{-N^*}\prod_{\langle ij\rangle} u_{ij}$ is a constant.

4.1.2 Determination of the critical temperature

One would now expect, again in analogy with the Ising model, that the fundamental duality relation (4.7) in conjunction with an assumption on the uniqueness of the phase transition determines the critical temperature of the random-bond Potts model on the square lattice.

To see that this is indeed the case we shall reproduce an argument first given by Kinzel and Domany [30]. Consider for simplicity the case where the couplings are drawn from a binary probability distribution

$$P(u) = p\delta(u - u_1) + (1 - p)\delta(u - u_2), \tag{4.8}$$

and where the two possible values u_1^* and u_2^* of the couplings in the dual system are given by Eq. (4.6). By taking the logarithm of the fundamental duality relation (4.7) we find the following identity between the singular parts of the free energy in the original system and its dual[1]

$$F_{\rm s}(u_1, u_2, p) = F_{\rm s}(u_1^*, u_2^*, p) = F_{\rm s}(u_2^*, u_1^*, 1-p), \tag{4.9}$$

where in the last step we have used a simple symmetry property of the function $F_{\rm s}$.

Assuming that the phase transition is unique this identity can be used to determine the value of p and the relation between the couplings u_1 and u_2 for which it occurs. To fix the value of p we examine the percolation limit ($u_1 \to 0$, $u_2 \to \infty$), where Eq. (4.9) becomes

$$F_{\rm s}(0, \infty, p) = F_{\rm s}(0, \infty, 1-p). \tag{4.10}$$

Since by assumption there is precisely one singularity as a function of p this must occur at $p = 1/2$, the percolation concentration. Inserting this value in Eq. (4.9) and invoking once more the uniqueness assumption we furthermore see that the couplings have to satisfy $u_1^* = u_2$.

It is clear that this argument generalises to more complicated probability distributions than (4.8). In fact, any discrete or continuous distribution $P(u)$ will locate the system at its critical point provided that it is self-dual

$$P(u) = P(u^*), \tag{4.11}$$

and that the random bonds are chosen in a spatially uncorrelated fashion. Any particular realisation of the randomness is then self-dual 'on average'.

4.2 Duality of cylindrical two-point correlators

In addition to the duality relation (4.7) for the partition function a similar relation can be established for the spin-spin correlation function with both spins situated on the boundary. We shall first consider the case of $\mathcal{L}$ having the geometry of a cylinder. This is particularly relevant to the construction of the transfer matrices in Chapter 5, and will eventually allow us to devise a very efficient way of extracting the magnetic scaling dimension.

As usual we define the local order parameter as [5]

$$M_a(r) = \left(\delta_{\sigma(r),a} - \frac{1}{q}\right), \qquad a = 1, \ldots, q. \tag{4.12}$$

[1] The constant qC appearing in Eq. (4.7) may seem problematic since it apparently depends on the thermodynamic parameters u_{ij} and hence may contribute non-trivially to F_s. Fortunately this is not so. Indeed it is easy to see, using Eq. (4.6) and the Euler relation, that for *any* self-dual realisation of the random bonds we have $qC = q^{(N-N^*)/2}$, which depends only on geometric factors.

In the high temperature phase all components of the order parameter vanish, whilst in the ordered (low temperature) phase the Z_q symmetry is spontaneously broken and one of the components, say $a = 1$, has a positive expectation value. A simple calculation now shows that the correlation function $G_{aa}(r_1, r_2) = \langle M_a(r_1) M_a(r_2) \rangle$ is proportional to the probability that the points r_1 and r_2 belong to the same cluster.

In a cylindrical geometry the graphs for which two points r_1 and r_2, that reside at opposite ends of the cylinder, are connected correspond to dual graphs where clusters are forbidden to wrap around the cylinder. This is equivalent to computing the dual partition function with twisted boundary conditions

$$\sigma \to (\sigma + 1) \bmod q \tag{4.13}$$

across a *seam* running from r_1 to r_2. By permuting the Potts spin states the shape of this seam can be deformed at will as long as it connects r_1 and r_2. Duality thus maps the correlation function onto a disorder operator

$$\langle M_a(r_1) M_a(r_2) \rangle = \left\langle \prod_{\text{seam}} \exp(-K^* \delta_{\sigma_i \sigma_j}) \right\rangle_{Z^*} , \tag{4.14}$$

where $Z^* = Z^*(\{K^*\})$ is the dual partition function with periodic boundary conditions. The factors of $\exp(-K^* \delta_{\sigma_i \sigma_j})$ explicitly delete the seam-crossing edges (ij) otherwise implied by the factors of $\exp(K^* \delta_{\sigma_i \sigma_j})$ in Z^*, *cfr.* Eq. (4.1).

4.3 Duality of planar n-point boundary correlation functions

As regards the n-point boundary correlation functions of the Potts model defined on a *planar* graph $\mathcal{L}$ a dramatic development has taken place during the last year. A novel diagrammatic approach which enables such correlators to be related to ratios of dual partition functions under fixed boundary conditions (*alias* generalised surface tensions) was introduced by F. Y. Wu, and the well-known result for $n = 2$ was explicitly generalised to $n = 3$ [35]. In an attempt to carry through an analogous calculation for $n = 4$ we have suggested that an appropriate generalisation of Wu's method would be to extract one linear relationship between the b_n different partition functions with fixed values of the boundary spins for each way of connecting the n boundary points by means of a so-called auxiliary graph [36]. For reasons of planarity these graphs have to be *well-nested*, but for $n \geq 4$ the number c_n of well-nested graphs is less than the number of unknowns b_n, whence the problem appeared insoluble.

Further progress was made when Wu and Huang realised that for $n \geq 4$ the c_n equations thus obtained could be supplemented by certain sum rule identities. The number of such identities is precisely $b_n - c_n$, hence providing a complete solution of the problem [37]. Finally, a conjecture made in Ref. [37] for the case of general n was proved by Lu and Wu [38].

In this section we first present the unadorned diagrammatic method, illustrated for the simplest non-trivial case of $n = 2$. We then present a proof that

this method is insufficient for solving the case $n = 4$. Next, the sum rules are discussed along with the solution for general n. We conclude with some observations on the presently unsolved case of $n \geq 3$ with *cylindrical* boundary conditions.

4.3.1 The diagrammatic method

Consider the q-state Potts model defined on any lattice $\mathcal{L}$ which can be embedded in the plane so as to have a unique[2] boundary $\partial\mathcal{L}$. Let i and j be two sites on $\partial\mathcal{L}$ occupied by Potts spins σ_i and σ_j. The (generalised) two-point correlator giving the probability that $\{\sigma_i, \sigma_j\}$ are in the states $\{\sigma, \sigma'\}$ is defined by

$$P_2(\sigma, \sigma') = \langle \delta_{\sigma_i,\sigma} \delta_{\sigma_j,\sigma'} \rangle, \tag{4.15}$$

and its diagonal element is related to the conventional two-point correlator Γ_2 through

$$\Gamma_2(\sigma_i, \sigma_j) \equiv \langle q\delta_{\sigma_i,\sigma_j} - 1 \rangle = q^2 P_2(\sigma, \sigma) - 1. \tag{4.16}$$

The object of this subsection is to express P_2 through ratios of dual partition functions under fixed boundary conditions.

By permutation symmetry of the Potts spins ($\sigma = 1, 2, \ldots, q$) the partition function can be written as

$$Z = qZ_{11} + q(q-1)Z_{12},$$

where $Z_{\sigma\sigma'} = Z(\sigma_i = \sigma, \sigma_j = \sigma')$ denotes a restricted partition function with fixed values of the boundary spins. Similarly, for the dual model we define restricted partition functions with a fixed value of the dual spin that is situated in the infinite region exterior to $\mathcal{L}$ and that interact with all the dual spins adjacent to the boundary $\partial\mathcal{L}$. For purposes that will become clear shortly we shall imagine that there are *two* such exterior spins: one interacting with all the spins adjacent to the portion of the boundary traversed when going clockwise from i to j, and another interacting with the remaining boundary spins as shown in Fig. 4.2. For the moment these two spins are identical, and invoking the permutation symmetry once more we have the identity $Z^* = qZ^*_{11}$. From the fundamental duality relation (4.7) we then find

$$Z_{11} + (q-1)Z_{12} = qCZ^*_{11}. \tag{4.17}$$

To obtain another relation involving Z_{11} and Z_{12} we consider augmenting $\mathcal{L}$ by an *auxiliary graph*, which in this elementary example simply consists of an additional bond of strength K located outside $\partial\mathcal{L}$ that connects the sites i and j. Clearly, such a bond separates the two exterior dual spins just introduced,

[2] The cylinder can also be embedded in the plane, but will then have both an inner and an outer boundary. It is thus excluded from the subsequent discussion.

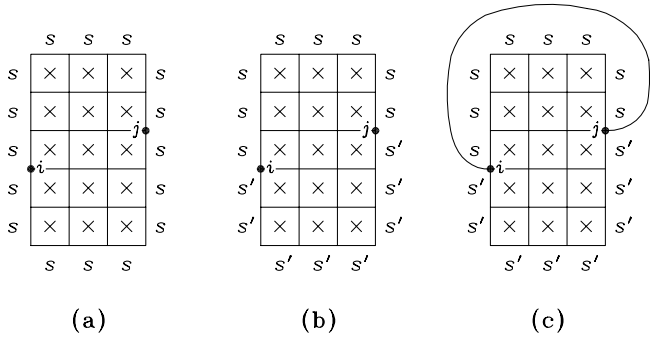

Figure 4.2: A simple example of a planar lattice $\mathcal{L}$ for which the duality transformation of the two-point function applies. Dual spins inside the lattice boundary $\partial\mathcal{L}$ are shown as crosses. In situation (a) all dual spins adjacent to $\partial\mathcal{L}$ interact with a unique dual spin s in the exterior region. When computing Z^*_{12} there are two such exterior spins, s and s'. In (c) sites i and j are connected through a bond of strength K. Any such connectivity among the boundary spins is known as an *auxiliary graph*. Taken from Ref. [35].

and these in turn now interact through a bond of strength K^*. The partition functions for the modified system are

$$\tilde{Z} = q\mathrm{e}^K Z_{11} + q(q-1)Z_{12}, \qquad \tilde{Z}^* = q\mathrm{e}^{K^*} Z^*_{11} + q(q-1)Z^*_{12},$$

and Eq. (4.7) now yields

$$\mathrm{e}^K Z_{11} + (q-1)Z_{12} = C(\mathrm{e}^K - 1)[\mathrm{e}^{K^*} Z^*_{11} + (q-1)Z^*_{12}], \tag{4.18}$$

since when including the auxiliary graph we have increased N^* by one.

Solving Eqs. (4.17) and (4.18) we find

$$Z_{11} = C[Z^*_{11} + (q-1)Z^*_{12}], \qquad Z_{12} = C[Z^*_{11} - Z^*_{12}], \tag{4.19}$$

which, of course, does not depend on the dummy variable K. For the correlators one then finally obtains

$$P_2(\sigma,\sigma') = \frac{1}{q^2}\left[1 + (q\delta_{\sigma,\sigma'} - 1)\frac{Z^*_{12}}{Z^*_{11}}\right], \qquad \Gamma_2(\sigma_i,\sigma_j) = (q-1)\frac{Z^*_{12}}{Z^*_{11}}. \tag{4.20}$$

In particular for $T \neq T_\mathrm{c}$ and $|i-j| \gg a$ the exponential decay $\langle\sigma_i\sigma_j\rangle \sim \mathrm{e}^{-|i-j|/\xi}$ is related to the surface tension $\tau \equiv -\lim_{|i-j|\to\infty} |i-j|^{-1} \ln(Z^*_{12}/Z^*_{11})$ in the dual model through $\xi = 1/\tau$.

4.3.2 Inadequacy of equations for $n = 4$

This very elegant method of obtaining duality identities for the boundary correlation functions works equally well for $n = 3$ [35]. There are now five restricted

partition functions to be determined, *viz.* Z_{111}, Z_{123}, Z_{211}, Z_{121} and Z_{112}, and accordingly one can find five independent relations. The first of these is trivial and corresponds to Eq. (4.17) above. The other four are obtained by connecting either of the two boundary spins through an auxiliary bond as before (three equations) and by connecting all three spins to an auxiliary (non-dual) spin placed outside $\partial\mathcal{L}$. It is then tempting to conjecture that the method generalises straightforwardly to the case of $n \geq 4$ points [35].

This is however not so, as we shall now demonstrate. The essence of the argument is there are in general b_n quantities $Z_{\sigma_1,\sigma_2,\ldots,\sigma_n}$ to be determined, but there are only c_n topologically distinct auxiliary graphs each furnishing a single independent relation among these quantities [36]. Here, c_n is the number of well-nested n-point connectivities (see Eq. (5.8) for a precise definition), whilst b_n is the total number of n-point connectivities, including the non-well nested ones. For $n \leq 3$ these two numbers are identical, but for $n \geq 4$ we have $b_n > c_n$. Further details on well-nestedness are given in Chapter 5.

As the argument is admittedly somewhat laborious the reader who is willing to trust his geometrical intuition may wish to skip the rest of this subsection. Before doing so, however, his attention is drawn to Eq. (4.28) which shows that the result for the conventional four-point correlator

$$\Gamma_4(\sigma_i,\sigma_j,\sigma_k,\sigma_l) = q^4 P_4(\sigma,\sigma,\sigma,\sigma) - 1 = q^4 \frac{Z_{1111}}{Z} - 1 \tag{4.21}$$

can still be deduced by rather elementary means.

Consider then the q-state Potts model on a two-dimensional planar graph $\mathcal{L}$ having a free boundary $\partial\mathcal{L}$. As a simple generalisation of the situation shown in Fig. 4.2 for the two-point case we let i, j, k and l be four sites on $\partial\mathcal{L}$, following one another in a clockwise fashion, and we define exterior dual spins s_1, s_2, s_3 and s_4 so that all boundary spins between sites l and i of $\mathcal{L}$ interact with a spin in state s_1, boundary spins between i and j interact with s_2, spins between j and k with s_3, and finally spins between k and l with s_4. The partition function with the four Potts spins σ_i, σ_j, σ_k and σ_l fixed in definite states is called $Z_{\sigma_i\sigma_j\sigma_k\sigma_l}$, and similarly the dual partition function for fixed exterior dual spins is denoted by $Z^*_{s_1s_2s_3s_4}$.

Up to the q-fold permutation symmetry of the Potts spin labels there exist 15 different boundary conditions for $Z_{\sigma_i\sigma_j\sigma_k\sigma_l}$ out of which we can form five combinations

$$\begin{aligned}
Z_4 &= Z_{1111}, \\
Z_3 &= Z_{2111} + Z_{1211} + Z_{1121} + Z_{1112}, \\
Z_{2\mathrm{p}} &= Z_{2211} + Z_{2121} + Z_{2112}, \\
Z_{\mathrm{p}} &= Z_{1123} + Z_{1213} + Z_{1231} + Z_{2113} + Z_{2131} + Z_{2311}, \\
Z_0 &= Z_{1234},
\end{aligned} \tag{4.22}$$

which are symmetric under permutations of the four sites i, j, k and l. We introduce them here in order to simplify the notation in subsequent equations.

Following the strategy outlined by Wu [35] we should extract equations relating the $Z_{\sigma_i\sigma_j\sigma_k\sigma_l}$ and the $Z^*_{s_1s_2s_3s_4}$ by running through all possible ways of connecting the sites i, j, k and l by auxiliary bonds and using the fundamental duality relation (4.7). Each time a bond is added one of the exterior dual spins s is separated from its neighbours and thus allowed to take a different value. We remark here that since the auxiliary bonds have to be drawn outside $\partial\mathcal{L}$ they must necessarily be well-nested in the sense of Blöte and Nightingale [44].

First, using the 'empty' connection (*i.e.*, introducing no auxiliary bonds) we find the relation

$$Z_4 + (q-1)[Z_3 + Z_{2\mathrm{p}}] + (q-1)(q-2)Z_{\mathrm{p}} + (q-1)(q-2)(q-3)Z_0 = qCZ^*_4, \quad (4.23)$$

corresponding to Wu's Eq. (19) for $n = 3$ (or to Eq. (4.17) given above for the simplest case of $n = 2$).

Next, consider adding a bond between sites k and l. Summing over the two 'free' sites, i and j, we obtain a reduction to the two-point case:

$$\begin{aligned} Z_{11} &= \sum_{i=1}^{q}\sum_{j=1}^{q} Z_{ij11} \qquad (4.24)\\ &= Z_{1111} + (q-1)[Z_{2111} + Z_{1211} + Z_{2211}] + (q-1)(q-2)Z_{2311}. \end{aligned}$$

Among the exterior dual spins, s_4 has been separated from $s_1 = s_2 = s_3$, so that $Z^*_{11} = Z^*_{1111}$ and $Z^*_{12} = Z^*_{1112}$. Using the known result for Z_{11}, Wu's Eq. (12), along with the duality we find that

$$Z_{1111} + (q-1)[Z_{2111} + Z_{1211} + Z_{2211}] + (q-1)(q-2)Z_{2311} = C[Z^*_{1111} + (q-1)Z^*_{1112}]. \quad (4.25)$$

The equation is one out of a set of six obtained by connecting two of the sites i, j, k and l with a bond. Further equations can be found by considering Z_{12} instead of Z_{11}, but they can be shown to be linear combinations of Eq. (4.23) and the six equations just obtained.

Another set of four equations can be found by letting one of the sites i, j, k and l be 'free' and connecting the remaining three. This corresponds to a reduction to the three-point case. For example, adding a bond between sites j and k and another between k and l, we find that $Z_{111} = \sum_{i=1}^{q} Z_{i111} = Z_{1111} + (q-1)Z_{2111}$. The exterior dual spins now satisfy $s_2 = s_1$, and from Wu's Eq. (25) for Z_{111} and the duality we obtain

$$Z_{1111} + (q-1)Z_{2111} = \frac{C}{q}\{Z^*_{1111} + (q-1)[Z^*_{2211} + Z^*_{1121} + Z^*_{1112}] + (q-1)(q-2)Z^*_{1123}\}. \quad (4.26)$$

Again, the equations found by reducing to Wu's expressions for Z_{123}, Z_{211}, Z_{121} or Z_{112} do not contain any new information.

The remaining equations can be found by connecting all of the sites i, j, k and l with auxiliary bonds of strength K. One way of doing this corresponds to Wu's Eq. (24) and implies connecting all the four sites to a common new point n outside the boundary of $\mathcal{L}$. All of the exterior dual spins are now separated from one another. Summing over n we can express the partition function $\tilde{Z}$ of this system through the $Z_{\sigma_i\sigma_j\sigma_k\sigma_l}$ and relate it by duality to its corresponding dual partition function $\tilde{Z}^*$. The result is

$$\begin{aligned}
&Z_4(e^{4K}+q-1)+(q-1)Z_3(e^{3K}+e^K+q-2)+\\
&(q-1)Z_{2\mathrm{p}}(2e^{2K}+q-2)+(q-1)(q-2)Z_{\mathrm{p}}(e^{2K}+2e^K+q-3)+\\
&(q-1)(q-2)(q-3)Z_0(4e^K+q-4)=\\
&\quad\frac{C}{q^2}(e^K-1)^4\Big\{Z_4^*e^{4K^*}+(q-1)[Z_3^*+Z_{2\mathrm{p}}^*-Z_{2121}^*]e^{2K^*}+\\
&\quad(q-1)Z_{2121}^*+(q-1)(q-2)[Z_{\mathrm{p}}^*-Z_{1213}^*-Z_{2131}^*]e^{K^*}+\\
&\quad(q-1)(q-2)[Z_{1213}^*+Z_{2131}^*]+(q-1)(q-2)(q-3)Z_0^*\Big\}\,. \qquad (4.27)
\end{aligned}$$

From this equation the quantity Z_4 can be found by substituting the duality relation $e^{K^*}=1+q/(e^K-1)$ on the right-hand side, expanding, and comparing the terms multiplying $(e^K-1)^4$ on both sides:

$$\begin{aligned}
Z_4=\frac{C}{q^2}\Big\{&Z_4^*+(q-1)[Z_3^*+Z_{2\mathrm{p}}^*]+\\
&(q-1)(q-2)Z_{\mathrm{p}}^*+(q-1)(q-2)(q-3)Z_0^*\Big\}\,. \qquad (4.28)
\end{aligned}$$

The other way of connecting all four sites is to connect two of them to a new point m and the remaining two to another new point n. For instance, connecting each of i and j to m with a bond of strength K, and similarly k and l to n, we find that s_2 and s_4 are separated whilst $s_3=s_1$. Summing over m and n as before we arrive at

$$\begin{aligned}
&\Sigma_{11}^2\{Z_4+(q-1)Z_{2211}\}+\Sigma_{11}\Sigma_{21}\{(q-1)Z_3+\\
&(q-1)(q-2)[Z_{1123}+Z_{2311}]\}+\Sigma_{21}^2\{(q-1)[Z_{2112}+Z_{2121}]+\\
&(q-1)(q-2)[Z_{\mathrm{p}}-Z_{1123}-Z_{2311}]+(q-1)(q-2)(q-3)Z_0\}=\\
&\quad\frac{C}{q}(e^K-1)^4\Big\{Z_4^*e^{4K^*}+(q-1)[Z_{1211}^*+Z_{1112}^*]e^{2K^*}+\\
&\quad(q-1)Z_{2121}^*+(q-1)(q-2)Z_{1213}^*\Big\}\,, \qquad (4.29)
\end{aligned}$$

where we have defined $\Sigma_{11}=e^{2K}+(q-1)$ and $\Sigma_{21}=2e^K+(q-2)$. A similar equation can be obtained by connecting i and l to m, and j and k to n.

Exactly at this point we get into trouble. For the 15 unknown quantities appearing on the right-hand side of Eq. (4.22) we have obtained 14 independent equations, namely Eq. (4.23), six equations of the type (4.26), four equations of the type (4.27), Eq. (4.27), and two equations of the type (4.29). Clearly, what is missing is a third equation of the type (4.29) in which i and k are connected

to m, and j and l are connected to n. But the auxiliary bonds making such a connection would necessarily intersect, thus violating the planarity of the graph $\mathcal{L}$. In other words, it would be impossible to define the four exterior dual spins.

This problem has to do with the connectedness [44] of the four points. Generally, for an n-point correlation the number of partition functions $Z_{\sigma_1\sigma_2\ldots\sigma_n}$ with fixed values of the n boundary spins equals the number of ways b_n in which n points can be interconnected, as is easily seen by interpreting $\sigma_i = \sigma_j$ as a connection between sites i and j. But the number of equations obtainable using the method of Wu is only equal to the number c_n of well-nested n-point connectivities [44], which is in general less than b_n. For $n = 1, 2, 3, 4, 5, \ldots$ we have $b_n = 1, 2, 5, 15, 52, \ldots$, whilst $c_n = 1, 2, 5, 14, 42, \ldots$.

In an attempt to obtain more equations one could imagine introducing connections between spins in the infinite face, thus converting any spin of valence four or higher into interconnected spins of valence three.[3] In the case of the four-point function this would mean discarding Eq. (4.27) and in addition to the two equations of type (4.29) considering two new equations obtained from these by connecting spins m and n with a fifth auxiliary bond. The first of these, corresponding to Eq. (4.29), looks like

$$\begin{aligned}
&Z_4(\Sigma_{11}^2 + \Sigma_4) + (q-1)Z_3(\Sigma_{11}\Sigma_{21} + \Sigma_3) + (q-1)Z_{2211}(\Sigma_{11}^2 + \Sigma_{2\mathrm{p}}) + \\
&(q-1)(q-2)[Z_{1123} + Z_{2311}](\Sigma_{11}\Sigma_{21} + \Sigma_{\mathrm{p}}) + \\
&(q-1)(q-2)[Z_{\mathrm{p}} - Z_{1123} - Z_{2311}](\Sigma_{21}^2 + \Sigma_{\mathrm{p}}) + \\
&(q-1)[Z_{2121} + Z_{2112}](\Sigma_{21}^2 + \Sigma_{2\mathrm{p}}) + (q-1)(q-2)(q-3)Z_0(\Sigma_{21}^2 + \Sigma_0) = \\
&\quad \frac{C}{q^3}(e^K - 1)^5 \Big\{ Z_4^* e^{5K^*} + (q-1)[Z_{1211}^* + Z_{1112}^*]e^{3K^*} + \\
&\quad (q-1)[Z_{2111}^* + Z_{1121}^* + Z_{2211}^* + Z_{2112}^*]e^{2K^*} + (q-1)Z_{2121}^* e^{K^*} + \\
&\quad (q-1)(q-2)[Z_{\mathrm{p}}^* - Z_{2131}^*]e^{K^*} + (q-1)(q-2)Z_{2131}^* + \\
&\quad (q-1)(q-2)(q-3)Z_0^* \Big\}\,,
\end{aligned} \tag{4.30}$$

where we have defined $\Sigma_4 = (e^K - 1)(e^{4K} + q - 1)$, $\Sigma_3 = (e^K - 1)(e^{3K} + e^K + q - 2)$, $\Sigma_{2\mathrm{p}} = (e^K - 1)(2e^{2K} + q - 2)$, $\Sigma_{\mathrm{p}} = (e^K - 1)(e^{2K} + 2e^K + q - 3)$, and $\Sigma_0 = (e^K - 1)(4e^K + q - 4)$. Now, subtracting the sum of Eqs. (4.30) from the sum of Eqs. (4.29) we regain, after some tedious algebra, $2(e^K - 1)$ times Eq. (4.27). But, on the other hand, subtracting the difference of Eqs. (4.30) from the difference of Eqs. (4.29) we find the result *zero*, *i.e.*, out of the four equations where all the four spins are connected to auxiliary bonds there are still only three that are linearly independent.

Although the method thus fails to isolate the $Z_{\sigma_i\sigma_j\sigma_k\sigma_l}$, which would enable us to construct the generalised correlation function $P_4(\sigma, \sigma', \sigma'', \sigma''')$ as defined in Wu's Eq. (1), the ordinary correlation function $\Gamma_4(\sigma_i, \sigma_j, \sigma_k, \sigma_l)$ (see Wu's Eq. (2)) has been fixed by our determination of Z_4 in Eq. (4.28). By finding linear combinations of the above set of equations for which the left-hand sides

[3] This was suggested to us by F. Y. Wu.

are symmetric in the sense of Eq. (4.22) one could pursue the more modest goal of obtaining the remaining four of these symmetrised combinations.

Clearly, Eqs. (4.23) and (4.27) are already in this form. The sum of the six equations of type (4.26) yields

$$6Z_4 + 3(q-1)Z_3 + 2(q-1)Z_{2\mathrm{p}} + (q-1)(q-2)Z_{\mathrm{p}} = C\left\{6Z_4^* + (q-1)[Z_3^* + Z_{2\mathrm{p}}^* - Z_{2121}^*]\right\}, \tag{4.31}$$

while the four equations of type (4.27) sum up to

$$4Z_4 + (q-1)Z_3 = \tag{4.32}$$

$$\frac{C}{q}\left\{4Z_4^* + (q-1)(q-2)[Z_{\mathrm{p}}^* - Z_{1213}^* - Z_{2131}^*] + (q-1)[2Z_3^* + 2(Z_{2\mathrm{p}}^* - Z_{2121}^*)]\right\}. \tag{4.33}$$

As before, employing the two-point reduction Z_{12} or the three-point reductions Z_{123}, Z_{211}, Z_{121} and Z_{112} merely provides us with linear combinations of what we already know.

But again, the two equations of the type (4.29) lead us into trouble. When taking their sum it can easily be seen that the left-hand side cannot be written in terms of the symmetrical combinations (4.22). To symmetrise we would again need the non-well nested equation, which however leaves the exterior dual spins undefined.

In conclusion we have shown that, although the ordinary correlation function is fairly easily determined, the method of Ref. [35] does not permit a complete evaluation of the generalised correlation function for $n \geq 4$ points, due to the appearance of non-well nested connectivities.

4.3.3 Sum rule identities and general solution

Despite this apparent shortage of equations the general problem of mapping the planar n-point correlator under duality can be solved. This was made clear when Wu and Huang subsequently realised that the c_n equations furnished by drawing well-nested auxiliary graphs [36] must be supplemented by certain new sum rule identities [37]. The number of such identities is exactly $b_n - c_n$, thus allowing for a complete determination of the b_n quantities $Z_{ij\ldots kl}$.

The idea is surprisingly simple, at least for $n = 4$ which we consider first. Imagine expanding the restricted partition functions as sums over graphical configurations in the random cluster model. In Eq. (4.3) the number of clusters $C(G)$ now excludes those that are connected to any of the four boundary spins, which are taken to have fixed values. The contributions to the 'non-well nested' quantity Z_{1212} can then be written as a sum of three terms

$$Z_{1212} = T_1 + T_2 + T_3, \tag{4.34}$$

where T_1 are the graphs where sites i and k belong to the same cluster, T_2 are those where j and l belong to the same cluster, and finally T_3 are the graphs

where all four boundary spins i, j, k and l belong to different clusters. Note that it is of course possible for two boundary spins to be in the same fixed state without belonging to the same cluster, but not conversely!

From the same line of reasoning it is then found that

$$Z_{1213} = T_1 + T_3, \qquad Z_{2131} = T_2 + T_3, \qquad Z_{1234} = T_3, \tag{4.35}$$

and we thus infer the *sum rule* identity

$$Z_{1212} = Z_{1213} + Z_{2131} - Z_{1234}. \tag{4.36}$$

This provides the fifteenth equation missed so far, and indeed allows for a complete determination of all the Z_{ijkl}.

In general, the existence of one sum rule identity for each non-well nested $Z_{ij\cdots kl}$ originates from the planarity of the random cluster graphs. The identities then follow from the principle of inclusion-exclusion [45]. All identities thus obtained are distinct due to the uniqueness of the random cluster expansion.

Having obtained explicit expressions for the fifteen Z_{ijkl} in the case $n = 4$ the authors of Ref. [37] were then able to conjecture the structure of the solution for general n. Very recently this conjecture has been proven by Lu and Wu [38].

4.3.4 Cylindrical boundary conditions

The question now arises how these results can be generalised to other geometries (*i.e.*, $\mathcal{L}$ non-planar). Here the case of the four-point correlator with two points placed as nearest neighbours on each end of a cylinder is especially interesting, since it is related to the energy-energy correlation of the Potts model. If one could sort out how such a correlation function maps under duality and then go on to compute it by imposing certain boundary conditions on the dual lattice transfer matrix, one would have devised a useful alternative way of computing the thermal scaling dimension x_t from what is essentially a free energy. The existence of such a method would be particularly desirable in the context of the random-bond Potts model, since here the lack of self-averaging leads to highly non-trivial problems about using generalisations of standard conformal field theory predictions to extract the scaling dimensions from the transfer matrix (Lyapunov) spectrum [46, 47]. These aspects will be discussed further in Chapter 6, where we shall also shown how the duality relation for the two-point function in a cylindrical topology, Eq. (4.14), leads to a very powerful method of determining the *magnetic* exponent x_h.

Writing, for general n,

$$Z_{x_1 x_2 \cdots x_n} = C_n(q) \sum_{y_1 y_2 \cdots y_n} M(x_1, x_2, \ldots, x_n | y_1, y_2, \ldots, y_n) Z^*_{y_1, y_2, \cdots, y_n} \tag{4.37}$$

with $C_n(q) = q^{-n+(N-N^*)/2}$ [37], duality can be viewed as a linear transformation with coefficients equal to the elements of some matrix $\mathbf{M}$. The dimension of $\mathbf{M}$ is $q^n \times q^n$, but taking into account the permutation symmetry between

the states of the n Potts spins the effective dimensionality is in fact only $b_n \times b_n$. Applying the duality transformation twice one ends up with the identity transformation

$$\mathbf{M}^2(x_1, x_2, \ldots, x_n | x_1', x_2', \ldots, x_n') = q^{n+1}\delta(x_1, x_2')\delta(x_2, x_3') \cdots \delta(x_n, x_1'), \tag{4.38}$$

apart from a trivial constant and a cyclic permutation of the indices[4].

This *reciprocal inversion relation* [37] is valid for any geometry that can be embedded in the plane, at does not depend on the existence of auxiliary graphs. To be specific, let us imagine the cylinder embedded in the plane, in which case its boundaries are topologically equivalent to those of an annulus. The physically interesting case corresponds to boundary spins i and j residing on the exterior boundary $\partial\mathcal{L}_1$, and spins k and l on the interior boundary $\partial\mathcal{L}_2$.

At first sight Eq. (4.38) provides us with $15 \times 15 = 225$ equations for the 15 unknown quantities Z_{ijkl}. However, a detailed study reveals that a lot of these are in fact linearly dependent. This information can be supplemented with relations obtained from $c_2 \times c_2 = 4$ auxiliary graphs, but despite a thorough investigation undertaken by Wu and the present author this still does not appear sufficient to determine the duality relations. A simple derivation in the spirit of Eq. (4.28) is impossible since the topology at hand does not allow for auxiliary graphs interconnecting spins on $\partial\mathcal{L}_1$ and $\partial\mathcal{L}_2$. Another worrying observation is that the result (4.14) for the equivalent $n = 2$ case reveals that the physically interesting boundary condition has to do with a *seam* connecting $\partial\mathcal{L}_1$ and $\partial\mathcal{L}_2$, and it is not obvious how this can be related to boundary condition on the dual spins adjacent to $\partial\mathcal{L}_{1,2}$.

Resumo en Esperanto

Ni montras ke la Potts-modelo ekvivalentas al la grapolmodelo (4.3) en kiu ĉiu konfiguracio estas eĝperkola grafikaĵo kiel tiu de Fig-o 4.1. Ĉiu sendependa 'grapolo' portas faktoron de q. Ĉi-reprezentiĝo faciligas la diskuton de dualeco (4.7). Ajna aŭtoduala distribuo (4.11) de la aleatoraj kuplaĵoj lokas la sistemon sur la kritan punkton. La dupunkta korelacia funkcio en cilindra geometrio dualas al orlo da frustritaj eĝoj (4.14).

Nova diagrama metodo de Wu [35] permesas la dualigon de n-punkta *randa* korelacio por ebena reto (Fig-o 4.2). Por $n \geq 4$ estas malpli da diagramoj (t.e. ekvacioj) ol nekonataj $Z_{\sigma_1,\sigma_2,\ldots,\sigma_n}$, kaj necesas suplementi ilin per sumreguloj (4.36). La analoga problemo por malebena (ekz-e cilindra) reto restas nesolvita.

[4] The permutation arises when both the x_i and the y_i variables follow one another in a clockwise fashion. By orienting the y_i anti-clockwise this asymmetry could be circumvented.

Chapter V

Construction of the transfer matrices

Although a large amount of high-precision results has been obtained by combining transfer matrix (TM) techniques with finite-size scaling for almost any conceivable type of pure statistical mechanics system (see, e.g., Ref. [48] for a review) the use of TMs in the study of disordered systems seems to have attracted rather little interest as compared with the complementary approach of Monte Carlo simulations.

A straightforward way of setting up the TMs for the q-state Potts model is to use the traditional spin basis where the state of a row of L spins is labeled by the q^L basis states $\{\sigma_1, \sigma_2, \ldots, \sigma_L\}$, $\sigma_i = 1, \ldots, q$. Whilst this approach is highly efficient for $q = 2, 3$ it has two major shortcomings in the general case. First, the dimension of the matrices grows exponentially with q, in particular making inaccessible the regime of $q > 4$ which is our main concern. Second, the restriction to integer values of q is unnecessary and in fact makes it difficult to compare numerical results with analytical calculations in the $(q-2)$-expansion [24, 25, 26].

Both these shortcomings can be remedied by writing the TMs in the connectivity basis introduced by Blöte and Nightingale [44]. In this representation the dimension of the TMs is independent of q which enters only as a *continuous* parameter. In fact, the number of basis states is asymptotically $\sim 4^L$ (or $\sim 5^L$ upon imposition of a magnetic field) with a rather small coefficient of proportionality, in practice making this basis the preferred choice for all but the Ising model ($q = 2$).

We have generalised these TMs to include quenched bond randomness, and also devised an alternative method of accessing the magnetic properties through the introduction of a seam along the strip. Furthermore, in the percolation limit the TMs are found to simplify in a manner that makes calculations for rather large strip widths feasible. For convenience these results are presented along with a review of the relevant parts of Ref. [44] thus making our description of the Potts model TMs self-contained.

Before turning to the details on this construction, let us briefly recapitulate how numerical results about the universal behaviour of a system at a second

order phase transition can be extracted from the transfer matrices.

5.1 Numerical results from transfer matrices

Although our transfer matrices will apply to the random cluster model with any inhomogeneous distribution of the bond strengths and the cluster weights[1] we shall mainly be concerned with the random bond Potts model (RBPM) where the quenched random couplings K_{ij} are drawn from the symmetric binary distribution[2]

$$P(K) = \frac{1}{2}[\delta(K - K_1) + \delta(K - K_2)]. \tag{5.1}$$

Here the ratio between strong and weak bonds $R = K_2/K_1$ measures the strength of the randomness. For the special choice

$$\left(\mathrm{e}^{K_1} - 1\right)\left(\mathrm{e}^{K_2} - 1\right) = q \tag{5.2}$$

the model is on average self-dual, as was discussed in detail in Chapter 4. Assuming that the phase transition is unique the model is therefore at its critical point [30].

5.1.1 Eigenvalue and Lyapunov spectra

One could thus hope that the transfer matrices constructed in accordance with Eqs. (5.1) and (5.2) do somehow contain exact numerical information about the universal behaviour of the RBMP at its critical point. That this is indeed the case for any *pure* system that is conformally invariant constitutes one of the remarkable triumphs of the application of conformal field theory (CFT) to the description of two-dimensional critical phenomena.

Namely, it is well-known that in the pure case ($R = 1$) the operator content of the CFT underlying the model is related to the eigenvalue spectrum $\{\lambda_i(L)\}$, $i = 0, 1, 2, \ldots$, of the TM for a strip of width L through [49]

$$f_i(L) - f_0(L) = \frac{2\pi x_i}{L^2} + \cdots, \tag{5.3}$$

where $f_i(L) = -\frac{1}{L}\ln\lambda_i(L)$ are the generalised free energies per site (in units of $k_{\mathrm{B}}T$) and x_i the scaling dimensions of the corresponding operators.

Similarly the central charge (also known as the conformal anomaly) c, measuring the number of bosonic degrees of freedom of the CFT, is related to the

[1] We do not know of any model, though, where the latter generalisation is of any physical relevance.

[2] Other self-dual distributions of the random bonds than that of Eq. (5.1) have also been investigated in order to check our results. In particular, we have found the trinary distribution introduced in Sect. 6.2 useful, since it gives us a clearer idea about the length scale associated with the random impurities.

finite-size corrections to the customary free energy through [50, 51]

$$f_0(L) = f_0(\infty) - \frac{\pi c}{6L^2} + \cdots . \tag{5.4}$$

Both these results can be derived by conformally mapping the plane onto the strip through the logarithmic transformation $w(z) = \frac{L}{2\pi} \ln z$.

In the random case the TMs are no longer constant but depend on the particular realisation of the random bonds within each row of the strip. Accordingly the concept of eigenvalues generalises to that of Lyapunov exponents. Starting with some suitable initial vector of unit norm $|\mathrm{v}_0\rangle$, the leading Lyapunov exponent can be found by the Furstenberg method [52]

$$\Lambda_0(L) = \lim_{m\to\infty} \frac{1}{m} \ln \left\| \left(\prod_{j=1}^{m} \mathcal{T}_j \right) |\mathrm{v}_0\rangle \right\| , \tag{5.5}$$

where $\mathcal{T}_j$ is the TM acting between rows $j-1$ and j. The average free energy per site is given as before by $f_0(L) = -\frac{1}{L}\Lambda_0(L)$. Higher exponents are found by iterating a set of n vectors $\{|\mathrm{v}_k\rangle\}_{k=0}^{n-1}$, where a given $|\mathrm{v}_k\rangle$ is orthogonalised to the set $\{|\mathrm{v}_l\rangle\}_{l=0}^{k-1}$ after each multiplication by $\mathcal{T}_j$ [53]. Surprisingly, this method works even for a non-hermitian TM, and it is numerically shown to be independent of the choice of the initial vectors.

When some symmetry (*e.g.*, spin reversal or duality) is manifest in $\mathcal{T}_j$ the orthogonalisation can be circumvented by iterating vectors which belong to definite irreducible components of that symmetry, but the S_q permutational symmetry inherent in the Potts model has been lost through the mapping to the random cluster model which forms the backbone of our TMs; see Sect. 5.2.

As to the extraction of physical information from the spectra, Eq. (5.4) is supposed to retain its validity provided that c is replaced by the effective central charge c', that in the standard replica formalism is the derivative of $c(n)$ with respect to the number of replicas at $n = 0$ [23]. The question to which extent Eq. (5.3) also remains valid is by no means trivial and we shall dedicate a fair part of the discussion in Chapter 6 to it.

5.2 Mapping to the random cluster model

Introducing an imaginary 'ghost site' with fixed spin $\sigma_0 = 1$ the partition function for the Potts model can be written as

$$Z = \sum_{\{\sigma\}} \left(\prod_{\langle ij \rangle} \exp(K_{ij}\delta_{\sigma_i\sigma_j}) \right) \left(\prod_{\langle i0 \rangle} \exp(H_i\delta_{\sigma_i\sigma_0}) \right) , \tag{5.6}$$

where $\prod_{\langle ij \rangle}$ is the usual product over pairs of nearest neighbour sites and each site i has been connected to the ghost site 0 with a similar notation. The reduced magnetic field H_i, here taken to be site-dependent, now enters at the

same footing as the reduced exchange couplings K_{ij}. It should be pointed out, however, that a random coupling to the ghost site is not a true random field, since the latter would try to force different sites into *different* Potts states and not just into the particular state of the ghost site with a site-dependent probability. To avoid any confusion we shall therefore specialise to the case of a homogeneous field $H_i \equiv H$.

The site variables can now be traded for bond variables through the mapping to the random cluster model introduced by Kasteleyn and Fortuin [31]. Details on the case of $H = 0$ have already been given in Sect. 4.1.1. In terms of the variables $u_{ij} = \mathrm{e}^{K_{ij}} - 1$ and $v = \mathrm{e}^{H} - 1$ the appropriate generalisation to the case of $H > 0$ reads

$$Z = q^N \sum_{G \subseteq \mathcal{L}} \sum_{G_0 \subseteq \mathcal{L}_0} \left(\prod_{\langle ij \rangle \in G} \frac{u_{ij}}{q} \right) \left(\prod_{\langle i0 \rangle \in G_0} \frac{v}{q} \right) q^{l(G \cup G_0)}, \tag{5.7}$$

where $\mathcal{L}$ denotes the set of all nearest neighbour bonds, $\mathcal{L}_0$ the bonds from each of the N Potts spin to the ghost site, and $l(G \cup G_0)$ is the number of independent loops on the combined graph $G \cup G_0$.

The usual construction of the transfer matrix $\mathcal{T}$ for a strip of width L seems to be obstructed by the non-local factor $l(G \cup G_0)$, but this can be taken into account by choosing a basis containing information about which sites of a given row are interconnected through the part of the lattice below[3] that row (including connections via the ghost site). This leads us to the concept of connectivity states, which we consider next.

5.3 The connectivity states

In order to determine the number of loop closures induced by appending a new row of L sites along with the corresponding L connections to the ghost site to the top of $G \cup G_0$, we need information about how the sites in the top row of $G \cup G_0$ were previously interconnected. This information is comprised in the *connectivity state* $(i_1 i_2 \ldots i_L)$, where $i_t = 0$ if site t is connected to 0 within the combined graph $G \cup G_0$ and, otherwise, $i_r = i_s$ is a (non-unique) positive integer if and only if sites r and s are connected within G.

Whilst this 'index representation' is useful for determining whether a newly appended bond does or does not close a loop, and thus will allow us to explicitly construct the single-bond TMs in the next subsection, a one-to-one mapping to the set of consecutive integers $\{1, 2, \ldots\}$ is clearly needed to define a 'number representation' which will enable us to label the entries of the TM and thus to perform actual computations. These representations and the mapping were supplied by Ref. [44] as were the determination of the number of connectivity states (d_L with and c_L without a magnetic field). We shall review the necessary details and also give details on the construction of the inverse of the mapping just mentioned.

[3] Unlike what is commonly the case we take the TM to act in the positive y-direction.

5.3.1 Enumeration and ordering of the connectivities

Consider first the case of $H = 0$ where all ghost bonds carry zero weight ($v = 0$). The connectivity states then have all $i_t > 0$ and can be recursively ordered by noting that the index representation is *well-nested*, *i.e.*, for $r < s < t < u$

$$(i_r = i_t) \wedge (i_s = i_u) \Rightarrow i_s = i_t . \tag{5.8}$$

It follows that if we define the *cut function* $\rho(i_1 i_2 \ldots i_L)$ to be the smallest $t > 1$ such that $i_1 = i_t$, if such a t exists, and $L + 1$ otherwise, the left $(i_2 i_3 \ldots i_{\rho-1})$ and right $(i_\rho i_{\rho+1} \ldots i_L)$ parts of the index representation are both well-nested. A complete ordering of the well-nested sequences is now induced by applying the cut function first to the whole sequence, then recursively to its right and finally to its left part.

More precisely, the mapping from the index to the number representation is effected by

$$\sigma(i_1 i_2 \ldots i_L) = \begin{cases} 1 \text{ if } L \leq 1 \\ c_{L,k-1} + [\sigma(i_k \ldots i_L) - 1] c_{k-2} + \sigma(i_2 \ldots i_{k-1}) \text{ otherwise,} \end{cases} \tag{5.9}$$

where $k = \rho(i_1 i_2 \ldots i_L)$ and $c_{n,l} = \sum_{i=2}^{l} c_{i-2} c_{n-i+1}$ with

$$c_n \equiv c_{n,n+1} = \frac{(2n)!}{n!(n+1)!} \tag{5.10}$$

giving the number of well-nested n-point connectivities [44, 54]. Explicit values are shown in Table 5.1.

To consider the general case of $v \neq 0$ we remark that the subsequence of non-zero indices is still well-nested. A complete ordering of an index representation $(i_1 i_2 \ldots i_L)$ with precisely s zero indices is then induced by first ordering according to the value of s, then lexicographically ordering the zeros, and finally using the ordering of the well-nested subsequence $(i_{p_1} i_{p_2} \ldots i_{p_{L-s}})$ given by Eq. (5.9). The lexicographic ordering is carried out by

$$\psi(i_1 i_2 \ldots i_L) = \begin{cases} 1 \text{ if } L = 1 \text{ or } s = L \\ \psi(i_2 i_3 \ldots i_L) \text{ if } i_1 \neq 0 \\ \binom{L-1}{s} + \psi(i_2 i_3 \ldots i_L) \text{ if } i_1 = 0, \end{cases} \tag{5.11}$$

and the mapping to the number representation is finally

$$\tau(i_1 i_2 \ldots i_L) = d_{L,s-1} + [\psi(i_1 i_2 \ldots i_L) - 1] c_{L-s} + \sigma(i_{p_1} i_{p_2} \ldots i_{p_{L-s}}), \tag{5.12}$$

where $d_{n,l} = \sum_{i=0}^{l} \binom{n}{i} c_{n-i}$ with

$$d_n \equiv d_{n,n} = \sum_{i=0}^{n} \binom{n}{i} c_{n-i} \tag{5.13}$$

giving the number of general n-point connectivities. Again, explicit values are presented in Table 5.1.

L	c_L	d_L	$d_L - c_L$	Lc_L	b_L
1	1	2	1	1	1
2	2	5	3	4	2
3	5	15	10	15	5
4	14	51	37	56	15
5	42	188	146	210	52
6	132	731	599	792	203
7	429	2,950	2,521	3,003	877
8	1,430	12,235	10,805	11,440	4,140
9	4,862	51,822	46,960	43,758	21,147
10	16,796	223,191	206,395	167,960	115,975
11	58,786	974,427	915,641	646,646	678,570
12	208,012	4,302,645	4,094,633	2,496,144	4,213,597

Table 5.1: The number of connectivity states for a Potts model transfer matrix of width L with (d_L) and without (c_L) an external magnetic field. Also shown is the size of the magnetic sector when using a ghost site ($d_L - c_L$) and a seam (Lc_L). For large strip widths the seam is advantageous. The number c_L of well-nested L-point connectivities should be compared to the total number of L-point connectivities b_L which increases *faster* than exponentially as a function of L.

5.3.2 The inverse mapping

To construct the inverse mapping, *i.e.*, the one taking us from the number to the index representation, we solve $\tau = \tau(i_1 i_2 \dots i_L)$ for the indices $(i_1 i_2 \dots i_L)$ by performing the following steps. First, the number of zero indices is found as $s = \max\{s | d_{L,s-1} < \tau\}$. Second, perform a slightly modified integer division by writing

$$\tau - d_{L,s-1} = Q c_{L-s} + R, \tag{5.14}$$

where the remainder R is restricted to take its values in the interval $[1, c_{L-s}]$. From Eq. (5.12) we infer that $\psi = Q + 1$ and $\sigma = R$. Third, the position of the first (leftmost) zero index is given by $i_{L-k} = 0$, where $k = \max\{k | \binom{k}{s} < \psi\}$. This procedure of finding the zero indices is then iterated with

$$\psi \to \psi^{(1)} \equiv \psi - \binom{k}{s} \tag{5.15}$$

until $\psi^{(s')} = 1$, and the remaining $s - s'$ zeros are filled in from the right: $i_{s'+1} = \dots = i_{s-1} = i_s = 0$.

It remains to deduce the subsequence of non-zero indices by inverting $\sigma = \sigma(i_{p_1} i_{p_2} \dots i_{p_l})$ with $l = L - s$. After initialising $i_{p_1} = p_1$ we proceed by recursion as follows. First, choose

$$k = \min\{k | c_{l,k-1} + c_{k-2} c_{l-k+1} \geq \sigma\}. \tag{5.16}$$

If $k \leq l$ we have then found a connection: $i_{p_1} = i_{p_k}$. This procedure of finding the connections is now iterated on the left $(i_{p_2}, \ldots, i_{p_{k-1}})$ and the right $(i_{p_k}, \ldots, i_{p_l})$ parts of the remaining sequence. If $k \geq 2$ the assignment $i_{p_2} = p_2$ is performed. By (modified) integer division we then write

$$\sigma - c_{l,k-1} = Q c_{k-2} + R \tag{5.17}$$

with $R \in [1, c_{l-k+1}]$, and pass over the left part of the sequence with

$$\sigma \to \sigma^{(1)}_{\text{left}} \equiv R, \quad l \to l^{(1)}_{\text{left}} \equiv k-2, \tag{5.18}$$

and the right part with

$$\sigma \to \sigma^{(1)}_{\text{right}} \equiv Q+1, \quad l \to l^{(1)}_{\text{right}} \equiv l-k+1. \tag{5.19}$$

The recursion stops when for any sequence $l^{(m)} \leq 2$. If then $l^{(m)} = 2$ and the sequence is $(i_{p_a}, i_{p_{a+1}})$ we perform the assignment $i_{p_{a+1}} = i_{p_a}$ if $\sigma^{(m)} = 1$ and $i_{p_{a+1}} = p_{a+1}$ if $\sigma^{(m)} = 2$.

Any way of constructing the index representation $(i_1 i_2 \ldots i_L)$ will of course reflect the above-mentioned arbitrariness as to the actual values of the non-zero indices, but the particular procedure just outlined is easily seen to ensure that all indices are $\leq L$. This invariant is useful since then any given site t can be *disconnected* from the rest by assigning $i_t = L+1$.

5.4 The single-bond transfer matrices

The amount of computer time necessary for building up a long strip by repeated application of the transfer matrix $\mathcal{T}$ can be enormously reduced by decomposing the latter as a product of sparse matrices, each corresponding to the addition of a single bond to $\mathcal{L}$.

Specifically we write $\mathcal{T} = \mathcal{T}^0 \mathcal{T}^{\text{h}} \mathcal{T}^{\text{v}}$, where

$$\mathcal{T}^{\text{v}} = \mathcal{T}^{\text{v}}_L \cdots \mathcal{T}^{\text{v}}_2 \mathcal{T}^{\text{v}}_1 \tag{5.20}$$

is connecting each of the L spin sites in the uppermost row of the strip to a new spin site situated *vertically* above it, and

$$\mathcal{T}^{\text{h}} = \mathcal{T}^{\text{h}}_{L,1} \cdots \mathcal{T}^{\text{h}}_{2,3} \mathcal{T}^{\text{h}}_{1,2} \tag{5.21}$$

is finishing the new row of $\mathcal{L}$ by appending *horizontal* bonds between each of the nearest-neighbour dangling ends created by $\mathcal{T}^{\text{v}}$. The matrix $\mathcal{T}^{\text{h}}_{L,1}$ imposes periodic boundary conditions by interconnecting the newly added spins at sites L and 1. Finally

$$\mathcal{T}^0 = \mathcal{T}^0_L \cdots \mathcal{T}^0_2 \mathcal{T}^0_1 \tag{5.22}$$

furnishes the bonds of $\mathcal{L}_0$ from each of the new spin sites to the ghost site. Each of these single-bond TMs is implicitly understood to depend on the particular

realisation of the bond randomness and, in the case of $\mathcal{T}_i^0$, the possibly random coupling to the ghost site pertaining to the site in question (but see Sect. 5.2).

Upon addition of one single bond the summation over graphs in Eq. (5.7) is augmented by a sum over the two possible states of this new degree of freedom, viz. the bond added to $\mathcal{L}$ ($\mathcal{L}_0$) can be either *present* or *absent* in G (G_0). Correspondingly each column of the TM has at most two distinct non-zero entries.

Consider first adding a vertical bond by action of $\mathcal{T}_l^{\rm v}$, $l \in \{1, \ldots, L\}$. If the bond is 'present' any given connectivity state $(i_1 i_2 \ldots i_L)$ of the L uppermost spin sites will be left unchanged. In case of an 'absent' bond site l will be disconnected, and the number representation of the new connectivity state can be found by assigning $i_l = L + 1$ and using Eq. (5.12). Interpreting the factor of q^N in Eq. (5.7) as an extra factor of q going with each vertical bond we see that the non-zero entries in $\mathcal{T}_l^{\rm v}$ corresponding to a column with a given connectivity number are a diagonal contribution of u_{ij} and a possibly off-diagonal contribution of q. In particular the vertical bonds do not induce any loop closures.

Similarly the TM of a horizontal bond $\mathcal{T}_{l,l+1}^{\rm h}$ has a diagonal entry of 1 for each column, corresponding to the bond being absent. The other non-zero entry corresponds to a present bond, and its value depends on whether a loop is being closed or not. Given the connectivity state $(i_1 \ldots i_l i_{l+1} \ldots i_L)$ of some column in the TM this is determined by comparing i_l and i_{l+1}: if they are equal we get an additional diagonal contribution of $u_{l,l+1}$ corresponding to a loop closure, whereas if they are different there is an off-diagonal entry with value $u_{l,l+1}/q$. In the latter case the connectivity number is found by assigning the value $\min\{i_l, i_{l+1}\}$ to all indices that were formerly equal to either i_l or i_{l+1} and applying Eq. (5.12). (The reason why we copy the *minimum* index is to ensure the proper handling of spins connected to the ghost site.)

Finally, the TM of a ghost bond $\mathcal{T}_l^{\rm h}$ has the same form as in the case of a horizontal bond if we make the substitutions $u_{l,l+1} \to v_l$ and $i_{l+1} \to 0$.

5.4.1 An example: $L = 3$ and $v = 0$

As a simple illustration of this decomposition of the transfer matrices we shall explicitly construct the single-bond TMs in the non-magnetic sector ($v = 0$) for a strip of width $L = 3$. In this case there are $c_3 = 5$ connectivity states, which are ordered according to Eq. (5.9) as

$$\sigma(111) = 1, \quad \sigma(112) = 2, \quad \sigma(121) = 3, \quad \sigma(122) = 4, \quad \sigma(123) = 5. \tag{5.23}$$

Consider first adding the first vertical bond of strength $u_1^{\rm v}$. The task is to find out what are the possible transitions among the connectivity states and, for each allowed transition, what is the value of the appropriate Boltzmann factor. For each connectivity state the bond may be either 'present' or 'absent'. The former case will lead simply to a diagonal element $u_1^{\rm v}$, whereas the latter will imply the disconnection of site 1 and hence the following non-trivial transitions between connectivities

$$1 \to 4, \quad 2 \to 5, \quad 3 \to 5, \quad 4 \to 4, \quad 5 \to 5. \tag{5.24}$$

In each of these cases site 1 has been liberated and must at the present stage be regarded as an isolated site. Accordingly it will carry a weight of q. The single-bond TM is therefore

$$\mathcal{T}_1^{\mathrm{v}} = \begin{bmatrix} u_1^{\mathrm{v}} & 0 & 0 & 0 & 0 \\ 0 & u_1^{\mathrm{v}} & 0 & 0 & 0 \\ 0 & 0 & u_1^{\mathrm{v}} & 0 & 0 \\ q & 0 & 0 & u_1^{\mathrm{v}}+q & 0 \\ 0 & q & q & 0 & u_1^{\mathrm{v}}+q \end{bmatrix}. \tag{5.25}$$

Expressions for $\mathcal{T}_2^{\mathrm{v}}$ and $\mathcal{T}_3^{\mathrm{v}}$ can be found from analogous considerations, or alternatively by a cyclic shift of the site labels leading to the following permutations of the connectivity states

$$\{1,2,3,4,5\} \to \{1,3,4,2,5\}, \tag{5.26}$$

which can be used to obtain $\mathcal{T}_2^{\mathrm{v}}$ from $\mathcal{T}_1^{\mathrm{v}}$ and again to get $\mathcal{T}_3^{\mathrm{v}}$ from $\mathcal{T}_2^{\mathrm{v}}$.

The matrix $\mathcal{T}_{1,2}^{\mathrm{h}}$ pertains to the horizontal bond of strength $u_{1,2}^{\mathrm{h}}$ connecting sites 1 and 2. For any connectivity state there is a diagonal contribution of 1 corresponding to this bond being 'absent'. A 'present' bond, on the other hand, will lead to the following transitions

$$1 \to 1, \quad 2 \to 2, \quad 3 \to 1, \quad 4 \to 1, \quad 5 \to 2. \tag{5.27}$$

In the two cases where the connectivity state does not change, sites 1 and 2 were connected beforehand and the Boltzmann weight is $u_{1,2}^{\mathrm{h}}$. The three off-diagonal transitions correspond to a merger of two clusters and must hence be weighted by $u_{1,2}^{\mathrm{h}}/q$. Hence

$$\mathcal{T}_{1,2}^{\mathrm{h}} = \begin{bmatrix} 1+u_{1,2}^{\mathrm{h}} & 0 & u_{1,2}^{\mathrm{h}}/q & u_{1,2}^{\mathrm{h}}/q & 0 \\ 0 & 1+u_{1,2}^{\mathrm{h}} & 0 & 0 & u_{1,2}^{\mathrm{h}}/q \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{5.28}$$

Once again, $\mathcal{T}_{2,3}^{\mathrm{h}}$ and $\mathcal{T}_{3,1}^{\mathrm{h}}$ can be found by permuting the basis states according to Eq. (5.26).

Finally, since $v = 0$ in this example, $\mathcal{T}_0$ is just the identity matrix.

5.5 Magnetic properties

It is well known, at least in the case of a pure system, that physically interesting quantities like the central charge c as well as the thermal (x_T) and the magnetic (x_H) scaling dimensions can be extracted from the transfer matrix spectrum; this has already been mentioned in Sect. 5.1.1. Consider for the moment the case of vanishing magnetic field, $H = 0$. Since connections to the ghost site are

then generated with zero weight ($v = 0$) such connections can only be present in any row if they were already there in the preceding row. In particular, noting that in the numbering of connectivities induced by Eq. (5.12) the non-ghost connectivities precede the others, we see that the TM assumes the following block form [44]

$$\mathcal{T} = \begin{bmatrix} \mathcal{T}^{11} & \mathcal{T}^{12} \\ 0 & \mathcal{T}^{22} \end{bmatrix}, \tag{5.29}$$

where superscript 2 (1) refers to the (non-)ghost connectivities.

The largest and the next-largest eigenvalues of $\mathcal{T}$ turn out to be the largest eigenvalue of block $\mathcal{T}^{11}$ and $\mathcal{T}^{22}$ respectively, and from the corresponding (reduced) free energies per site $f_0^{ii}(L) = -\frac{1}{L}\lambda_0^{ii}$ $(i = 1, 2)$ for a strip of width L the magnetic scaling dimension can be found from the CFT formula [49]

$$f_0^{22}(L) - f_0^{11}(L) = \frac{2\pi x_H}{L^2} + \cdots . \tag{5.30}$$

Physically this relation to x_H can be understood by noting that by acting repeatedly with $\mathcal{T}^{22}$ on some initial (row) state $|\mathrm{v}_0\rangle \neq 0$ one measures the decay of clusters extending back to row 0. This must have the same spatial dependence as the spin-spin correlation function and hence be related to x_H [44]. Analogously $\mathcal{T}^{11}$ measures the decay of two-point correlations between *pairs* of spins being interconnected within the random cluster model. This is nothing but the energy-energy correlation in the strip geometry, and accordingly we expect that

$$f_1^{11}(L) - f_0^{11}(L) = \frac{2\pi x_T}{L^2} + \cdots , \tag{5.31}$$

where $f_1^{11}(L)$ is the next-largest eigenvalue of $\mathcal{T}^{11}$.

We have checked the results for x_H by constructing a realisation of the TM in the presence of a *seam* spanning the length of the cylinder. Our algorithm also merits attention on its own right since it improves the asymptotic number of basis states necessary for finding $f_0^{22}(L)$ from $d_L - c_L \sim 5^L$ (the dimension of $\mathcal{T}^{22}$) to $Lc_L \sim L4^L$. In practice, however, with the strip widths L accessible using present-day computers the two algorithms perform more or less equally fast (see Table 5.1 for a comparison).

Let us recall from Sect. 4.2 that the spin-spin correlator for the Potts model on a cylinder is mapped onto a disorder operator under duality. Since the relation between the transfer matrix spectrum and the scaling dimensions of physical operators pertains to the limit of an infinitely long cylinder, and the square lattice wrapped on such a cylinder is in fact self-dual, we need not distinguish between the lattice and its dual. The task then is to compute the (dual) partition function with *twisted boundary conditions*

$$\sigma \to (\sigma + 1) \bmod q \tag{5.32}$$

across a seam spanning the length of the cylinder. The permutational symmetry between the Potts spin states implies that the shape of this seam can be

deformed at will as long as its end-points remain fixed. Obviously the boundary condition (5.32) is equivalent to restraining the clusters of the random cluster model from wrapping around the cylinder.

The construction of the TM in the presence of a seam is facilitated by the following observation: If no cluster is allowed to wrap the cylinder, each graph contributing to the partition function can be associated with a function $s(j)$ of the row number j, such that $s(j) = k \in \{1, \ldots, L\}$ means that in row j no horizontal bond connecting sites k and $k+1 \pmod L$ is present. For obvious reasons we shall refer to s as the *virtual seam*. We can then write the TM in a basis which is the direct product of the L possible values of the virtual seam and the customary c_L non-ghost connectivities. The virtual seam is initialised by assigning to it a definite value in row 0, viz. $s(0) = L$ for all graph configurations of that row.

The single-bond TM of a vertical bond is diagonal in s, but a present horizontal bond *not* inducing a loop closure may alter the value of the virtual seam. Let us recall from Sect. 5.4 that to find the connectivity state $(i_1 \ldots i_l i_{l+1} \ldots i_L)$ giving the row label of $T^{\mathrm{h}}_{l,l+1}$ that corresponds to the off-diagonal entry with value $u_{l,l+1}/q$ we would join the two distinct clusters formerly labeled by either i_l or i_{l+1}. But such a merger would ruin the invariant stated above, unless we move the virtual seam at the same time. On the other hand, if $i_l = i_{l+1}$ and $s(j) = l$ we must explicitly prevent a cluster from wrapping the cylinder by leaving out that extra diagonal contribution which would otherwise by implied by the condition $i_l = i_{l+1}$. In this case the virtual seam is not moved.

5.6 The percolation limit

In the random bond Potts model the couplings $u_{ij} \geq 0$ are quenched random variables, and the critical point can be accessed by drawing them from the symmetric binary distribution $P(u) = \frac{1}{2}[\delta(u-u_1) + \delta(u-u_2)]$, where $u_1 u_2 = q$. For details, see Sect. 4.1.2. Bond percolation can be studied in the limit $u_1 \to 0$, $u_2 \to \infty$ of infinitely weak and strong bonds respectively. In this limit considerable simplifications occur in the TM, rendering computations with rather large strip widths feasible.

In the percolation limit all single-bond TM have only one non-zero entry per column. Recall from Sect. 5.4 that in the general case there are two such entries of which one is diagonal and the other is 'non-trivial'. In the case of the strong vertical bonds and the weak horizontal bonds only the diagonal entries survive, so that the matrices $\mathcal{T}^{\mathrm{v}}_{\mathrm{strong}} = u_2 \mathbf{1}$ and $\mathcal{T}^{\mathrm{h}}_{\mathrm{weak}} = \mathbf{1}$ both become trivial. On the other hand, a weak vertical bond corresponds to a TM having one non-trivial entry of q per column, whilst a strong horizontal bond is represented by a TM that is u_2 times a non-trivial matrix with entries of 1's and $1/q$'s.

The factors of u_2 multiplying both $\mathcal{T}^{\mathrm{v}}_{\mathrm{strong}}$ and $\mathcal{T}^{\mathrm{h}}_{\mathrm{strong}}$ are innocuous albeit infinite, since of the $2L$ single-bond matrices constituting the entire $\mathcal{T}$ there will on average be L strong ones, hence L factors of u_2. On the level of the specific

free energy this amounts to an infinite additive constant

$$f_0^{11}(L) = -\ln u_2 + \tilde{f}_0^{11}(L) \tag{5.33}$$

independent of the strip width L. In particular, the central charge c can be extracted from the *finite* quantity $\tilde{f}_0^{11}(L)$.

As we shall see in Chapter 6 this quantity can be found by measuring the asymptotic growth of the norm of $\left(\prod_{j=1}^{m} \mathcal{T}_j^{11}\right) |\mathrm{v}_0\rangle$, where $|\mathrm{v}_0\rangle$ is some largely arbitrary initial vector. In the percolation limit the TM turn out to be so sparse that after a very few iterations the resulting vector has only one non-zero component. Computationally this means that it is sufficient to store the row index of that non-zero component as well as its norm. Both time and memory requirements are thus enormously reduced, allowing us to access larger system sizes.

The disadvantage of this projective quality of the percolation point TM is that neither the thermal nor the magnetic scaling dimensions can be found from the Lyapunov spectrum. In the case of x_H an initial vector in the $\mathcal{T}^{22}$ sector will rapidly decay to zero, thus invalidating the procedure for finding $\tilde{f}_0^{22}(L)$, and the alternative of using a seam is obstructed by the fact that disallowing the entry in the horizontal bond TM that corresponds to a cluster wrapping the cylinder is incompatible with the argument of pulling out an overall factor of u_2 from the TM. Very recently, an analogous problem encountered when constructing the TM for simple bond percolation has been resolved [55].

Resumo en Esperanto

La grapolmodelo (5.6) estas la elirpunkto de tre efika reprezentiĝo de la transfermatrico por la aleatora Potts-modelo. Anstataŭ la tradicia spinbazo uzatas bazo de ĉiuj eblaj konektecoj de L punktoj plenumantaj (5.8). Tiel la dimensio de transfermatrico por rubando de larĝo L reduktiĝas de q^L al $c_L \sim 4^L$ (vidu Tab-on 5.1). Alia avantaĝo estas ke q nun aperas kiel *kontinua* parametro. La magnetaj proprecoj perdiĝas en la proceso, sed eblas retrovi ilin tra kuplado de ĉiuj spinoj al 'fantoma spino' $\sigma_0 \equiv 1$.

Ni detale konstruas la konektostatojn kaj nombre ordigas ilin. Malkombinado de la transfermatrico en magrajn matricojn por la unuopaj eĝoj kompletigas la optimumiĝon de nia algoritmo. En la perkola limo okazas simpliĝoj kiuj permesas kalkulaĵojn por eĉ pli larĝaj rubandoj.

Chapter VI

Numerical results

Some of the most powerful predictions of conformal field theory pertain to the relation between the structure of the transfer matrix spectra and the critical exponents of physical operators. Using these and other techniques we determine the central charge and various scaling dimensions of the random-bond Potts model. The generalisation from eigenvalue to Lyapunov spectra turns out to imply that certain well-established predictions must be modified. We conclude the chapter with a few considerations on the presently unknown CFT underlying the RBPM.

6.1 Softening of the transition

Before attempting to determine the universality classes of the RBPM it is essential to make sure that quenched bond randomness indeed renders the phase transitions second-order. For $q > 4$ the pure system has a first-order transition for which the free energy per site is expected to scale like [44]

$$f_0(L) = f_0(\infty) + aL^{-d}\exp(-L/\xi), \tag{6.1}$$

where ξ is the bulk correlation length and a is an amplitude depending on q. In Fig. 6.1 we show plots of the function

$$\begin{aligned} \lambda(L) &\equiv \ln[f_0(L) - f_0(\infty)] + d\ln L \\ &\sim \text{const} - L/\xi \end{aligned} \tag{6.2}$$

for various values of q and the randomness strength R. These plots are rather sensitive to the value of $f_0(\infty)$, but although this is known exactly only for the pure model [19] it can nevertheless be determined with sufficient accuracy from the parabolic fits described in Sect. 6.2 below.

For $q = 8$ the finite correlation length of the pure system ($\xi \sim 70$) is seen to be rendered effectively infinite ($\xi \sim 10^3$) upon imposition of the randomness, whilst the transition of the Ising model ($q = 2$) simply stays second order. Despite the simplicity of these plots we also find a fair agreement with the

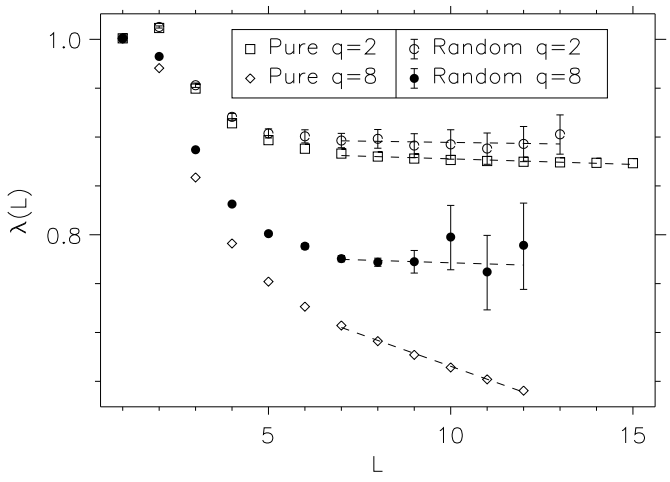

Figure 6.1: Plots of $\lambda(L)$, normalised to $\lambda(1) = 1$, showing that bond randomness renders the phase transition second order. The random systems have $R \equiv K_2/K_1 = 2$.

recently found analytical values of ξ for the pure systems [56]; near $q = 4$ these assume the simple form

$$\xi \simeq \frac{\sqrt{2}}{8} \exp\left(\frac{\pi^2}{\sqrt{q-4}}\right). \tag{6.3}$$

Another criterion for distinguishing between first and second-order phase transitions is the values of the (effective) exponents x_H and x_T as found from Eq. (5.30) and (5.31) respectively. Generally speaking, for pure systems with $q > 4$ these equations give rise to rather poor fits which however have extrapolated values of the effective exponents that are in the vicinity of, and slightly below, zero, whereas when randomness is added the fits are much better and yield exponents in the interval $]0, 2[$. In view of the problems justifying such fits in the random case (see below) this evidence for a softening of the transition is however not to be taken too seriously.

6.2 Free energy and central charge at the random fixed point

6.2.1 Free energy

The free energy per site $f_0^{11}(L)$ for the RBPM on long strips of width L is readily found from Eq. (5.5) applied to the $\mathcal{T}^{11}$ sector of the TM. We have performed

extensive simulations for various values of q and the randomness strength R, though in most cases $R = 2$ was found to describe the random FP adequately.

Representative samples of our data are shown in Table 6.1. For each run a normalised initial vector $|\mathrm{v}_0\rangle$ was prepared by choosing its components randomly, and after discarding the results of the first 2,000 multiplications by $\mathcal{T}_i^{11}$ in order to eliminate transients, data collection was made for each 200 iterations until a strip of a total length of $m = 10^5$ had been built up. For $q > 2$ a total of 100 independent runs were made for $1 \leq L \leq 8$, and 3 runs for $9 \leq L \leq 12$, whilst for the Ising model ($q = 2$) we were able to make 100 runs for $1 \leq L \leq 13$ by using the conventional spin basis. Final results and error bars were extracted by computing the mean and the standard deviation for the totality of patches of length 200.

It is not *a priori* obvious that the Lyapunov exponents found from Eq. (5.5) are independent of the norm used. The standard norm in both the spin basis and the connectivity basis is given by the square root of the sum of the squared components, and these two are of course not identical. To impose the spin basis norm on the connectivity basis each term in the sum must be weighted by a factor q^C, where C is the number of clusters in the relevant connectivity state. We have checked the consistency of our results by comparing the first few Lyapunov exponents obtained from imposing the two different norms on the connectivity basis, and we find that not only are the results identical but there is even a complete agreement of the first three significant digits of the error bars. For $q = 2$ we found that the results using the spin basis and the connectivity basis were consistent, but that the error bars obtained using the spin basis were slightly smaller.

Our results for the free energies of the random-bond Ising model agree with, and are more precise than, those of de Queiroz[1] [57].

6.2.2 Central charge

Values of the effective central charge c' can be extracted from Eq. (5.4) by employing various fitting procedures. In spite of the relatively slow convergence of both two-point fits $(L, L+1)$ and straight-line least-squares fits against $1/L^2$ [57], iterating such fits yields quite good results in the pure model. When randomness is added this is no longer so, since rather substantial error bars on the first estimates prevent us from efficiently iterating the fits.

A better scheme is to include the leading correction to the scaling of Eq. (5.4), which in the pure case has been shown numerically to take the form [44]

$$f_0^{11}(L) = f_0(\infty) - \frac{\pi c'}{6L^2} + \frac{A}{L^4} + \cdots . \tag{6.4}$$

[1] Actually they differ by a constant since the Hamiltonian in Ref. [57] is defined as $-\sum_{\langle ij \rangle} K_{ij} s_i s_j$ with $s_i = \pm 1$, as opposed to our Eq. (2.25). Since $s_i s_j = 2\delta_{\sigma_i \sigma_j} - 1$ there is a free energy difference of $2\overline{K_{ij}}$, which for $R = 2$ equals 0.91407.

L	$q=2$	$q=3$	$q=4$	$q=8$	$q=16$	$q=64$
1	2.17460 (12)	2.62881 (13)	2.96193 (13)	3.80035 (16)	4.68198 (18)	6.54635 (24)
2	1.95329 (8)	2.26650 (9)	2.49558 (9)	3.06980 (10)	3.67393 (11)	4.95619 (14)
3	1.90971 (7)	2.19534 (7)	2.40405 (7)	2.92819 (8)	3.48241 (9)	4.67423 (11)
4	1.89550 (6)	2.17203 (6)	2.37431 (6)	2.88328 (7)	3.42387 (8)	4.59557 (10)
5	1.88895 (5)	2.16182 (6)	2.36126 (6)	2.86392 (6)	3.39934 (7)	4.56442 (9)
6	1.88568 (5)	2.15649 (5)	2.35449 (5)	2.85395 (6)	3.38683 (6)	4.54893 (8)
7	1.88377 (4)	2.15328 (5)	2.35040 (5)	2.84798 (5)	3.37948 (6)	4.53974 (7)
8	1.88250 (4)	2.15113 (4)	2.34782 (4)	2.84424 (5)	3.37479 (5)	4.53394 (7)
9	1.88164 (4)	2.14993 (24)	2.34624 (25)	2.84172 (11)	3.37186 (31)	4.53017 (39)
10	1.88098 (4)	2.14858 (23)	2.34504 (23)	2.84011 (26)	3.36918 (29)	4.52653 (25)
11	1.88048 (4)	2.14804 (22)	2.34386 (22)	2.83851 (24)	3.36768 (29)	4.52465 (34)
12	1.88017 (3)	2.14744 (20)	2.34314 (21)	2.83765 (24)	3.36639 (26)	4.52316 (34)
13	1.87991 (3)					

Table 6.1: Critical free energies per site, $-f_0^{11}(L)$, for $R=2$ and various values of q. The figures in parentheses indicate the error bar on the last quoted digits.

One then performs either three-point fits $(L, L+1, L+2)$ or parabolic least-squares fits against $1/L^2$, and because of the much faster convergence no iteration is needed [57]. Although a correction proportional to $1/L^4$, due to the operator $T\overline{T}$, must necessarily be present in every system that is conformally invariant [4] it can of course not be guaranteed to be the dominant one in general.

In Table 6.2 the results of parabolic fits including the data points for $L_0 \leq L \leq L_{\max}$ have been shown as a function of L_0. It is seen that L_0 must be chosen large enough to justify the omission of higher terms in the series (6.4), and small enough to minimise error bars. From the special cases of the Ising model and of the percolation point (see Sect. 6.3 below) we concluded that the choice $L_0 = 3$ is optimal.

Apart from the results shown in Table 6.2 we have also performed some runs for $q = 1.5$, finding, as expected from the Harris criterion [8], no difference between the results for the pure and the random model.

In the intermediate regime $2 \leq q \leq 4$ our results compare favourably to those of the $(q-2)$-expansion, at least up to $q = 3$. On the other hand, it is evident from Fig. 6.2 that the difference between c for the pure model and c' for the random one is of the same order of magnitude as our error bars, and only near $q = 4$, where the expansion is expected to break down anyway, are our results able to distinguish between the two different behaviours. Exactly at $q = 2$ the randomness is marginal and logarithmic corrections to the finite-size scaling forms, Eqs. (5.3) and (5.4), are expected. Whilst this issue has recently attracted considerable interest in the case of the critical exponents [58] the corrections to the central charge are much weaker [59] and accordingly our result is consistent with that of the pure Ising model.

6.2.3 crossover effects

In Fig. 6.3 we have displayed our results for c' as a function of $\log_{10} q$ for selected values of $q \in [1.5, 64]$. We have juxtaposed the results for two strengths of the randomness, namely weak randomness ($R = 2$, closed circles on Fig. 6.3) and strong randomness ($R = 10$, open circles). For small values of q both randomness strengths give rise to the same c', as witnessed by the overlap of the $q = 4$ data points. However, for larger q the $R = 2$ curve flattens out and grows slower than logarithmically. Sample runs show that the same is true for larger values of R, the difference being that the range of q-values for which the growth is logarithmic is extended as R is increased. This is illustrated by the $R = 10$ curve's staying above, but very close to, the percolative result $\sim \log q$ (see Sect. 6.3 below) for the whole range of q-values shown on the plot. Another way to state this is that for fixed q and varying R, the quantity c' is an increasing function of R that eventually reaches a plateau as R becomes large enough. It then appears from Fig. 6.3 that for $q \leq 64$ the randomness strength $R = 10$ is sufficient to reach this plateau.

These findings are interpreted as follows. According to the $(q-2)$-expansion

L_0	$q=2$ $R=2$	$q=3$ $R=2$	$q=4$ $R=2$	$q=8$ $R=2$	$q=8$ $R=10$	$q=16$ $R=2$	$q=16$ $R=10$	$q=64$ $R=2$	$q=64$ $R=10$
1	0.563 (1)	0.927 (1)	1.184 (1)	1.787 (1)	1.731 (3)	2.330 (1)	2.322 (4)	3.120 (1)	3.476 (5)
2	0.508 (2)	0.825 (3)	1.042 (2)	1.515 (3)	1.586 (10)	1.864 (3)	2.101 (10)	2.194 (4)	3.150 (13)
3	**0.499(3)**	**0.800(6)**	**1.003(6)**	1.441 (7)	**1.521(23)**	1.752 (8)	**2.052(25)**	2.065 (10)	**3.034(30)**
4	0.500 (6)	0.813 (14)	0.994 (14)	1.424 (15)	1.548 (52)	1.750 (17)	2.089 (57)	2.157 (22)	3.079 (68)
5	0.505 (11)	0.842 (30)	1.005 (31)	1.426 (33)	1.622 (113)	1.785 (38)	2.203 (125)	2.305 (47)	3.209 (148)
6	0.500 (20)	0.818 (62)	0.963 (63)	1.360 (67)	1.587 (228)	1.794 (78)	2.196 (251)	2.384 (93)	3.213 (300)

Table 6.2: Effective central charge c' extracted from parabolic fits with $L_0 \leq L \leq L_{\max}$, as described in the text. Error bars on the last quoted digit are shown in parentheses. The choice $L_0 = 3$ appears to be optimal, provided that the strength of the randomness R is large enough (see text), and the corresponding values of c', shown in bold face, should be regarded as our results.

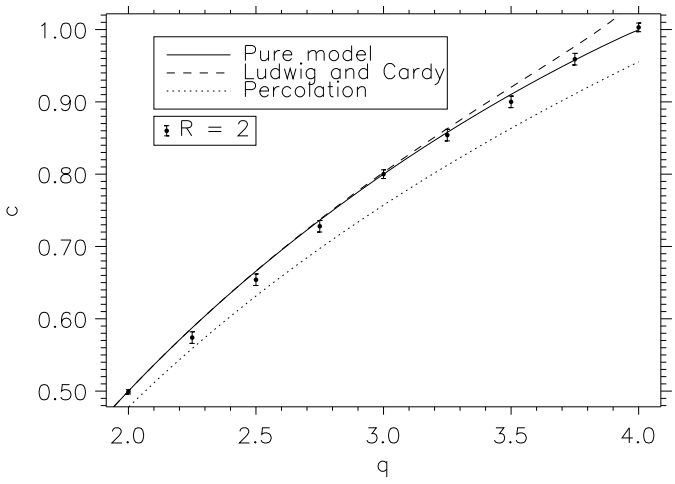

Figure 6.2: The effective central charge c' as a function of q for $2 \leq q \leq 4$. The perturbative results by Ludwig and Cardy [23] do not differ appreciably within the range of q-values where the expansion is supposed to be valid. Accordingly the numerical data are unable to distinguish between pure and non-trivial random behaviour. They are also quite close to, but distinguishable from, the percolation point values.

the randomness strength R^* corresponding to the random FP is an increasing function of q. Assuming this FP to persist as we enter the regime $q > 4$ (see Fig. 3.2) we now claim that the monotonicity of $R^*(q)$ also holds true when the $(q-2)$-expansion breaks down. From the RG flows given in Fig. 3.2 we see that any initial value of $R \in]1, \infty[$ will eventually flow to the random FP as the system is viewed on larger and larger length scales. However, if we start out very far from R^* the onset of the asymptotic scaling given by Eq. (5.4) may be deferred to much larger length scales than the strip widths L numerically accessible for our TMs. We therefore expect poor scaling for strip widths $L \leq L_{\max}$. Conversely, if we choose the strength of the randomness as $R \sim R^*$ the resulting value of c' is expected to be more or less independent of the precise choice of R and equal to the true value of the central charge. But in our simulations we find that this is precisely accomplished by choosing R as an increasing function of q. Further justification for this interpretation is found from the phenomenological RG treatment in Sect. 6.6 below.

A heuristic argument explaining that the "effective" $c'(R, q)$ obtained for small values of R is less than the "correct" value $c'(R^*, q)$ associated with the random FP is readily furnished, at least for large q. Namely, from Zamolodchikov's c-theorem [60] we know that there exists a function $c(\{K\})$ of the

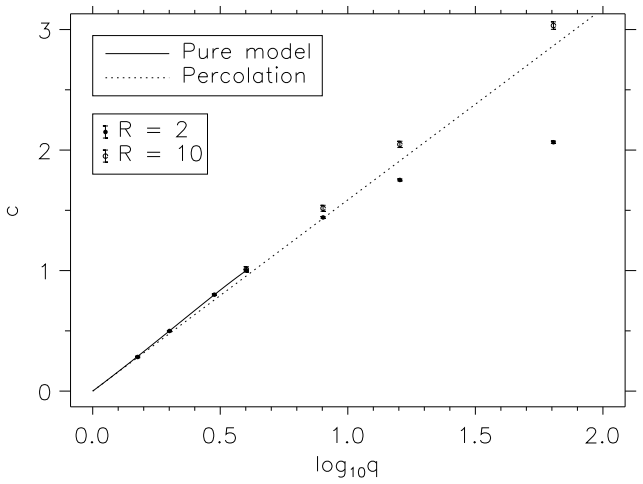

Figure 6.3: Effective central charge as a function of $\log_{10} q$ for $1.5 \leq q \leq 64$. For large q the data for $R = 10$ are supposed to represent the true behaviour at the random fixed point, as argued in the text. For $q = 4$ the $R = 2$ and the $R = 10$ data points overlap; this is barely visible on the graph.

couplings that decreases along the RG flow and equals the central charge at the fixed points. As a corollary the curves of constant c are orthogonal to the RG flow. In particular, for large q where the RG flow is known from the mapping to the RFIM (see Eqs. (3.27) and (3.28)), it is evident from Fig. 3.2 that $c'(R, q)$ is equal to $c'(R^*, q^*)$ for some $q^* < q$. Since our numerical results indicate that $c'(R^*, q)$ is an increasing function of q the proposition follows.

Very recently, the complicated issue of crossover effects in the RBPM has been addressed by Picco [27]. The results of this author are consistent with the picture given above.

6.2.4 Trinary randomness

To check our results for c' we have also made 100 independent runs for each of the strip widths $1 \leq L \leq 8$ where the random bonds were drawn from the *trinary* distribution

$$\begin{aligned} P(K) &= p[\delta(K - K_1) + \delta(K - K_2)] \\ &+ (1 - 2p)\delta(K - K^*), \end{aligned} \tag{6.5}$$

where K_1 and $K_2 = 1000 K_1$ satisfy the criterion (3.3) and $(\exp K^* - 1)^2 = q$. Here $p \ll 1$ is the strength of the randomness. Of course this realisation of

the randomness also preserves self-duality, and hence the model is again at its critical point [30].

Numerical results for c' using trinary randomness are shown in Table 6.3 and they are consistent with the binary results given above, again provided that p is increased as we go to larger and larger q. In particular it is reassuring to verify that we seem to probe the true random behaviour when $2/p$ (the length scale associated with this randomness) is comparable to the correlation length of the pure system (6.3).

6.2.5 Zamolodchikov's c-theorem

An interesting question is whether the asymptotic value of c' is approached from above or below when the system is viewed on larger and larger length scales. For models exhibiting reflection positivity Zamolodchikov's c-theorem [60] ensures that the convergence is from above. In particular the condition of positivity holds true for unitary models, whereas for a random model it may well fail to be fulfilled. Indeed, in the case of the RBPM a perturbative calculation [23] suggests that the convergence may be from below in some cases.

In order to discuss this point the parabolic fits versus $1/L^2$ employed above are no longer appropriate. Apart from speeding up the rate of convergence to a point where information about its direction becomes obliterated due to error bars, the inclusion of higher-order corrections to the finite-size scaling form (5.4) may have the effect of reversing this direction. E.g., in the case of the pure Ising model it is found [57] that the estimators obtained from parabolic fits converge from below, whereas the corresponding linear fits (*i.e.*, without the $1/L^4$ correction) yield estimators that converge from above in accordance with the theoretical prediction.

In Table 6.4 we show the results of such linear least-squares fits for several values of q. The randomness strength R was chosen in accordance with the considerations given above. It appears that in all cases the finite-size estimators converge towards the asymptotic values of Table 6.2 from above.

6.2.6 Alternative representation of the transfer matrices

We remark that values of c' similar to ours have recently been reported by Picco [61]. For $q = 8$ this author found $c' = 1.45 \pm 0.06$ which agrees with our result of respectively $c' = 1.52 \pm 0.02$ for binary randomness of strength $R = 10$, and $c' = 1.51 \pm 0.04$ for trinary randomness of strength $p = 0.10$. Our observation that c' appears to be an increasing function of R, eventually reaching a plateau as R becomes large enough, was confirmed by Ref. [61] that used binary randomness of strength $R = 10$ throughout. Strong evidence was also given that $c'(q)$ grows roughly logarithmically with q in the regime $q \in [5, 256]$, but a further discussion of what this implies will be deferred to Sect. 7.1 below.

It is worthwhile to compare the TM algorithm used in Ref. [61] to ours. It

p	$q=2$	$q=4$	$q=8$	$q=16$	$q=32$	$q=64$
0.05	0.522 (25)	1.030 (26)	1.477 (27)	1.780 (29)	1.920 (30)	1.974 (31)
0.10	0.519 (35)	1.032 (36)	1.510 (38)	1.915 (40)	2.251 (42)	2.552 (43)
0.15			1.539 (46)	1.996 (48)	2.416 (50)	2.817 (52)

Table 6.3: Effective central charge c' obtained using a trinary distribution of the random bonds. There is a fraction p of very weak and very strong bonds respectively, the remaining fraction $1-2p$ being critical.

L_0	$q=2$ $R=2$	$q=3$ $R=2$	$q=4$ $R=2$	$q=8$ $R=10$	$q=16$ $R=10$	$q=64$ $R=10$
2	0.5662 (6)	0.9300 (7)	1.1903 (7)	1.723 (3)	2.313 (3)	3.469 (4)
3	0.535 (1)	0.878 (1)	1.117 (1)	1.656 (5)	2.207 (6)	3.311 (7)
4	0.521 (2)	0.848 (3)	1.075 (3)	1.605 (10)	2.148 (11)	3.206 (13)
5	0.514 (3)	0.836 (5)	1.051 (5)	1.585 (19)	2.125 (20)	3.163 (24)
6	0.512 (4)	0.839 (9)	1.043 (9)	1.595 (34)	2.144 (38)	3.174 (45)
7	0.510 (6)	0.840 (18)	1.022 (18)			
8	0.509 (9)	0.812 (33)	1.011 (34)			
9	0.503 (14)					
10	0.500 (23)					

Table 6.4: Effective central charge c' extracted from linear fits of $f_0^{11}(L) - f_0^{11}(\infty)$ versus $1/L^2$, with $L_0 \leq L \leq L_{\max}$. In all cases the approach towards the asymptotic values of Table 6.2 is from above. Error bars on the last quoted digit are shown in parentheses.

was found that the number of distinct entries in the pure model TM in the spin basis is

$$b_L = \sum_{i_2=1}^{2} \sum_{i_3=1}^{m_3} \sum_{i_4=1}^{m_4} \cdots \sum_{i_L=1}^{m_L} 1, \tag{6.6}$$

where $m_i = \max(i_2, i_3, \ldots, m_{i-1})$, and L designates the strip width as usually. Further taking into account the 2^L different realisations of the binary randomness in each strip, recursion relations between the different elements of the TM were found by computing a total of $(b_L)^2 2^L$ polynomials. Since this number of polynomials increases rapidly with L high-precision computations could only be performed up to $L_{\max} = 6$. The number of iterations used to determine $f_0(L = 6)$ was similar to ours, whereas more iterations were used for the smaller strip widths.

Evidently this algorithm also has the advantage that q enters only as a parameter, thus making accessible any value of q for the simulations[2]. However, for large L it performs inefficiently, as we will now show. The numbers b_L of Eq. (6.6) are by no means unfamiliar. Indeed, they are nothing but the total number of L-point connectivities, including the non-well nested ones [36]. Alternatively they can be viewed as the number of ways that L objects can be partitioned into indistinguishable parts [37]. With m_ν parts of ν objects each $(\nu = 1, 2, \ldots)$ this can be rewritten as

$$b_L = \sum_{m_\nu=0}^{\infty}{}' \prod_{\nu=1}^{\infty} \frac{L!}{(\nu!)^{m_\nu}\, m_\nu!}, \tag{6.7}$$

where the primed summation is constrained by the condition $\sum_{\nu=1}^{\infty} \nu m_\nu = L$. From this representation the generating function can be immediately inferred

$$\exp(\mathrm{e}^t - 1) = \sum_{n=0}^{\infty} \frac{b_n t^n}{n!}. \tag{6.8}$$

Explicit values, found by Taylor expansion of the left-hand side, are shown in Table 5.1. Asymptotically the b_L are seen to grow faster than L^L whereas the well-nested connectivities only grow as $\sim 4^L$.

6.3 The percolation limit

In the case of the binary randomness (3.2) the percolation limit is reached by letting $(\mathrm{e}^{K_1} - 1) \to 0$ and $(\mathrm{e}^{K_2} - 1) \to \infty$ whilst maintaining the self-duality criterion (3.3). The partition function of the random cluster model is then dominated by one graph only, *viz.* the one that covers all of the strong bonds

[2] A minor drawback, however, is that the number of polynomials depends on the randomness distribution, which in practice limits the method to the case of binary randomness.

and none of the weak ones. (Note in particular that the limits $R \to \infty$ and $q \to \infty$ do not commute.) Expressed in terms of the free energy per site this reads

$$f_0^{\mathrm{perc}} = -\frac{B}{N} \ln(\mathrm{e}^{K_2} - 1) - \frac{C}{N} \ln q, \tag{6.9}$$

where B is the number of strong bonds and C is the number of clusters in the dominant graph.

The quenched average over the randomness must be taken on the level of the free energy. Evidently, with the chosen distribution of the randomness, $\overline{B} = N$ whence the first term is simply a trivial, albeit infinite, constant. (Incidentally this is the same constant that was pulled out in Eq. (5.33).) On the other hand, the average number of percolation clusters is related to a derivative in the *pure* Q-state Potts model [5]

$$\overline{C} = \frac{\partial}{\partial Q} \ln Z(Q) \bigg|_{Q=1}, \tag{6.10}$$

thus determining the effective central charge $c'(q)$ at percolation as

$$c'(q) = \ln q \, \frac{\partial c(Q)}{\partial Q} \bigg|_{Q=1}. \tag{6.11}$$

An alternative argument for this relation is furnished by the observation that the replicated model is simply the Potts model with q^n states; differentiating $c(q^n)$ with respect to the number of replicas n and taking the limit $n \to 0$ one recovers the result (6.11). The central charge of the pure model is given by an expression due to Kadanoff [4, 62]

$$c = \frac{(2-3y)(1+y)}{(2-y)}, \tag{6.12}$$

where $\sqrt{Q} = 2\cos(\pi y/2)$ and $0 \le y \le 1$, and taking the appropriate derivative of this we finally arrive at

$$c'(q) = \frac{5\sqrt{3}}{4\pi} \ln q. \tag{6.13}$$

As described in Sect. 5.6 the single-bond TMs in the percolation limit have only one non-zero entry per column, equal to either q, 1 or $1/q$. Taken together with their projective quality and Eq. (5.5) for the largest Lyapunov exponent it is clear that the free energy, and hence the central charge, must be explicitly proportional to $\ln q$. So it suffices to do the numerics for one value of $q \neq 1$. Because of the simple form of these TMs we were able to average $\tilde{f}_0^{11}(L)$ of Eq. (5.33) over 100 strips of length $m = 10^5$ for the range $1 \le L \le 19$. Consequently the factor of proportionality could be determined quite accurately as 0.688 ± 0.003, in excellent agreement with $\frac{5\sqrt{3}}{4\pi} \simeq 0.689$.

It is evident from the mapping between bond percolation and the pure $Q = 1$ Potts model that the critical exponents of the two models are identical: $x_T = \frac{5}{4}$ and $x_H = \frac{5}{48}$ [4]. Since all correlation functions at percolation can only take the values 0 and 1, it is also clear that different moments of a given correlation function all have the same scaling dimension. Thus, in the notation of Ludwig [25], $x_n = x_1$ for all $n > 1$. The pure model represents the other trivial extreme case of multiscaling behaviour: $x_n = nx_1$.

6.4 The cumulant expansion

The concept of multiscaling was briefly introduced in Sect. 2.1 where we considered the simple example of the random-bond Ising chain [7]. We recall that although the free energy was found to be self-averaging—essentially a consequence of the central limit theorem—this was not so for the correlation functions. In particular $\overline{\langle s_1 s_R \rangle^2}$ and $\overline{\langle s_1 s_R \rangle}^2$ are expected to scale with different exponents.

In Sect. 5.5 we related the spin-spin correlation function $G(m)$ on a strip of the RBPM to the free energy in the presence of a seam of frustrated bonds (or with a ghost site). Taking the logarithm of Eq. (4.14) and exploiting the self-duality of the lattice we have

$$\Delta f(L) \equiv f_0^{22}(L) - f_0^{11}(L) = \frac{1}{mL} \ln G(m), \tag{6.14}$$

and in the pure system, according to conformal symmetry [49], this decays along the strip as $2\pi x_H / L^2$, *cfr.* Eq. (5.30). When randomness is present $\Delta f(L)$ is a fluctuating quantity, and since free energies are supposed to be normally distributed these fluctuations are $\mathcal{O}(1/\sqrt{m})$. Consequently $\ln G$ is a self-averaging quantity and G is not [7], exactly as in the simple example given above.

In the *multiscaling* scenario of Ludwig [25] different moments $\overline{G(m)^n}$ scale with dimensions x_n which, as opposed to what is the case in the pure model, are not necessarily linear in n. (In this notation $x_H \equiv x_1$.) For $n_1 > n_2$ we have $x_{n_1} \geq x_{n_2}$ and $x_{n_1}/n_1 \leq x_{n_2}/n_2$ (convexity); pure and percolative behaviour are thus realisations of the two possible extremes of multiscaling.

Since translational invariance is one of the basic assumptions of conformal symmetry [4], the latter only refers to the averaged quantities $\overline{G(m)^n}$ and not to the $G(m)^n$ themselves. These averages cannot be computed directly in a numerical experiment because of the lack of self-averaging; this can however be circumvented by performing a cumulant expansion

$$\ln \overline{G^n} = n \overline{\ln G} + \frac{1}{2} n^2 \overline{(\ln G - \overline{\ln G})^2} + \cdots , \tag{6.15}$$

where each term on the right-hand side *is* self-averaging and can be directly extracted from the statistical fluctuations in $\Delta f(L)$ between the patches of length 200 into which we have divided our strip.

Quite generally for a stochastic variable X we have

$$\langle \exp X \rangle = \exp \sum_{j=1}^{\infty} \frac{1}{j!} k_j, \tag{6.16}$$

where explicit expressions for the six first cumulants k_i in terms of the moments m_i of X are given by [63]

$$\begin{aligned}
k_1 &= m_1 \\
k_2 &= m_2 - m_1^2 \\
k_3 &= m_3 - 3m_2m_1 + 2m_1^3 \\
k_4 &= m_4 - 4m_3m_1 - 3m_2^2 + 12m_2m_1^2 - 6m_1^4 \\
k_5 &= m_5 - 5m_4m_1 - 10m_3m_2 + 20m_3m_1^2 \\
&+ 30m_2^2m_1 - 60m_2m_1^3 + 24m_1^5 \\
k_6 &= m_6 - 6m_5m_1 - 15m_4m_2 + 30m_4m_1^2 - 10m_3^2 \\
&+ 120m_3m_2m_1 - 120m_3m_1^3 + 30m_2^3 - 270m_2^2m_1^2 \\
&+ 360m_2m_1^4 - 120m_1^6
\end{aligned}$$

We have computed these six cumulants of $\Delta f(L)$ for various values of R and q, based on 100 independent strips of length $m = 10^5$ and width $1 \leq L \leq 7$. Sample results for $R = 2$ and $q = 3, 8$ are shown in Table 6.5.

For $q = 3$ the cumulant expansion converges well. The magnitude of the higher cumulants decreases very rapidly, especially for $L \geq 3$, and reliable estimates for the left-hand side of Eq. (6.15) can be obtained simply by summing the first 3 or 4 cumulants, at least when n is not too large. Performing parabolic least-squares fits using Eq. (5.30) with an $1/L^4$ correction we thus expect to extract quite accurate values of x_n at the random FP.

As q increases the convergence is slower. This is witnessed by the $q = 8$ results of Table 6.5 decreasing noticeably slower, both for a definite cumulant as a function of L (vertically) and for a definite L as a function of the cumulant number (horizontally). The approximation of leaving out the higher cumulants in the sum (6.15) thus becomes increasingly difficult to justify, and eventually the cumulant expansion breaks down. This problem is enhanced by the fact that for $q > 8$ we expect a randomness strength of $R = 2$ to be insufficient in order to access the true behaviour at the random FP. We are thus forced to increase R, whence the fluctuations become even more violent and the cumulant expansion accordingly ill-behaved.

6.4.1 Results for the magnetic exponent x_1

Our results for x_1 are shown in Fig. 6.4. Since error bars on the individual cumulants are related to the magnitude of the higher cumulants the question of how to assign a final error bar to x_1 becomes a delicate one. We have addressed

	L	1. cumulant	2. cumulant	3. cumulant	4. cumulant	5. cumulant	6. cumulant
$q = 3$	1	-1.039786 (242)	0.060716	-0.000791	0.000830	-0.004583	-0.000778
	2	-0.253209 (146)	0.012391	-0.000347	0.000279	0.000059	0.000386
	3	-0.106163 (113)	0.004963	-0.000246	0.000063	-0.000102	0.000046
	4	-0.057901 (95)	0.002784	-0.000143	0.000006	0.000034	0.000003
	5	-0.036521 (84)	0.001810	-0.000105	0.000001	-0.000002	-0.000003
	6	-0.025172 (76)	0.001289	-0.000075	0.000008	-0.000001	-0.000002
	7	-0.018426 (69)	0.000968	-0.000069	0.000002	0.000002	-0.000001
$q = 8$	1	-1.380171 (289)	0.104382	0.004069	0.014001	0.019452	-0.013889
	2	-0.326484 (177)	0.028366	-0.001683	-0.000432	0.000157	-0.003145
	3	-0.132560 (138)	0.014908	-0.001822	0.000356	0.000221	-0.000083
	4	-0.071296 (115)	0.010129	-0.001610	0.000319	-0.000959	0.002323
	5	-0.044886 (102)	0.007880	-0.001619	0.000252	-0.000082	-0.000456
	6	-0.031045 (92)	0.006450	-0.001538	0.000607	-0.000184	-0.001096
	7	-0.022851 (84)	0.005401	-0.001505	0.000237	0.000234	-0.000280

Table 6.5: The first six cumulants of $-\Delta f(L)$ for $1 \leq L \leq 7$ and $R = 2$. The error bar on the first cumulant (shown in parentheses) is related to the second cumulant; error bars on the higher cumulants are not shown.

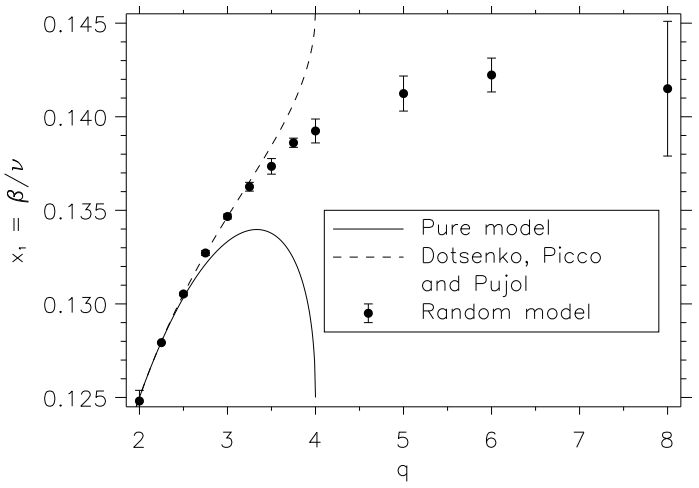

Figure 6.4: Magnetic exponent $x_1 = \beta/\nu$ for $R = 2$ as obtained from the cumulant expansion. x_1 is an increasing function of q, continuously connecting onto the perturbative results near $q = 2$. For $q > 8$ the cumulant expansion begins to break down. The agreement with the third-order perturbative result by Dotsenko, Picco and Pujol [26] is excellent up to about $q = 3$.

this issue by averaging the estimates for x_1 obtained from various parabolic least-squares fits. More precisely, the average is calculated from 4 values, namely fits with $L_0 = 3$ or 4 and including the first 3 or 4 cumulants on the right-hand side of Eq. (6.15). The consistency of these 4 values is regarded as a check of the validity of the expansion.

In particular, for $q = 3$ we find $x_1(3) = 0.13467 \pm 0.00013$ which is 10 standard deviations above the value $x_1^{\rm pure}(3) = \frac{2}{5} \simeq 0.13333$ of the pure three-state Potts model [4] and at the same time in perfect agreement with the result $x_1(3) = 0.13465 + \mathcal{O}(\epsilon^4)$ of the $(q-2)$-expansion [26]. The Monte Carlo result $x_1(3) = 0.1337 \pm 0.0007$ of Picco [64] was not able to distinguish convincingly between pure and random behaviour.

For $q = 4$ our result is $x_1(4) = 0.1396 \pm 0.0005$, in nice agreement with Picco's preliminary result $x_1(4) \sim 0.139$ [65] and decidedly different from the corresponding pure value of $x_1^{\rm pure}(4) = \frac{1}{8}$.

As discussed at length in Sect. 2.5.1 a major motivation for this work was to determine whether the impurity softened transitions for $q > 4$ do or do not have the critical exponents of the pure Ising model. The data of Fig. 6.4 clearly show a smooth continuation of the perturbative results [25, 26] exhibiting no singularity whatsoever at $q = 4$. Our result $x_1(8) = 0.1415 \pm 0.0036$ is comfortably away from the pure Ising value and provides a striking piece of evidence for both our

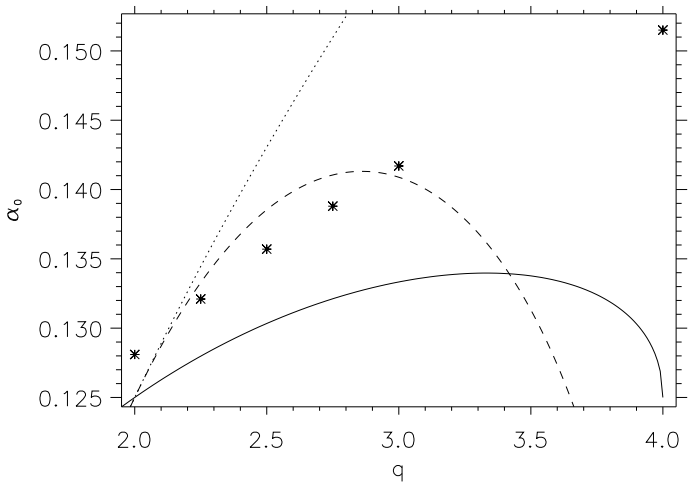

Figure 6.5: Comparison of our numerical results for α_0 (asterisks) with various orders of the $(q-2)$-expansion: Zeroth order (solid line), first order (dotted line), and second order (dashed line). The numerical results are for $R = 2$, and the error bars are smaller than the size of the symbols.

phase diagram and the FP structure of the $(q-2)$-expansion.

All the results quoted for x_1 were computed using $R = 2$. We have checked that other values of R yield results consistent herewith, provided that R is chosen neither too small, in which case the crossover length $\xi_X \sim \exp(1/2Aw^2)$ found from Eq. (3.27) becomes too large for the random FP to be reached, nor too large, in which case the cumulant expansion breaks down. The same holds true when the random bonds are drawn from the trinary distribution (6.5) with various values for the dilution parameter p.

6.4.2 Other multiscaling exponents

Because of the positive sign of the second cumulant the values of x_1 are invariably smaller than those one would have obtained without the cumulant expansion (*i.e.*, using only the first cumulant). The latter, however, determine a universal exponent α_0 that describes the asymptotic decay of the spin-spin correlation function in a *fixed* sample at criticality. In terms of the multiscaling exponents this reads

$$\alpha_0 = \left.\frac{\mathrm{d}x_n}{\mathrm{d}n}\right|_{n=0} \tag{6.17}$$

Ludwig's original first-order expansion around $q = 2$ [25] has recently been

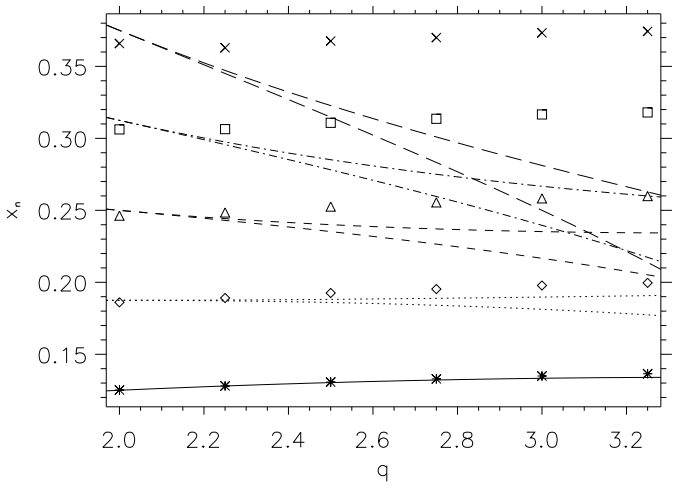

Figure 6.6: Numerical results for the multiscaling exponents x_n for $n = 1$ (asterisks), $n = 1.5$ (diamonds), $n = 2$ (triangles), $n = 2.5$ (squares), and $n = 3$ (crosses). The comparison to the results of the $(q-2)$-expansion is made to both first and second order (lower and upper curves respectively). As usual the numerical results exhibit logarithmic corrections exactly at $q = 2$. Evidently the radius of convergence of the perturbative expansion decreases rapidly as a function of n.

extended to second order by Lewis [66] with the result

$$\alpha_0 = x_1^{\text{pure}} + \frac{9}{32}\left(\frac{2}{3}\epsilon + \left(\frac{11}{12} - \frac{2K}{3} - \frac{\alpha}{12}\right)\epsilon^2\right) + \mathcal{O}(\epsilon^3). \tag{6.18}$$

Here ϵ is the usual Coulomb gas parametrisation of the deviation from $q = 2$ (for $q = 3$ we have $\epsilon = 2/15$), and $K = 6\ln 2$ and $\alpha = -33 + 29\pi/\sqrt{3}$ are constants. In Fig. 6.5 we show the comparison between both the first- and the second-order results with our numerical values in the interval $q \in [2, 4]$. For small fractional values of $(q-2)$ the agreement is quite good if one takes into account the logarithmic corrections expected exactly at $q = 2$. On the other hand, it is clear that the radius of convergence of the perturbative results for α_0 is smaller than in the case of the magnetic exponent x_1.

Actually, Eq. (6.18) is merely a special case of the second-order result for the general multiscaling exponent x_n [66]

$$x_n = nx_1^{\text{pure}} - \frac{9}{32}n(n-1)\left(\frac{2}{3}\epsilon + \left(\frac{11}{12} - \frac{2K}{3} + \frac{\alpha}{24}(n-2)\right)\epsilon^2\right) + \mathcal{O}(\epsilon^3). \tag{6.19}$$

A comparison with the numerical results as found from the cumulant expansion (6.15) is provided by Fig. 6.6. As anticipated the evaluation of higher moments of the spin-spin correlator leads to convergence problems in both the $(q-2)$-expansion and the cumulant expansion. However, for the data given on the figure the error bars are believed to be of the order of the symbol size. It should be obvious that Eq. (6.19) cannot pretend to be very accurate in the physically interesting case of $q = 3$, not even for $n = 2$.

6.4.3 Replica interpretation of the breakdown

Before concluding this section we should like to give a heuristic argument that the cumulant expansion breaks down for large q. In a replica formulation we can imagine the central charge $c(n)$ as a function of the number of replicas n. In this notation the central charge of the pure and the random systems are $c(1)$ and $c'(0)$ respectively, where the prime denotes a differentiation with respect to n. The partition function of the replicated strip is then

$$\begin{aligned} \overline{Z^n} &= \int \exp(-nmLf) P(f)\, df \\ &= \exp\left(-mL\overline{f} + \frac{\pi m c(n)}{6L} - \dots\right), \end{aligned} \tag{6.20}$$

where $P(f)$ is the probability distribution of the free energy. Differentiating this expression twice with respect to n and taking the replica limit $n \to 0$ we infer that the second cumulant of f contains a term that is proportional to $c''(0)$. The cumulant expansion is thus expected to break down if $c(n)$ has a large curvature at $n = 0$.

For $2 \le q \le 4$ the replicas are weakly coupled, since $c(1) \simeq c'(0)$ [23]. Hence $c''(0) \ll 1$. But when $q = 4 + \epsilon$ the transition of the pure system goes first order so that the function $c(n)$ starts out with slope $c'(0) = 1$ and somehow curves down to assume the value $c(1) = 0$. Consequently $c''(0) = \mathcal{O}(1)$ and the higher cumulants begin to contribute significantly to the sum (6.15). Finally, for $q \gg 4$ we are in the strong coupling regime. We still have $c(1) = 0$ and as our numerical data indicate that $c'(0) \sim \ln q$ it follows that $c''(0) \gg 1$. This means that the cumulant expansion must break down.

One may speculate whether the transition actually becomes first-order whenever $q^n > 4$. Clearly this is the case for the pure Potts model [19], but a similar statement is true when N Ising models are coupled by their local energy density. Namely, in this case an RG analysis [15] implies a fluctuation-driven first-order transition whenever $N > 2$, that is to say for $2^N > 4$. If this conjecture is correct one would then suppose the function $c(n)$ to vanish for $n \ge n_0$, where $q^{n_0} = 4$. Evidently such a scenario is in accordance with our observation that $c''(0) \gg 1$ for $q \gg 4$.

6.5 The thermal exponent

Because of the rather striking success of the cumulant expansion for x_1 one would now expect the thermal exponent x_T to be similarly related to the fluctuations of $\Delta f_T(L) = f_1^{11}(L) - f_0^{11}(L)$. Surprisingly, this seems not to be the case. Computing the equivalent of α_0, i.e., using only the first cumulant, we find the following results for different values of q: $\alpha_0^T(2) = 1.028 \pm 0.001$, $\alpha_0^T(3) = 0.91 \pm 0.01$, $\alpha_0^T(4) = 0.81 \pm 0.02$ and $\alpha_0^T(8) = 0.65 \pm 0.01$. As remarked above the results using more cumulants can only be lower.

This is bad news since the quenched correlation length exponent ν can be shown quite rigorously to satisfy the bound [67, 68]

$$\nu \geq \frac{2}{d}, \tag{6.21}$$

or, in our notation, $x_T \geq 1$. Though the proof of Ref. [67] refers to the divergence of the correlation length as the critical point is approached, and hence strictly speaking does not apply to the system under consideration since we work exactly *at* the critical point, the RBPM is among the simplest physical systems for which Eq. (6.21) is believed to be valid [68]. The point is strengthened by noting that the $(q-2)$-expansion yields $x_T = 1.02 + \mathcal{O}(\epsilon^3)$ at $q = 3$ [24]. It is therefore difficult to have confidence in the cumulant expansion for the thermal exponent, and independent methods of assessing x_T must be devised.

At this point we note that although the RG equation (3.29) seems to warrant an effective exponent of $x_T^{\text{eff}} = 1 - Aw^2$ for q large, this argument is only superficially true. Indeed, near $q = \infty$ the RG flows must extend to infinite w before reaching the random FP, and consequently an expansion valid for weak randomness is not to be trusted.

The alternative method for finding x_T that comes closest to the spirit of Refs. [67, 68] is that of finite-size scaling off the critical point. This is discussed at length in the next subsection, and for the moment we concentrate on less "obvious" possibilities.

6.5.1 Duality relations

One of the key points in the construction of the cumulant expansion was the realisation that the spin-spin correlation function was mapped onto a surface tension under duality, and hence could be expressed in terms of the *largest* Lyapunov exponent of a TM with twisted boundary conditions. Reinterpreting the latter as a free energy the self-averaging property was evident, and the cumulant expansion correspondingly behaved quite well if the fluctuations were not too large. In Sect. 4.3 we have seen that under duality four-point correlation functions are similarly mapped onto (generalised) surface tensions. Presently these duality relations have only been worked out for planar graphs, but there is some hope that they may be extended to the case of cylindrical boundary conditions as well. Taking two of the points as nearest neighbours on either end

of the cylinder we would then recover the energy-energy correlator, and if the corresponding boundary conditions can be implemented in the TM x_T follows from a cumulant expansion.

6.5.2 The Furstenberg method revisited

Next, we discuss the method of iterating orthogonal vectors in order to extract the second Lyapunov exponent [53] in more physical terms. The energy-energy correlator (Green's function) can be written as

$$\langle E(r_1)E(r_2)\rangle = \frac{\mathrm{Tr}\, E(r_1)E(r_2)\exp(-\mathcal{H})}{\mathrm{Tr}\exp(-\mathcal{H})}. \tag{6.22}$$

Now imagine building up the strip by repeated action with the random TMs on some initial state situated at $r = -\infty$. When we reach r_1 the system is in a state $|\mathrm{a}_0\rangle$ on which we act with the energy operator to define $|\mathrm{b}_0\rangle = E(r_1)|\mathrm{a}_0\rangle$. After n further iterations these states become

$$\begin{aligned} |\mathrm{a}_n\rangle &= \mathcal{T}_n\cdots\mathcal{T}_2\mathcal{T}_1|\mathrm{a}_0\rangle \\ |\mathrm{b}_n\rangle &= \mathcal{T}_n\cdots\mathcal{T}_2\mathcal{T}_1|\mathrm{b}_0\rangle. \end{aligned} \tag{6.23}$$

Defining a new state $|\tilde{\mathrm{b}}_n\rangle$ by orthogonalising $|\mathrm{b}_n\rangle$ with respect to $|\mathrm{a}_n\rangle$

$$|\tilde{\mathrm{b}}_n\rangle = |\mathrm{b}_n\rangle - \frac{\langle \mathrm{a}_n|\mathrm{b}_n\rangle}{\langle \mathrm{a}_n|\mathrm{a}_n\rangle}|\mathrm{a}_n\rangle \tag{6.24}$$

we find that

$$\begin{aligned} \frac{\langle\tilde{\mathrm{b}}_n|\tilde{\mathrm{b}}_n\rangle}{\langle \mathrm{a}_n|\mathrm{a}_n\rangle} &= \frac{\langle \mathrm{b}_n|\mathrm{b}_n\rangle}{\langle \mathrm{a}_n|\mathrm{a}_n\rangle} - \frac{\langle \mathrm{a}_n|\mathrm{b}_n\rangle\langle \mathrm{b}_n|\mathrm{a}_n\rangle}{\langle \mathrm{a}_n|\mathrm{a}_n\rangle^2} \\ &= \langle E(r_1)E(r_2)\rangle - \langle E(r_1)\rangle\langle E(r_2)\rangle. \end{aligned} \tag{6.25}$$

Thus the process of orthogonalisation corresponds precisely to subtracting off the disconnected part of the correlation function.

When $n \gg 1$ the states $|\mathrm{b}_n\rangle$ and $|\mathrm{a}_n\rangle$ are almost identical due to contamination and have a huge norm $\sim \Lambda_0^n$. The idea of orthogonalising them is therefore numerically extremely unsound. Fortunately a simple calculation shows that orthogonalising after n_1 iterations and then again after $n-n_1$ further iterations is equivalent to orthogonalising only once, as above. Hence, by induction, we are allowed to orthogonalise after each iteration, leaving us with the method of Benettin *et al.* [53]. Similar observations hold true for the higher Lyapunov spectrum.

At this point an objection may be raised. Since

$$\langle \mathrm{a}_n|\mathrm{a}_n\rangle = \langle \mathrm{a}_0|\mathcal{T}_1^\dagger\mathcal{T}_2^\dagger\cdots\mathcal{T}_n^\dagger\mathcal{T}_n\cdots\mathcal{T}_2\mathcal{T}_1|\mathrm{a}_0\rangle, \tag{6.26}$$

where the dagger denotes transposition, the correlation function (6.25) corresponds to a realisation of the randomness that is always symmetric around the

midpoint of r_1 and r_2. From the above physical picture leading to Eq. (6.25) it seems that what we really ought to compute is

$$\frac{\langle \tilde{\mathrm{b}}'_n | \tilde{\mathrm{b}}_n \rangle}{\langle \mathrm{a}'_n | \mathrm{a}_n \rangle}, \tag{6.27}$$

where the (transposed) TMs used to obtain the states on the left implement a different realisation of the randomness than that used to obtain the states on the right.

Numerically we are now facing the problem of computing the average of huge numbers that are no longer necessarily positive. As discussed in connection with Eq. (6.15) we do not expect the correlation function to be self-averaging, and because of possible negative values of Eq. (6.27) the subterfuge of averaging its logarithm will not help us out. Trial runs seem to indicate that for sufficiently small values of q and R (such as $q = 3$ and $R = 2$) the matrix elements appearing in Eq. (6.27) computed for the usual samples of length 200 may have either sign, but that their quotient is invariably positive. The corresponding result for x_T is roughly equal to that obtained from the cumulant expansion. Unfortunately, for larger values of q or R rare events of negative quotients begin to occur, and any attempt of averaging Eq. (6.27) without resorting to logarithms is hampered by such large fluctuations as to render the results insufficiently accurate at the very best. Computations along these lines, though physically appealing, must therefore be abandoned on numerical grounds.

6.5.3 Conformal sum rule

Yet another possibility of determining at least an approximate value of x_T is through the conformal sum rule [69, 70] that for an n-fold replicated system reads

$$\frac{c(n)}{12} = \frac{\sum_i d_i(n) x_i \mathrm{e}^{-2\pi x_i}}{1 + \sum_i d_i(n) \mathrm{e}^{-2\pi x_i}}, \tag{6.28}$$

where the sum runs over all operators in the theory, including the descendants of the Verma module with their appropriate multiplicities, and $d_i(n)$ are the multiplicities pertaining to the permutational symmetry of the n replicas and the q Potts states. For the magnetic operator $d_i(n) = n(q-1)$, since there are $(q-1)$ independent order parameters, and in the case of the energy operator $d_i(n) = n$. In the pure system this yields quite accurate estimates for x_T if the exact values of $c(1)$ and x_H are inserted along with the first descendant of the latter. Differentiating and going to the replica limit we find that for a random system

$$\frac{c'(0)}{12} = x_1(q-1)\mathrm{e}^{-2\pi x_1} - \frac{x_2}{2}(q-1)^2 \mathrm{e}^{-2\pi x_2} + \cdots, \tag{6.29}$$

so that for values of x_1, x_2 and x_T near those of the Ising model the term with x_T enters only as a small correction. Consequently, at the very best only x_2

can be determined with some confidence from our previous results for $c'(0)$ and x_1. Its value appears to be consistent with that obtained from the cumulant expansion.

6.5.4 Exact partition function zeroes

Finally we should like to mention that preliminary studies of exact partition function zeros for small $L \times L$ lattices with quenched bond randomness hints at an interesting new method of estimating x_T. Although the different realisations of the quenched bond randomness in general lead to a considerable scatter in the positions of such zeros, it turns out that the zeros that are closest to the real axis only exhibit a very weak dependence on the realisation. But these zeros are precisely those that fix x_T through their finite-size scaling. Results along these lines, both for zeros of the Lee-Yang and of the Fisher type, will be published elsewhere [41].

6.6 Phenomenological renormalisation

In view of the difficulties encountered in our attempts to extract x_T directly at the critical point we turn our attention to the method of *phenomenological renormalisation* [71], which is closer in spirit to the ideas that lead to the bound (6.21).

The magnetic correlation length can be found from the TM spectra through

$$\xi(L,T)^{-1} = \ln\left(\frac{\Lambda_0^{11}}{\Lambda_0^{22}}\right) = L(f_0^{22} - f_0^{11}), \tag{6.30}$$

and we note that this quantity would be self-averaging in the random case. Motivated by the form $\xi \sim (T - T_c)^{-\nu}$ of the divergence of the correlation length in the infinite system we make the finite-size scaling ansatz

$$\xi(L,T) = L\phi\big((T - T_c)L^{1/\nu}\big). \tag{6.31}$$

For pure systems, then, one traditionally scans through the vicinity of T_c to find an effective $T_c(L)$ as the solution of

$$\frac{\xi(L, T_c(L))}{L} = \frac{\xi(L-1, T_c(L))}{L-1}, \tag{6.32}$$

and computes an approximant $\nu(L)$ from

$$1 + \frac{1}{\nu(L)} = \left.\frac{\ln\big(\mu(L,T)/\mu(L-1,T)\big)}{\ln\big(L/(L-1)\big)}\right|_{T=T_c(L)}, \tag{6.33}$$

where the derivatives

$$\mu(L,T) \equiv \frac{\partial \xi(L,T)}{\partial T} = L^{1+\frac{1}{\nu}}\phi'\big((T - T_c)L^{1/\nu}\big) \tag{6.34}$$

are found by numerical differentiation. As $L \to \infty$ we have $T_c(L) \to T_c$ and $\nu(L) \to \nu$.

L	$\mu(L)$	$\nu(L)$
1	1.087 (1)	–
2	4.229 (3)	1.041 (1)
3	10.426 (8)	0.816 (2)
4	19.682 (18)	0.827 (3)
5	31.867 (33)	0.863 (5)
6	46.994 (53)	0.885 (7)
7	65.020 (79)	0.904 (9)

Table 6.6: Phenomenological renormalisation for the thermal scaling dimension $x_T = 2 - 1/\nu$ at $q = 8$ and $R = 10$. For each strip width L the 100 independent strips of length $m = 10^5$ are divided into patches of length 200. Within each patch the *exact* $\mu(L, T_c)$ is computed, based on evaluations of $\xi(L, K_1', K_2')$ at $K_1' = K_1(1 \pm \epsilon_K)$ and $K_2' = RK_1'$, where K_1 is found from Eq. (3.3). Final results and error bars are obtained as the mean value and the standard deviation over the totality of patches.

6.6.1 Efficient numerical implementation

In the random case the extracted values of $\xi(L, T)$ are hampered by substantial error bars, and the method just outlined becomes by far too inefficient. Fortunately the very costly idea of scanning for $T_c(L)$ can be discarded, since the exact T_c of the RBPM is known from Eq. (3.3). Consequently the derivatives (6.34) and the approximants (6.33) are evaluated at the exact T_c, whence the only remaining source of errors is that of statistical fluctuations over the different realisations of the randomness.

Naively one would now find the derivative (6.34) by subtraction of the free energies evaluated at $T = T_c(1 \pm \epsilon)$, where $\epsilon \ll 1$. Although this method yields reasonable results for $\epsilon \sim 10^{-2}$ it is way too inaccurate, since it involves the subtraction of almost identical quantities (with error bars). A superior strategy is to divide the strip into patches of length 200, calculate *exact*[3] values of $\mu(L)$ for each of those, and finally average over the totality of such patches. In this way one exploits the fact that $\xi\big(L, T_c(1-\epsilon)\big)$ and $\xi\big(L, T_c(1+\epsilon)\big)$ are strongly correlated when the realisation of the randomness is kept fixed. In practice we found that this trick leads to a reduction of the error bars with a factor ~ 120.

Sample results obtained by using these prescriptions are shown in Table 6.6. It is seen that although the convergence is still too slow for reliable extrapolations to the limit of an infinite system to be made, the conflict with the bound (6.21) appears to be resolved.

We have found that the convergence of the estimators $\nu(L)$ can be significantly sped up if one performs the numerical differentiation (6.34) by going

[3] Since we are now faced with differentiating a quantity that is known with full machine precision (10^{-16}) we can concentrate on minimising the rounding and truncation errors. This is accomplished by choosing $\epsilon = 10^{-5}$ [72].

	L	$\mu(L)$	$\nu(L)$
$R = 6$	2	1.898 (1)	–
	3	4.456 (2)	0.905 (1)
	4	8.172 (6)	0.902 (2)
	5	12.974 (11)	0.933 (4)
	6	18.883 (18)	0.945 (6)
	7	25.891 (27)	0.955 (8)
$R = 10$	2	1.832 (1)	–
	3	3.917 (2)	1.144 (2)
	4	6.815 (5)	1.081 (4)
	5	10.486 (9)	1.074 (6)
	6	14.948 (15)	1.059 (8)
	7	20.198 (22)	1.050 (11)

Table 6.7: Phenomenological renormalisation going perpendicularly off the critical surface for $q = 8$ and $R = 6$ and 10 respectively. The data collection was done as before.

perpendicularly off the self-duality criterion in (K_1, K_2) space instead of maintaining the condition $K_2 = RK_1$. Indeed, one may imagine that there is another exponent associated with motion *along* the critical surface, and maintaining $K_2 = RK_1$ one would then measure an admixture of this spurious exponent, in particular for large R. A simple calculation using Eq. (3.3) shows that one should then evaluate $\xi(L, K_1', K_2')$ at

$$\begin{aligned} K_1' &= K_1 \left(1 \pm \epsilon \frac{R\mathrm{e}^{K_1}}{q\mathrm{e}^{K_2}} (\mathrm{e}^{K_2} - 1)^2\right) \\ K_2' &= K_2(1 \pm \epsilon). \end{aligned} \tag{6.35}$$

The sample results shown in Table 6.7 exhibit a conspicuous improvement over those of Table 6.6. Not only is the convergence faster, but it is even seen that the estimators $\nu(L)$ form a monotonically increasing sequence for low values of R and a monotonically decreasing one for high R. The extrapolated ν is pinched between those two sequences and consequently quite accurately determined.

6.6.2 Results for the exponent ν

Plots of the estimators $\nu(L)$ for $3 \leq L \leq 7$ and several values of R are shown in Figs. 6.7 and 6.8 for $q = 8$ and $q = 64$ respectively. It is seen that the curves for $\nu(L)$ and $\nu(L-1)$ intersect at a unique value of R that seemingly converges quite fast to a definite value R^* as L increases. We interpret R^* as the randomness strength at the critical FP and the corresponding value of ν as the correct critical exponent. It is tempting to conjecture that the curves $\nu(L)$ approach ν on the entire interval $R \in]1, \infty[$ as $L \to \infty$. From the graphs it seems that the convergence is faster for large q.

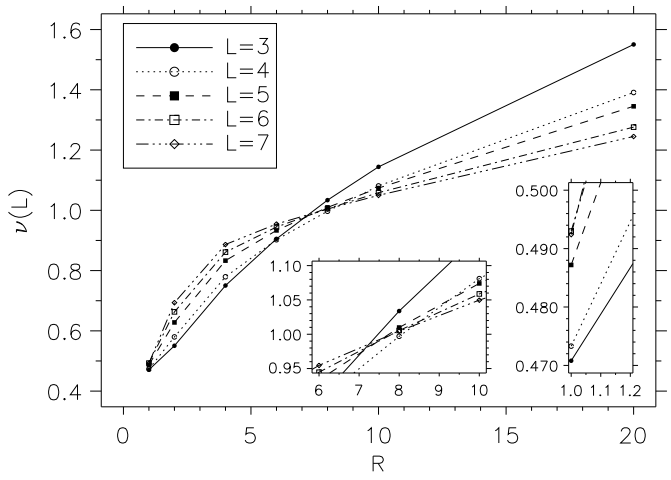

Figure 6.7: Estimants $\nu(L)$ for the thermal exponent at $q = 8$ as obtained from phenomenological renormalisation applied to strips of width L and $L-1$. In the pure system ($R \to 1$, see rightmost inset) the estimants converge to $\nu(\infty) = \frac{1}{2}$. Curves for neighbouring system sizes intersect at values of ν and R that converge to those at the random fixed point as $L \to \infty$. In this case $\nu = 1.01 \pm 0.02$ and $R^* \approx 9$ (see left inset). Error bars are no larger than the size of the symbols.

The values of ν and R^* corresponding to this scenario are shown in Table 6.8. In accordance with the phase diagram (Fig. 3.2) R^* is a slowly, supposedly logarithmically[4] , increasing function of q. For $q = 2$ the deviation from $\nu = 1$ can be ascribed to logarithmic corrections [58], and for $q = 3$ our result for ν is in agreement with the $(q-2)$-expansion [24] though the possibility of replica symmetry breaking cannot be ruled out [73]. Also for $q > 4$, our values for ν are numerically consistent with unity, indicating that, unlike what is the case for the magnetic exponent, the thermal one displays only a weak q-dependence.

From Figs. 6.7 and 6.8 a remarkable feature about the pure system ($R = 1$) is apparent. For $q = 8$ the estimators $\nu(L)$ seem to converge to $\nu = \frac{1}{2}$ whilst for $q = 64$ the extrapolated value is $\nu \simeq 0$. The former value is hardly surprising since, as we also remarked above, a first-order transition is expected to exhibit scaling with trivial effective exponents (in this case: $x_T = 0$). On the other hand, $\nu \simeq 0$ for $q = 64$ has to do with the length scales of the system. Namely,

[4] This supposition constitutes the simplest possibility allowed by the phase diagram of Fig. 3.2 in which $R^* \to \infty$ as $\ln q \to \infty$.

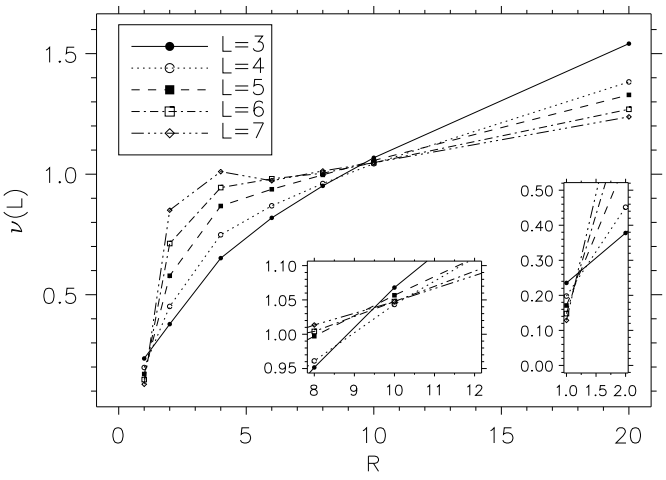

Figure 6.8: Phenomenological renormalisation at $q = 64$. The curves intersect at larger angles than before, allowing for a rather accurate determination $\nu = 1.02 \pm 0.03$ in spite of the large fluctuations. Error bars are comparable to the size of the symbols. From the rightmost inset it is seen that $\nu \to 0$ in the pure systems, as explained in the text. The left inset is a magnification of the region around $R^* \approx 10$.

from the asymptotic behaviour of the correlation length of the pure system [56]

$$\xi \sim \frac{2}{\ln q} \text{ as } q \to \infty \tag{6.36}$$

we infer that $\xi \sim 1 \ll L$ at the transition point of the $q = 64$ system. But away from the transition point we also expect $\xi = \mathcal{O}(1)$, since the lattice spacing (unity) is the smallest length scale in the system. After all there is a ferromagnetic interaction between nearest-neighbour spins. We thus conclude that ξ is roughly temperature-independent in this case. In order for this to be consistent

q	ν	R^*
2	1.12 (3)	7 (1)
3	1.04 (4)	8 (1)
8	1.01 (2)	9 (1)
64	1.02 (3)	10 (1)

Table 6.8: Values of the critical exponent ν and the randomness strength R^* at the random fixed point as obtained from phenomenological renormalisation.

with the asymptotic behaviour of the finite-size scaling ansatz (6.31)

$$\phi(x) \sim x^{-\nu} \text{ for } x \ll 1 \tag{6.37}$$

we must then have $\nu \simeq 0$. This is to be contrasted to the case of $q = 8$ where $\xi \gg L$ so that we clearly "see" the phase transitions in our strips of width L.

6.6.3 The criticism by Pázmándi *et al.*

Very recently the bound (6.21) was challenged by Pázmándi *et al.* [74]. These authors claimed that the standard method of averaging over the disorder in finite-size (FS) systems introduces a new diverging length scale into the problem, whence the resulting $\nu_{\rm FS}$ may be unrelated to the true exponent ν governing the divergence of the correlation length in the infinite system. In particular ν can be less than $\frac{2}{d}$, and if this is the case the standard method is liable to yield exactly $\nu_{\rm FS} = \frac{2}{d}$. Ref. [74] then went on to suggest a noise reduction procedure that professedly would allow one to access the true ν. For each realisation of the binary randomness (3.2) used in the disorder average there will be a fraction p of weak bonds (K_1). The noise due to the fluctuations of p around its average value $\overline{p} = \frac{1}{2}$ can then be reduced by adjusting the couplings (K_1, K_2) for that particular realisation to the values they would assume at the critical point of an infinite system with a (mean) fraction of weak bonds equal to p.

To implement this we are faced with the task of finding, to lowest order in $(p - \frac{1}{2})$, the two-dimensional critical surface in the space (K_1, K_2, p) from our knowledge of its one-dimensional intersection with the plane $p = \frac{1}{2}$, *viz.* Eq. (3.3). Let the fraction of weak bonds in a particular realisation be

$$p = \frac{1}{2}(1 + \epsilon_p), \tag{6.38}$$

where $\epsilon_p \ll 1$. The symmetry $p \leftrightarrow 1 - p$ ensures that, to first order in ϵ_p, we must still go perpendicularly off the self-dual curve in the (K_1, K_2) subspace, as in Eq. (6.35). Since an increase in the number of weak bonds must be offset by an increase of the K-s in order to keep the coupling to the energy density constant, the correct prescription is

$$\begin{aligned} K_1 &\rightarrow K_1 \left(1 + \epsilon_K \frac{R\mathrm{e}^{K_1}}{q\mathrm{e}^{K_2}}(\mathrm{e}^{K_2} - 1)^2\right) \\ K_2 &\rightarrow K_2(1 + \epsilon_K) \end{aligned} \tag{6.39}$$

for some $\epsilon_K > 0$. Demanding that the combined change in p and in (K_1, K_2) must leave the coupling to the energy density invariant furnishes a relation between ϵ_p and ϵ_K

$$\epsilon_K = \epsilon_p \frac{K_2 - K_1}{K_1 \frac{R\mathrm{e}^{K}}{q\mathrm{e}^{K_2}}(\mathrm{e}^{K_2} - 1)^2 + K_2} \tag{6.40}$$

q	1. gap	2. gap	3. gap	4. gap	5. gap
3	0.899 (4)	1.877 (13)	1.885 (12)	2.045 (24)	2.050 (23)
4	0.817 (5)	1.811 (9)	1.818 (8)	2.043 (23)	2.049 (24)
5	0.754 (6)	1.771 (6)	1.779 (6)	2.058 (24)	2.065 (25)

Table 6.9: Scaling dimensions corresponding to the first five gaps in the Lyapunov spectrum of $\mathcal{T}^{11}$ for $R = 2$. The parabolic least-squares fits included the first three cumulants of the probability distribution, and error bars were extracted based on the fits with $L_0 = 4, 5$ and 6.

and the derivatives (6.34) are now evaluated at these values of the parameters by going perpendicularly off the critical surface. To first order, of course, Eq. (6.35) still gives the correct way of doing so.

Our confidence in the results of Table 6.8 is increased by observing that the implementation of this novel averaging procedure does *not* alter our results. Indeed, trial runs for $q = 8$, where the discrepancy between the x_T extracted from the Lyapunov spectrum and phenomenological RG respectively is large, render the values of the estimators $\mu(L)$ unchanged within (small) error bars. It is thus concluded that even though our results for ν are conspicuously close to satisfying the bound (6.21) with equality, this is not due to an artifact in the averaging procedure.

6.7 The higher Lyapunov spectrum

Although the second Lyapunov exponent of $\mathcal{T}^{11}$ fails to yield the thermal scaling dimension x_T in the standard way it is hard to believe that the Lyapunov spectrum is not in some way related to the operator content of the CFT underlying the RBPM. In the case of the pure three-state Potts model, for example, it is well known [4] that the first five gaps of the Z_2-even sector are related to the energy density ϵ, its first descendants $L_{-1}\epsilon$ and $\overline{L}_{-1}\epsilon$, the stress tensor T and its conjugate $\overline{T}$. To wit, the scaling dimensions of these operators can be found from the gaps through Eq. (5.3), and we have verified this using our connectivity basis TMs.

In view of the bound (6.21) it is problematic to associate the first gap with the energy density in the random case, but it is nevertheless a beguiling question whether such concepts as descendants and the stress tensor are preserved by the randomness. To investigate this issue we have computed the first few Lyapunov exponents of $\mathcal{T}^{11}$ for $1 \leq L \leq 8$, averaging over 100 runs as usual. The scaling dimensions corresponding to the first five gaps for $q = 3, 4, 5$ and $R = 2$ are shown in Table 6.9. Self-averaging was ensured by utilising the cumulant expansion, and parabolic least-squares fits included the first three cumulants.

It is quite remarkable that even if the scaling dimension corresponding to the first gap may not be equal to x_T our data give strong reasons to believe

that it has a descendant, and that this descendant has the expected degeneracy of two. And even though the scaling dimensions in general depend on q those corresponding to the fourth and the fifth gaps are constant within error bars and very close to 2, as is expected for the stress tensor of a conformally invariant system [4]. Preliminary data for even higher Lyapunov exponents seem to hint at descendants at level two, but since we have found that in the pure system higher and higher eigenvalues require larger and larger system sizes before the asymptotic scaling form (5.3) is valid, massive computations are needed to establish reliable results for all but the first few scaling dimensions.

Another interesting feature of our data for the higher Lyapunov exponents is that the Harris criterion seems to be valid in a very complete sense. Namely, trial runs for $q = 1$ seem to indicate that although individual cumulants exhibit a pronounced dependence of R, their sum is virtually independent of the strength of the randomness in the whole range $R \in [1, 2]$. It thus appears that all exponents x_i that we can extract numerically from the Lyapunov spectrum, using Eq. (5.3) and the cumulant expansion, obey the Harris criterion. Since the connection between these exponents and the scaling dimensions of the underlying CFT is not completely known (witness x_T) this may well turn out to be a non-trivial observation.

Resumo en Esperanto

La nombraj rezultoj akiritaj per niaj transfermatricoj prezentiĝas. Kiel antaŭdirite [14] la malordo ŝanĝas la ordon de la Potts-modela faztransiro por $q > 4$ (Fig-o 6.1). La skalumo de la libera energio laŭ sistemgrando fiksas la valoron de la (efika) centra ŝargo c' (6.4). Por $q < 4$ la nombra precizeco apenaŭ sufiĉas por distingi inter la pure modelo kaj perturba kalkulo en potencoj de $(q - 2)$ (Fig-o 6.2). Aliflanke, por $q > 4$, $c'(q)$ kreskas proksimume logaritme kaj tre proksimas al la ekzakta rezulto (6.13) por la perkola limo (Fig-o 6.3). Skalŝanĝaj efikoj subtilas kaj estas zorge pridiskutataj.

Manke de aŭtoaveraĝo ne eblas senpere mezuri la magnetan korelacion. Anstataŭe ni proponas la kumulantan serion (6.15) en kiu la dekstreflankaj termoj ja estas aŭtoaveraĝaj kvankam la maldekstra flanko (la dezirata grando) ne. Por malgrandaj q la serio rapide konverĝas kaj donas tre precizajn valorojn por la magneta eksponento (Fig-o 6.4). La diverĝo por pli grandaj q okazas pro forta kuplaĵo inter la replikoj. La ekzisto de ne-nula dua (kaj pli altaj) kumulanto implicas multiskalumon, t.e. ke eksponentoj por diversaj potencoj de la korelacio ĝenerale ne simple rilatas inter si.

La termika eksponento kaŝas kelkajn surprizojn. Ĝiaj valoroj ŝajne malobeas al (6.21), kaj la kumulanta serio nur pligrandigas tiun problemon. Ni serĉas alternativajn manierojn fiksi x_T kaj i.a. diskutas pli fizikan interpreton (6.25) de la Furstenberga metodo [52, 53]. Finfine ni trovas x_T per fenomenologia renormigo, sed necesas iom modifi la metodon por atingi sufiĉe rapidan konverĝon por malpuraj sistemoj. La rezultaj valoroj de x_T estas superaj de (sed tre proksimaj al) unu por ĉiu q.

Kvankam x_T ŝajne ne simple rilatas al la Lyapunova spektro, tiu lasta enhavas du gravajn ecojn antaŭviditajn por sistemo kun konforma simetrio. Unue la eksponentoj havas entjeran interspacon (supozeble implicante la ekziston de Verma-modulo kun descendaj operatoroj) kaj due ekzistas duoble degenera operatoro kun skaluma dimensio du (la streĉo-energia tensoro) sendependa de q. La malobeo de la unua Lyapunova interspaco al (6.21) tiel restas ankoraŭ pli mistera.

Chapter VII

Discussion and outlook

Values for some of the critical exponents of the RBPM have previously been obtained by other authors. We comment on the relation between their results and ours, and conclude Part I with a discussion of various open questions as well as other models relevant to the issue of whether impurities soften first-order phase transitions.

7.1 Central charge

In a recent paper by Picco [61] it has been suggested that for $q = 2^p$ the effective central charge at the random FP is $c' = \frac{p}{2}$, and that this class of models thus behaves as p decoupled Ising models. Even without referring to our values of the magnetic exponent we should like to point out that all the data show is that $c' \propto \ln q$ with a constant of proportionality that is very close to $\frac{1}{2 \ln 2} \simeq 0.721$. But this constant is *also* very close to that of the percolation point, *viz.* $\frac{5\sqrt{3}}{4\pi} \simeq 0.689$. Indeed, these two numbers differ by less than 5%, and since our error bars and those of Picco are in the 2% and the 4% range respectively, there is no irrefutable way of numerically distinguishing between percolative, Ising-like or indeed some other, presently unknown, behaviour of the central charge. A similar observation is valid for $2 \leq q \leq 4$ where our numerical data as displayed in Fig. 6.2 are compatible, within error bars, with both the values at the pure and the random FP (but not, in this case, with those at the percolation point).

7.2 Magnetic exponent

On the other hand, our results for the magnetic exponent should leave no doubt that the correct CFT describing the RBPM cannot be that of a number of decoupled Ising models. In particular, the non-Ising value at $q = 8$ is in sharp disagreement with the Monte Carlo results of Ref. [20].

One possible explanation of this discrepancy would seem to be that these authors define a non-standard order parameter through

$$\rho = L^{-d}\langle \max(N_1, N_2, \ldots, N_q)\rangle, \tag{7.1}$$

where N_i is the number of spins in state i, which is related to our local order parameter defined in Eq. (4.12) by $N_i = \sum_r (M_i(r) + q^{-1})$. The site label r runs over a hypercube of side L with $24 \leq L \leq 84$. Note that ρ may also be written as

$$\rho = L^{-d} \lim_{k\to\infty} \left(\sum_{i=1}^{q} \langle N_i^k \rangle \right)^{1/k}. \tag{7.2}$$

Expressed in terms of the local order parameter, $\langle N_i^k \rangle$ gives a sum of terms each of the form

$$\sum_{r_j} \langle M_i(r_1)^{k_1} M_i(r_2)^{k_2} \ldots M_i(r_n)^{k_n} \rangle, \tag{7.3}$$

where $k_1 + k_2 + \cdots = k$. As $k \to \infty$ at fixed L, it is clear that at least some of the k_j must grow large. In the pure system, this should not matter, since each term will scale in the same manner. But when multiscaling is present, the scaling behaviour of the k_j power of the local order parameter may be different. Indeed, in the limit of $k \to \infty$ one would expect ρ to scale with dimension $\lim_{k\to\infty} x_k/k$, which is *less* than x_1 by convexity.

Though appealing, this explanation must now be dismissed since more recent Monte Carlo simulations using the same non-standard order parameter [27, 28] are also at odds with Ref. [20]. For $q = 8$ these new results are $x_1 = 0.153(3)$ [28] and $x_1 = 0.151(4)$ [27]. Using a different method the authors of Ref. [28] offered the alternative estimate $x_1 \sim 0.145$, which is closer to the value $x_1 = 0.142(4)$ found by us [46, 47]. A very likely explanation of our value's being somewhat too small is that the breakdown of the cumulant expansion for large values of q forced us to use a rather small randomness strength ($R = 2$) in the calculation of x_1, although we knew from both phenomenological renormalisation and the results for the central charge that $R \sim 10$ would be closer to the random fixed point at $q = 8$. Incidentally, $R = 10$ was used in both Ref. [27] and [28]. Despite of these details the conclusion is very clear: Both $x_1(q = 4) \sim 0.139$ and $x_1(q = 8) \sim 0.15$ are most definitely non-Ising exponents.

Another criticism of Ref. [20], which does not apply to Refs. [27, 28], is that the realisations of the binary randomness considered were confined to those for which the number of strong and weak bonds in each of the two lattice directions were equal. Though this restriction is clearly innocuous in the limit $L \to \infty$ this may not be so as far as the finite-size scaling is concerned. From trial runs where similar restrictions were imposed to the bond distributions of the TMs we found that seemingly harmless noise reductions schemes can influence the output substantially.

Finally, the mapping to the RFIM [46] illustrated that for large q typical configurations consist of large clusters of spins in the same Potts state, separated by domain walls. Whilst our very long strips are guaranteed to accommodate large regions in which all q values of the order parameters are realised, it is not clear that this should be the case in the much smaller square geometries of Ref. [20]. Indeed it seems likely that one would find Ising exponents if the geometry under consideration typically can accommodate at most two different large clusters.

7.3 Lyapunov spectrum of the even sector

We now turn our attention to the thermal exponent. If the phenomenological RG scheme is to be trusted the values of x_T only exhibit a weak dependence on q, although the $(q-2)$-expansion gives us reason to believe that there is some variation [24]. It is interesting that x_T stays so close to unity even at very high q, but presently we do not have any arguments to explain this finding. Unfortunately the method employed was unable to resolve the slight deviations from unity, and it is indeed a challenge to future research to find more accurate ways of assessing x_T for disordered systems. Our results on the higher Lyapunov spectrum are nothing if not intriguing, and we believe that a great effort must be made to understand why the first gap in the spectrum fails to be related to x_T in the standard way, even though the higher gaps show clear indications of a conformal field theory underlying the RBPM.

A very interesting issue to be addressed by future research is that of the dynamical universality class of the RBPM. In particular it would be interesting to see whether the dynamic critical exponent z does or does not agree with the Ising value of $z \approx 2$, or whether, in analogy with the RFIM, there is logarithmically slow dynamics due to the pinning of domain walls by impurities.

7.4 Other softening scenarios

Other types of randomness are also of interest to the question whether a first-order phase transition is softened due to impurities. In Part I of this thesis we have studied the effect of quenched bond randomness in a flat, regular lattice. A somewhat different scenario is obtained by investigating the pure q-state Potts model on lattices with quenched *connectivity* disorder. In Ref. [75, 76] MC simulations of the $q = 8$ model on two-dimensional Poissonian random lattices (Voronoi tessellations) with toroidal topology unambiguously showed that the first-order nature of the transition was not modified.

An argument why this must be so, at least for large q, can readily be given. For simplicity consider the model on the dual Delaunay random lattice, which by construction is a triangulation of space [77]. As on the regular lattice, at large q there are only two important states in the equivalent random cluster model: the empty lattice, which contributes a term $q^{N_{\text{vertices}}}$ to the partition function, and

the full lattice, contributing a factor $u^{N_{\rm edges}}$, where $u = e^K - 1$. Since for any triangulation $2N_{\rm edges} = 3N_{\rm vertices}$, the transition occurs when $u \sim q^{2/3}$ for *any* geometry. If we now consider that part of the lattice within a large hypercube of side L, the fluctuations in the difference of the energies of these two states inside this region will come solely from the edges which penetrate the boundary. On average, the difference in the energies of these two states will still be zero, but there will be fluctuations of the order of the square root of the number of bonds which penetrate the boundary, which will therefore be $\mathcal{O}(L^{(d-1)/2})$. These are very much smaller than the analogous fluctuations which are present when random bonds are added: these are $\mathcal{O}(L^{d/2})$, which leads to the Imry-Ma type of argument [10, 13, 14]. For $d = 2$ Voronoi tessellations the fluctuations are thus always smaller than the domain wall energy $\mathcal{O}(L^{d-1})$, and we conclude that such randomness is strongly irrelevant (at least for large q), in agreement with the results of Ref. [75, 76].

Yet another kind of randomness is obtained by studying the Potts model on quenched random gravity graphs, for which MC simulations for $q = 10$ have provided strong evidence for a softening scenario similar to ours [78]. However, in this case the *curvature* is random and hence when the lattice is embedded in the plane, it is fractal. Although the argument about compensation of the bulk energies when $u \sim q^{2/3}$ works for any triangulation, the number of boundary edges may well scale in a different manner for these lattices. Whilst it would be interesting to study this in detail, it is clear that this problem is quite different from ours, and neither our arguments nor those of Refs. [13, 14] can be directly applied.

Resumo en Esperanto

La rilato inter niaj rezultoj kaj tiuj de aliaj aŭtoroj estas pridiskutata. Unu el niaj ĉefaj rezultoj estas la ne-Isingaj valoroj por la magneta eksponento x_h kiam $q > 4$. Ref-o [20] tiel estas refutita.

Fine ni komentas pri aliaj situacioj kiuj koncernas la demandon ĉu malordo ŝanĝas la unuaordon faztransiron de la pura sistemo. Ni montras ke Potts-modelo sur reto kun frostita *konekta* malordo havu neŝanĝitajn valorojn de la kritaj eksponentoj [75, 76].

Part II

Frustration

Chapter VIII

Geometrical frustration

In Chapter 1 we have given some intuitive arguments why discrete lattice models of the Ising type can be expected to display a second order phase transition at some critical temperature T_c, that is of the order of the exchange coupling J. These considerations were epitomised by the equation $F = E - TS$, showing that a system that has a finite number of minimum energy states will tend to favour *one* of these states at low temperatures whilst being in a paramagnetic state at high temperatures. Although this would suggest that T_c is always non-zero, we have also seen from our survey of the random-field Ising model in Sect. 2.4 that disorder may drive the critical temperature all the way down to zero.[1]

Other types of systems with $T_c = 0$ exist, however, and they are the subject of this part of the present thesis. These are models that exhibit *geometrical frustration*, *i.e.*, they have an infinite number of degenerate ground states. More precisely, the systems with which we shall be concerned are *fully frustrated* in the sense that their ground states have a finite residual entropy per degree of freedom. A large class of such systems can be shown to have critical ground state ensembles, so that local operators constructed from the fundamental (spin) degrees of freedom display power law correlations at large distances. Evidently in such cases the criticality cannot be said to arise from the conflict between entropy and energy, but is indeed driven by entropy alone.

8.1 Critical ground state ensembles

To get an idea about the physics underlying the various fully frustrated systems, let us briefly consider some of the models which have been found over the years to display this type of behaviour.

[1] Another example of a system with $T_c = 0$ is the one-dimensional Ising model, as is easily seen from simple phase space considerations.

8.1.1 Triangular antiferromagnetic Ising model

The simplest of these models, and the first that was solved [79], is the isotropic antiferromagnetic Ising model on the triangular lattice.[2] The frustration of this model originates from the fact that the three spins on each elementary triangle of the lattice cannot all be pairwise antiparallel. Indeed, any configuration in which the tree spins on any elementary triangle are not all in the same state (up or down) is one of minimum energy.

An argument that the model is in fact fully frustrated can readily be given. Consider dividing the triangular lattice into the conventional three sublattices A, B and C, such that any spin on sublattice A has as its nearest neighbours three spins on sublattice B and three on sublattice C. Choosing all the B-spins to be up and all the C-spins to be down it is seen that any spin configuration on sublattice A will lead to a minimum energy state. Thus the residual entropy per spin has the lower bound $s > \frac{1}{3}\ln 2 \simeq 0.231$. It follows from Wannier's evaluation of the free energy [79] that the exact expression is

$$s_{\mathrm{TAI}} = \frac{2}{\pi}\int_0^{\pi/3} \ln(2\cos\omega)\,\mathrm{d}\omega \simeq 0.323. \tag{8.1}$$

An analytical evaluation of the spin-spin correlation function that is asymptotically exact along the three main directions of the lattice [81] shows that the ground state ensemble is critical and with exponents that differ from those of the ferromagnetic Ising model. These findings were also among the first to be placed into the general framework of *loop models* that shall be an important ingredient in the remainder of this thesis [82, 83].

8.1.2 Six-vertex model

Another fully frustrated model, crucial to the development of statistical mechanics around 1970, is the celebrated six-vertex model [84]. Originally introduced as a model of the hydrogen bonds in an ice crystal, this model is defined by placing arrows on the bonds of the square lattice subject to the constraint that the divergence of arrows at every vertex be zero. In the context of ice the vertices are taken to represent oxygen atoms. A hydrogen atom then lives on each lattice bond, and the direction of the arrow indicates to which oxygen atom it is bonded. The demand of local charge neutrality (the so-called *ice rule*) is equivalent to the above constraint on the arrows. The ice rule implies that every vertex can be in one of six configurations, and the fact that each arrow must be shared by the two vertices to which its bond is adjacent suggests that the six-vertex model is fully frustrated. Lieb's exact solution [85] implies that this is indeed the case:

$$s_{6\mathrm{V}} = \frac{3}{2}\ln\left(\frac{4}{3}\right) \simeq 0.432\,. \tag{8.2}$$

[2] A near-perfect experimental realisation can be found in the yavapaiite layered structure of anhydrous alums such as $\mathrm{RbFe(SO_4)_2}$ [80].

For various special choices of the vertex weights the six-vertex model can be recast as a loop model. The case where each of the four vertices with non-zero polarisation of the arrows carry half the weight of the two unpolarised vertices was considered in [86, 87], whereas the equal-weighted model emerges as special cases of a variety of models [88, 89, 90].

8.1.3 Potts vertex antiferromagnets

In both the examples mentioned this far the fully frustrated state can be viewed as the ground state ensemble of some more general model, in which a temperature variable T is taken to control the density of violations of the constraint associated with the frustration. This generalisation makes close contact with the zero-temperature critical points alluded to in the introduction, since a fully frustrated state is typically critical. We shall now present a few more examples of systems exhibiting this behaviour, and at the same time get acquainted with a class of combinatorial problems which are closely related to the two-flavour fully packed loop (FPL^2) model that is considered in detail in the subsequent chapters.

For any regular lattice of coordination number q one can define a q-state antiferromagnetic Potts *vertex* model through the Hamiltonian

$$\mathcal{H} = J \sum_{\mathbf{x}} \sum_{i<j=1}^{q} \delta(\sigma_i(\mathbf{x}), \sigma_j(\mathbf{x}))\,. \tag{8.3}$$

Here $\mathbf{x}$ labels the sites of the lattice, and Potts spins $\sigma_i(\mathbf{x}) = 1, 2, \ldots, q$ are taken to live on the bonds $i = 1, 2, \ldots, q$ adjacent to site $\mathbf{x}$. (Note that this is in contrast to the traditional definition of a Potts model, where the spins live on the vertices and interact along the bonds.) Evidently, in the limit of vanishing temperature the only allowed states are those for which the q spins surrounding any given site are in states that are all different. Associating each Potts state with a *colour* the $T = 0$ ground state encompasses the following combinatorial problem: "In how many ways can one colour the bonds of a q-fold coordinated lattice, using q different colours, subject to the constraint that all neighbouring bonds must be coloured differently?" It should hardly come as a surprise that the ground state ensemble thus defined is indeed fully frustrated. More importantly, in the colouring problems examined this far it was found to be *critical*.

The two-dimensional lattices for which this colouring problem has been addressed are the honeycomb lattice with $q = 3$ [91, 92] and the square lattice which has $q = 4$ [93, 90]. Investigations of the six-colouring problem on the triangular lattice are currently being undertaken.

Quite generally the q-colouring problem can be recast as a loop model. The idea is to define the loops as alternating sequences of a colour *pair*. These loops are then necessarily closed, and fully packed in the sense that every site of the lattice is visited by a loop. When $q \geq 4$, there is more than one independent

colour pair, and consequently more than one flavour of loops can be defined. The various flavours are all fully packed and mutually excluding.

Since loop models are seen to be ubiquitous in the study of geometrically frustrated systems, and indeed in all of statistical mechanics, we now turn to a general presentation of this subject.

8.2 Loop models

8.2.1 Introduction

Lattice models of loops have emerged as an important paradigm in two-dimensional critical phenomena. They allow for a determination of the scaling properties of different types of random walks which are used to model conformations of different phases of polymers [94]. For instance, the solution of the O(n) loop model has lead to exact results for conformational exponents of swollen and dense polymers [95], as well as polymers at the theta point [96]. The theta point is the tricritical point which governs the transition between the swollen and the collapsed phase of polymers in solution [94]. Examples of conformational exponents are γ, which describes the scaling of the number of polymer conformations with the number of monomers $\mathcal{N}$, and ν, determining the scaling of the linear size of a polymer, as measured by the radius of gyration, with $\mathcal{N}$. In Chapter 12 we calculate for the first time the exact value of γ for polymers on the square lattice, in the *compact phase*. Compact polymers completely fill the lattice and are of direct relevance to statistical studies of protein folding [97, 98].

Further motivation for studying loop models comes from the Fortuin-Kasteleyn construction which maps many discrete spin models (*e.g.*, Q-state Potts) to random cluster models. Since cluster boundaries in two dimensions form loops this naturally leads to a loop model representation. This random geometrical description of two-dimensional lattice models then provides a setting in which a general theory of their scaling limits can be sought. It is one of the goals of Part II of this thesis to outline a specific proposal for such a theory in the form of an effective field theory of fluctuating loops. This field theory is constructed following the Coulomb gas recipe [99] with some important new ingredients added [89]. It describes the fluctuations of a random surface for which the loops are contour lines.

Scaling limits of many (but not all) two-dimensional lattice models are described by conformally invariant field theories [100, 101]. This observation has lead to exact results for critical exponents and other universal quantities, and to a classification of critical points based on their symmetry properties with respect to the group of conformal transformations. An obvious question which is often difficult to answer is: "Given a particular lattice model, how does one *construct* the conformal field theory of its scaling limit?" Loop models provide examples for which the scaling limit can be constructed in a *physically* transparent way. This is accomplished by mapping a loop model to an interface model, where the

loops are simply equal-height contours. An explicit coarse graining procedure is then implemented for the height model, and it leads to a well known conformal field theory—the *Liouville field theory*.

Interesting examples of loop models are also provided by one-dimensional quantum models, spin chains in particular, where loops appear as world lines of the spin. This mapping of spins to loops has recently been used to formulate very efficient numerical schemes for simulating spin chains and ladders. These *loop algorithms* allow one to simulate much bigger system sizes and lower temperatures than by using more traditional algorithms with local updates (see Ref. [102] and references therein). The loop representation of quantum spin chains also gives an illuminating stochastic-geometrical view of their quantum fluctuations [103]. For example, the spin-spin correlation function is related to the probability that two points on the space-time lattice belong to the same loop. This insight might lead to a *practical* theory of plateau transitions in the Integer Quantum Hall Effect, *i.e.*, one that would allow for a calculation of the correlation length exponent and other universal quantities which have been measured in experiments. Namely, the Chalker-Coddington network model [104], which is believed to be in the same universality class as the plateau transitions, was recently mapped to an $SU(n \to 0)$ quantum spin chain [105]. It remains to be seen if this spin chain has a tractable loop-model representation.

In the bigger picture, loop models are of interest as simple examples where the fundamental constituents are non-local, extended objects as opposed to point-like objects such as particles and spins. Fluctuating geometries of this sort are used to model flux lines in superconductors, domain walls in magnets, and crystalline interfaces, to name a few experimentally relevant systems.

The extended nature of loops turns out to have profound consequences when one attempts to write down an effective continuum description of these models, say, following Landau's dictum of expanding the free energy (Euclidean action) in powers of the order parameter and its derivatives. Namely, terms which are geometrical in origin and non-perturbative in nature, and hence cannot be inferred from symmetry arguments alone, appear in the action. On the other hand, exactly *because* these geometrical terms are present the values of the *effective* coupling constants of the field theory are completely determined, a rather remarkable occurrence.

Usually in an effective description provided by a field theory, coupling constants are phenomenological parameters fixed by auxiliary information about observable quantities, such as the response functions or the related correlation functions. The Coulomb gas approach to two-dimensional critical phenomena is an example of an effective theory wherein the electromagnetic coupling constant (*i.e.*, the "magnitude of the unit charge") is determined from an exact solution of the model; typically it suffices to calculate the exact value of a single critical exponent. Our construction of an effective field theory of loop models closely parallels the Coulomb gas method with the important difference that the coupling constants are determined without recourse to any exact information about the model. For the FPL^2 model at hand no such information is available

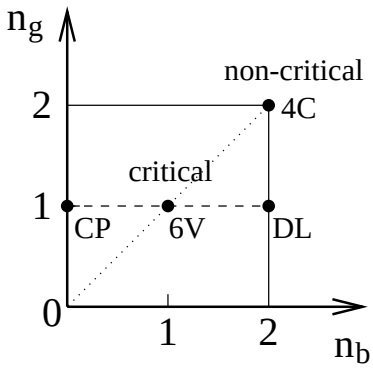

Figure 8.1: Phase diagram of the two-flavour fully packed loop model on the square lattice. The loop model is critical for loop fugacities $0 \leq n_{\rm b}, n_{\rm g} \leq 2$. Particular points in the critical phase map to previously studied models: 6V – equal weighted six-vertex model [106], DL – dimer loop model [107], 4C – four-colouring model [93]. The dashed line is the fully packed loop model studied numerically in Ref. [88]; the point CP along this line corresponds to the problem of compact polymers. Finally, the dotted line is the loop model for which an effective Liouville field theory was constructed in Ref. [89].

anyway, and moreover there are indications that the model is not exactly solvable [88]. On one level our theory can be viewed as a trick that allows one to calculate critical exponents in two-dimensional loop models without doing the "hard work" of exactly solving the model. On a deeper level it shows that lattice models of loops lead to continuum theories that are *geometrical* in nature, *i.e.*, devoid of any couplings that depend on the microscopic details.

8.2.2 The FPL^2 model

In Chapter 10 we study in detail the two-flavour fully packed loop (FPL^2) model on the square lattice. This is a statistical model which describes two flavours of loops that occupy the bonds of the square lattice, subject to certain close packing constraints to which we shall return shortly. The phase diagram of this model is described by two variables, $n_{\rm b}$ and $n_{\rm g}$, which are the loop fugacities of the two flavours; see Fig. 8.1. The phase diagram of the FPL^2 model has three important features that we wish to emphasize from the outset:
i) For loop fugacities that fall into the region $0 \leq n_{\rm b}, n_{\rm g} \leq 2$ of the phase diagram the model is critical, *i.e.*, it exhibits a power-law distribution of loop sizes. The novel feature is that every point in the critical region defines a *different* universality class characterised by an infinite set of geometrical critical exponents. All previously studied loop models (*e.g.*, Q-state Potts, O(n) models) exhibit a *line* of fixed points.
ii) The effective field theory of the FPL^2 model in the critical region describes a

fluctuating two-dimensional interface in five dimensions, which is characterised by *three* elastic constants. We calculate these three couplings exactly as a function of the two loop fugacities. It is important to note that all previously solved loop models are characterised by a single elastic constant.
iii) From the field theory of the FPL^2 model we calculate for the first time *exact* results for the conformational exponents of compact polymers on the square lattice. Furthermore, a particular line of fixed points in the phase diagram of the FPL^2 model can be identified with *interacting* compact polymers ($n_{\rm b} = 0, n_{\rm g} \leq 2$). We find that along this line the exponent γ changes continuously, whilst ν stays constant.

The organisation of the remainder of this thesis is as follows. In Chapter 9 we review the scaling theory of compact polymers which provides our main motivation for introducing the two-flavour fully packed loop model on the square lattice. The subsequent chapters are devoted to the study of this model using field-theoretical techniques and numerical transfer matrix calculations.

The FPL^2 model is mapped to an interface model in Chapter 10. For the interface model we construct the scaling limit in terms of a Liouville field theory, in Chapter 11. In Chapter 12 we make use of the field theory to calculate the central charge and the infinite set of geometrical exponents associated with loops, in the critical region of the loop model. A short description of the non-critical region based on the field theory is given afterwards.

Following the field-theoretical treatment of the FPL^2 model, in Chapters 13 and 14 we describe the construction of transfer matrices for different boundary conditions. They are used to determine the central charge, the first few geometrical exponents, and the residual entropy; the numerical results are in excellent agreement with the theoretical predictions. Finally, in Chapter 15, we present some general observations regarding compact polymers and the Coulomb gas description of conformal field theories. We also comment on the dimer-loop model [107] and the three-state Potts antiferromagnet [106], in light of our solution of the fully packed loop model on the square lattice.

Resumo en Esperanto

Krita konduto ne necese ekestas pro konflikto inter energio kaj entropio. Geometrie frustritaj sistemoj provizas ekzemplojn de kritaj statoj je temperaturo nula kies origino estas pure entropia. Aparte interesas nin kolorigaj problemoj de la tipo: "En kiom da manieroj eblas kolorigi la eĝojn de regula reto per q koloroj, tiel ke la aro da eĝoj tuŝantaj iun verticon enhavas precize q malsamajn kolorojn?" Uzanta paron da koloroj por difini buklon kiel sekvencon da alternaj koloroj oni atingas modelon de fermitaj bukloj kiuj vizitas ĉiun verticon de la reto.

Nia ĉefa motivo por studi tiajn modelojn estas la problemo de kompakta polimero sur la kvadrata reto. La kvarkoloriga problemo sur tiu reto nature difinas *du* specojn da bukloj, po kun sia pezo (kemia potencialo). Kiam ambaŭ

buklopezoj apartenas al la intervalo $[0, 2]$ la buklomodelo estas krita. Ĝia faz-diagramo estas tre riĉa kaj enhavas plurajn punktojn kiujn ni identigas kun modeloj antaŭe konsideritaj (Fig-o 8.1).

Chapter IX

Compact polymers

Novel universality classes for lattice polymers can be found by studying compact polymers, or Hamiltonian walks, which are self-avoiding random walks that visit *all* the sites of the underlying lattice; see Fig. 9.1. They have been used as simple models of polymer melts [108] and appear in statistical studies of protein folding [97, 98]. Unlike dilute and dense polymers (to be defined below) whose scaling properties were calculated exactly from the O(n) loop model [109], compact polymers defied a similar treatment until recently. Numerical transfer matrix calculations [110], a Bethe-ansatz solution [91], and a Coulomb gas theory [92] of the fully packed loop model on the *honeycomb* lattice, all conclude that compact polymers define a new universality class of critical behaviour. Here we study compact polymers on the *square* lattice. We calculate exact scaling exponents and find them to be distinct from the honeycomb case. This was first reported in Ref. [88] on the basis of numerical transfer matrix results.

The lattice dependence of critical properties distinguishes the compact polymer problem from its dilute and dense counterparts in a crucial way. It places them into the class of geometrically frustrated critical systems[1]. A physically relevant measure of frustration for compact polymers is the number of *contacts* per monomer. Contacts are realised by monomer pairs where the two monomers are nearest neighbors on the lattice but are not adjacent along the polymer chain. In lattice models of proteins hydrophobic interactions among the amino acids occur at contacts [97, 98]. For the square model studied here the number of contacts per monomer is *two*, whilst on the honeycomb lattice it is *one*; see figure 9.2. The triangular lattice, for which the compact polymer problem has not yet been solved, has four contacts per monomer.

At first sight it is somewhat surprising that the critical exponents for dense polymers, which by definition cover a finite fraction $0 < f < 1$ of the vertices, are lattice independent whilst those of compact polymers define new lattice dependent universality classes. A key to the understanding of this fact lies in the

[1] Another example is the antiferromagnetic three-state Potts model which has a zero-temperature critical point on the square [106] and the Kagomé [111] lattices characterised by different critical exponents.

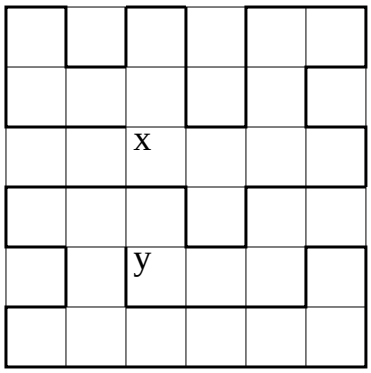

Figure 9.1: Compact polymer on the square lattice; **x** and **y** are the positions of the chain ends.

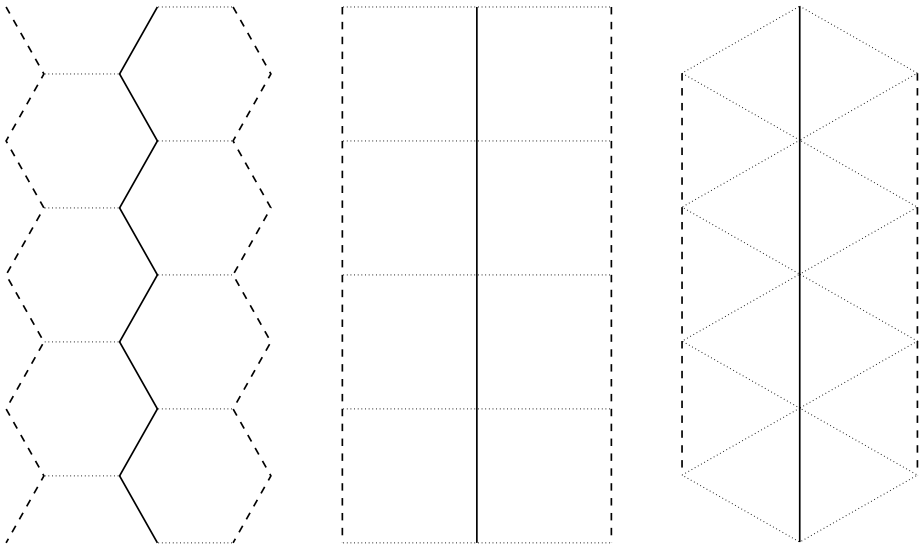

Figure 9.2: Number of contacts per monomer for compact polymers on the honeycomb (left), square (middle) and triangular (right) lattices. A macroscopically straight portion of the compact polymer (shown in thick solid linestyle) is surrounded on either side by other portions of the polymer (dashed lines) that are close to it in terms of the distance on the lattice, but far away in terms of the distance along the monomer chain. Contacts, *i.e.*, empty lattice edges connecting the different portions of the polymer, are shown as dotted lines. It is seen that the number of contacts is one per (thick solid) monomer on the honeycomb lattice, two on the square lattice, and four on the triangular lattice.

construction of the height representation for the equivalent interface model (see Chapter 10). Dense polymers, even for $f = 0.999$, are adequately described by a scalar height [112], whereas compact polymers ($f = 1$) necessitate a vectorial height, the dimensionality of which depends on the underlying lattice [92, 93]. In spite of this, certain *contact exponents* do turn out to be independent of $f > 0$. We shall return to the issue of universality in Chapter 15.

Compact polymers are crucial for the description of protein folding. We conclude this chapter by a concise review of basic questions addressed by this intriguing and very active field of research, and we state what can be learned about proteins from the study of two-dimensional statistical mechanics models.

9.1 Conformational exponents

9.1.1 The exponents ν and γ

In order to study the scaling properties of compact polymers we focus our attention on the two most widely studied conformational exponents, ν and γ. If $R = \sqrt{\langle(\mathbf{r} - \langle\mathbf{r}\rangle)^2\rangle}$ is the radius of gyration of the polymer then

$$R \sim \mathcal{N}^{\nu}, \tag{9.1}$$

where $\mathcal{N}$ is the number of monomers. Since compact polymers visit all the sites of a lattice, they are space-filling and we conclude that $\nu = 1/2$. This simple result will serve as an important check on our field theoretical calculations where it will be recovered.

In order to define the conformational exponent γ we introduce $C(\mathcal{N})$, the number of compact polymers (Hamiltonian walks) on a square lattice with $\mathcal{N}$ sites. Since a compact polymer fills the lattice, boundary conditions (free, periodic, *etc.*) play an important role. Following Duplantier and Saleur [112], we define γ in a way that is insensitive to the boundaries. Namely, if we introduce the quantity $C_\circ(\mathcal{N})$, the number of compact-polymer *rings*, then we can expect

$$\frac{C(\mathcal{N})}{C_\circ(\mathcal{N})} \sim \mathcal{N}^{\gamma}, \tag{9.2}$$

where γ does not depend on the choice of boundary conditions. Therefore, in order to calculate γ we need to solve the hard combinatorial problem of counting the number of open and closed compact polymers on the square lattice. Following de Gennes we do this by mapping the counting problem to the calculation of a correlation function in a particular statistical model at the critical point.

9.1.2 Scaling theory for γ

Consider the quantity $Z(\mathbf{x}, \mathbf{y}; \mathcal{N})$, the number of compact polymer conformations that start at the vertex $\mathbf{x}$ of the $\sqrt{\mathcal{N}} \times \sqrt{\mathcal{N}}$ square lattice, and end at $\mathbf{y}$ (see Fig. 9.1); we consider the limit $1 \ll |\mathbf{x} - \mathbf{y}| \ll \sqrt{\mathcal{N}}$, where $\mathbf{x}$ and $\mathbf{y}$ are

chosen far from the boundaries of the lattice. For this quantity we can write down the scaling form [112]:

$$Z(\mathbf{x},\mathbf{y};\mathcal{N}) = C_0(\mathcal{N})|\mathbf{x}-\mathbf{y}|^{-2x_1} f\left(\frac{|\mathbf{x}-\mathbf{y}|}{\mathcal{N}^{1/2}}\right) , \tag{9.3}$$

where $f(u)$ is a scaling function with the property $f(u) \to$ const. as $u \to 0$, and x_1 is a geometrical exponent related to γ. Integrating $Z(\mathbf{x},\mathbf{y};\mathcal{N})$ over all end-points $\mathbf{y}$ and comparing the result to Eq. (9.2), the scaling relation

$$\gamma = 1 - x_1 \tag{9.4}$$

follows.

To calculate the geometrical exponent x_1 we introduce in the next chapter the two-flavour fully packed loop model on the square lattice. The fact that we need *two* loop flavours follows from the simple observation that the bonds not covered by the compact polymer also form loops whose number is *unconstrained*. For the loop model we then construct an effective field theory in which $Z(\mathbf{x},\mathbf{y};\mathcal{N})$ becomes a two-point correlation function. The asymptotics of this function can be calculated exactly and we find $x_1 = -5/112$, from which

$$\gamma = 117/112 = 1.0446\ldots \tag{9.5}$$

follows. This is to be compared to the mean-field theory value $\gamma_{\mathrm{MF}} = 1$ [113], which is also the result obtained for compact polymers on the honeycomb lattice [91].

The conformational exponent γ was measured directly from enumerations of conformations of chains with lengths up to 30 in Ref. [97], and the value $\gamma = 1.01(5)$ was reported. More recently, from a numerical transfer matrix study of the fully packed loop model on the square lattice the geometrical exponent $x_1 = -0.0444(1)$ was determined [88], in excellent agreement with the exact result.

9.1.3 Connective constant

Another quantity of interest is the connective constant κ which determines the leading, exponential with system size, scaling of the number of compact polymers. For a system with a surface its scaling form reads [114]

$$C(\mathcal{N}) \sim \kappa^{\mathcal{N}} \kappa_{\mathrm{s}}^{\mathcal{N}^{(d-1)/d}} \mathcal{N}^{\gamma-1} . \tag{9.6}$$

Here κ_{s} is the surface connective constant; it appears due to the space-filling nature of compact polymers. Both the value $\kappa = 1.475(15)$ found in Ref. [97], and the estimate $\kappa \simeq 1.472$ obtained from transfer matrix calculations [115] similar to ours, seem in favour of the mean-field result $\kappa_{\mathrm{MF}} = \frac{4}{e} = 1.4715\ldots$ [113].[2] In Sect. 14.4 we report the very accurate numerical value

$$\kappa = 1.472801(10), \tag{9.7}$$

[2] Very recently the field theory of Ref. [113] has been improved [116] yielding, however, unchanged values for γ_{MF} and κ_{MF}.

which shows that the connective constant for compact polymers also deviates slightly from the mean-field result.

In the following chapters we elaborate on the calculation of γ for compact polymers, in the process unveiling an extremely rich phase diagram of the associated loop model. As remarked earlier, it contains a two-dimensional region of fixed points, which we characterise in detail by calculating the central charge and the geometrical exponents associated with loops for each point on the critical manifold.

9.2 Protein folding

Recently, compact polymers on two and three-dimensional lattices have become the model of choice for protein folding studies [117]. Here the focus is on the effect of non-specific and non-local hydrophobic interactions among the amino acids, on the folding process, and on the formation of secondary structure (helices and sheets). These investigations have been almost exclusively numerical, and an analytical theory of protein conformations that takes into account self-avoidance and compactness, as well as specific sequence information, would be of considerable interest [117].

The field theory of compact polymers on the square lattice[3] developed in the following chapters should be regarded as a first step towards this goal. Being a loop model our theory automatically takes self-avoidance into account, and compactness is modeled through the fully packing constraint. Chain connectivity, *i.e.*, the fact that there is a unique polymer covering the whole lattice, corresponds simply to taking one of the loop weights (fugacities) to zero. As it stands the model describes homopolymer folding and does not implement the restrictions due to specific sequences of amino acids. However, there is some hope that this effect may be incorporated in a similarly transparent way through the imposition of quenched randomness. We shall return to this point in Chapter 15.

To appreciate the connection between our results [118] and the questions raised by the protein folding community we begin by reviewing some basic properties of proteins, based on Refs. [119, 117]. Of course one may question the relevance of two-dimensional model systems to a problem that is essentially three-dimensional. However, from numerous numerical studies it has become clear that in most situation the physics of proteins is adequately captured in two dimensions. As stressed in Ref. [117] the advantage of exact model studies as ours is that they have few implicit biases and arbitrary parameters and that, in the spirit of universality, their predictions are often independent of our limited

[3] As mentioned earlier the compact polymer problem on the honeycomb lattice has previously been studied [91, 92]. However, within the context of protein folding this model is somewhat unphysical since there is only *one* contact per monomer. Indeed, since hydrophobic interactions in lattice models of proteins occur at contacts, the square and honeycomb problem describe different physical situations. (We thank Terry Hwa for pointing this out to us.)

knowledge of exact microscopic details.

Apart from the conformational exponents mentioned in Sect. 9.1 above and a definite prediction on the cooperativity of protein folding thermodynamics (see Sect. 9.2.5 below) it is our hope that the FPL^2 model will eventually enable us to make many further quantitative predictions for two-dimensional model proteins. The discussion of some of the things that we have in mind is deferred to Chapter 15.

9.2.1 Classification of biopolymers

Heteropolymers, long chain molecules made of different species of monomers, are ubiquitous in biological systems. Here they are known as biopolymers, and can be divided into several classes such as polysaccharides, nucleic acids (DNA, RNA), and proteins. Proteins are distinguished from other biopolymers by the fact that in their native aquatic environment they fold to (almost) unique compact structures. The prediction of the three-dimensional folded state from a particular sequence of monomers (amino acids) is known as the *protein folding problem*. This is an extremely hard problem involving both frustration and disorder, and the rugged nature of the phase space leads to slow dynamics and glassy behaviour. It is therefore quite legitimate to study the simpler problem of homopolymer collapse separately.

Proteins mediate many important biological processes such as catalysis (enzymes) and ionic transport (hemoglobin, chlorophyll). They are also present in virus shells and prions. The monomers building up proteins are the amino acids

$$\mathrm{NH_2 - C_\alpha HR - COOH} \tag{9.8}$$

consisting of the amine group $\mathrm{NH_2}$, the alpha-carbon group $\mathrm{C_\alpha HR}$, and the acidic group COOH. There are 20 different amino acids, each characterised by a particular residue R in the alpha-carbon group. The size of R ranges from a single hydrogen atom (glycine) to $\mathrm{CH_2 - C - CH - NH - C_6H_4}$ (tryptophan). A useful classification of the amino acids consists of first dividing them into hydrophobic (non-polar) and hydrophilic (polar), and then subdividing the latter according to their ionic charge (neutral, positive or negative). The hydrophobic amino-acid residues are responsible for the compactness of the folded state.

9.2.2 Formation of the native state

The full protein is formed by polycondensation of the amino acids: a water molecule is formed by splitting off an $\mathrm{OH^-}$ ion from the acid group of one amino acid and an $\mathrm{H^+}$ ion from the amine group of a second amino acid. As a result the two amino acids are now tied together by the *peptide bond* CONH, which is strongly planar. Symbolically the protein can then be represented by its backbone chain

$$\ldots - \mathrm{C_\alpha - C - N - C_\alpha - C - N - C_\alpha - C - N} - \ldots . \tag{9.9}$$

Typically proteins consist of 100–500 amino acids, thus placing them into the class of mesoscopic systems. In model studies, however, they can for most purposes be regarded as infinitely long, and they are then in principle tractable using the methods of statistical mechanics. At room temperature the soft degrees of freedom are the so-called ϕ-ψ angles, which are just the torsion angles along the backbone.

A remarkable fact about protein folding is that the folded state is essentially unique, *i.e.*, the conformational entropy in the native state is zero. Furthermore, this statement appears to be independent of detailed knowledge of the sequential information. Since the sole restriction of compactness does not prevent the conformational entropy from being quite large, *cfr.* our evaluation of the connective constant κ in Eq. (9.7), it is seen that the steric exclusion in a compact chain cannot by itself account for the uniqueness of the folded state. The entropy of the remaining degrees of freedom in the compact state must somehow be overcome by the forces of folding.

Another surprise is that the time needed for the folding is very large—of the order of one second, which should be compared to the time scale of 10^{-15} seconds associated with the atomic motion. However, even this astronomic separation of time scales does not permit the protein to find its unique native state by a 'trial and error' search among its huge number of possible configurations, not even if the search could somehow be restricted to the compact states. Indeed, the question how the protein 'selects' an efficient folding pathway remains one of the great mysteries of protein folding. The large folding time has to do with a rugged energy landscape with high free energy barriers and an exponentially large number of metastable states. Incidentally, this is true even for homopolymers.

9.2.3 Structure: Helices and sheets

The structure of a globular protein in its native state is conventionally classified in a three-state hierarchy. *Primary structure* is simply the sequential information for the amino acids along the backbone. This structure is present even in the denaturated (swollen) phase. *Secondary structure* is the local ordering in the folded state occurring due to the tendency of maximising the number of hydrogen bonds between the remnants of the acidic $(C-O)$ and the amine $(H-N)$ groups once the polycondensation has taken place.

The secondary structure can be further divided into the celebrated α-helices and β-sheets (see Fig. 9.3), which are quasi one- and two-dimensional structures respectively. The average number of turns (for the helices) and strands (for the sheets) is approximately as shown on the figure.

Finally, *tertiary structure* is the full three-dimensional structure of the globular protein, encompassing the compact packing of the secondary structure. Very large proteins sometimes consist of several globular domains, which are interconnected by only a few amino acids. The global arrangement of domains is then referred to as *quaternary structure*.

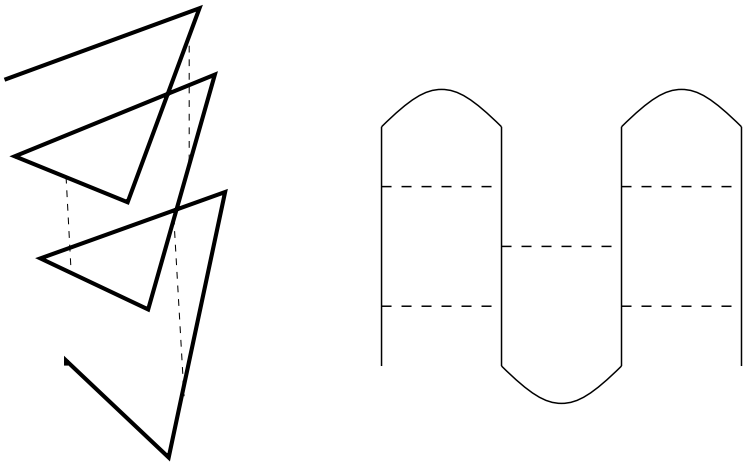

Figure 9.3: α-helix (left) and β-sheet (right). The dashed lines denote the hydrogen bonds. Taken from Ref. [119].

9.2.4 Interactions driving the folding

Over the years much effort has been made to answer the question which force(s) dominate the folding process. According to the *framework model* local interactions are determinant. This point of view was first set forth by Mirsky and Pauling in 1936 [120], who proposed that the local force of hydrogen bonding leads to the helical propensities (*cfr.* Fig. 9.3). The idea then is that local interactions drive the formation of secondary structure before the collapse, and that the tertiary structure is formed subsequently by close-packing of helices and sheets. Thus, symbolically, primary $\rightarrow$ secondary $\rightarrow$ tertiary.

The observation made by Kauzmann in 1954 [121] that hydrophobicity alone can drive folding put global properties into focus anew. Namely, in its native aquatic environment the protein globule tends to have the hydrophilic groups among its amino acids on the surface, thus shielding the hydrophobic core from the surrounding water. Two decades later these two points of view were synthesised into the commonly held belief that hydrophobicity is responsible for the compactness of the folded state, whereas hydrogen bonding leads to the formation of secondary structure.

An alternative view is contained in the *collapse model*, promoted by Dill and coworkers [117]. According to this model folding is driven by global properties *alone*, and the collapse drives concurrent formation of secondary structure. Clearly, our FPL^2 model is based on this point of view. The model property leading to the formation of a single compact chain is the loop fugacity's being taken to zero, and this is clearly expressing a non-local interaction.

9.2.5 Protein folding thermodynamics

A question within protein folding thermodynamics that has attracted much attention is whether cooperativity is a one-state or a two-state process [117]. In the language of biochemistry, cooperativity refers to a sigmoidal transition from the denatured to the native state, *i.e.*, one that is accompanied by a peak in thermodynamic quantities such as the heat capacity. Generally, this issue is discussed in terms of heteropolymers where hydrophobic contacts are energetically favoured, but it is also of interest to answer the question within the context of homopolymer collapse.

In the FPL^2 model denatured states are identified with finite-temperature excitations that violate the fully-packing constraint. To see in detail how this can be implemented in the colouring representation we refer to our calculation of the thermal exponent x_T in Sect. 12.2.4 below. Each violation corresponds an increase of the energy. For a polymer near the collapse transition, *i.e.*, a loop configuration in thermal equilibrium at a small but non-zero temperature, one can then imagine plotting the population of the states at various energies as a function of temperature. One-state and two-state behaviour refers to the number of peaks in the population-versus-energy distribution upon approach of the collapse transition.

It can be argued [117] that two-state behaviour corresponds to the density of states' having a gap separating native states from the denatured ones. In view of this, our identification of compact polymers with a *critical* model amounts to a definite prediction on the cooperativity of protein folding thermodynamics. Namely, since in a critical state there is no energy gap separating the first excited (non-compact) state from the native (compact) ones, at least in the large chain limit, we conclude that homopolymer collapse in two dimensions is a one-state process.

Resumo en Esperanto

Kompaktaj polimeroj difinas novajn retdependajn klasojn da krita konduto, malsamajn ol tiuj de densaj kaj ŝvelaj polimeroj. Ni enkondukas la konformaciajn eksponentojn ν (9.1) kaj γ (9.2). Por kompakta ĉeno $\nu = 1/2$ (triviale), kaj $\gamma = 117/112$ kalkuleblas tra la duspeca buklomodelo (9.4) kun buklopezo nula por unu el la specoj. Transfermatricoj donas tre precizan nombran valoron por la konekta konstanto κ (9.7).

Globetaj proteinoj en akvo kompaktas, kaj tiel niaj rezultoj—kvankam en nur du dimensioj—koncernas la studadon de proteinfaldado. Ni resumas tiun problemon [119, 117]. La buklomodelo nature pritraktas la spacan ekskludon kaj kompaktecon de proteinoj, sed ankoraŭ ne la hidrofoban interagon kiun kaŭzas specifa sekvenco da aminoacidoj. Ĝi tre bele montras kiel neloka interago sufiĉas por krei kompaktan strukturon.

Ĉar la buklomodelo estas krita ni konkludas ke homopolimera kolapso en du dimensioj estas unu-stata proceso.

Chapter X

The FPL^2 model and its height representation

Last chapter conveyed our motivation for studying compact polymers on the square lattice, and it is now time to explicitly introduce the loop model which contains the polymer problem as a special case. Interpreting loops as height contours allows for an interfacial representation, the continuum limit of which is described by the field theory of Chapter 11. Eventually, in Chapter 12, this construction will enable us to calculate *exact* values of the conformational exponents.

10.1 Four-colouring model and its loop generalisation

The two-flavour fully packed loop (FPL^2) model on the square lattice was introduced in Ref. [93] as the loop representation of the four-colouring model [122]. It is the natural generalisation of the fully packed loop model on the honeycomb lattice, which is the loop representation of the three-colouring model [92]. In general, a q-colouring model on a q-fold coordinated lattice is given by edge colourings of the lattice with q different colours; an edge colouring of a graph is one where no two bonds that share a common vertex are coloured equally. The colouring model is mapped to a loop model by choosing $[q/2]$ colour-pairs[1]; each pair defines strings of alternating colour that necessarily form loops (unless they terminate at the boundary). In this way we end up with a loop model with $[q/2]$ flavours of loops.

10.1.1 Partition function

To define the FPL^2 model we first specify the allowed loop configurations $\mathcal{G}$. In $\mathcal{G}$ every bond of the square lattice belongs to one and only one loop of either flavour, and loops of the same flavour are not allowed to cross. Representing the two flavours by solid (black) and hatched (grey) line segments respectively this fully packing constraint allows each vertex of the square lattice to have one

[1] The symbol $[x]$ denotes the integer part of x.

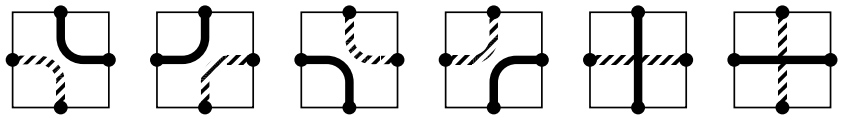

Figure 10.1: The six vertex configurations of the FPL^2 model that are allowed by the fully packing constraint. Black and grey loop segments are shown here as solid and hatched lines respectively. Each vertex is adjacent to four edges, here shown as filled circles, that are referred to as "dangling" if they are not connected to an edge of a neighbouring vertex. Note that the two rightmost vertices explicitly permit the two flavours to cross.

of the six appearances depicted in Fig. 10.1. Each loop is assigned a fugacity depending on its flavour: n_{b} for black loops and n_{g} for grey loops. The partition function of the FPL^2 model is then

$$Z = \sum_{\mathcal{G}} n_{\mathrm{b}}^{N_{\mathrm{b}}} n_{\mathrm{g}}^{N_{\mathrm{g}}} \ , \tag{10.1}$$

where N_{b} and N_{g} are the respective number of black and grey loops. The fully packed loop model of Batchelor *et al.* [88] is obtained by setting the loop fugacity of the grey loops to unity. In the limit $n_{\mathrm{b}} \to 0$ we recover the compact polymer problem.

10.1.2 Compact polymer limit

If we define a restricted partition function of the FPL^2 model, to which only configurations with a single black loop segment propagating between points $\mathbf{x}$ and $\mathbf{y}$ contribute, then $Z(\mathbf{x}, \mathbf{y}; \mathcal{N})$ in Eq. (9.3) is obtained in the limit $n_{\mathrm{b}} \to 0$, $n_{\mathrm{g}} \to 1$. The first limit discards all configurations with black loops present, leaving only the black Hamiltonian walk (compact polymer) between $\mathbf{x}$ and $\mathbf{y}$, whilst the second ensures that all walks are weighted equally. We could also consider weighting different Hamiltonian walks differently by setting $n_{\mathrm{g}} \neq 1$. This situation can be interpreted as describing interacting compact polymers, and, as will be shown later, it leads to a continuously varying exponent γ. A similar property of interacting oriented polymers in the swollen phase was suggested by Cardy from a field-theoretical calculation [123]. Recent numerical studies of the interacting oriented self-avoiding walk by Trovato and Seno [124], though, seem to be at odds with Cardy's prediction of an exponent γ that varies continuously with the interaction strength.

10.1.3 Phase diagram

Some idea of the phase diagram of the FPL^2 model as a function of n_b and n_g can be obtained by examining the extreme limits of the loop fugacities. Namely, for $n_\mathrm{b}, n_\mathrm{g} \to \infty$ all loops have the minimum length of four, *i.e.*, they each surround a single plaquette of the square lattice. There are no large loops in the system and the model is non-critical, or in other words, the average loop length is finite. On the other hand, in the critical phase of the loop model, which is our main interest, in a typical configuration one finds loops of all sizes characterised by a power-law distribution. This leads to an average loop length which diverges with the system size. Such is the case in the other extreme limit of loop fugacities, $n_\mathrm{b}, n_\mathrm{g} \to 0$, when the loops cover the whole lattice.

Other previously studied models that are particular points in the phase diagram of the FPL^2 model are the four-colouring model, the dimer loop model, and the equal-weighted six-vertex model; see Fig. 8.1.

For $(n_\mathrm{b}, n_\mathrm{g}) = (2, 2)$ the loop fugacity of each loop can be evenly (1+1) distributed among the two ways of colouring the bonds occupied by the loop with two colours in an alternating fashion:

$$\begin{aligned} &\mathbf{ABAB}\ldots\mathbf{AB} && \text{for black loops,} \\ &\mathbf{CDCD}\ldots\mathbf{CD} && \text{for grey loops.} \end{aligned} \tag{10.2}$$

More precisely, starting from a given vertex a black loop can be realised as either an $\mathbf{ABAB}\ldots$ or a $\mathbf{BABA}\ldots$ sequence (and similarly for the grey loops), thus giving rise to a loop fugacity of $n_\mathrm{b} = n_\mathrm{g} = 2$. This is then the symmetric four-colouring model ($\mathbf{A}, \mathbf{B}, \mathbf{C}$, and $\mathbf{D}$ are the colours) studied by Baxter [125].

In the dimer loop model black and white dimers are placed on the square lattice so that every vertex is covered by one of each [107]. If we identify the dimer-covered bonds with the black loops then this model is mapped to the $(n_\mathrm{b}, n_\mathrm{g}) = (2, 1)$ FPL^2 model. And finally $(n_\mathrm{b}, n_\mathrm{g}) = (1, 1)$ constitutes the equal-weighted six-vertex model [126], the allowed vertices being those of Fig. 10.1.

10.2 Height representation

The critical phase of the FPL^2 model can be described in terms of an effective field theory, following the general procedure discussed in Ref. [99]. The idea is to think of loops as contours of a scalar field, which we refer to as the height. Depending on the loop model in question the height can have one or more components. If the number of components is $D_\perp$ then the effective field theory of the loop model describes a fluctuating two-dimensional interface in $D_\perp + 2$ dimensions.

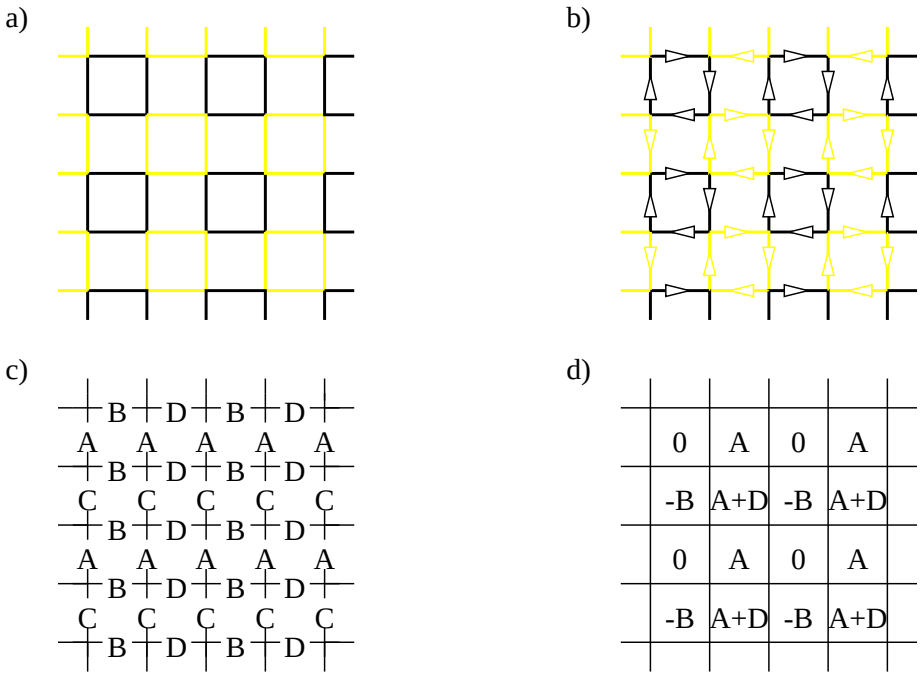

Figure 10.2: Mapping of the FPL^2 model to an interface model. (a) $\to$ (b): Transform the loop configuration into an *oriented* loop configuration by choosing the orientation of each loop independently and randomly. (b) $\to$ (c): Every bond in the oriented loop configuration is in one of four states, depending on its flavour and direction; these four states are represented by three-vectors $\mathbf{A}, \mathbf{B}, \mathbf{C}$, and $\mathbf{D}$. (c) $\to$ (d): The microscopic height $\mathbf{z}$ of the interface model changes from plaquette to neighbouring plaquette by $\mathbf{A}, \mathbf{B}, \mathbf{C}$, or $\mathbf{D}$ depending on the state of the bond between the two plaquettes. The change in $\mathbf{z}$ is positive going clockwise around even vertices and counterclockwise around odd ones.

10.2.1 Oriented loop model

To introduce the heights we first map the loop model to an *oriented* loop model, as shown in Fig. 10.2. The orientation of every loop is chosen randomly and independently. Every non-oriented loop configuration is thus transformed into an oriented one ($\mathcal{G}'$); the number of oriented configurations that correspond to the same non-oriented loop configuration is simply $2^{N_\mathrm{b}+N_\mathrm{g}}$.

Next, for each loop we redistribute its weight (fugacity), n_b or n_g depending on whether it is black or grey, between the two possible orientations. For the black loops we do this by assigning to the clockwise orientation the *phase factor* $\exp(\mathrm{i}\pi e_\mathrm{b})$, and the opposite phase, $\exp(-\mathrm{i}\pi e_\mathrm{b})$, to a counter-clockwise oriented black loop. Similarly for grey loops the clockwise oriented ones are assigned a weight $\exp(\mathrm{i}\pi e_\mathrm{g})$ whilst the counter-clockwise loops are weighted with

$\exp(-\mathrm{i}\pi e_{\mathrm{g}})$. The loop fugacities are related to the newly introduced parameters e_{b} and e_{g} by

$$\begin{aligned} n_{\mathrm{b}} &= 2\cos(\pi e_{\mathrm{b}}), \\ n_{\mathrm{g}} &= 2\cos(\pi e_{\mathrm{g}}), \end{aligned} \tag{10.3}$$

since the partition function of the original (non-oriented) model, as given by Eq. (10.1), must be recovered by independently summing over the two possible orientations for each loop. Note that for $0 \leq n_{\mathrm{b}}, n_{\mathrm{g}} \leq 2$ the parameters e_{b} and e_{g} are real, whilst for $n_{\mathrm{b}}, n_{\mathrm{g}} > 2$ they are purely imaginary. As discussed in more detail in Sect. 12.4 this is the crucial property that leads to a critical state of the loop model in the former and a non-critical one in the latter case.

10.2.2 Definition of microscopic heights

Now that the loops are oriented we can interpret them as contours of a height field; the orientation is necessary as it determines the direction of increasing height. The systematic construction of the *microscopic heights* sets out from the observation that every bond of the square lattice is in one of four possible states: it can be coloured black or grey, and oriented from an even to an odd site, or from odd to even. "Even" and "odd" refer here to the two sublattices of the bipartite square lattice; every even site is surrounded by four nearest neighbouring odd sites, and *vice versa*.

The four possible bond-states are represented by four vectors—which are the colours in the four-colouring representation—$\mathbf{A}, \mathbf{B}, \mathbf{C}$ and $\mathbf{D}$; see Fig. 10.2c. The microscopic heights $\{\mathbf{z}\}$ are defined on the dual lattice and the change in height when going from one plaquette centre to the next is given by $\mathbf{A}, \mathbf{B}, \mathbf{C}$ or $\mathbf{D}$, depending on the state of the bond which is crossed; Fig. 10.2d. For the height to be uniquely defined the four vectors must satisfy the constraint

$$\mathbf{A} + \mathbf{B} + \mathbf{C} + \mathbf{D} = 0. \tag{10.4}$$

This means that the microscopic heights live in a *three-dimensional* vector space, which we take to be $\mathbf{Z}^3$. In other words the oriented FPL^2 model maps to a model of a two-dimensional interface in five spatial dimensions.

By reasons of symmetry the four vectors are chosen so as to point from the centre to the vertices of a regular tetrahedron. With a suitable choice of coordinates they are represented by three-vectors:

$$\begin{aligned} \mathbf{A} &= (-1,+1,+1), \quad & \mathbf{B} &= (+1,+1,-1), \\ \mathbf{C} &= (-1,-1,-1), \quad & \mathbf{D} &= (+1,-1,+1). \end{aligned} \tag{10.5}$$

This is the same normalisation as the one used in Ref. [93].

10.2.3 Local redistribution of loop weights

Mapping the loop model to an oriented loop model also allows for a *local* redistribution of the loop weights. This is important since it leads to a local field

theory for the heights. As we will find out shortly, though local, this field theory is somewhat unconventional due to the non-local, extended nature of the fundamental microscopic objects it purports to describe.

To redistribute the phase factors associated with oriented loops we assign a phase $\exp(-i\pi e_{\rm b}/4)$ to a vertex of the square lattice if a black loop makes a left turn at that vertex, the opposite phase $\exp(+i\pi e_{\rm b}/4)$ if it makes a right turn, and the weight 1 if it continues straight. The total vertex weight $\lambda(\mathbf{x})$ is a product of the phase factor originating from the black loop and an equivalent one from the grey loop passing through the same vertex $\mathbf{x}$. The partition function of the FPL^2 model, Eq. (10.1), can now be rewritten as a sum over oriented loop configurations (*i.e.*, colouring configurations)

$$Z = \sum_{\mathcal{G}'} \prod_{\mathbf{x}} \lambda(\mathbf{x}) \ . \tag{10.6}$$

Once the height at a single point is fixed $\mathcal{G}'$ is in a one-to-one correspondence with the configurations of the microscopic heights, and the summand in the above equation is the appropriate weight. In the critical phase of the FPL^2 model the interface described by Eq. (10.6) is *rough*, and the field theory is constructed so as to correctly reproduce its long-wavelength fluctuations.

10.3 Continuum description

10.3.1 Spectrum of electromagnetic charges

The mapping from oriented loop configurations, which are equivalent to edge colourings, to microscopic height configurations is one to many. In particular, two height configurations corresponding to the same edge colouring can have their heights shifted with respect to each other by a *global* shift $\mathbf{m} \in \mathcal{R}$. The set $\mathcal{R}$ forms a three-dimensional Bravais lattice, *i.e.*, it is closed under integral linear combinations, and its elements are the *magnetic* charges in the Coulomb gas representation of the FPL^2 model. The lattice reciprocal to the lattice of magnetic charges, $\mathcal{R}^*$, defines the *electric* charges $\mathbf{e} \in \mathcal{R}^*$, with the property $\mathbf{e} \cdot \mathbf{m} = 2\pi m, m \in \mathbf{Z}$.

The construction of the lattice $\mathcal{R}$ for the FPL^2 model follows the usual prescription for height models, and has been carried out in detail in Ref. [93]. For the sake of completeness we outline this construction below.

10.3.2 Ideal states

It is convenient to first identify the *flat* states (also referred to as the *ideal* states), *i.e.*, those colouring states which minimise the variance of the microscopic height $\mathbf{z}$. From the height mapping described above it follows that these states have all of their plaquettes coloured with two colours only; an example is shown in Fig. 10.2c. This leads to a colouring state that is periodic, with the same 2×2 colouring pattern repeated throughout the lattice. There are twenty

four flat/ideal states for the colouring representation of the FPL^2 model, corresponding to the number of permutations of four different colours. Namely, an ideal state is completely specified by listing the colours of the bonds around a single site (say the origin), starting from the left horizontal bond and proceeding clockwise. To each flat state we assign a *coarse grained height* $\mathbf{h} = \langle \mathbf{z} \rangle$, which is the average microscopic height over a 2×2 unit cell of the colouring.

The flat states form a three-dimensional graph, which we refer to as the ideal state graph, $\mathcal{I}$. Namely, starting from any ideal state four other ideal states can be reached by exchanging a pair of colours that form a plaquette. For example, by exchanging the colours **A** and **B** in Fig. 10.2c all the **ABAB** plaquettes are turned into **BABA** plaquettes to give a new ideal state. Under these plaquette flips only the microscopic heights at the centres of the affected plaquettes are changed. In this way the ideal states form a four-fold coordinated graph in height space, where each vertex is indexed by a colour permutation, and its position in $\mathbb{R}^3$ is given by the coarse grained height $\mathbf{h}$. Bonds are associated with transpositions of two colours; they lie along the direction defined by the difference of the two colour vectors, and have a length of $1/\sqrt{2}$ if the normalisation in Eq. (10.5) is chosen.

The ideal state graph $\mathcal{I}$ is instrumental in the continuum description of the loop model. The basic assumption being made is that the dominant contributions to the partition function consist of bounded fluctuations around the ideal states, a point that will be clarified though the explicit construction of the (Liouville) field theory of the continuum limit in Chapter 11. In order to have a well-defined continuum limit local operators of the microscopic height must then be defined so as to be uniform in the ideal states. We turn to this issue next, and at the same time find occasion to discuss the details of the explicit construction of $\mathcal{I}$.

10.3.3 Definition of local operators

Labeling the ideal states by the colour configuration $(\sigma_1(\mathbf{x}), \sigma_2(\mathbf{x}), \sigma_3(\mathbf{x}), \sigma_4(\mathbf{x}))$ around a fixed vertex $\mathbf{x}$ local operators that are uniform in the ideal states can be defined as functions of the colours $\sigma_i(\mathbf{x})$. Following Ref. [93] we define the *staggered spin* $\mathbf{S}(\mathbf{x})$, the *row-staggered spin* $\mathbf{R}(\mathbf{x})$, the *parity* $P(\mathbf{x})$ and the *cross-staggered spin* $\mathbf{Q}(\mathbf{x})$ as follows:

$$\begin{aligned} \mathbf{S}(\mathbf{x}) &= \sigma_1(\mathbf{x}) + \mathrm{i}\sigma_2(\mathbf{x}) - \sigma_3(\mathbf{x}) - \mathrm{i}\sigma_4(\mathbf{x}), \\ \mathbf{R}(\mathbf{x}) &= \sigma_1(\mathbf{x}) - \sigma_2(\mathbf{x}) + \sigma_3(\mathbf{x}) - \sigma_4(\mathbf{x}), \\ P(\mathbf{x}) &= \sigma_1(\mathbf{x}) \cdot [\sigma_2(\mathbf{x}) \times \sigma_3(\mathbf{x})], \\ \mathbf{Q}(\mathbf{x}) &= [\sigma_1(\mathbf{x}) - \sigma_3(\mathbf{x})] \times [\sigma_2(\mathbf{x}) - \sigma_4(\mathbf{x})]. \end{aligned} \tag{10.7}$$

In Table 10.1 we show the explicit values of these four operators in each of the ideal states and at the same time give detailed information about how to construct the ideal state graph $\mathcal{I}$. The first column lists the colour configuration $(\sigma_1(\mathbf{x}), \sigma_2(\mathbf{x}), \sigma_3(\mathbf{x}), \sigma_4(\mathbf{x}))$ of the 24 different ideal states. Each state is related

State	$\Delta\mathbf{h}$	$\mathbf{S}$	$\mathbf{R}$	P	$\mathbf{Q}$
ABCD		$(0,+2+2\mathrm{i},+2-2\mathrm{i})$	$(-4,0,0)$	$+4$	$(-8,0,0)$
BACD	$(\mathbf{B}-\mathbf{A})/4=(+1/2,0,-1/2)$	$(+2-2\mathrm{i},+2+2\mathrm{i},0)$	$(0,0,-4)$	-4	$(0,0,+8)$
BCAD	$(\mathbf{C}-\mathbf{A})/4=(0,-1/2,-1/2)$	$(+2-2\mathrm{i},0,-2-2\mathrm{i})$	$(0,+4,0)$	$+4$	$(0,+8,0)$
CBAD	$(\mathbf{C}-\mathbf{B})/4=(-1/2,-1/2,0)$	$(0,-2+2\mathrm{i},-2-2\mathrm{i})$	$(-4,0,0)$	-4	$(+8,0,0)$
CABD	$(\mathbf{A}-\mathbf{B})/4=(-1/2,0,+1/2)$	$(-2-2\mathrm{i},-2+2\mathrm{i},0)$	$(0,0,-4)$	$+4$	$(0,0,-8)$
ACBD	$(\mathbf{A}-\mathbf{C})/4=(0,+1/2,+1/2)$	$(-2-2\mathrm{i},0,+2-2\mathrm{i})$	$(0,+4,0)$	-4	$(0,-8,0)$
DCBA	$(\mathbf{A}-\mathbf{D})/4=(-1/2,+1/2,0)$	$(0,-2-2\mathrm{i},+2-2\mathrm{i})$	$(+4,0,0)$	$+4$	$(+8,0,0)$
CDBA	$(\mathbf{C}-\mathbf{D})/4=(-1/2,0,-1/2)$	$(-2+2\mathrm{i},-2-2\mathrm{i},0)$	$(0,0,-4)$	-4	$(0,0,+8)$
CBDA	$(\mathbf{B}-\mathbf{D})/4=(0,+1/2,-1/2)$	$(-2+2\mathrm{i},0,-2-2\mathrm{i})$	$(0,-4,0)$	$+4$	$(0,-8,0)$
BCDA	$(\mathbf{B}-\mathbf{C})/4=(+1/2,+1/2,0)$	$(0,+2-2\mathrm{i},-2-2\mathrm{i})$	$(+4,0,0)$	-4	$(-8,0,0)$
BDCA	$(\mathbf{D}-\mathbf{C})/4=(+1/2,0,+1/2)$	$(+2+2\mathrm{i},+2-2\mathrm{i},0)$	$(0,0,-4)$	$+4$	$(0,0,-8)$
DBCA	$(\mathbf{D}-\mathbf{B})/4=(0,-1/2,+1/2)$	$(+2+2\mathrm{i},0,+2-2\mathrm{i})$	$(0,-4,0)$	-4	$(0,+8,0)$
DBAC	$(\mathbf{A}-\mathbf{C})/4=(0,+1/2,+1/2)$	$(+2+2\mathrm{i},-2+2\mathrm{i},0)$	$(0,0,+4)$	$+4$	$(0,0,+8)$
BDAC	$(\mathbf{B}-\mathbf{D})/4=(0,+1/2,-1/2)$	$(+2+2\mathrm{i},0,-2+2\mathrm{i})$	$(0,+4,0)$	-4	$(0,-8,0)$
BADC	$(\mathbf{A}-\mathbf{D})/4=(-1/2,+1/2,0)$	$(0,+2+2\mathrm{i},-2+2\mathrm{i})$	$(+4,0,0)$	$+4$	$(+8,0,0)$
ABDC	$(\mathbf{A}-\mathbf{B})/4=(-1/2,0,+1/2)$	$(-2+2\mathrm{i},+2+2\mathrm{i},0)$	$(0,0,+4)$	-4	$(0,0,-8)$
ADBC	$(\mathbf{D}-\mathbf{B})/4=(0,-1/2,+1/2)$	$(-2+2\mathrm{i},0,+2+2\mathrm{i})$	$(0,+4,0)$	$+4$	$(0,+8,0)$
DABC	$(\mathbf{D}-\mathbf{A})/4=(+1/2,-1/2,0)$	$(0,-2+2\mathrm{i},+2+2\mathrm{i})$	$(+4,0,0)$	-4	$(-8,0,0)$
DACB	$(\mathbf{C}-\mathbf{B})/4=(-1/2,-1/2,0)$	$(+2-2\mathrm{i},0,+2+2\mathrm{i})$	$(0,-4,0)$	$+4$	$(0,-8,0)$
ADCB	$(\mathbf{A}-\mathbf{D})/4=(-1/2,+1/2,0)$	$(0,+2-2\mathrm{i},+2+2\mathrm{i})$	$(-4,0,0)$	-4	$(+8,0,0)$
ACDB	$(\mathbf{C}-\mathbf{D})/4=(-1/2,0,-1/2)$	$(-2-2\mathrm{i},+2-2\mathrm{i},0)$	$(0,0,+4)$	$+4$	$(0,0,+8)$
CADB	$(\mathbf{C}-\mathbf{A})/4=(0,-1/2,-1/2)$	$(-2-2\mathrm{i},0,-2+2\mathrm{i})$	$(0,-4,0)$	-4	$(0,+8,0)$
CDAB	$(\mathbf{D}-\mathbf{A})/4=(+1/2,-1/2,0)$	$(0,-2-2\mathrm{i},-2+2\mathrm{i})$	$(-4,0,0)$	$+4$	$(-8,0,0)$
DCAB	$(\mathbf{D}-\mathbf{C})/4=(+1/2,0,+1/2)$	$(+2-2\mathrm{i},-2-2\mathrm{i},0)$	$(0,0,+4)$	-4	$(0,0,-8)$

Table 10.1: The 24 ideal states are labeled by the colour configuration around a fixed vertex. Colour transpositions correspond to bonds $\Delta\mathbf{h}$ of the ideal state graph $\mathcal{I}$. Also shown are the values of various operators defined in the text.

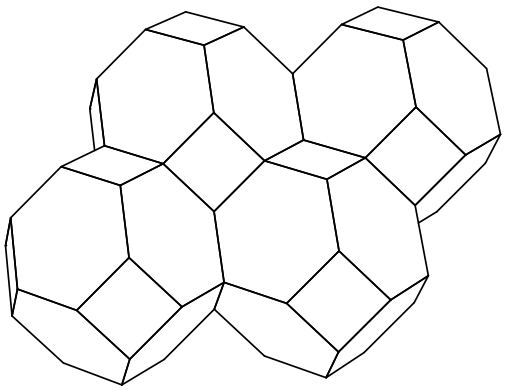

Figure 10.3: The ideal state graph of the FPL^2 model in the four-colouring representation.

to the following one by a transposition of two of the σ_i, corresponding to an exchange of the two colours around one fourth of the lattice plaquettes and an according modification of the microscopic heights. Considering the heights on the remaining three fourths of the plaquettes to be fixed during this process, the change in the coarse-grained height $\mathbf{\Delta h}$ is then one fourth of the difference between the concerned colour vectors (see column 2). Each value of $\mathbf{\Delta h}$ furnishes a link between two nodes on $\mathcal{I}$. By subsequently translating the 24 links thus obtained so as to fill space periodically we trace out Fig. 10.3. Note that as a result of this construction distinct nodes on $\mathcal{I}$ may correspond to the same ideal state.

10.3.4 The repeat lattice

The ideal state graph is a tiling of $\mathbb{R}^3$ with truncated octahedra; this regular polyhedron is better known as the Wigner-Seitz cell [2] of a body-centred cubic (bcc) lattice (see Fig. 10.3). A single truncated octahedron in $\mathcal{I}$ has twenty four vertices corresponding to the twenty four different ideal states. The set of vertices in $\mathcal{I}$ representing the same ideal state[2] form the *repeat lattice* $\mathcal{R}$, which is face-centred cubic (fcc) with a conventional cubic cell of side 4.

For this reason heights differing by a vector in $\mathcal{R}$ can be identified

$$\mathbf{h} \equiv \mathbf{h} + \mathcal{R}, \tag{10.8}$$

and the local operators $\mathcal{O}(\mathbf{x})$ just defined do by construction display the same

[2] We emphasize that two different vertices in $\mathcal{I}$ that correspond to the same ideal state do not necessarily occupy identical positions on the two octahedra to which they belong.

periodicity. They can therefore be expanded as a Fourier series

$$\mathcal{O}(\mathbf{x}) = \sum_{\mathbf{e}\in\mathcal{R}^*} \tilde{\mathcal{O}}_{\mathbf{e}} \exp\left(\mathrm{i}\mathbf{e}\cdot\mathbf{h}(\mathbf{x})\right), \tag{10.9}$$

where $\mathcal{R}^*$ is the lattice reciprocal to the lattice of magnetic charges $\mathcal{R}$. It is a body-centred cubic (bcc) lattice with conventional cubic cell of side π. The scaling dimension of $\mathcal{O}(\mathbf{x})$ is equal to the scaling dimension of the most relevant term(s) in this expansion, and this in turn is associated with the shortest vector(s) $\mathbf{e}$ for which $\tilde{\mathcal{O}}_{\mathbf{e}}$ is non-zero.

We are now ready to discuss the four different operators of Eq. (10.7) in turn. Here Table 10.1 proves useful for determining the most relevant reciprocal lattice vector occurring in the expansion (10.9). Another aid is to infer from the definitions (10.7) the various symmetries under permutations of the colours. For each operator the task is to find the *shortest* vectors in the sublattice of $\mathcal{R}^*$ that constitutes the height periods of the given operator.

The staggered spin $\mathbf{S}(\mathbf{x})$ assumes different values for all twenty four ideal states, and hence serves as a kind of order parameter in the model. Quite naturally the operator taking the highest number of different values on $\mathcal{I}$ must have the shortest electric charges in its Fourier expansion. The latter is therefore dominated by the eight vectors of type

$$\mathbf{e}_{\mathbf{S}} = \pi\left(\frac{1}{2},\frac{1}{2},\frac{1}{2}\right) \tag{10.10}$$

since these are the shortest vectors in the bcc-lattice $\mathcal{R}^*$ (which has a conventional cubic cell of side π).

The row-staggered spin $\mathbf{R}(\mathbf{x})$ and the parity $P(\mathbf{x})$ turn out to have the same periodicity. Their expansions are dominated by the next-shortest vectors in $\mathcal{R}^*$, which are the six vectors of type

$$\mathbf{e}_{\mathbf{R}} = \mathbf{e}_P = \pi(1,0,0). \tag{10.11}$$

Finally, the expansion of the cross-staggered spin $\mathbf{Q}(\mathbf{x})$ is dominated by the third-shortest vectors in $\mathcal{R}^*$, which are the twelve vectors of type

$$\mathbf{e}_{\mathbf{Q}} = \pi(1,1,0). \tag{10.12}$$

10.3.5 Coarse graining the microscopic height

To obtain the continuum description of the FPL^2 model we coarse grain the microscopic height over domains of ideal states. Fig. 10.4 illustrates this process which can be thought of as a two-step continuum limit: In (a) $\to$ (b) the identification of ideal state domains implies taking the continuum limit of the space variable $\mathbf{x}$ whereas $\mathbf{h}\in\mathcal{I}$ remains discrete. Then, in (b) $\to$ (c), a coarse graining of $\mathbf{h}$ promotes it to a continuously varying *height field* $\mathbf{h}(\mathbf{x})$.

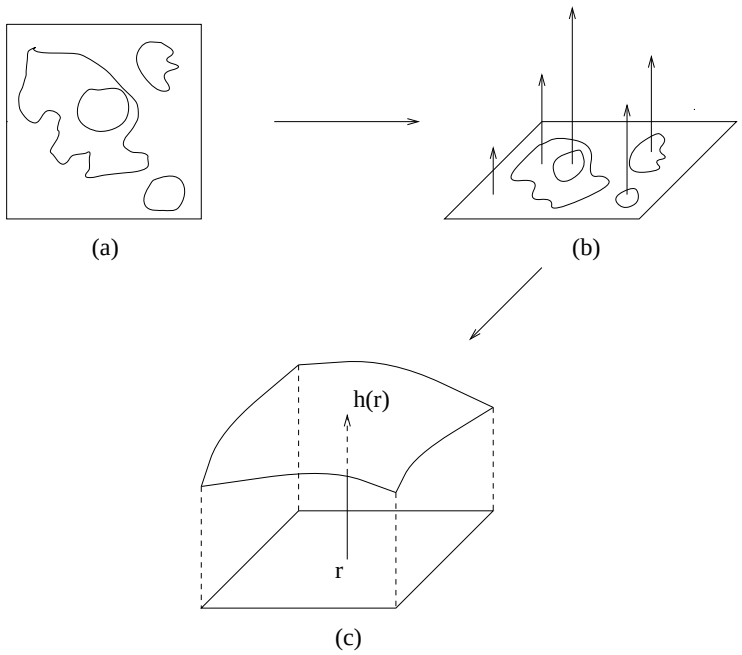

Figure 10.4: The coarse graining procedure. In (a) both the space variable $\mathbf{x} \in \mathbf{Z}^2$ labeling the vertices of the square lattice and the microscopic height $\mathbf{z} \in \mathbf{Z}^3$ are discrete. Dividing the particular colouring configuration into ideal state domains means that, when passing from (a) to (b), $\mathbf{x}$ may be coarse grained so as to constitute a continuous variable. At the same time $\mathbf{z}$ has been averaged over the ideal states, but the resulting coarse grained height $\mathbf{z} \in \mathcal{I}$ is still discrete. Finally, in (c) we take the continuum limit of the height, thus promoting it to the height field $\mathbf{h}(\mathbf{x})$. Taken from Ref. [93].

According to Eq. (10.8) we should consider the height field to be *compactified* on the 'three-torus' $\mathbb{R}^3/\mathcal{R}$. The phase space of the height is not simply connected, thus allowing for topological defects (vortices) with topological charges that take their values in $\mathcal{R}$ [127]. These defects are associated with *magnetic* charges in the Coulomb gas representation of the FPL^2 model. *Electric* charges on the other hand are associated with vertex operators $\exp(\mathrm{i}\mathbf{e}\cdot\mathbf{h})$. If we take the height to live in $\mathbb{R}^3/\mathcal{R}$ then vertex operators are well defined only for values of the electric charge $\mathbf{e} \in \mathcal{R}^*$.

We emphasize that the basic assumption being made is that most of the entropy in the FPL^2 model is contained in bounded fluctuations around the ideal states [93]. The crucial property that selects out the ideal states is that they are states of maximum entropy. To make this statement more precise consider what is the minimum change that can be made to a given colouring configuration. Clearly, at least one of the colours has to be changed, say $\mathbf{C} \to \mathbf{B}$

at some vertex $\mathbf{x}$. But then, in order not to violate the colouring constraint, the *other* **B**-bond at $\mathbf{x}$ must be changed to a **C**-bond. Now this implies a mismatch at the neighbouring vertex $\mathbf{x}'$, and further updates must be made. The end result is that the colours **C** and **B** along the entire loop passing through $\mathbf{x}$ defined by these two colours are transposed; *cfr.* Fig. 11.1. By saying that ideal states are states of maximum entropy we mean that they maximise the number of loops of alternating colour and hence allow for the maximum number of loop flips just described.

These considerations form one of the main ingredients of the construction of the field theory for the FPL^2 model, which is the subject of the next chapter.

Resumo en Esperanto

Ni konstruas interfacan reprezentiĝon por la buklomodelo (10.1). La difino de bukloj kiel alternaj sekvencoj de du koloroj ebligas direkti la buklojn kaj distribui ilian pezon inter la du direktojn (10.3). Ĉar ekz-e dekstruma buklo sur la kvadrata reto turnas kvarfoje pli dekstren ol maldekstren oni povas eĉ fari lokan redistribuon (10.6), kiu finfine implicos *lokan* kampteorion.

La eĝaj koloroj difinas diferencojn inter mikroskopaj altoj sur la duala reto se oni reprezentas ĉiun koloron per vektoro. Por plenumi (10.4) la vektoroj estu tri-dimensiaj (10.5). Kontinua limo de la mikroskopaj altoj konstruiĝas tra la 24 idealaj statoj (Fig-o 10.2c). Ili makroskope ebenas (Fig-o 10.2d) kaj estas entropie selektitaj ĉar ili ebligas la maksimuman nombron da ŝanĝiĝoj. La averaĝa alto en la idealaj statoj formas kovron de $\mathbb{R}^3$ per tranĉitaj okedroj (Fig-o 10.3). Verticoj sur malsamaj okedroj povas korespondi al la sama ideala stato, kaj ni do kompaktigas la averaĝan alton $\mathbf{h}$ rilate al la koresponda tri-dimensia krado $\mathcal{R}$ (10.8).

Fig-o 10.4 ilustras la manieron trovi la kontinuan limon por ajna stato. Unue dividu la staton en regionojn de idealaj statoj (a). Al ĉiu regiono atribuu la avaraĝan alton de la koncerna ideala stato (b). Finfine, pasu al la kontinua limo por la alto (c). Tiu procedo donas kulombgasan reprezentiĝon por la bukloj. Elektraj kaj magnetaj vektoraj ŝargoj apartenas al la kradoj $\mathcal{R}^*$ kaj $\mathcal{R}$ respektive. Fizike, elekta ŝargo korespondas al vertica operatoro $\exp(\mathrm{i}\mathbf{e}\cdot\mathbf{h})$ kaj magneta ŝargo al topologia difekto (vortico), kiu povas ekesti ĉar la alto estas kompaktigita sur la nekonektita objekto $\mathbb{R}^3/\mathcal{R}$.

Chapter XI

Liouville field theory

In constructing an effective field theory of the FPL^2 model one should aim at describing large-scale properties of loops. The kind of questions we expect it to answer are ones that do not refer to the microscopic details of the lattice model. For example, from the effective field theory we will calculate the asymptotics of the probability that two points lie on the same loop, when the separation between the points is large compared to the lattice spacing. From this and related quantities the conformational exponents of compact polymers can be extracted.

The field theory of the FPL^2 model is defined by the Euclidean action for the coarse-grained height $\mathbf{h}$. Consider a typical configuration of the oriented FPL^2 model which is equivalent to the colouring model. It consists of domains of ideal states. To each ideal state domain we assign a coarse-grained height, defined earlier as the average microscopic height over the domain. In the continuum limit we assume that this height is a smoothly varying function of the basal plane coordinates (x^1, x^2). The partition function that takes into account only the large-scale fluctuations of the height can be written as a functional integral,

$$Z_> = \int \mathcal{D}\mathbf{h}\ \exp(-S[\mathbf{h}]), \tag{11.1}$$

where S is the Euclidean action of a Liouville field theory with imaginary couplings [89]. The Liouville action contains three terms,

$$S = S_{\rm E} + S_{\rm B} + S_{\rm L}\ . \tag{11.2}$$

Each one has a concrete geometrical interpretation in the FPL^2 model, which we describe next.

11.1 Elastic term

The first term in the effective action for the FPL^2 model describes the elastic fluctuations of the interface. It gives less weight to configurations that deviate

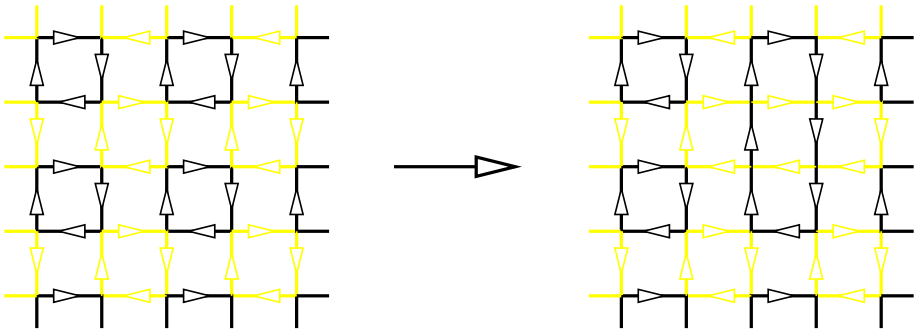

Figure 11.1: A loop flip changes one oriented loop configuration into another. Here the bond states **C** and **B** are exchanged along a single **BC** plaquette (*cfr.* Fig. 10.2c).

from the flat states, by penalising finite gradients of the height. This term is entropic in origin. Namely, in order to change the colour of a particular bond in the four-colouring representation of the loop model, say $\mathbf{C} \to \mathbf{B}$, all the **C**'s and **B**'s have to be interchanged along the **CB** loop which contains the chosen bond. This transformation we call a *loop flip*; see Fig. 11.1. The ideal states *maximise* the number of loops of alternating colour and consequently they have the largest entropy of loop flips.

In its most general form the elastic term in the effective action can be written as a gradient expansion,

$$S_{\mathrm{E}} = \frac{1}{2} \int d^2\mathbf{x} K^{ij}_{\alpha\beta} \partial_i h^\alpha \partial_j h^\beta \ , \tag{11.3}$$

where higher powers of the height gradients and higher derivatives of the height are less relevant at large scales. The stiffness tensor $K^{ij}_{\alpha\beta}$ nominally has 36 components; the indices $i, j = 1, 2$ are for the basal plane coordinates, whilst $\alpha, \beta = 1, 2, 3$ label the three components of the height. Summation over repeated indices is assumed throughout.

The number of independent non-zero components of the stiffness tensor (*i.e.*, elastic constants) is actually only *three*, once all the symmetries of the FPL^2 model are taken into account. The relevant symmetry transformations, that is the ones that become the symmetries of the effective action, are the ones that leave the weights of oriented loop configurations unchanged. First, there are the lattice symmetries, translations and rotations, which cut the number of independent elastic constants down to six. The terms that are allowed in S_{E} are scalars under rotations in the basal plane $\{(x^1, x^2)\}$, and they are necessarily of the form $\boldsymbol{\partial} h^\alpha \cdot \boldsymbol{\partial} h^\beta$, where $\boldsymbol{\partial} = (\partial_1, \partial_2)$ is the usual gradient. Second, the FPL^2 model possesses colour symmetries,

$$\mathbf{A} \leftrightarrow \mathbf{B} : e_{\mathrm{b}} \leftrightarrow -e_{\mathrm{b}} \text{ and } z_1 \leftrightarrow z_3 \tag{11.4}$$

and

$$\mathbf{C} \leftrightarrow \mathbf{D} : e_{\rm g} \leftrightarrow -e_{\rm g} \text{ and } z_1 \leftrightarrow -z_3 \ , \tag{11.5}$$

which interchange the colours and at the same time transform the microscopic heights. [We recall that (z_1, z_2, z_3) are the components of the microscopic height.] Taking into account the colour symmetries the elastic contribution to the action takes on the form:

$$\begin{aligned} S_{\rm E} = \frac{1}{2} \int d^2\mathbf{x} \ \{ K_{11}[(\boldsymbol{\partial} h^1)^2 + (\boldsymbol{\partial} h^3)^2] \ &+ \\ 2K_{13}(\boldsymbol{\partial} h^1 \cdot \boldsymbol{\partial} h^3) \ &+ \ K_{22}(\boldsymbol{\partial} h^2)^2 \} \ . \end{aligned} \tag{11.6}$$

Furthermore, by introducing a change of coordinates in height space,

$$H^1 = \frac{1}{2}(h^1 - h^3) \ , \quad H^2 = h^2 \ , \quad H^3 = \frac{1}{2}(h^1 + h^3) \tag{11.7}$$

$S_{\rm E}$ becomes diagonal,

$$S_{\rm E} = \frac{1}{2} \int d^2\mathbf{x} \ g_\alpha (\boldsymbol{\partial} H^\alpha)^2 \ . \tag{11.8}$$

The three coupling constants g_α ($\alpha = 1, 2, 3$) are linearly related to the three elastic constants,

$$g_1 = 2(K_{11} - K_{13}), \quad g_2 = K_{22}, \quad g_3 = 2(K_{11} + K_{13}). \tag{11.9}$$

The appearance of *three* elastic constants is rather intriguing from the viewpoint of loop models that have been solved previously. The Q-state Potts, the O(n), and the honeycomb FPL models are all characterised by a *single* coupling constant, which has been determined case by case from their exact solutions. Below we will show that all three couplings in Eq. (11.8) can be calculated exactly from the *loop ansatz* introduced in Ref. [89].[1] The ansatz states that the operator which enforces the complex weights assigned to oriented loops is *marginal* in the renormalisation group sense. This property of the field theory is intimately related to the random geometry of loops; we elaborate on this important point in Sect. 11.3.2.

11.2 Boundary term

The mapping of the loop model to an oriented loop model with local complex weights $\lambda(\mathbf{x})$ (Eq. (10.6)) fails for loops that experience the boundary. For example, if we define the FPL^2 model on a cylinder then loops that wind around the cylinder will not be weighted properly. The winding loop has an equal

[1] The coupling constant g for all the loop models known to date can be calculated using this method, therefore dispensing with the need for an exact solution.

number of left and right turns and hence it will be assigned a weight one. Summing over the two orientations gives a weight two, and not the correct $n_{\rm b}$ or $n_{\rm g}$, depending on the flavour. To correctly weight these loops one introduces a boundary term into the effective action,

$$S_{\rm B} = \frac{\rm i}{4\pi} \int d^2\mathbf{x}\ (\mathbf{e}_0 \cdot \mathbf{h})\mathcal{R}\ ; \tag{11.10}$$

$\mathcal{R}$ is the scalar curvature and $\mathbf{e}_0$ is the *background* electric charge, which is to be determined. Since we are only concerned with the situation where the lattice on which the FPL^2 model is defined is flat, the scalar curvature vanishes everywhere except at the boundary.

To determine $\mathbf{e}_0$ we consider the FPL^2 model on the cylinder. The scalar curvature of the cylinder is proportional to the difference of two delta functions situated at the two far ends of the cylinder:

$$\mathcal{R} = 4\pi\left[\delta(+\infty) - \delta(-\infty)\right]. \tag{11.11}$$

Therefore $S_{\rm B}$ has the effect of placing vertex operators $\exp(\pm{\rm i}\mathbf{e}_0 \cdot \mathbf{h})$ at $x^2 = \pm\infty$; here x^2 is the coordinate along the length of the cylinder. These vertex operators assign an additional weight $\exp({\rm i}\mathbf{e}_0 \cdot (\mathbf{h}(+\infty) - \mathbf{h}(-\infty))$ to oriented loop configurations on the cylinder. Now, in order for $\mathbf{h}(+\infty) - \mathbf{h}(-\infty)$ to be non-zero there must be at least a single winding loop present. If this winding loop is black, then the height difference is $\mathbf{A}$ or $\mathbf{B}$ depending on its orientation; similarly if the loop is grey the height difference is $\mathbf{C}$ or $\mathbf{D}$. Furthermore if the background charge is chosen so as to satisfy

$$\begin{aligned} \mathbf{e}_0 \cdot \mathbf{A} &= \pi e_{\rm b} \qquad & \mathbf{e}_0 \cdot \mathbf{B} &= -\pi e_{\rm b} \\ \mathbf{e}_0 \cdot \mathbf{C} &= \pi e_{\rm g} \qquad & \mathbf{e}_0 \cdot \mathbf{D} &= -\pi e_{\rm g} \end{aligned} \tag{11.12}$$

then the winding loops will be assigned their proper weights. This is again seen by summing over the two possible orientations of the winding loop. In the normalisation chosen for the colour vectors, Eq. (10.5), the unique solution of the system of linear equations in Eq. (11.12) is

$$\mathbf{e}_0 = -\frac{\pi}{2}(e_{\rm g} + e_{\rm b}, 0, e_{\rm g} - e_{\rm b}). \tag{11.13}$$

This calculation of the *vector* background charge generalises the scalar case studied previously [99].

11.3 Liouville potential

The elastic term and the boundary term make up the usual Coulomb gas approach to two-dimensional critical phenomena. Recently it has been argued that this description is incomplete [89] and that an extra term $S_{\rm L}$ must be added to the effective action. To see this consider a large loop in the bulk, one that

does not experience the boundary. Without the extra term this loop would be weighted exclusively by the bulk term $S_{\rm E}$. There are two problems with this: $S_{\rm E}$ is real whilst an oriented loop should be weighted by a complex phase, and, $S_{\rm E}$ does not distinguish between the two orientations of a loop which are assigned different weights. We conclude that an extra *bulk* term is necessary!

The most general form of a bulk term is

$$S_{\rm L} = \int d^2\mathbf{x}\; w[\mathbf{h}(\mathbf{x})]\;, \tag{11.14}$$

where $\exp(-w[\mathbf{h}(\mathbf{x})])$ is the scaling limit of $\lambda(\mathbf{x})$ in Eq. (10.6). In this sense $S_{\rm L}$ is energetic in origin, as opposed to $S_{\rm E}$, which we argued in Sect. 11.1 accounts for the entropy of edge colourings.

Microscopically, the vertex weight λ can be written in terms of the colours of the bonds around the particular vertex as $\lambda = \exp(-w)$ where

$$\begin{aligned}
w(\mathbf{B},\mathbf{C},\mathbf{A},\mathbf{D}) &= 0,\\
w(\mathbf{B},\mathbf{D},\mathbf{A},\mathbf{C}) &= 0,\\
w(\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D}) &= \mp\mathrm{i}\frac{\pi}{4}(e_{\rm g}+e_{\rm b}),\\
w(\mathbf{B},\mathbf{A},\mathbf{C},\mathbf{D}) &= \mp\mathrm{i}\frac{\pi}{4}(e_{\rm g}-e_{\rm b}),\\
w(\mathbf{A},\mathbf{B},\mathbf{D},\mathbf{C}) &= \mp\mathrm{i}\frac{\pi}{4}(e_{\rm b}-e_{\rm g}),\\
w(\mathbf{B},\mathbf{A},\mathbf{D},\mathbf{C}) &= \mp\mathrm{i}\frac{\pi}{4}(-e_{\rm b}-e_{\rm g})\;;
\end{aligned} \tag{11.15}$$

the top sign is for even vertices whilst the bottom sign applies to odd vertices of the square lattice. Here we adopt the notation $(\sigma_1,\sigma_2,\sigma_3,\sigma_4)$ for the ordering of the colours around a vertex by listing the colours clockwise from the leftmost bond. The operator w is completely specified by the values it takes on the six edge colourings listed above since it does not change under cyclic permutations of its arguments.

By explicitly going through the six colour configurations listed above it is easily checked that

$$w(\mathbf{x}) = \frac{\mathrm{i}}{16}\;\mathbf{e}_0\cdot\mathbf{Q}(\mathbf{x}), \tag{11.16}$$

where the cross-staggered operator (*cfr.* Sect. 10.3.3) is defined by

$$\mathbf{Q}(\mathbf{x}) = \pm[\sigma_1(\mathbf{x})-\sigma_3(\mathbf{x})]\times[\sigma_2(\mathbf{x})-\sigma_4(\mathbf{x})]. \tag{11.17}$$

Since $\mathbf{Q}(\mathbf{x})$ is manifestly invariant under 90° rotations of the colours around $\mathbf{x}$, Eq. (11.16) is seen to hold true for any distribution of the colours around a given vertex.

In order to find the coarse-grained version of $w(\mathbf{x})$ we express it as a function of the height field $\mathbf{h}(\mathbf{x})$, following the general procedure outlined in Sec. 10.3.4.

First note that the microscopic operator $w(\mathbf{x})$ is *uniform* in each of the ideal states of the four colouring model. As such it defines a function on the ideal state graph $w(\mathbf{h})$, where $\mathbf{h} \in \mathcal{I}$ is the coarse-grained height. Furthermore, it is a periodic function of $\mathbf{h}$ and it can therefore be written as a Fourier sum:

$$w(\mathbf{h}) = \sum_{\mathbf{e}\in\mathcal{R}_w^*} \tilde{w}_\mathbf{e} \exp(\mathrm{i}\mathbf{e}\cdot\mathbf{h}) \ . \tag{11.18}$$

The electric charges appearing in the sum take their values in the sub-lattice $\mathcal{R}_w^* \subset \mathcal{R}^*$, which is the lattice reciprocal to the lattice of *periods* of $w(\mathbf{h})$. In the continuum limit the coarse-grained height $\mathbf{h}$ is promoted into the height field $\mathbf{h}(\mathbf{x})$, and the scaling limit of the operator w is obtained by replacing $\mathbf{h}$ by $\mathbf{h}(\mathbf{x})$ in Eq. (11.18). Therefore $w[\mathbf{h}(\mathbf{x})]$ is a sum of vertex operators,

$$w[\mathbf{h}(\mathbf{x})] = \sum_{\mathbf{e}\in\mathcal{R}_w^*} \tilde{w}_\mathbf{e} \exp(\mathrm{i}\mathbf{e}\cdot\mathbf{h}(\mathbf{x})) \ , \tag{11.19}$$

of which only the most relevant one(s) are kept in the effective action. Since the relevance of an operator is determined by its scaling dimension we turn to this calculation next.

11.3.1 Dimensions of charge operators

In the Coulomb gas formalism operators are associated with either electric or magnetic charges. Electric operators are vertex operators $\exp(\mathrm{i}\mathbf{e}\cdot\mathbf{h})$ and they appear as the scaling limits of microscopic operators in the FPL^2 model that can be expressed as local functions of the colours; the loop-weight operator is one example.

Magnetic operators on the other hand cannot be expressed as local functions of the height but can be thought of as a constraint on the height field that generates a topological defect of strength $\mathbf{m}$. If $\mathbf{x}$ is the position of the defect core then the net height increase around any loop that encloses $\mathbf{x}$ is $\mathbf{m}$ (assuming that no other defects are encircled). Geometrical exponents for loops in the FPL^2 model are given by dimensions of electric and magnetic operators in the associated Coulomb gas.

For an operator that has total electromagnetic charge $(\mathbf{e},\mathbf{m})$, where $\mathbf{e} = (e_1, e_2, e_3)$ and $\mathbf{m} = (m^1, m^2, m^3)$, the scaling dimension is the sum of the electric and magnetic dimensions,[2]

$$2x(\mathbf{e},\mathbf{m}) = \frac{1}{2\pi}\left[\frac{1}{g_\alpha}E_\alpha(E_\alpha - 2E_{0\alpha}) + g_\alpha(M^\alpha)^2\right] \ , \tag{11.20}$$

where

$$E_1 = e_1 - e_3 \ , \quad E_2 = e_2 \ , \quad E_3 = e_1 + e_3 \tag{11.21}$$

[2]The derivation of Eq. (11.20) is an exercise in Gaussian integration and is reviewed in Appendix A.

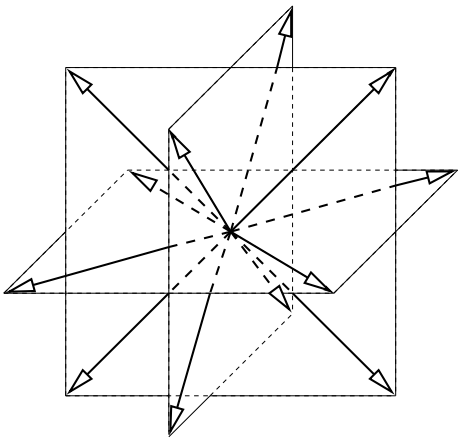

Figure 11.2: The twelve root vectors of the Lie algebra $su(4)$. Taken from Ref. [87].

and

$$M^1 = \frac{1}{2}(m^1 - m^3) \ , \quad M^2 = m^2 \ , \quad M^3 = \frac{1}{2}(m^1 + m^3) \tag{11.22}$$

are the electric and magnetic charge vectors in the basis in which the elastic term in the action is diagonal. Since the magnetic charges are given by height differences they must transform according to Eq. (11.7), whilst the electric charges transform in a dual fashion (*cfr.* their appearance in the vertex operators).

11.3.2 Loop ansatz

With the dimension formula in hand, we can settle the issue of the most relevant operators appearing in the Fourier expansion of $w(\mathbf{h})$; see Eq. (11.19). There are twelve vertex operators to choose from corresponding to the twelve (110)-type vectors in the bcc lattice $\mathcal{R}^*$. These are the shortest vectors in the lattice $\mathcal{R}^*_w$; see Fig. 11.2. To find which of these electric charges minimise $x(\mathbf{e}, 0)$ (Eq. (11.20)) it is convenient to first consider the simpler case of the FPL^2 model for $n_\mathrm{b} = n_\mathrm{g}$.

For the FPL^2 model with equal fugacities for the black and grey loops the effective action is considerably simplified. Namely, in this case the cyclic permutation of the colours,

$$\begin{aligned} (\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}) &\leftrightarrow (\mathbf{B}, \mathbf{C}, \mathbf{D}, \mathbf{A}) \ : \\ (z_1, z_2, z_3) &\leftrightarrow (-z_1, z_3, -z_2) \end{aligned} \tag{11.23}$$

does not change the vertex weight λ, and is thus an additional symmetry of the action S. This symmetry implies that $K_{13} = 0$ and $K_{22} = K_{11}$ in Eq. (11.6). Consequently there is only one elastic constant, $K \equiv K_{11}$. This then simplifies

the formula for the dimension of an electromagnetic charge,

$$2x(\mathbf{e},\mathbf{m}) = \frac{1}{2\pi K}\mathbf{e}\cdot(\mathbf{e}-2\mathbf{e}_0) + \frac{K}{2\pi}\mathbf{m}^2 \ , \tag{11.24}$$

where from Eq. (11.13) it follows that the background charge in this case has only one non-zero component, $\mathbf{e}_0 = -\pi(e_\mathrm{b},0,0)$. Now it is a simple matter to check that of the twelve (110)-type vectors in the lattice of electric charges $\mathcal{R}^*$, the four charges

$$\begin{aligned} \mathbf{e}^{(1)} &= (-\pi,0,+\pi), & \mathbf{e}^{(2)} &= (-\pi,0,-\pi), \\ \mathbf{e}^{(3)} &= (-\pi,+\pi,0), & \mathbf{e}^{(4)} &= (-\pi,-\pi,0) \end{aligned} \tag{11.25}$$

are degenerate in dimension and they minimise $2x(\mathbf{e},\mathbf{0})$. These are therefore the electric charges of the vertex operators that are kept in the action. From the viewpoint of conformal field theories endowed with Lie algebra symmetries the twelve shortest vectors in $\mathcal{R}^*$ are the so-called root vectors; see Fig. 11.2. The connection with conformal field theory is discussed further in the following section.

Now we turn to the *loop ansatz* which states that the operator $w(\mathbf{h})$ is exactly marginal in the renormalisation group sense. This is the statement that the loop weight does not renormalise at large scales. The geometrical meaning of this becomes obvious when one realises that the number of loops inside a domain of size ρ, whose linear size is *comparable* to ρ, is thermodynamically conjugate to the loop weight at scale ρ. Thus the loop ansatz states that the number of large loops does not grow with scale (more precisely it is sufficient to assume that it does not grow faster than any power of the scale). The analogous statement can be proven rigorously for critical percolation where it is the source of hyperscaling [128].

The assumption that there is of order one loop at every scale is linked to the variance of the height difference between two points in the basal plane, separated by a macroscopic distance $|\mathbf{x}|$. Namely, if we assume that when going from one point to the other there is of order one contour loop that is crossed at every *scale*, and further assuming that the directions of these contours are independent from scale to scale, it follows from the law of large numbers that the variance of the height difference grows as the number of contours crossed, that is as $\log(|\mathbf{x}|)$.[3] This of course is nothing but the large $|\mathbf{x}|$ behaviour of $\left\langle (H^\alpha(\mathbf{x}) - H^\alpha(\mathbf{0}))^2 \right\rangle$ calculated in the Gaussian model of Eq. (11.8).

The loop ansatz, or in other words the marginality hypothesis for the loop weight operator, simply translates into a statement about its scaling dimension:

$$x(\mathbf{e}^{(i)},\mathbf{0}) = 2 \quad i = 1,2,3,4 \ . \tag{11.26}$$

This, using the dimension formula Eq. (11.24), leads to a formula for the single elastic constant K.

[3] Proof: Let $\rho(r)$ be the density of loops on scale r. By assumption $\int_r^{kr} \rho(r')\,\mathrm{d}r'$ is independent of r, whence upon differentiation $\rho(r)/\rho(kr) = k$. It follows that $\rho(r) \propto 1/r$.

In the general case $n_{\rm b} \neq n_{\rm g}$, the scaling dimensions of the four electric charges identified above are

$$\begin{aligned} x(\mathbf{e}^{(1)},\mathbf{0}) &= \pi\,\frac{1-e_{\rm b}}{g_1}, \\ x(\mathbf{e}^{(2)},\mathbf{0}) &= \pi\,\frac{1-e_{\rm g}}{g_3}, \\ x(\mathbf{e}^{(3)},\mathbf{0}) = x(\mathbf{e}^{(4)},\mathbf{0}) &= \frac{\pi}{4}\left(\frac{1-2e_{\rm b}}{g_1}+\frac{1}{g_2}+\frac{1-2e_{\rm g}}{g_3}\right); \end{aligned} \tag{11.27}$$

the last two remain degenerate in dimension. The dimensions of the first two charges are also equal due to the "duality" transformation of the FPL^2 model which exchanges the two flavours, $n_{\rm b} \leftrightarrow n_{\rm g}$. This transforms the microscopic heights $z_2 \to -z_2$ and $z_3 \to -z_3$ (and similarly for the appropriate components of the height field). Furthermore, the elastic constants K_{11} and K_{22} in Eq. (11.6) are unchanged, whilst $K_{13} \to -K_{13}$. Finally, from Eq. (11.9) it follows that the duality transformation exchanges the couplings $g_1 \leftrightarrow g_3$ thus rendering $\mathbf{e}^{(1)}$ and $\mathbf{e}^{(2)}$ degenerate in dimension, as the FPL^2 model is self-dual.

Unlike the case of $n_{\rm b} = n_{\rm g}$, the loop ansatz in the general case requires that at least *two* of the electric charges $\mathbf{e}^{(i)}$ $(i = 1,2,3,4)$ remain marginal, thus enforcing the non-renormalisability of the two fugacities $n_{\rm b}$ and $n_{\rm g}$. If we now further *assume* that these charges are unrelated by the "duality" transformation described above, it follows that in fact all four are marginal. The three couplings are then simply calculated by setting the right hand sides of Eq. (11.27) equal to 2. We find:

$$\begin{aligned} g_1 &= \frac{\pi}{2}(1-e_{\rm b}), \\ g_3 &= \frac{\pi}{2}(1-e_{\rm g}), \\ \frac{1}{g_2} &= \frac{1}{g_1}+\frac{1}{g_3}\,. \end{aligned} \tag{11.28}$$

One final comment is in order. The relation $1/g_2 = 1/g_1 + 1/g_3$ comes as somewhat of a surprise, as it was not anticipated on symmetry grounds. Of course, since a particular point in the critical region of the FPL^2 model is determined by two parameters, $n_{\rm b}$ and $n_{\rm g}$, one relation between the three couplings is to be expected. It is therefore an interesting open question whether a critical loop model can be constructed in which g_2 would be unconstrained.[4]

With the values of the couplings g_1, g_2, and g_3 in hand, as well as the formula for the scaling dimensions of charged operators, Eq. (11.20), we are fully equipped to calculate critical exponents of the FPL^2 model. In particular, in the next chapter we calculate the formulae for the central charge and the geometrical exponents associated with loops as a function of the loop fugacities, $n_{\rm b}$ and $n_{\rm g}$, for the whole critical region of the model.

[4] This possibility was suggested to us by D. Huse.

11.4 Relation to Conformal field theory

The Liouville field theory proposed for the effective theory of the FPL^2 model in the critical region is conformally invariant. Each point in the critical phase is characterised by the central charge and the scaling dimensions of primary fields, which are associated with electric and magnetic charges in the Coulomb gas. For generic values of the loop fugacities the background charge $\mathbf{e}_0$ is not commensurate with the electric charges that make up the lattice $\mathcal{R}^*$. This implies that amongst the electric operators there will be many (an infinite number, in fact) that have negative dimensions, signaling the non-unitary nature of the conformal field theory.

To see this, note that the operator product expansion of two vertex operators with respective electric charges $\mathbf{E}_1$ and $\mathbf{E}_2$ reads [42]

$$e^{i\mathbf{E}_1\cdot\mathbf{h}(\mathbf{x})}e^{i\mathbf{E}_2\cdot\mathbf{h}(\mathbf{y})} \sim |\mathbf{x}-\mathbf{y}|^{E_1^\alpha E_2^\alpha/2\pi g_\alpha} e^{i(\mathbf{E}_1+\mathbf{E}_2)\cdot\mathbf{h}(\mathbf{x})} + \cdots , \tag{11.29}$$

where we are working in the basis (11.21) for notational simplicity. The main point here is that we generate a new electric operator with a charge that is the *sum* of the electric charges we started out from. Therefore, the complete set of electric operators available in the Coulomb gas theory have charges that can be formed as integral linear combinations of $\mathbf{e} \in \mathcal{R}^*$ and the background charge $\mathbf{e}_0$. If the latter is not commensurate with $\mathcal{R}^*$ there exists non-zero electric charges of arbitrarily small length. In particular, the scaling dimension (11.20) can be made negative.

Non-unitary CFT's appear in many other critical geometrical models, critical percolation being the best known example.

11.4.1 Screening charges and the loop ansatz

Liouville field theory provides the Euclidean action for the Coulomb gas description of conformal field theories proposed by Dotsenko and Fateev [62, 129]. As such it contains the so-called screening charges which are the vertex operators that make up the Liouville potential. In the original formulation these charges were introduced on formal grounds so as to ensure the existence of non-vanishing four-point correlation functions in the theory. In order for the modified Gaussian model (the modification is the addition of the boundary term to the gradient-square action) to stay conformal, these vertex operators are necessarily marginal, *i.e.*, their scaling dimension is two.

Here we have found a physical interpretation of the screening charges. Their role in loop models is to ensure that the number of large loops from scale to scale stays of order one; this translates into the statement that the loop fugacities do not flow under the action of the renormalisation group.

The fact that we have a concrete physical interpretation of the screening charges directly leads to the calculation of the elastic constants in the Liouville field theory. In the traditional Coulomb gas approach these coupling constants are calculated by comparing with formulae derived from an exact solution of

the model. Once these constants are known marginal vertex operators that play the role of screening charges can be written down. Our construction basically reverses this procedure, and by doing so *makes no reference to an exact solution.*

11.4.2 Lie algebra symmetries

Perhaps the most pleasing aspect of loop models is that they offer a concrete *geometrical* interpretation of many of the usual CFT concepts. For instance, the conformal field itself is just the height field $\mathbf{h}(\mathbf{x})$, but a lot more is true [87]. Here we just mention some of the main points, since the introduction of the necessary formalism [42] would take us too far afield.

The four-colouring model, which is contained as the point $(n_{\mathrm{b}}, n_{\mathrm{g}}) = (2, 2)$ in the phase diagram of the FPL^2 model, was originally proposed by Read [122] as a critical lattice model whose continuum limit is endowed with an $su(4)$ Lie algebra symmetry. From the theory of simple Lie algebras it is well known [130] that the symmetry algebra of a conformal field theory can be larger than the Virasoro algebra. This happens if the repeat lattice is equal, up to a scale factor, to the root lattice of some simple Lie algebra and the compactification radius (or, in other words, the[5] elastic constant) has the correct value, so as to render the vertex operators associated with the root vectors currents of conformal dimension one. In the case of the four-colouring model this is precisely the case, and the continuous symmetry of the effective field theory can be shown to be given by the $su(4)_{k=1}$ Kac-Moody algebra [87]. In particular, the root lattice of the $su(4)$ algebra is an fcc-lattice, as is the repeat lattice $\mathcal{R}$ entering our height construction.

Not only does the four-colouring model have a Kac-Moody symmetry algebra, but its stress-energy tensor can also be shown to have the Sugawara form [87]. This implies that the four-colouring model is the free-field representation of the $SU(4)_{k=1}$ Wess-Zumino-Witten (WZW) model. An equivalent programme can be carried through for the six-vertex model and the three-colouring model on the honeycomb lattice, which turn out to be free field representations of the $SU(2)_{k=1}$ and the $SU(3)_{k=1}$ WZW models, respectively. An interesting open question, which we are currently pursuing, is whether the six-colouring model on the triangular lattice does similarly represent the $SU(6)_{k=1}$ WZW model.

When moving away from the $(n_{\mathrm{b}}, n_{\mathrm{g}}) = (2, 2)$ point in the FPL^2 phase diagram by means of a non-zero background charge and perturbation by an exactly marginal operator, the $SU(4)_{k=1}$ symmetry disappears, since the background charge changes the compactification radius.

To complete our 'geometrical' identification we note that $\mathcal{R}^*$ is the weight lattice of the $su(4)$ Lie algebra. $\mathcal{R}_w^*$ is the root lattice, and its twelve shortest vectors are the roots; see Fig. 11.2. As expected for an $su(4)$ algebra the quotient of the weight lattice and the root lattice is Z_4.

[5] For $n_{\mathrm{b}} = n_{\mathrm{g}}$ there is only one elastic constant, as mentioned in connection with Eq. (11.23).

Resumo en Esperanto

La Liouville kampteorio por la interfaca reprezentiĝo de la buklomodelo havas tri termojn en sia efika ago. *Elasta termo* (11.6) kontrolas la devion de la entropie selektitaj idealaj statoj. Post konsidero de simetrioj restas tri malsamaj elastaj konstantoj; ĉiuj ĝisnunaj buklomodeloj nur havis unu tian konstanton. *Randa termo* (11.10) asignas la ĝustan pezon al bukloj kiuj ĉirkaŭas la punkton transfinian. Ĝi kuplas la altokampon al la skalara kurbiĝo $\mathcal{R}$ tra fona elektra ŝargo (11.13). Finfine, la *Liouville potencialo* (11.14) asignas la ĝustan pezon al ĉiuj ceteraj bukloj. En la kontinua limo ĝi esprimeblas kiel serio da verticaj operatoroj (11.19), inter kiuj ni serĉas la plej modifivajn.

La skaluma dimensio (11.20) por operatoro kun *vektora* elektromagneta ŝargo estas eltrovita en Apendico A. Montriĝas ke el la verticaj operatoroj, kiuj korespondas al la dekdu malplej longaj elektraj ŝargoj en la krado $\mathcal{R}_w^*$, nur kvar estas degeneraj laŭ la linio $n_\mathrm{b} = n_\mathrm{g}$, kaj ni argumentas ke ili estu degeneraj en la tuta krita regiono. Fakte ili estu *ekzakte marĝenaj* (11.26) por ke la buklopezoj ne renormiĝu je larĝaj longoskaloj; ili do estas geometria interpreto de la kutimaj ŝirmoŝargoj [62, 129]. La marĝena kondiĉo fiksas la elastajn konstantojn (11.28), kaj la modelo estas solvita.

Chapter XII

Critical exponents

Some of the critical exponents of the FPL^2 model have already been alluded to in Chapter 9. A supposedly complete characterisation of the critical behaviour in the region $0 \leq n_{\rm b}, n_{\rm g} \leq 2$ is furnished by the values of the central charge c and an infinity of geometrical scaling dimensions associated with height defects in the interface representation. We calculate these universal quantities in turn and conclude the Chapter with a short description of the termination of critical behaviour as one of the fugacities becomes larger than two; this is based on the field theory developed in Chapter 11.

12.1 Central charge

We first turn to the calculation of the central charge in the critical region. Exactly at the point $(n_{\rm b}, n_{\rm g}) = (2, 2)$ the background charge vanishes, $\mathbf{e}_0 = \mathbf{0}$, and the action consists only of the elastic term $S_{\rm E}$ given by Eq. (11.8). Since this is then simply a theory of three free massless bosonic fields we conclude that, in this case, $c = 3$ [93].

For a general value of the background charge this generalises to [62, 129]

$$c = 3 + 12x(\mathbf{e}_0, \mathbf{0}) \; . \tag{12.1}$$

One way to rationalise the factor of 12 is to compare the coefficients of the finite-size corrections in the well-known formulae [49, 50, 51]

$$f_0(\infty) - f_0(L) \;=\; \frac{\pi c}{6L^2} + \cdots \tag{12.2}$$

$$f_i(L) - f_0(L) \;=\; \frac{2\pi x_i}{L^2} + \cdots \; , \tag{12.3}$$

where $f_{0,i}(L)$ is the free energy density on a cylinder of circumference L, the subscript 0 referring to the vacuum and i to the case when an operator of scaling dimension x_i is inserted. The physical meaning of Eq. (12.1) is that the presence of the background charge—$+\mathbf{e}_0$ and $-\mathbf{e}_0$ at the two ends of the cylinder—lowers the free energy and with it the central charge.

Now using the dimension formula, Eq. (11.20), and inserting the values of the couplings g_α from Eq. (11.28), we arrive at

$$c = 3 - 6\left(\frac{e_{\rm b}^2}{1-e_{\rm b}} + \frac{e_{\rm g}^2}{1-e_{\rm g}}\right), \tag{12.4}$$

where we recall that $n_{\rm b} = 2\cos(\pi e_{\rm b})$ and similarly for $n_{\rm g}$. In Table 14.1 the numerically calculated values of the conformal charge are compared to the above formula, and excellent agreement is found.

12.2 Geometrical scaling dimensions

12.2.1 Two-string dimension

In addition to the central charge, the Coulomb gas representation of the loop model provided by the Liouville field theory, Eq. (11.2), allows for the evaluation of various geometrical scaling dimensions. As an example of such a quantity, consider the probability $G_2(r)$ that two points separated by a distance r lie on the same, say, black loop. In the critical phase we expect this probability to decay as $G_2(r) \sim r^{-2x_2}$, which defines the scaling dimension x_2. Since a black loop is represented as a sequence of alternating **A** and **B**-coloured edges it follows from the colouring constraint that the microscopic heights **z** just outside this loop differ by integer multiples of **C** and **D** only. In other words, a black loop is a *contour* loop for the component of the height along the direction perpendicular to both **C** and **D**, *i.e.*, the $(1,0,-1)$ direction in height space. Similarly the grey loops are contour loops for the height component along the $(1,0,1)$ direction.

It has been argued that the scaling dimension governing the probability that two points belong to the same contour loop of a random Gaussian surface equals $1/2$, independent of the stiffness [131]. Thus, for $(n_{\rm b}, n_{\rm g}) = (2,2)$ when $\mathbf{e}_0 = \mathbf{0}$ and the effective field theory is Gaussian, we expect $x_2 = 1/2$. For other values of the fugacities the Gaussian theory is modified by the background charge and the same argument cannot be made.

A more illuminating way of making contact with the interface representation is to view $G_2(r)$ as a two-string correlation function associated with defect configurations where two black strings emanating from the origin annihilate one another at a distant point **r**; see Fig. 12.1b. This can be accomplished by rewriting $G_2(r)$ as $Z(r)/Z$, where Z is the partition function defined by Eq. (10.6), and $Z(r)$ is similarly defined but with the summation restricted to those configurations $\mathcal{G}'_{\mathbf{r}}$ where an oriented black loop passes through the points **0** and **r**. Now consider reversing the direction of one half of the loop, so that instead of having one oriented loop passing through **0** and **r** we have two oriented loop segments directed from **0** to **r** [99]. This corresponds to the introduction of defect configurations at these two points, where we have violated the edge-colouring constraint. At **0** we find a (**C**, **D**, **A**, **A**) configuration of colours which

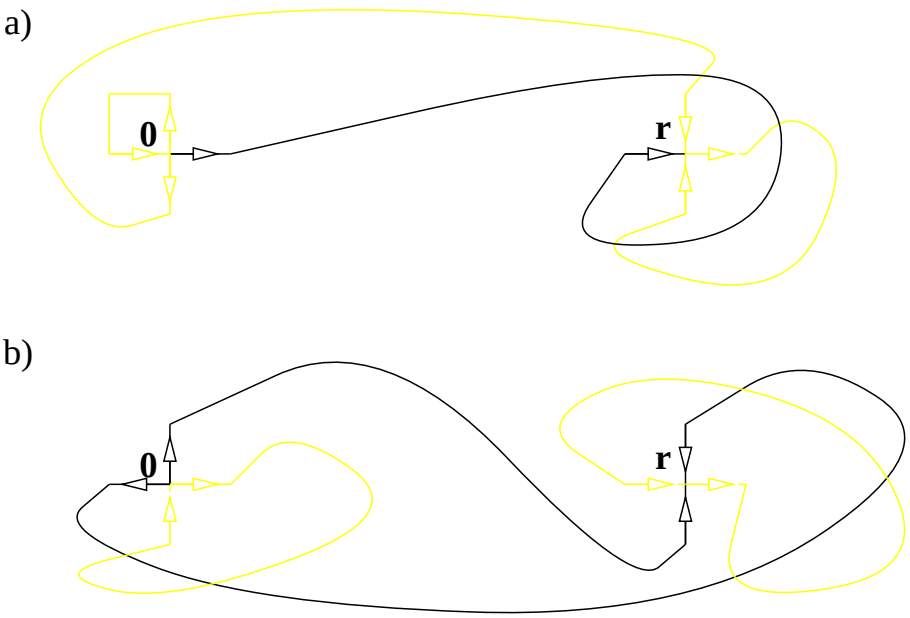

Figure 12.1: Defect configurations used to calculate the geometrical exponents x_1 (a) and x_2 (b) in the FPL^2 model. In (a) there is a single oriented black loop segment and a single oriented grey loop segment propagating from $\mathbf{0}$ to $\mathbf{r}$, whilst in (b) there are two oriented black loop segments between $\mathbf{0}$ and $\mathbf{r}$.

in the height language corresponds to a vortex of strength

$$\mathbf{m}_2 = \mathbf{A} - \mathbf{B} = (-2, 0, 2)\ . \tag{12.5}$$

The strength of the vortex (its Burgers charge) at $\mathbf{0}$ is calculated as the total height change around $\mathbf{0}$. Similarly, at $\mathbf{r}$ we have the corresponding antivortex $(\mathbf{B}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ of strength $-\mathbf{m}_2$ as illustrated in Fig. 12.1b.

In order to calculate x_2 for general values of the loop fugacities we have to take into account the effect of the complex phase factors associated with oriented loops. Namely, when one or more, say, black strings are associated with a vortex-antivortex configuration, spurious phase factors $\exp(\pm i\pi e_{\rm b})$ will arise whenever a black loop segment winds around one of the vortex cores [99]; for example, in Fig. 12.1b one of the two black strings winds once around the point $\mathbf{r}$. The spurious winding phase can be removed by inserting the vertex operator $\exp(\mathrm{i}\mathbf{e}_{\rm b} \cdot \mathbf{h})$ at the positions of both vortex cores. Since a black loop has alternating $\mathbf{A}$ and $\mathbf{B}$ colours the electric charge $\mathbf{e}_{\rm b}$ must satisfy

$$\begin{aligned} \mathbf{e}_{\rm b} \cdot \mathbf{A} &= +\pi e_{\rm b}, \qquad & \mathbf{e}_{\rm b} \cdot \mathbf{C} &= 0, \\ \mathbf{e}_{\rm b} \cdot \mathbf{B} &= -\pi e_{\rm b}, \qquad & \mathbf{e}_{\rm b} \cdot \mathbf{D} &= 0. \end{aligned} \tag{12.6}$$

Similarly, if there are grey strings propagating between two vertices the spurious phase factors associated with winding configurations are corrected with vertex operators whose electric charge $\mathbf{e}_{\rm g}$ is determined by

$$\begin{aligned} \mathbf{e}_{\rm g}\cdot\mathbf{A} &= 0, \qquad \mathbf{e}_{\rm g}\cdot\mathbf{C} = +\pi e_{\rm g}, \\ \mathbf{e}_{\rm g}\cdot\mathbf{B} &= 0, \qquad \mathbf{e}_{\rm g}\cdot\mathbf{D} = -\pi e_{\rm g}. \end{aligned} \tag{12.7}$$

Using Eq. (10.5) for the colour-vectors we find,

$$\mathbf{e}_{\rm b} = -\frac{\pi}{2}(e_{\rm b},0,-e_{\rm b}), \qquad \mathbf{e}_{\rm g} = -\frac{\pi}{2}(e_{\rm g},0,e_{\rm g})\ . \tag{12.8}$$

Going back to the two-string operator we conclude that it has total electromagnetic charge $(\mathbf{e}_{\rm b},\mathbf{m}_2)$.

Finally, from the general expression for the scaling dimension of an electromagnetic operator, Eq. (11.20), it follows that

$$2x_2 = 2x(\mathbf{e}_{\rm b},\mathbf{m}_2) = (1-e_{\rm b}) - \frac{e_{\rm b}^2}{1-e_{\rm b}}. \tag{12.9}$$

In Table 14.5 exact values of x_2 calculated from this formula are compared to numerical results, and excellent agreement is found.

Interestingly the expression for x_2 is *independent* of $e_{\rm g}$, *i.e.*, it is not affected by the fugacity of grey loops. This observation conforms to our understanding of the scaling of compact polymers. The compact polymer problem is recovered in the limit $n_{\rm b}\to 0$ in which case there is a single black loop on the lattice. Since the loop fills space its Hausdorff dimension is necessarily $D=2$. Scaling tells us that [132]

$$D = 2 - x_2 \tag{12.10}$$

from which the result $x_2 = 0$ follows, *independent* of the fugacity of grey loops. The fact that our formula reproduces this simple result in the $n_{\rm b}=0$ ($e_{\rm b}=1/2$) case provides a non-trivial check on its validity.

12.2.2 One-string dimension

The scaling dimension x_1 corresponding to one black and one grey string propagating between two points on the lattice, can be computed in a way that is completely analogous to the case of two black strings discussed above. (Note that the fully packing constraint ensures that if there is a single black string between two points then these points are also connected by a grey string; see Fig. 12.1a.) Choosing one point on the even sublattice and the other on the odd, leads to the appearance of the defect configuration $(\mathbf{A},\mathbf{C},\mathbf{C},\mathbf{D})$ on both sites of the square lattice. These in turn correspond to vortices in the height representation with topological charges $\pm\mathbf{m}_1$, where

$$\mathbf{m}_1 = \mathbf{C}-\mathbf{B} = (-2,-2,0)\ . \tag{12.11}$$

Since strings of both flavours are now present the compensating electric charge is $\mathbf{e}_\mathrm{b} + \mathbf{e}_\mathrm{g} = \mathbf{e}_0$. Hence

$$\begin{aligned} 2x_1 &= 2x(\mathbf{e}_0, \mathbf{m}_1) \\ &= \frac{1}{4}\left[(1-e_\mathrm{b}) + (1-e_\mathrm{g})\right] \\ &+ \frac{(1-e_\mathrm{b})(1-e_\mathrm{g})}{(1-e_\mathrm{b})+(1-e_\mathrm{g})} - \left[\frac{e_\mathrm{b}^2}{1-e_\mathrm{b}} + \frac{e_\mathrm{g}^2}{1-e_\mathrm{g}}\right]. \end{aligned} \tag{12.12}$$

There are of course several different ways of choosing the defect configurations (in this case, eight), but it should hardly come as a surprise that they all lead to the same expression for the scaling dimension.

Unlike x_2, x_1 depends on both loop fugacities. Going back to our original motivation, the compact polymer problem ($n_\mathrm{b} = 0 \Rightarrow e_\mathrm{b} = 1/2$), x_1 determines the value of the conformational exponent $\gamma = 1 - x_1$, which describes the scaling of the number of compact polymers with the number of monomers. We see that depending on e_g there will be a continuum of γ's. How do we interpret this?

First note that the problem of counting the number of conformations of a single compact polymer is the case $n_\mathrm{g} = 1$ ($e_\mathrm{g} = 1/3$) which simply assigns equal weights to all conformations. Using Eq. (12.12) this choice leads to $x_1 = -5/112$ and to the result $\gamma = 117/112$ advertised in the abstract. Changing n_g (e_g) away from $n_\mathrm{g} = 1$, on the other hand, has the effect of favouring certain compact polymer conformations over others depending on the number of loops formed by the uncovered (grey) bonds. In this sense the weight assigned to grey loops can be thought of as an interaction between the monomers of the compact polymer, albeit a peculiar non-local one. Nonetheless, it is interesting that this interaction changes the scaling properties of the compact polymer leading to a continuously varying exponent γ (more on this in Chapter 15).

12.2.3 Many-string dimensions

The dimensions x_1 and x_2 given above are contained in a more general set of string dimensions $x_{s_\mathrm{b},s_\mathrm{g}}$ governing the probability of having s_b black loop segments and s_g grey loop segments propagating between two points on the lattice [109]. More precisely, we consider two microscopic regions centred around points separated by a macroscopic distance, one region being the source and the other the sink of the oriented loop segments. Since the defect configurations obtained by violations of the edge colouring constraint must necessarily give rise to an *even* number of strings we will only consider the case when $s_\mathrm{b} + s_\mathrm{g}$ is even.

Consider first the case $s_\mathrm{b} = 2k_\mathrm{b}$ and $s_\mathrm{g} = 2k_\mathrm{g}$. The appropriate magnetic charge is obtained by combining k_b vortices with charge $\mathbf{A} - \mathbf{B} = (-2, 0, 2)$, and k_g vortices with charge $\mathbf{C} - \mathbf{D} = (-2, 0, -2)$. The defect with charge $\mathbf{A} - \mathbf{B}$ acts as a source of two black segments, whilst $\mathbf{C} - \mathbf{D}$ is associated with two grey loop segments. We also need to introduce the electric charge $\mathbf{e}_\mathrm{b} + \mathbf{e}_\mathrm{g}$ to compensate for the extra winding phase associated with the black and grey loop segments.

The total electromagnetic charge is therefore

$$[\mathbf{e}_{2k_\mathrm{b},2k_\mathrm{g}}, \mathbf{m}_{2k_\mathrm{b},2k_\mathrm{g}}] = \\ [\mathbf{e}_\mathrm{b}(1-\delta_{k_\mathrm{b},0}) + \mathbf{e}_\mathrm{g}(1-\delta_{k_\mathrm{g},0}), -2(k_\mathrm{b}+k_\mathrm{g}, 0, k_\mathrm{g}-k_\mathrm{b})], \tag{12.13}$$

and from the dimension formula, Eq. (11.20), we find

$$\begin{aligned} 2x_{2k_\mathrm{b},2k_\mathrm{g}} &= (1-e_\mathrm{b})k_\mathrm{b}^2 + (1-e_\mathrm{g})k_\mathrm{g}^2 \\ &- \left[\frac{e_\mathrm{b}^2}{1-e_\mathrm{b}}(1-\delta_{k_\mathrm{b},0}) + \frac{e_\mathrm{g}^2}{1-e_\mathrm{g}}(1-\delta_{k_\mathrm{g},0})\right]. \end{aligned} \tag{12.14}$$

This formula generalises Eq. (12.9).

Similarly, for $s_\mathrm{b} = 2k_\mathrm{b} - 1$ and $s_\mathrm{g} = 2k_\mathrm{g} - 1$ the electromagnetic charge is

$$[\mathbf{e}_{2k_\mathrm{b}-1,2k_\mathrm{g}-1}, \mathbf{m}_{2k_\mathrm{b}-1,2k_\mathrm{g}-1}] = [\mathbf{e}_0, -2(k_\mathrm{b}+k_\mathrm{g}-1, 1, k_\mathrm{g}-k_\mathrm{b})]; \tag{12.15}$$

the magnetic charge is obtained by combining $k_\mathrm{b} - 1$ defects of charge $\mathbf{A} - \mathbf{B}$, $k_\mathrm{g} - 1$ defects of charge $\mathbf{C} - \mathbf{D}$, and a single defect of charge $\mathbf{C} - \mathbf{B}$ which produces the remaining single black and grey strings originating from the same vertex. The scaling dimension is found to be

$$\begin{aligned} 2x_{2k_\mathrm{b}-1,2k_\mathrm{g}-1} &= \frac{1}{4}\left[(1-e_\mathrm{b})(2k_\mathrm{b}-1)^2 + (1-e_\mathrm{g})(2k_\mathrm{g}-1)^2\right] \\ &+ \frac{(1-e_\mathrm{b})(1-e_\mathrm{g})}{(1-e_\mathrm{b})+(1-e_\mathrm{g})} - \left[\frac{e_\mathrm{b}^2}{1-e_\mathrm{b}} + \frac{e_\mathrm{g}^2}{1-e_\mathrm{g}}\right]. \end{aligned} \tag{12.16}$$

This generalises the expression given in Ref. [89] and correctly reduces to (12.12) for $k_\mathrm{b}, k_\mathrm{g} = 1$.

12.2.4 Thermal dimension

We now turn our attention to the thermal scaling dimension. The FPL^2 model can be thought of as the zero-temperature limit of a more general model where we allow for thermal excitations that violate the close packing constraint; see Sect. 8.1.3. In this sense the temperature variable is thermodynamically conjugate to the constraint that every vertex be visited by (say) a black loop. An appropriate defect configuration for computing x_T within the FPL^2 model is therefore $(\mathbf{C}, \mathbf{D}, \mathbf{C}, \mathbf{D})$. This is a vortex of strength

$$\mathbf{m}_T = 2(\mathbf{C}+\mathbf{D}) = (0, -4, 0), \tag{12.17}$$

and since no strings terminating in the bulk are generated there is no compensating electric charge. The scaling dimension is then

$$2x_T = 2x(\mathbf{0}, \mathbf{m}_T) = 4\frac{(1-e_\mathrm{b})(1-e_\mathrm{g})}{(1-e_\mathrm{b})+(1-e_\mathrm{g})}\ . \tag{12.18}$$

The exact values of x_T quoted in Table 14.2 are calculated using this formula. The numerical results are in excellent agreement.

12.2.5 Boundary-string dimensions

The simplest example of a string operator that cannot be accessed within the formalism presented above is that of one black and no grey strings propagating between two vertices of the square lattice. Since this configuration has an odd number of strings connecting two sites of the lattice these two sites necessarily reside on the *boundary*.

If we define the FPL^2 model on the cylinder, as will be the case when we construct its transfer matrix in Sec. 13.1, a single black string can be enforced to run along the length of the cylinder if its circumference is chosen *odd*. Taking our cue from the formulae derived above for the bulk string operators we *guess* the formula

$$X = \frac{1}{8} + \frac{1-e_{\rm b}}{8} - \frac{1}{2}\frac{e_{\rm b}^2}{1-e_{\rm b}} \tag{12.19}$$

from the numerical results shown in Table 14.3. X is the scaling dimension of the boundary operator which corresponds to a single black string.

The Coulomb gas interpretation of the second and third term in Eq. (12.19) is rather apparent when one compares them to Eq. (12.9). The second term can be rationalised as coming from a magnetic charge $(-1,0,1)$ which is half the charge $\mathbf{m}_2$ in Eq. (12.5), associated with two black strings; this is saying that we have a partial dislocation generated at the boundary. The third term is due to the compensating electric charge $\mathbf{e}_{\rm b}$ for a single black string, same as in the two-string case.

The first, constant term does not have an immediate interpretation. A possible scenario is that it is due to the boundary condition imposed on the height by virtue of having a cylinder of odd circumference. Namely, a translation along the periodic coordinate by an amount equal to the circumference (L) exchanges an even site for an odd site (and *vice versa*) resulting in a transformation of the height: $\mathbf{h}(x^1,x^2) = \mathrm{P}\mathbf{h}(x^1+L,x^2)$. Since $\mathrm{P}^2=1$ this boundary condition can be thought of as an insertion of a twist operator into the partition function. The twist operator has dimension 1/8 regardless of the stiffness of the interface [133].

The above considerations permit us to calculate the scaling dimension for the general case of an odd number of strings. For definiteness we consider the case of $s_{\rm b} = 2k_{\rm b}-1$ and $s_{\rm g} = 2k_{\rm g}$. The magnetic charge pertaining to this situation is found by combining $2k_{\rm b}-1$ defects of charge $\frac{1}{2}(\mathbf{A}-\mathbf{B})$ with $k_{\rm g}$ defects of charge $\mathbf{C}-\mathbf{D}$, totaling

$$\begin{aligned}[\mathbf{e}_{2k_{\rm b}-1,2k_{\rm g}}, \mathbf{m}_{2k_{\rm b}-1,2k_{\rm g}}] = \\ [\mathbf{e}_{\rm b} + \mathbf{e}_{\rm g}(1-\delta_{k_{\rm g},0}), (1-2k_{\rm b}-2k_{\rm g}, 0, 2k_{\rm b}-2k_{\rm g}-1)].\end{aligned} \tag{12.20}$$

Taking into account the contribution from the twist operator, *i.e.*, adding 1/8

to the result obtained from Eq. (11.20), the scaling dimension is then

$$\begin{aligned} 2x_{2k_{\rm b}-1,2k_{\rm g}} &= \frac{1}{8}+\frac{1}{4}\left[(1-e_{\rm b})(2k_{\rm b}-1)^2+(1-e_{\rm g})(2k_{\rm g})^2\right] \\ &- \left[\frac{e_{\rm b}^2}{1-e_{\rm b}}+\frac{e_{\rm g}^2}{1-e_{\rm g}}(1-\delta_{k_{\rm g},0})\right]. \end{aligned} \tag{12.21}$$

12.2.6 Complete spectrum of string dimensions

Finally, the results of Eqs. (12.14), (12.16) and (12.21) can be combined into a single equation for the scaling dimension of a string operator that corresponds to $s_{\rm b}$ black loop segments and $s_{\rm g}$ grey loop segments:

$$\begin{aligned} 2x_{s_{\rm b},s_{\rm g}} &= \frac{1}{8}\delta^{(2)}_{s_{\rm b}+s_{\rm g},1}+\frac{1}{4}\left[(1-e_{\rm b})s_{\rm b}^2+(1-e_{\rm g})s_{\rm g}^2\right] \\ &- \left[\frac{e_{\rm b}^2}{1-e_{\rm b}}(1-\delta_{s_{\rm b},0})+\frac{e_{\rm g}^2}{1-e_{\rm g}}(1-\delta_{s_{\rm g},0})\right] \\ &+ \delta^{(2)}_{s_{\rm b},1}\delta^{(2)}_{s_{\rm g},1}\frac{(1-e_{\rm b})(1-e_{\rm g})}{(1-e_{\rm b})+(1-e_{\rm g})}. \end{aligned} \tag{12.22}$$

Here we have defined $\delta^{(2)}_{i,j}\equiv\delta_{i=j\,(\mathrm{mod}\ 2)}$.

12.3 Scaling dimensions of other local operators

We now turn to the determination of the scaling dimensions of the microscopic operators defined in Sect. 10.3.3. The cross-staggered operator $\mathbf{Q}(\mathbf{x})$ has already been discussed in Sect. 11.3 in connection with the loop ansatz. The electric charges corresponding to the four most relevant vertex operators in its Fourier expansion turned out to play the role of the screening charges of the field theory. Consequently $\mathbf{Q}(\mathbf{x})$ is exactly marginal throughout the critical region:

$$2x_{\mathbf{Q}}\equiv 4. \tag{12.23}$$

The staggered spin $\mathbf{S}(\mathbf{x})$ was shown in Sect. 10.3.4 to have the eight vectors of type $\frac{\pi}{2}(1,1,1)$ as it shortest height periods. Explicit insertion in the dimension formula (11.20) reveals that four of these vectors minimise the scaling dimension:

$$\begin{aligned} \mathbf{e}_{\mathbf{S}}^{(1)} &= \frac{\pi}{2}(-1,+1,+1), & \mathbf{e}_{\mathbf{S}}^{(2)} &= \frac{\pi}{2}(-1,-1,+1), \\ \mathbf{e}_{\mathbf{S}}^{(3)} &= \frac{\pi}{2}(-1,+1,-1), & \mathbf{e}_{\mathbf{S}}^{(4)} &= \frac{\pi}{2}(-1,-1,-1). \end{aligned} \tag{12.24}$$

Their dimensions are

$$\begin{aligned} 2x(\mathbf{e}_{\mathbf{S}}^{(1)},\mathbf{0}) = 2x(\mathbf{e}_{\mathbf{S}}^{(2)},\mathbf{0}) &= \frac{1-2e_{\rm b}}{1-e_{\rm b}}+\frac{1}{4}\frac{2-e_{\rm b}-e_{\rm g}}{(1-e_{\rm b})(1-e_{\rm g})}, \\ 2x(\mathbf{e}_{\mathbf{S}}^{(3)},\mathbf{0}) = 2x(\mathbf{e}_{\mathbf{S}}^{(4)},\mathbf{0}) &= \frac{1-2e_{\rm g}}{1-e_{\rm g}}+\frac{1}{4}\frac{2-e_{\rm b}-e_{\rm g}}{(1-e_{\rm b})(1-e_{\rm g})}, \end{aligned} \tag{12.25}$$

and it is seen that $\mathbf{e}_{\mathbf{S}}^{(1)}$ and $\mathbf{e}_{\mathbf{S}}^{(2)}$ have the smallest dimension when $e_{\rm b} > e_{\rm g}$ and *vice versa*. Defining $e_> = \max\{e_{\rm b}, e_{\rm g}\}$ and $e_< = \min\{e_{\rm b}, e_{\rm g}\}$ the resulting scaling dimension for the staggered spin can be written

$$2x_{\mathbf{S}} = \frac{1-2e_>}{1-e_>} + \frac{1}{4}\frac{2-e_>-e_<}{(1-e_>)(1-e_<)}. \tag{12.26}$$

It remains to determine the dimension of the parity operator $P(\mathbf{x})$, which coincides with that of the row-staggered spin $\mathbf{R}(\mathbf{x})$. Of its six shortest $\pi(1,0,0)$ type electric charges the one generating the most relevant vertex operator turns out to be $\pi(-1,0,0)$. The corresponding scaling dimension is

$$2x_P = 2x_{\mathbf{R}} = \frac{1-2e_{\rm b}}{1-e_{\rm b}} + \frac{1-2e_{\rm g}}{1-e_{\rm g}}. \tag{12.27}$$

For the four-colouring model ($\mathbf{e}_0 = \mathbf{0}$) Eqs. (12.26) and (12.27) correctly reduce to the values previously found by Kondev and Henley [93]: $x_{\mathbf{S}} = 3/4$ and $x_P = x_{\mathbf{R}} = 1$.

12.4 Termination of critical behaviour

In the preceding chapters we have developed an effective description of the critical phase of the FPL^2 model in the form of a field theory. This theory has to break down at large values of the loop fugacity since in this case a typical state of the model will consist of small loops only, *i.e.*, a power-law distribution of loop sizes will be absent. That this indeed happens can be seen from the Liouville field theory itself as it carries the seeds of its own demise.

The mapping of the loop model to an oriented loop model for $n_{\rm b}, n_{\rm g} \leq 2$ works equally well for $n_{\rm b} > 2$ or $n_{\rm g} > 2$. From Eq. (10.3) it follows that in the latter case at least one of the parameters, $e_{\rm b}$ or $e_{\rm g}$, will be pure imaginary. This affects the Liouville potential which for $n_{\rm b} > 2$ or $n_{\rm g} > 2$ becomes a *relevant* perturbation to the (modified) Gaussian action $S_{\rm E} + S_{\rm B}$.

To understand how this comes about we consider the simple case provided by the $n_{\rm b} = n_{\rm g}$ FPL^2 model, discussed in Sect. 11.3.2. Namely, as we increase the value of the loop fugacity we expect small loops to be favoured and the stiffness K of the interface to grow. In the critical phase this is offset by the decrease in the background charge in a way that leaves the Liouville potential marginal. Now when the loop fugacity exceeds 2 the background charge $\mathbf{e}_0 = -\pi(e_{\rm b}, 0, 0)$ becomes pure imaginary and the dimension of the Liouville potential

$$x_{\rm L} = \frac{\pi}{2}\,\frac{1-e_{\rm b}}{K} \tag{12.28}$$

can no longer stay marginal; here $x_{\rm L} \equiv x(\mathbf{e}^{(i)}, 0)$, where the charges $\mathbf{e}^{(i)}$ are given in Eq. (11.25), and their scaling dimensions are calculated from Eq. (11.24). In fact, assuming that the stiffness K continues to increase with the loop fugacity

for $n_{\rm b} = n_{\rm g} > 2$, x_L turns complex with a real part that is *smaller* than two, rendering the Liouville potential *relevant*.

If we make the usual assumption of no intervening fixed points, the relevant Liouville potential will generate a finite correlation length and the loop model will no longer be critical. The correlation length has the physical interpretation of the average size of a loop in the system. This scenario has been confirmed for the fully packed loop model on the honeycomb lattice, from the Bethe ansatz solution of this model [134].

A different view of the non-critical region of the FPL^2 model is provided by the locking potential $V(\mathbf{h})$. Namely, the discrete nature of the microscopic heights can be taken into account in the field theory by a negative potential in height space that is peaked around the flat, ideal states. As such, this potential is uniform on the ideal state graph and can therefore be expanded in a Fourier series. Examination of the most relevant vertex operators in this series [93] reveals that they are the same as the ones for the loop-weight (Liouville) potential, $w(\mathbf{h})$. Therefore, just like $w(\mathbf{h})$, the locking potential in the non-critical region of the phase diagram is a *relevant* perturbation. Thus, it will lock the interface in one of the ideal, flat states. In this smooth phase the height fluctuations are bounded (as opposed to being logarithmically divergent) which is just another way of saying that large contour loops are exponentially suppressed. On the other hand, in the critical region of the FPL^2 model the locking potential is marginal as it would be for an interface model *at* the roughening transition [135, 136]. This might indicate that the whole critical region of the FPL^2 model can be understood as a manifold of essential singularities in some more general model, as was the case for the honeycomb FPL model [92, 137].

Resumo en Esperanto

Ni kalkulas la centran ŝargon c kaj diversajn skalumajn dimensiojn por la buklomodelo. Tri liberaj bosonaj kampoj donas $c = 3$, sed tiun rezulton ŝanĝas la fona elektra ŝargo (12.4). Vorticaj difektoj en la interfaca reprezentiĝo generas direktitajn kordojn (Fig-o 12.1). Kiam la difekta ŝargo (vidu ekz-e 12.5) estas konata eblas kalkuli la korespondan geometrian skaluman dimension. Ekv-o (12.22) estas la ĝenerala rezulto por $s_{\rm b}$ nigraj kaj $s_{\rm g}$ grizaj kordoj. La termika dimensio (12.18) korespondas al malobeo de la kondiĉo ke ĉiu vertico estu vizitata de kaj nigra kaj griza buklo.

Kiam unu el la buklopezoj superas du la Liouville potencialo fariĝas modifiva perturbo. Ĝi generas finian longon de korelacio kaj igas la interfacon glata. Aliflanke, en la tuta krita regiono la interfaca modelo estas *ĉe* aspriĝa transiro.

Chapter XIII

Loop model transfer matrices

Estimants of the central charge and a number of the geometrical scaling dimensions obtained from transfer matrix calculations provide the conclusive evidence in favour of the analytical results presented in Sects. 12.1 and 12.2. Before turning to a discussion of our numerical results, which is the subject of Chapter 14, we describe the particular representation of the transfer matrix used to obtain them.

13.1 Construction of the FPL^2 transfer matrix

To construct the transfer matrix for the FPL^2 model on a cylinder of circumference L we write the partition function as

$$Z^{(M)} = \sum_{\mathcal{G}_M} n_{\mathrm{b}}^{N_{\mathrm{b}}} n_{\mathrm{g}}^{N_{\mathrm{g}}}, \tag{13.1}$$

where the length of the cylinder M has been explicitly indicated. Periodic boundary conditions are imposed in the horizontal direction, whereas the bottom and the top row of the cylinder have open boundary conditions and hence terminate in L dangling edges. We recall that the restriction of the summation to the set of fully-packed graph configurations $\mathcal{G}_M$ implies that locally the vertices are constrained to have one of the six appearances shown in Fig. 10.1. In the first four possible vertices the loop segments do not cross, whilst in the last two vertices the two flavours intersect. The global constraint that all loops be closed in the limit of an infinite system means that loop segments cannot terminate in the bulk but only at the dangling edges in the top and bottom rows.

A typical loop configuration for a cylinder with $L = 6$ and $M = 12$ is shown in Fig. 13.1. The horizontal numbering pertains to the vertices, whilst in the vertical direction it is more convenient to label each row by the number of the vertex immediately below it. Accordingly the labels 0 and M refer to the bottom and the top row of dangling edges respectively. We shall soon see that the inclusion in $\mathcal{G}_M$ of one or more strings running between the dangling edges of row

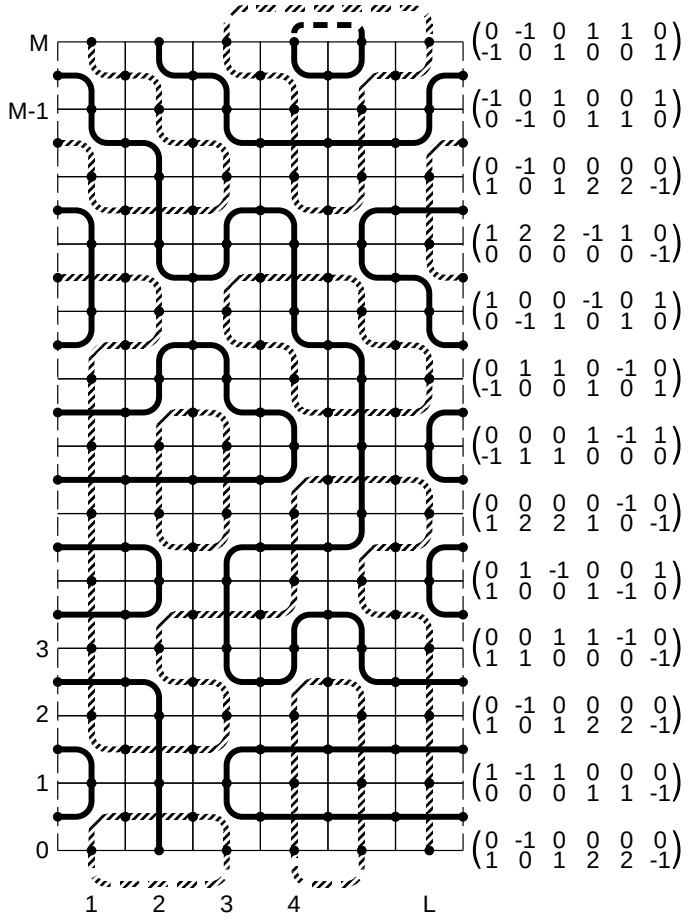

Figure 13.1: A typical loop configuration for $L = 6$ and $M = 12$. The dashed lines along the left and the right boundaries illustrate the periodic boundary conditions. Horizontally the vertices are numbered from 1 to L, whilst vertically the rows are labeled by the number of the vertex immediately below them. This particular configuration is constrained to having precisely one string of each flavour spanning the length of the cylinder, and hence it contributes to the geometrical exponent x_1. To the right we show the index representation of the connectivity state pertaining to each row (see Sect. 13.4 for details). Any valid configuration can be interpreted as a "jigsaw puzzle" assembled from the six "pieces" shown in Fig. 10.1. Note that when laying down the first row of this puzzle it must be stipulated how the dangling edges of row 0, which are not part of a string, are pairwise interconnected below that row. These implicit connections as well as their counterparts in row M have been depicted by dashed loop segments.

0 and M helps us access the geometrical exponents of the model. In particular, the configuration of Fig. 13.1 having one such string of each flavour furnishes a contribution to the scaling dimension x_1 which determines the conformational exponent $\gamma = 1 - x_1$.

13.2 Connectivity basis

The construction of a transfer matrix (TM) for Eq. (13.1) appears to be obstructed by the non-locality of N_i ($i = \mathrm{b}, \mathrm{g}$). The key to solving this problem is to write the TMs in a basis of *connectivity states* comprising information about how the dangling ends of row M are pairwise interconnected in the preceding rows and, if strings are present, information about the positions of such strings. In addition the connectivity states must keep track of the particular flavour of any loop or string segment terminating in row M. Our construction generalises the work of Blöte and Nightingale for the Q-state Potts model [44, 46] and that of Blöte and Nienhuis for the O(n) model [54] to take the extra flavour information into account, and our notation is consistent with that of these authors.

As mentioned in Chapter 5 it is essential to be able to represent a given connectivity state both in an *index representation* giving direct access to the flavour and connectedness information just mentioned, and in a *number representation* assigning an integer in the range $1, 2, \ldots, C_L^{(s_\mathrm{b}, s_\mathrm{g})}$ to the state under consideration. The latter representation enables us to enumerate the entries of the TM, whilst the former allows us to determine the number of loop closures when going from one connectivity state to another and hence the value of a particular entry in the TM. Here $C_L^{(s_\mathrm{b}, s_\mathrm{g})}$ is the number of distinct connectivity states for a cylinder of width L accommodating s_i strings of flavour $i = \mathrm{b}, \mathrm{g}$. The construction of these two representations, the mapping between them, and the evaluation of the $C_L^{(s_\mathrm{b}, s_\mathrm{g})}$ for $(s_\mathrm{b}, s_\mathrm{g}) = (0,0)$, (1,0), (1,1) and (2,0) is deferred to Sect. 13.4.

Designating the connectivity states by Greek letters we can write the partition function as a sum of restricted partition functions

$$Z^{(M)} = \sum_\beta Z_\beta^{(M)} = \sum_\beta \sum_{\mathcal{G}_M} \delta\big(\beta, \phi(\mathcal{G}_M)\big)\, n_\mathrm{b}^{N_\mathrm{b}} n_\mathrm{g}^{N_\mathrm{g}}, \tag{13.2}$$

where $\phi(\mathcal{G}_M)$ is the connectivity of the L dangling edges of row M, and $\delta(i,j)$ is the Kronecker delta. Now consider appending another row to the cylinder, giving us a total graph configuration $\mathcal{G}_{M+1} = \mathcal{G}_M \cup \mathcal{G}'$. Evidently the connectivity of the dangling edges of row $M+1$ is determined solely by that of the preceding row and by the appended subgraph $\mathcal{G}'$

$$\phi(\mathcal{G}_{M+1}) = \psi\big(\phi(\mathcal{G}_M), \mathcal{G}'\big). \tag{13.3}$$

Letting N_i' denote the number of loop closures induced by $\mathcal{G}'$ we arrive at the

relation

$$\begin{aligned} Z_\alpha^{(M+1)} &= \sum_{\mathcal{G}_{M+1}} \delta\big(\alpha, \phi(\mathcal{G}_{M+1})\big) n_{\rm b}^{N_{\rm b}+N'_{\rm b}} n_{\rm g}^{N_{\rm g}+N'_{\rm g}} \\ &= \sum_\beta \sum_{\mathcal{G}_M} \delta\big(\beta, \phi(\mathcal{G}_M)\big) n_{\rm b}^{N_{\rm b}} n_{\rm g}^{N_{\rm g}} \sum_{\mathcal{G}'|\mathcal{G}_M} \delta\big(\alpha, \psi(\phi(\mathcal{G}_M), \mathcal{G}')\big) n_{\rm b}^{N'_{\rm b}} n_{\rm g}^{N'_{\rm g}} \\ &= \sum_\beta T_{\alpha\beta} Z_\beta^{(M)}, \end{aligned} \tag{13.4}$$

where the transfer matrix is defined by

$$T_{\alpha\beta} = \sum_{\mathcal{G}'|\mathcal{G}_M} \delta\big(\alpha, \psi(\phi(\mathcal{G}_M), \mathcal{G}')\big) n_{\rm b}^{N'_{\rm b}} n_{\rm g}^{N'_{\rm g}}. \tag{13.5}$$

The notation $\mathcal{G}'|\mathcal{G}_M$ means that the summation is constrained to those subgraphs $\mathcal{G}'$ that fit the dangling edges of $\mathcal{G}_M$.

13.3 Single-vertex decomposition

A quintessential step in the practical implementation of the TM is its decomposition into matrices each corresponding to the addition of a single vertex,

$$\mathbf{T} = \mathbf{T}_L \cdot \mathbf{T}_{L-1} \cdot \ldots \cdot \mathbf{T}_1. \tag{13.6}$$

Here the single-vertex matrix $\mathbf{T}_i$, which adds the vertex at horizontal position i of the new row, has the advantage of being *sparse*, and we shall soon see that it has at most three non-zero entries per column. This property leads to a dramatical reduction of the time and storage requirements for the calculations.

13.3.1 First vertex in a new row

As was the case in the O(n) model [54], a minor complication arises due to the fact that the addition of the first vertex of a new row increases the number of dangling edges from L to $L+2$. This is illustrated in the left part of Fig. 13.2. Upon addition of further vertices the number of dangling edges is kept fixed at $L+2$, until the L'th vertex completes the row, and we are back at L dangling edges. Thus the dimensions of the single-vertex matrices are $C_{L+2} \times C_L$ for $\mathbf{T}_1$, $C_{L+2} \times C_{L+2}$ for $\mathbf{T}_2, \ldots, \mathbf{T}_{L-1}$, and $C_L \times C_{L+2}$ for $\mathbf{T}_L$.

In Fig. 13.2 we illustrate the action of $\mathbf{T}_1$ on $Z^{(M)}$ in detail. To ensure that row $M+1$, when completed, will have the same labels on its dangling edges as was the case in the preceding row, the solid dots illustrating the "active" dangling edges must be relabeled as shown in the lower left part of the figure. Shown to the right are the three possible choices of vertices fitting onto a black loop segment terminating at the dangling end 1 of $Z^{(M)}$. There are thus three

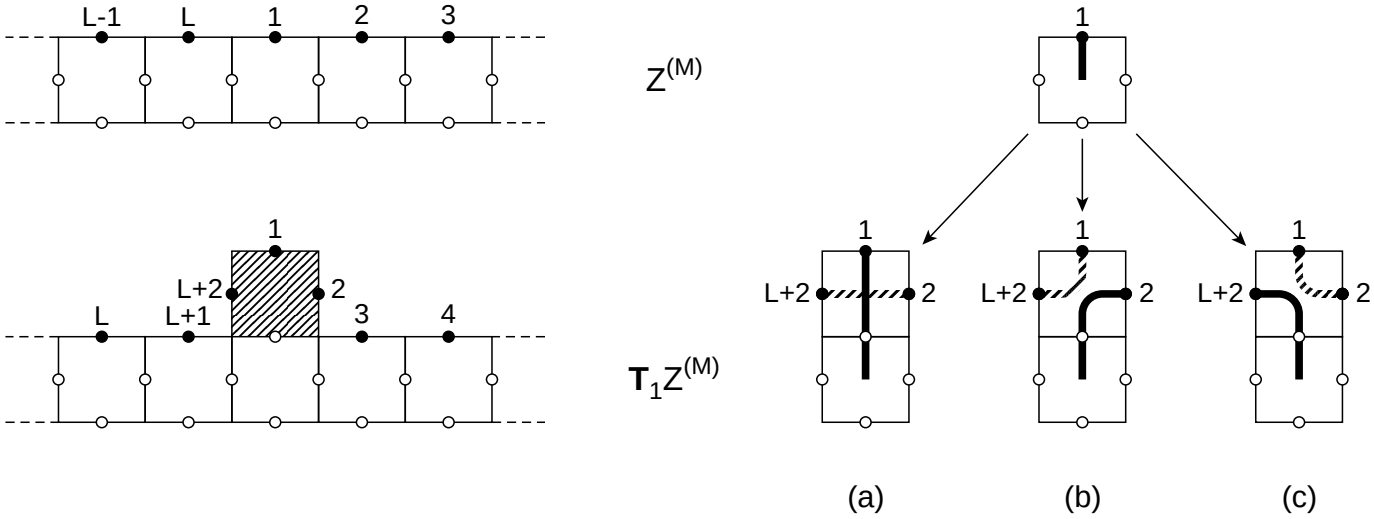

Figure 13.2: Adding the first vertex of the $(M+1)$'th row increases the number of dangling edges from L to $L+2$. The labeling of the "active" edges (filled circles) before and after addition of the new vertex (shaded) is as shown in the left part of the figure. The part of the lattice relevant for determining which of the vertices of Fig. 10.1 fit onto a given connectivity of row M, has been depicted in the right part of the figure. This information constitutes the *vertex rules*, and is explained in the text.

non-zero entries in each column of $\mathbf{T}_1$. Since no loop closures of either flavour can be induced ($N'_{\rm b} = N'_{\rm g} = 0$ in Eq. (13.5)) all these entries are unity. Similar considerations hold true when the loop segment to be fitted is grey, and the vertex rules can be read off from the figure by interchanging the two flavours.

13.3.2 Adding subsequent vertices

When acting with any one of the subsequent single-vertex TMs $\mathbf{T}_2, \ldots, \mathbf{T}_{L-1}$ the situation is as depicted in Fig. 13.3 for the case of $\mathbf{T}_2$. As the number of dangling edges is kept fixed no relabeling is needed, apart from the translation of labels 2 and 3 up on top of the newly added vertex. The vertex rules for the case where edge 2 of $\mathbf{T}_1 Z^{(M)}$ is black are shown in the right part of the figure; similar rules for the case where it is grey can be obtained by permuting the two flavours.

In situation (a) only one vertex fits onto the two dangling edges. The column of $\mathbf{T}_2$ determined by the number representation of the connectivity pertaining to the $L+2$ dangling ends that are active in the upper part of the figure thus has only one non-zero entry. Its value is either $n_{\rm b}$ or 1 depending on whether a black loop closure is induced ($i_2^{\rm b} = i_3^{\rm b}$) or not ($i_2^{\rm b} \neq i_3^{\rm b}$). In the index representation of the new connectivity state $i_2^{\rm g} = i_3^{\rm g}$ is set equal to a positive integer not assumed by any other $i_k^{\rm g}$. The new values of the black indices depend on whether a loop closure is induced or not. In the former case we simply set $i_2^{\rm b} = i_3^{\rm b} = 0$. In the latter, the two left-over black partners must be mutually connected before assigning $i_2^{\rm b} = i_3^{\rm b} = 0$.

Situations (b) and (c) correspond to two entries of each column of $\mathbf{T}_2$ taking the value unity, the others being zero. Since loop closures are out of the question the handling of these cases is simple. In (b) the two flavours cross, and the indices of sites 2 and 3 are interchanged. Case (c) is even simpler: it corresponds to a diagonal entry in $\mathbf{T}_2$.

When strings are present a few modifications of the above rules are necessary. In situation (a), if one of $i_2^{\rm b}$ and $i_3^{\rm b}$ equals -1 and the other is positive, the left-over partner to the non-string black segment must be made the new string. And if both $i_2^{\rm b}$ and $i_3^{\rm b}$ equal -1 the corresponding entry of $\mathbf{T}_2$ must be forced to zero, since two strings cannot be allowed to annihilate.

13.3.3 Completing the row

Finally, consider closing the $(M+1)$'th row through the action of $\mathbf{T}_L$, as depicted in Fig. 13.4. The labels $L+1$ and $L+2$ now disappear, and as far as the labeling goes the system is back in its original state. Each column of $\mathbf{T}_L$ has at most one non-zero entry per column, as witnessed by the vertex rules displayed in the right part of the figure. Once again, only half of the vertex rules are shown, and the other half is found by interchanging the two flavours.

In situation (a) no vertex of Fig. 10.1 can fit onto the three dangling edges at positions L, $L+1$ and $L+2$. The corresponding entry of $\mathbf{T}_L$ must therefore

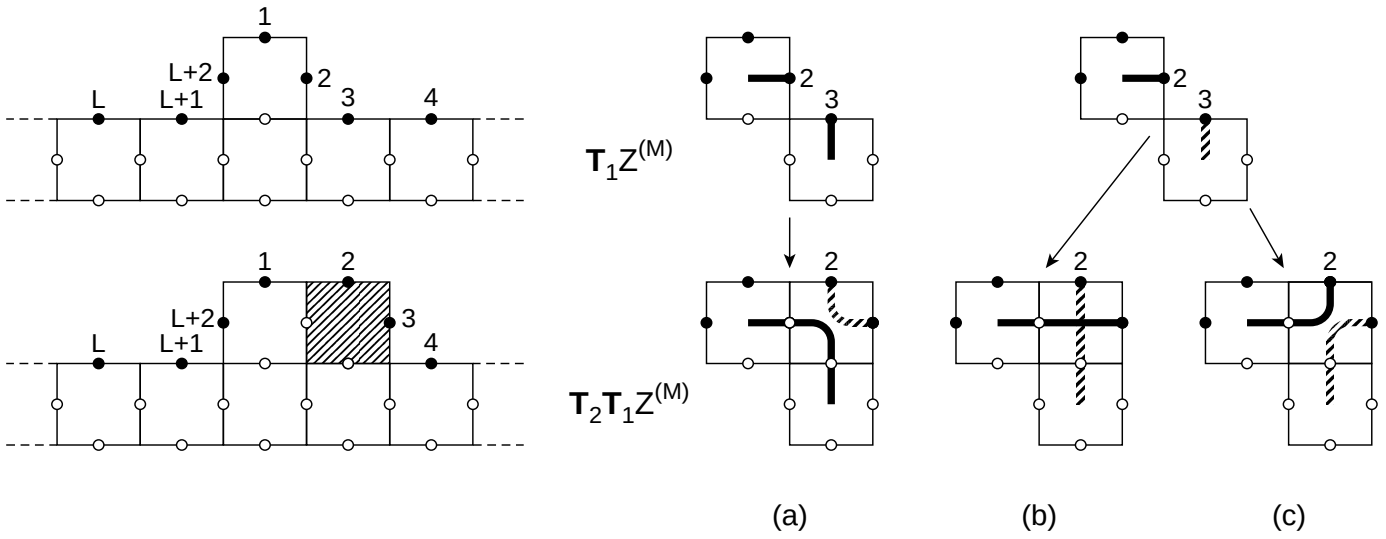

Figure 13.3: Addition of subsequent vertices keeps the number of dangling edges fixed at $L+2$. In the left part of the figure the system is shown before and after the addition of the second vertex (shaded). Vertex rules are displayed to the right. Situation (a) allows for the possibility of a black loop closure.

be forced to zero. Situations (b), (c) and (d) leave us to determine whether, for a given connectivity of the $L+2$ dangling edges, a black loop closure occurs or not. The handling in terms of the index representation is exactly as described above.

13.4 Enumeration of the connectivities

The implementation of the transfer matrix (TM) for the FPL^2 model on a cylinder of width L and length M requires an enumeration of the possible connectivity states of the L points on the dangling edges of row M. Each of these L points can either

1. be connected by $\mathcal{G}_M$ to one of the dangling edges of row 0 through a string of flavour $i = \mathrm{b}, \mathrm{g}$, or

2. be connected by $\mathcal{G}_M$ to one and only one other point in row M through a loop segment of flavour $i = \mathrm{b}, \mathrm{g}$.

A suitable representation of this information is furnished by a double state vector

$$\begin{pmatrix} i_1^{\mathrm{b}} i_2^{\mathrm{b}} i_3^{\mathrm{b}} \ldots i_L^{\mathrm{b}} \\ i_1^{\mathrm{g}} i_2^{\mathrm{g}} i_3^{\mathrm{g}} \ldots i_L^{\mathrm{g}} \end{pmatrix}, \tag{13.7}$$

which we refer to as the *index representation*. The indices i_k^{b} ($k = 1, 2, \ldots, L$) are defined as follows:

1. $i_k^{\mathrm{b}} = i_l^{\mathrm{b}}$ is a (non-unique) positive integer if and only if points k and l are interconnected through a black string.

2. $i_k^{\mathrm{b}} = 0$ if and only if point k touches a grey string or loop segment.

3. $i_k^{\mathrm{b}} = -1$ if and only if point k is connected to a dangling edge of row 0 through a black string.

A similar definition is true for the indices i_k^{g} provided that one reads "grey" instead of "black" and *vice versa*. Two index representations are said to be identical if they are so up to the arbitrariness of the choice of positive integers. Also note that if $i_k^{\mathrm{b}} \neq 0$ we have $i_k^{\mathrm{g}} = 0$ and conversely.

A restriction on those indices that take positive values follows from the fact that loops of the same flavour are not allowed to intersect. Namely, if $j < k < l < m$ the equalities $i_j^{\mathrm{b}} = i_l^{\mathrm{b}}$ and $i_k^{\mathrm{b}} = i_m^{\mathrm{b}}$ cannot both be true. So in addition to being pairwise these connectivities are also *well-nested* [44].[1] The same is true for the grey indices, whereas there are no such restrictions when

[1] Because of the reappearance of well-nestedness this section necessarily has some overlap with the discussion of the Potts model transfer matrices in Chapter 5. However, we have preferred to write this section so that it can be read independently.

both flavours are involved. Indeed, connectivity states with $i^{\rm b}_j = i^{\rm b}_l$ and $i^{\rm g}_k = i^{\rm g}_m$ are explicitly allowed by the last two vertices shown in Fig. 10.1.

In practice we are only interested in the first few eigenvalues of TMs having a definite number of strings of each flavour. The relevant sectors of the TM are denoted $\mathbf{T}^{(s_{\rm b},s_{\rm g})}$, where s_i is the number of strings of flavour $i = {\rm b}, {\rm g}$. The fully packing constraint means that we can only examine system sizes L that have the same parity as $s_{\rm b}+s_{\rm g}$. The various sectors have different physical interpretations and each requires a different enumeration of the connectivity states. Since the two flavours enter on an equal footing in the partition function, Eq. (13.1), we only need consider $s_{\rm b} \geq s_{\rm g}$. The $\mathbf{T}^{(0,0)}$ sector contains information about the free energy and the energy-like correlation length. The geometrical scaling dimensions x_1 and x_2 can be obtained from the $\mathbf{T}^{(1,1)}$ and the $\mathbf{T}^{(2,0)}$ sectors respectively. Finally the sector $\mathbf{T}^{(1,0)}$ gives the scaling dimension of the twist-like operator.

Whilst the index representation contains all information necessary for determining the value of a given entry in the TM it is obviously not suitable for labeling the entries. We therefore need another representation, the so-called *number representation*, in which the connectivities are labeled by the integers $1, 2, \ldots, C_L^{(s_{\rm b},s_{\rm g})}$, where $C_L^{(s_{\rm b},s_{\rm g})}$ is the number of different connectivity states in the relevant sector. The practical implementation of the TMs relies on the mapping from the index to the number representation and its inverse.

We shall now consider, one by one, the various sectors of the TM.

13.4.1 $\mathbf{T}^{(0,0)}$ sector

When no strings are present all the L dangling edges of row M are pairwise connected with either a black or a grey loop segment. In particular L must be even. For any particular connectivity we can then decompose L as $L = 2p_{\rm b}+2p_{\rm g}$, where $p_i \geq 0$ is the number of *pairs* of dangling edges covered by a flavour i loop segment. Since loops of different flavours are allowed to cross (see Fig. 10.1) the total number of connectivities is

$$C_L^{(0,0)} = \sum_{L=2p_{\rm b}+2p_{\rm g}} \binom{L}{2p_{\rm b}} c_{p_{\rm b}} c_{p_{\rm g}}, \tag{13.8}$$

where c_p is the number of pairwise well-nested connectivities of $2p$ points. The c_p's were first considered in the context of the Potts model [44], but were also found to play a central role in the TM formulation of the O(n) model [54]. We shall now briefly recall how they are evaluated.

Consider a well-nested pairwise connectivity of $2p$ points given by the index representation $(i_1 i_2 \ldots i_{2p})$. A recursion relation follows from observing that $i_1 = i_{2k}$ for precisely one integer $k \geq 1$. According to the well-nestedness criterion the sub-sequences $(i_2 i_3 \ldots i_{2k-1})$ and $(i_{2k+1} i_{2k+2} \ldots i_{2p})$ are both well-nested, and indices occurring in one of them do not occur in the other. Hence

L	4^L	$C_L^{(0,0)}$	$C_L^{(1,1)}$	$\tilde{C}_L^{(2,0)}$	L	$C_L^{(1,0)}$
2	16	2	2	1	1	1
4	256	10	24	12	3	6
6	4,096	70	300	150	5	50
8	65,536	588	3,920	1,960	7	490
10	1,048,576	5,544	52,920	26,460	9	5,292
12	16,777,216	56,628	731,808	365,904	11	60,984
14	268,435,456	613,470	10,306,296	5,153,148	13	736,164
16	4,294,967,296	6,952,660	147,232,800	73,616,400	15	9,202,050

Table 13.1: The number $C_L^{(s_\mathrm{b},s_\mathrm{g})}$ of FPL^2 connectivity states for L dangling edges accommodating s_i strings of flavour $i = \mathrm{b}, \mathrm{g}$. Only values of L with the same parity as $s_\mathrm{b} + s_\mathrm{g}$ are shown. When more than one string of any flavour is present further restrictions than the well-nestedness criterion apply, as described in the text. Accordingly the number $\tilde{C}_L^{(2,0)}$ is merely a useful upper limit on the true $C_L^{(2,0)}$. The efficiency of writing the TMs in the connectivity basis can be appreciated by comparing $C_L^{(0,0)}$ to 4^L, the latter being the dimensions of the TM written in the conventional colour basis, where every dangling end is labeled independently by **A**, **B**, **C** or **D**.

for $p \geq 1$

$$c_p = \sum_{k=1}^{p} c_{k-1} c_{p-k}, \tag{13.9}$$

and $c_0 = 1$. By means of the generating function $P(x) = \sum_{p=0}^{\infty} c_p x^p$ it is readily shown [54] that

$$c_p = \frac{(2p)!}{p!(p+1)!}, \tag{13.10}$$

and that asymptotically $c_p \sim 4^p$.

Using Eqs. (13.8) and (13.10) we can now compute explicit values for the $C_L^{(0,0)}$. These are shown for $2 \leq L \leq 16$ in Table 13.1.

For obvious reasons we shall call the function

$$\rho(i_1 i_2 \ldots i_{2p}) = k \tag{13.11}$$

defined by $i_1 = i_{2k}$ the *cut function* of the index representation $(i_1 i_2 \ldots i_{2p})$. A complete ordering of the well-nested sequences is now induced by applying the cut function first to the whole sequence, then recursively to its right and finally to its left part [44, 54]. Accordingly, the mapping from the index to the number

representation for a well-nested one-flavour connectivity is accomplished by

$$\sigma(i_1 i_2 \dots i_{2p}) = \begin{cases} 1 \text{ if } p \leq 1 \\ \sum_{l=1}^{k-1} c_{l-1} c_{p-l} + \sigma(i_2 \dots i_{2k-1}) \\ +[\sigma(i_{2k+1} \dots i_{2p}) - 1] c_{k-1} \text{ otherwise,} \end{cases} \tag{13.12}$$

where the c_p are given by Eq. (13.10).

To give a complete specification of the connectivity of any *one* flavour in the state (13.7) we need to keep track of the positions of those indices that are zero. For a fixed number of z zero indices this is accomplished by the lexicographic ordering

$$\psi(i_1 i_2 \dots i_L) = \begin{cases} 1 \text{ if } L = 1 \text{ or } z = L \\ \psi(i_2 i_3 \dots i_L) \text{ if } i_1 \neq 0 \\ \binom{L-1}{z} + \psi(i_2 i_3 \dots i_L) \text{ if } i_1 = 0, \end{cases} \tag{13.13}$$

assigning the lowest value to the sequence with all the zeros accumulated to the right.

The number representation of the two-flavour state (13.7) is now obtained by first ordering according to the number of indices i_k^{b} being zero, then lexicographically ordering the positions of these zero indices, and finally using the ordering (13.12), first on the well-nested subsequence of non-zero black indices and then on the corresponding grey subsequence. More precisely, the mapping from the index to the number representation in the $(s_{\mathrm{b}}, s_{\mathrm{g}}) = (0,0)$ sector is given by

$$\begin{aligned} \phi^{(0,0)} \begin{pmatrix} \mathbf{i}^{\mathrm{b}} \\ \mathbf{i}^{\mathrm{g}} \end{pmatrix} &= \sum_{k=p_{\mathrm{b}}+1}^{L/2} \binom{L}{2k} c_k c_{L/2-k} \\ &+ [\psi(\mathbf{i}^{\mathrm{b}}) - 1] c_{p_{\mathrm{b}}} c_{p_{\mathrm{g}}} + [\sigma(\tilde{\mathbf{i}}^{\mathrm{b}}) - 1] c_{p_{\mathrm{b}}} + \sigma(\tilde{\mathbf{i}}^{\mathrm{g}}), \end{aligned} \tag{13.14}$$

where $\mathbf{i}^{\mathrm{b}} = (i_1^{\mathrm{b}} i_2^{\mathrm{b}} \dots i_L^{\mathrm{b}})$ denotes the sequence of black indices and $\tilde{\mathbf{i}}^{\mathrm{b}}$ the subsequence of the p_{b} pairs of non-zero indices (and, of course, similarly for the grey flavour).

The inversion of Eq. (13.14), so as to furnish a mapping from the number to the index representation, is straightforward if we know how to invert the functions σ and ψ. Details on this have already been given in Sect. 5.3.2.

13.4.2 $\mathbf{T}^{(1,0)}$ sector

In the case of one black string spanning the length of the cylinder the number of dangling edges in row M can be written as $L = 2p_{\mathrm{b}} + 2p_{\mathrm{g}} + 1$, where the p_i have the same meaning as above. In particular L must be odd.

The presence of *one* string of either flavour does not impose any additional restrictions on the connectivity states of the subsequence of positive indices of that flavour. Indeed, if the position of the string is given by $i_r^{\mathrm{b}} = -1$ the non-zero subsequence of $(i_{r+1}^{\mathrm{b}} \dots i_L^{\mathrm{b}} i_1^{\mathrm{b}} \dots i_{r-1}^{\mathrm{b}})$ is still well-nested, and the arguments

given above apply. The number of connectivity states is therefore found by multiplying the L possible positions of the string by the number of $(s_{\rm b}, s_{\rm g}) = (0,0)$ states of the remaining $L-1$ points

$$C_L^{(1,0)} = LC_{L-1}^{(0,0)}. \tag{13.15}$$

Explicit values are shown in Table 13.1.

Similarly the mapping from the index to the number representation is found by first ordering after the position r of the string, and then after the value of $\phi^{(0,0)}$ taken of the remaining indices

$$\begin{aligned} \phi^{(1,0)} \begin{pmatrix} i_1^{\rm b} i_2^{\rm b} \ldots i_L^{\rm b} \\ i_1^{\rm g} i_2^{\rm g} \ldots i_L^{\rm g} \end{pmatrix} &= (r-1) C_{L-1}^{(0,0)} \\ &+ \phi^{(0,0)} \begin{pmatrix} i_1^{\rm b} \ldots i_{r-1}^{\rm b} i_{r+1}^{\rm b} \ldots i_L^{\rm b} \\ i_1^{\rm g} \ldots i_{r-1}^{\rm g} i_{r+1}^{\rm g} \ldots i_L^{\rm g} \end{pmatrix}. \end{aligned} \tag{13.16}$$

13.4.3 $\mathbf{T}^{(1,1)}$ sector

When one string of each flavour is present $L = 2p_{\rm b} + 2p_{\rm g} + 2$ must be even, and again it suffices to augment the considerations from the $\mathbf{T}^{(0,0)}$ case by some book-keeping as to the positions of the two strings. Explicit values of

$$C_L^{(1,1)} = L(L-1)C_{L-2}^{(0,0)}. \tag{13.17}$$

are shown in Table 13.1.

Letting r_i denote the position of the string of flavour $i = {\rm b}, {\rm g}$ we find that

$$\begin{aligned} \phi^{(2,0)} \begin{pmatrix} \mathbf{i}^{\rm b} \\ \mathbf{i}^{\rm g} \end{pmatrix} &= [(r_{\rm b}-1)(L-1) + (|r_{\rm g} - r_{\rm b}| - 1)] C_{L-2}^{(0,0)} \\ &+ \phi^{(0,0)} \begin{pmatrix} i_1^{\rm b} \ldots i_{r_{\rm b}-1}^{\rm b} i_{r_{\rm b}+1}^{\rm b} \ldots i_{r_{\rm g}-1}^{\rm b} i_{r_{\rm g}+1}^{\rm b} \ldots i_L^{\rm b} \\ i_1^{\rm g} \ldots i_{r_{\rm b}-1}^{\rm g} i_{r_{\rm b}+1}^{\rm g} \ldots i_{r_{\rm g}-1}^{\rm g} i_{r_{\rm g}+1}^{\rm g} \ldots i_L^{\rm g} \end{pmatrix} \end{aligned} \tag{13.18}$$

is the desired mapping from the index to the number representation.

A possible configuration of the system for $(s_{\rm b}, s_{\rm g}) = (1,1)$ is illustrated in Fig. 13.1, where the index representation of the connectivity state for each completed row is shown to the right of the figure.

13.4.4 $\mathbf{T}^{(2,0)}$ sector

Considering now the case of two black strings, it appears that the number of connectivity states for L even is given by

$$\tilde{C}_L^{(2,0)} = \binom{L}{2} C_{L-2}^{(0,0)}, \tag{13.19}$$

where we have simply divided Eq. (13.17) by 2 to take into account the indistinguishability of two strings of the *same* flavour. This is however not quite true,

since for $L \geq 4$ the number (13.19) includes certain disallowed basis states. For $L = 4$ these are

$$\begin{pmatrix} -1 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{13.20}$$

The reason why these states are not valid is that, by definition of the allowed vertices (see Fig. 10.1), black loop segments cannot cross a black string. In general, therefore, any configuration where the positions of two equal, positive black indices are separated by exactly one black string is not a valid one, even though the positive indices of each flavour satisfy the well-nestedness criterion. Accordingly, the true $C_L^{(2,0)}$ is less than the $\tilde{C}_L^{(2,0)}$ of Eq. (13.19).

We have not found it worthwhile to pursue the solution of this complication, since the numbers $\tilde{C}_L^{(2,0)}$ are already less than the $C_L^{(1,1)}$, and we need to diagonalise the transfer matrices $\mathbf{T}^{(1,1)}$ and $\mathbf{T}^{(2,0)}$ for the same values of L in order to determine the scaling dimensions x_1 and x_2 with the same numerical precision. Instead we found it efficient to construct all the $\tilde{C}_L^{(2,0)}$ basis states, list the number representations of those that are disallowed, and force the corresponding entries of $\mathbf{T}^{(2,0)}$ to zero.

With this proviso the mapping from the index to the number representation is

$$\begin{aligned} & \phi^{(2,0)} \begin{pmatrix} \mathbf{i}^{\mathrm{b}} \\ \mathbf{i}^{\mathrm{g}} \end{pmatrix} = [\psi(\mathbf{i}^{\mathrm{b}} + \mathbf{1}) - 1] C_{L-2}^{(0,0)} \\ + \quad & \phi^{(0,0)} \begin{pmatrix} i_1^{\mathrm{b}} \cdots i_{r_1-1}^{\mathrm{b}} i_{r_1+1}^{\mathrm{b}} \cdots i_{r_2-1}^{\mathrm{b}} i_{r_2+1}^{\mathrm{b}} \cdots i_L^{\mathrm{b}} \\ i_1^{\mathrm{g}} \cdots i_{r_1-1}^{\mathrm{g}} i_{r_1+1}^{\mathrm{g}} \cdots i_{r_2-1}^{\mathrm{g}} i_{r_2+1}^{\mathrm{g}} \cdots i_L^{\mathrm{g}} \end{pmatrix}, \end{aligned} \tag{13.21}$$

where r_1 and r_2 are the positions of the two black strings, and $\psi(\mathbf{i}^{\mathrm{b}} + \mathbf{1})$ means that we should lexicographically order the positions of the black indices that are -1.

Resumo en Esperanto

Uzante denove bazon de konektecoj ni konstruas la transfermatricojn por la duspeca buklomodelo. La nombro da statoj en la diversaj sektoroj kun kaj sen kordoj (Tab-o 13.1) estas multe malpli granda ol 4^L, la nombro da statoj en la kolora reprezentiĝo de la senkorda sektoro. Malkombinado en magrajn matricojn korespondajn al la unuopaj verticoj de tavolo implicas diferencon inter la ago de la unua matrico (Fig-o 13.2), la dua (Fig-o 13.3) kaj la lasta kiu kompletigas la tavolon (Fig-o 13.4).

Chapter XIV

Numerical results

The transfer matrices constructed in the last chapter can be used to extract values of the central charge and various geometrical scaling exponents for the FPL^2 model. We use these results to check the analytical formulae given in Chapter 12. Combining the analytical values for the central charge with numerical results also provides us with very accurate values for the residual entropy, a quantity which would otherwise only be available from an *exact* solution.

14.1 Central charge

The reduced free energy per vertex in the limit $M \to \infty$ of an infinitely long cylinder is given by

$$f_0^{(0,0)}(L) = \lim_{M\to\infty} \frac{1}{LM} \ln \mathrm{Tr}\; Z^{(M)} = -\frac{1}{L} \ln \lambda_0^{(0,0)}, \tag{14.1}$$

where $\lambda_0^{(s_\mathrm{b},s_\mathrm{g})}$ is the largest eigenvalue of $\mathbf{T}^{(s_\mathrm{b},s_\mathrm{g})}$. The partition function for a cylinder of length M is found by iterating the no-string TM

$$Z^{(M)} = \left(\mathbf{T}^{(0,0)}\right)^M Z^{(0)}. \tag{14.2}$$

It is well-known that conformal invariance relates the amplitude of the $1/L^2$ corrections to $f_0^{(0,0)}(\infty)$ to the central charge c [50, 51]. A further (non-universal) $1/L^4$ correction due to the operator $T\overline{T}$, where T denotes the stress tensor, must also be present in any conformally invariant system [4]. It is therefore found in a number of cases [44, 57, 47] that fits of the form

$$f_0^{(0,0)}(L) = f_0^{(0,0)}(\infty) - \frac{\pi c}{6L^2} + \frac{A}{L^4} \tag{14.3}$$

yield very rapidly converging estimates for c. An efficient application of this expression is to determine c from parabolic least-squares fits of the finite-size data against $1/L^2$ [57, 47].

n_b	n_g	$c(4,14)$	$c(6,14)$	$c(8,14)$	$c(10,14)$	Extrapolation	Exact
0.0	0.0	-2.8943	-2.8861	-2.9220	-2.9514	-3.0037	-3.0000
0.5	0.0	-1.8528	-1.7641	-1.7716	-1.7873	-1.8152	-1.8197
0.5	0.5	-0.7295	-0.6249	-0.6159	-0.6220	-0.6328	-0.6395
1.0	0.0	-1.0012	-0.9542	-0.9636	-0.9761	-0.9983	-1.0000
1.0	0.5	0.1341	0.1877	0.1924	0.1895	0.1843	0.1803
1.0	1.0	0.9918	0.9969	0.9986	0.9999	1.0004	1.0000
1.5	0.0	-0.3765	-0.3669	-0.3817	-0.3923	-0.4111	-0.4124
1.5	0.5	0.7652	0.7746	0.7729	0.7715	0.7690	0.7678
1.5	1.0	1.6215	1.5818	1.5778	1.5806	1.5856	1.5876
1.5	1.5	2.2541	2.1691	2.1581	2.1627	2.1709	2.1751
2.0	0.0	0.0706	0.0549	0.0342	0.0235	-0.0019	0.0000
2.0	0.5	1.2209	1.1937	1.1868	1.1861	1.1849	1.1803
2.0	1.0	2.0792	2.0002	1.9899	1.9937	2.0005	2.0000
2.0	1.5	2.7139	2.5919	2.5737	2.5781	2.5859	2.5876
2.0	2.0	3.1629	3.0121	2.9885	2.9936	3.0027	3.0000

Table 14.1: Estimants $c(L_0, L_\mathrm{max})$ for the central charge are obtained from parabolic least-squares fits against $1/L^2$ using the numerical data for $L_0 \leq L \leq L_\mathrm{max}$. The extrapolation in L_0 is described in the text.

In Table 14.1 the results of such fits including the data points for $L_0 \leq L \leq L_\mathrm{max}$ are shown as a function of L_0. Numerically we were able to access $L_\mathrm{max} = 14$, in which case the largest single-vertex TMs have dimension $\sim 7 \cdot 10^6$ (see Table 13.1). The extrapolation of the estimants $c(L_0, L_\mathrm{max})$ to the limit of infinite L_0 is assumed to take the form of a power law

$$c(L_0, L_\mathrm{max}) = c + k L_0^{-p}, \tag{14.4}$$

at least within an asymptotic regime of large enough L_0. As is evident from Table 14.1 the last three estimants usually exhibit monotonicity, thus allowing us to fix the constants c, k and p. When this was not the case, or whenever the power p thus obtained was to small to produce a reliable extrapolation the Ising-like value $p = 2$ was used by default to extrapolate the last two estimants. An error bar for this type of fit can be estimated from the variation among the individual estimants. The extrapolants are invariably in excellent agreement with our analytical results, the relative deviation being typically of the order 10^{-3}.

The results for c are shown for all integer and half-integer values of $n_i \in [0, 2]$. Because of the symmetric appearance of the two flavours in Eq. (13.1) only $n_\mathrm{b} \geq n_\mathrm{g}$ need be considered. For either $n_\mathrm{b} = 1$ or $n_\mathrm{g} = 1$ the FPL^2 model reduces to the simpler FPL model earlier considered by Batchelor *et al.* [88], and for $n_\mathrm{b} = n_\mathrm{g}$ we recover another special case recently investigated by Kondev [89].

14.2 Thermal scaling dimension

A further prediction of conformal invariance is that the finite-size scaling of the first gap in the eigenvalue spectrum of $\mathbf{T}^{(0,0)}$ is related to the thermal scaling dimension [49]

$$f_1^{(0,0)}(L) - f_0^{(0,0)}(L) = \frac{2\pi x_T}{L^2} + \cdots , \tag{14.5}$$

where $f_1^{(0,0)}$ is found from the next-largest eigenvalue of $\mathbf{T}^{(0,0)}$ through $f_1^{(0,0)} = -\frac{1}{L}\ln\lambda_1^{(0,0)}$. These computations were also carried through for even L up to $L_{\max} = 14$. In this case as well the convergence of the estimants can be considerably sped up by including a $1/L^4$ term in Eq. (14.5) and performing parabolic least-squares fits versus $1/L^2$.

The results for x_T as displayed in Table 14.2 again agree with those of the previously studied special cases [88, 89]. The data for $(n_{\rm b}, n_{\rm g}) = (0,0)$ merit a special comment. Monitoring the three leading eigenvalues $\lambda_0^{(0,0)}$, $\lambda_1^{(0,0)}$ and $\lambda_2^{(0,0)}$ as a function of n for $n_{\rm b} = n_{\rm g} \equiv n$ we found that $\lambda_1^{(0,0)}$ and $\lambda_2^{(0,0)}$ are exactly degenerate for all n down to $n \sim 0.20$. Hereafter $\lambda_1^{(0,0)}$ splits off from $\lambda_2^{(0,0)}$ and eventually becomes degenerate with $\lambda_0^{(0,0)}$ at $n = 0$. Because of this level crossing it thus seems very likely that near $(n_1, n_2) = (0,0)$ the thermal eigenvalue should be related to the gap $f_2^{(0,0)}(L) - f_0^{(0,0)}(L)$. Comparison with the exactly known result $x_T = 1/2$ [89] confirms this suspicion. A similar comment holds true near $(n_{\rm b}, n_{\rm g}) = (2,2)$, and again we find fair agreement with the exact result if we apply Eq. (14.5) to $\lambda_2^{(0,0)}$, and not to $\lambda_1^{(0,0)}$ (which in this case becomes two-fold degenerate).

For $n_{\rm b} < 2$ the extrapolants are again in excellent ($\sim 10^{-3}$ or better) agreement with our analytical results. For $n_{\rm b} = 2$ the slower convergence can be attributed to logarithmic corrections [59] arising from an enhanced number of marginal vertex operators. Indeed, of the twelve vertex operators corresponding to the shortest vectors in $\mathcal{R}_w^*$, Eq. (11.19), seven stay marginal when either $n_{\rm b} < 2$ or $n_{\rm g} < 2$. In the general case, when both $n_{\rm b} < 2$ and $n_{\rm g} < 2$, there are only four marginal vertex operators; this is the loop ansatz, Eq. (11.26).

14.3 Dimensions of string operators

14.3.1 One black string

We now turn our attention to the determination of the scaling dimensions associated with one or more strings spanning the length of the cylinder. The presence of one black string corresponds to a height mismatch in the ideal states, and the relevant scaling dimension X is therefore that of a *twist-like operator* [133]. We have calculated the leading eigenvalue of $\mathbf{T}^{(1,0)}$ for odd system sizes up to $L_{\max} = 13$ and determined the corresponding estimants $c_{\rm odd}(L_0, L_{\max})$ by the

n_b	n_g	$x_T(4,14)$	$x_T(6,14)$	$x_T(8,14)$	$x_T(10,14)$	Extrapolation	Ref. [88]	Exact
0.0	0.0	0.5712	0.5280	0.5121	0.5060	0.4987		0.5000
0.5	0.0	0.5704	0.5535	0.5452	0.5417	0.5366		0.5372
0.5	0.5	0.5916	0.5882	0.5845	0.5825	0.5789		0.5804
1.0	0.0	0.5826	0.5798	0.5765	0.5748	0.5708	0.573 (1)	0.5714
1.0	0.5	0.6204	0.6227	0.6218	0.6211	0.6199	0.6200 (5)	0.6206
1.0	1.0	0.66368	0.66600	0.66642	0.66654	0.66663	0.6666 (1)	0.66667
1.5	0.0	0.5965	0.6053	0.6060	0.6058	0.6054		0.6063
1.5	0.5	0.6493	0.6559	0.6574	0.6578	0.6585		0.6619
1.5	1.0	0.7782	0.7094	0.7108	0.7115	0.7130	0.713 (1)	0.7146
1.5	1.5	0.8950	0.7657	0.7674	0.7684	0.7702		0.7699
2.0	0.0	0.6167	0.6295	0.6338	0.6349	0.6356		0.6667
2.0	0.5	0.7481	0.6878	0.6913	0.6927	0.6945		0.7345
2.0	1.0	0.8741	0.7566	0.7552	0.7565	0.7588	0.76 (1)	0.8000
2.0	1.5	0.9436	0.8755	0.8284	0.8303	0.8337		0.8702
2.0	2.0	0.9996	0.9850	0.9400	0.9200	0.8876		1.0000

Table 14.2: The thermal scaling dimension x_T. The extrapolation of the estimants $x_T(L_0, L_{\max})$ is described in the text. For comparison we also show the numerical data for the case of either n_b or n_g being unity [88]. Due to level crossing the values of x_T for $(n_b, n_g) = (0,0)$ and $(2,2)$ are found from the gap $f_2^{(0,0)}(L) - f_0^{(0,0)}(L)$ rather than from $f_1^{(0,0)}(L) - f_0^{(0,0)}(L)$.

n_b	n_g	$X(3,13)$	$X(5,13)$	$X(7,13)$	$X(9,13)$	Extrapolation	Result	Exact
0.0	0.0	-0.05586	-0.06109	-0.06203	-0.06232	-0.06257	-0.06269 (31)	-0.06250
0.0	0.5	-0.06080	-0.06197	-0.06220	-0.06233	-0.06253		-0.06250
0.0	1.0	-0.06043	-0.06198	-0.06221	-0.06233	-0.06250		-0.06250
0.0	1.5	-0.05869	-0.06156	-0.06215	-0.06233	-0.06259		-0.06250
0.0	2.0	-0.05804	-0.06190	-0.06297	-0.06316	-0.06324		-0.06250
0.5	0.0	0.04674	0.04538	0.04558	0.04569	0.04587	0.04583 (16)	0.04591
0.5	0.5	0.04572	0.04585	0.04589	0.04588	0.04588		0.04591
0.5	1.0	0.04643	0.04622	0.04614	0.04607	0.04595		0.04591
0.5	1.5	0.04781	0.04675	0.04638	0.04622	0.04590		0.04591
0.5	2.0	0.04828	0.04664	0.04593	0.04573	0.04555		0.04591
1.0	0.0	0.11895	0.12278	0.12398	0.12438	0.12501	0.12497 (8)	0.12500
1.0	0.5	0.12346	0.12422	0.12458	0.12470	0.12489		0.12500
1.0	1.0	0.12465	0.12485	0.12495	0.12496	0.12498		0.12500
1.0	1.5	0.12584	0.12540	0.12529	0.12521	0.12508		0.12500
1.0	2.0	0.12652	0.12549	0.12513	0.12501	0.12490		0.12500
1.5	0.0	0.17106	0.18253	0.18453	0.18536	0.18662	0.18663 (25)	0.18687
1.5	0.5	0.18283	0.18468	0.18553	0.18585	0.18633		0.18687
1.5	1.0	0.18515	0.18588	0.18620	0.18632	0.18646		0.18687
1.5	1.5	0.18684	0.18684	0.18687	0.18684	0.18680		0.18687
1.5	2.0	0.18878	0.18796	0.18759	0.18741	0.18696		0.18687
2.0	0.0	0.2076	0.2296	0.2321	0.2340	0.2369	0.2392 (27)	0.2500
2.0	0.5	0.2283	0.2323	0.2342	0.2351	0.2371		0.2500
2.0	1.0	0.2325	0.2347	0.2358	0.2363	0.2383		0.2500
2.0	1.5	0.2357	0.2372	0.2379	0.2383	0.2400		0.2500
2.0	2.0	0.2402	0.2413	0.2417	0.2420	0.2435		0.2500

Table 14.3: Estimants $X(L_0, L_{\max})$ for the scaling dimension of the twist operator (see text).

usual parabolic fits to $f_0^{(1,0)}(L)$, cfr. Eq. (14.3). Estimants $X(L_0, L_{\max})$ are then defined by

$$X(L_0, L_{\max}) = \frac{c - c_{\text{odd}}(L_0, L_{\max})}{12}, \tag{14.6}$$

where the factor of 12 originates from a comparison of Eq. (14.3) with Eq. (14.5). For the central charge c of an even-sized system we use our analytical results, Eq. (12.4).

These estimants and their extrapolations are found in Table 14.3. Note that we can no longer limit the parameter values by $n_{\text{b}} \geq n_{\text{g}}$, as the condition $(s_{\text{b}}, s_{\text{g}}) = (1, 0)$ treats the two flavours asymmetrically. In the case of the FPL model ($n_{\text{b}} = 1$) it was found [88] that X was independent of n_{g}. It is evident from our numerical data that this n_{g}-independence in fact pertains to all $n_{\text{b}} \in [0, 2]$. Final results for X as a function of n_{b} have therefore been computed by averaging the available extrapolated scaling dimensions over n_{g}. For $n_{\text{b}} = 1$ the agreement with the result $X \approx 1/8$ found by Batchelor *et al.* [88] is excellent. Furthermore we are able to conjecture the general formula, Eq. (12.19), for X as a function of the loop fugacities.

14.3.2 One black and one grey string

When $(s_{\text{b}}, s_{\text{g}}) = (1, 1)$ the parity of L must again be even, and we can make parabolic fits for the gap $f_0^{(1,1)}(L) - f_0^{(0,0)}(L)$, as in Eq. (14.5), without resorting to the less accurate method of fitting for two central charges separately as above. The corresponding universal amplitude is identified with the scaling dimension x_1. The results, now for $L_{\max} = 12$, are shown in Table 14.4, and our values for the scaling dimension are once again in agreement with the analytical results, apart from $n_{\text{g}} = 2$ where logarithmic corrections are the most likely source of systematic errors [59].

14.3.3 Two black strings

Finally, the results for x_2 as obtained from parabolic fits for the gap $f_0^{(2,0)}(L) - f_0^{(0,0)}(L)$ are shown in Table 14.5. Again we have $L_{\max} = 12$. Just like in the case of X we find the extrapolated values of x_2 to be independent of n_{g}, and final results are obtained by averaging over this parameter.

14.4 Entropy

Apart from the various universal quantities, such as the central charge and the scaling dimensions, the transfer matrices also provide numerical values for the residual entropy per vertex, $s = f_0(\infty)$. In the limit $n_{\text{g}} \to 0$ of compact polymers this quantity is of interest to the protein folding community, due to the fact that native conformations of all globular proteins are compact [98].

n_b	n_g	$x_1(4,12)$	$x_1(6,12)$	$x_1(8,12)$	Extrapolation	Ref. [88]	Exact
0.0	0.0	-0.2433	-0.2447	-0.2470	-0.2500		-0.2500
0.5	0.0	-0.1328	-0.1295	-0.1303	-0.1313		-0.1323
0.5	0.5	-0.01713	-0.01228	-0.01217	-0.01217		-0.0131
1.0	0.0	-0.0440	-0.0423	-0.0430	-0.0439	-0.0444 (1)	-0.0446
1.0	0.5	0.0737	0.0763	0.0764	0.0765	0.0750 (3)	0.0761
1.0	1.0	0.16608	0.16646	0.16657	0.16663	0.1667 (1)	0.16667
1.5	0.0	0.0267	0.0271	0.0264	0.0255		0.0260
1.5	0.5	0.1466	0.1472	0.1472	0.1472		0.1483
1.5	1.0	0.2411	0.2395	0.2395	0.2394	0.242 (2)	0.2405
1.5	1.5	0.3196	0.3159	0.3156	0.3156		0.3162
2.0	0.0	0.0845	0.0848	0.0844	0.0839		0.1042
2.0	0.5	0.2070	0.2067	0.2071	0.2076		0.2295
2.0	1.0	0.3048	0.3021	0.3024	0.3028	0.307 (2)	0.3250
2.0	1.5	0.3882	0.3841	0.3842	0.3843		0.4044
2.0	2.0	0.4640	0.4618	0.4635	0.4657		0.5000

Table 14.4: Scaling dimension x_1, corresponding to one string of each flavour.

n_b	n_g	$x_2(4,12)$	$x_2(6,12)$	$x_2(8,12)$	Extrapolation	Result	Exact
0.0	0.0	0.0000	0.0000	0.0000	0.0000	0.0000 (0)	0.0000
0.0	0.5	0.0000	0.0000	0.0000	0.0000		0.0000
0.0	1.0	0.0000	0.0000	0.0000	0.0000		0.0000
0.0	1.5	0.0000	0.0000	0.0000	0.0000		0.0000
0.0	2.0	0.0000	0.0000	0.0000	0.0000		0.0000
0.5	0.0	0.1279	0.1355	0.1372	0.1389	0.1386 (2)	0.1386
0.5	0.5	0.1365	0.1371	0.1378	0.1387		0.1386
0.5	1.0	0.1377	0.1374	0.1379	0.1385		0.1386
0.5	1.5	0.1383	0.1375	0.1379	0.1384		0.1386
0.5	2.0	0.1392	0.1376	0.1379	0.1383		0.1386
1.0	0.0	0.2333	0.2447	0.2472	0.2504	0.2495 (5)	0.2500
1.0	0.5	0.2488	0.2477	0.2484	0.2493		0.2500
1.0	1.0	0.2514	0.2487	0.2490	0.2494		0.2500
1.0	1.5	0.2538	0.2497	0.2495	0.2492		0.2500
1.0	2.0	0.2573	0.2512	0.2504	0.2494		0.2500
1.5	0.0	0.3197	0.3377	0.3416	0.3466	0.3487 (26)	0.3506
1.5	0.5	0.3429	0.3425	0.3443	0.3466		0.3506
1.5	1.0	0.3486	0.3457	0.3466	0.3478		0.3506
1.5	1.5	0.3548	0.3497	0.3496	0.3494		0.3506
1.5	2.0	0.3636	0.3561	0.3547	0.3529		0.3506
2.0	0.0	0.3920	0.4202	0.4268	0.4353	0.446 (12)	0.5000
2.0	0.5	0.4244	0.4277	0.4323	0.4382		0.5000
2.0	1.0	0.4346	0.4348	0.4382	0.4426		0.5000
2.0	1.5	0.4468	0.4452	0.4474	0.4502		0.5000
2.0	2.0	0.4640	0.4618	0.4635	0.4657		0.5000

Table 14.5: Scaling dimension x_2, corresponding to two black strings.

Using our knowledge of the exact form of the finite-size corrections of order $1/L^2$, Eq. (14.3), we have obtained very accurate extrapolations to the limit of an infinite system.[1] After subtracting the $1/L^2$ correction a series of estimants $s(L, L_{\max})$ may be obtained by fitting the residual size dependence to a pure $1/L^4$ form. The remaining L-dependence of these estimants turns out to be well accounted for by a further $1/L^4$ fit, and in this way we arrive at a final value for s. The error bar on the final value can be estimated as its deviation from the most accurate extrapolant, $s(L_{\max} - 2, L_{\max})$.

The most accurate results are quite naturally found by employing this procedure on $f_0^{(0,0)}(L)$, and they are shown in Table 14.6. Results obtained by extrapolating the free energies for other sectors of the transfer matrix containing strings are consistent herewith but have error bars that are roughly 10 times larger. If the fugacity of one of the strings equals two the error bars are even larger, which is to be anticipated from the fact that logarithmic corrections to the scaling dimensions are larger than similar corrections to the central charge [59].

In the special case of the equal-weighted six-vertex model, $(n_{\mathrm{b}}, n_{\mathrm{g}}) = (1, 1)$, our value for s is in excellent agreement with the exact result due to Lieb [85],

$$s(1,1) = \frac{3}{2} \ln\left(\frac{4}{3}\right) \simeq 0.4315231\cdots, \tag{14.7}$$

and in the limit of two mutually excluding Hamiltonian walks, $(n_{\mathrm{b}}, n_{\mathrm{g}}) = (0, 0)$, we are able to conjecture the result

$$s(0,0) = \frac{1}{2} \ln(2) \simeq 0.3465735\cdots. \tag{14.8}$$

In fact, after having made this conjecture we discovered that the numerical values of $f_0^{(1,1)}(L)$, *i.e.*, the free energy per site in the sector where we enforce one string of each flavour, are *independent* of L for $4 \leq L \leq 12$, and equal to $\frac{1}{2}\ln(2)$ with full 16-digit machine precision. Since the free energy per site in the thermodynamic limit is unchanged by the introduction of a string defect, this observation lends credibility to the correctness of the above conjecture.

Our result in the compact polymer limit merits special attention. Traditionally the entropy is quoted in terms of the so-called connective constant $\kappa = \mathrm{e}^{s(1,0)}$; see Eq. (9.6). Early approximations due to Flory [138] and Huggins [139] yielded respectively

$$\kappa_{\mathrm{Flory}} = \frac{z-1}{e} \simeq 1.104 \tag{14.9}$$

and

$$\kappa_{\mathrm{Huggins}} = (z-1)\left(1 - \frac{2}{z}\right)^{z/2-1} = \frac{3}{2}. \tag{14.10}$$

[1] The logarithmic corrections to the free energy implied by the $\mathcal{N}^{\gamma-1}$ term in Eq. (9.6) does not pertain to the cylindrical geometry implicit in our transfer matrix calculations. A similar remark applies to the surface term $\kappa_{\mathrm{s}}\mathcal{N}^{(d-1)/d}$.

n_b	n_g	$f_0^{(0,0)}(4)$	$f_0^{(0,0)}(6)$	$f_0^{(0,0)}(8)$	$f_0^{(0,0)}(10)$	$f_0^{(0,0)}(12)$	$f_0^{(0,0)}(14)$	s
0.0	0.0	0.17328680	0.28881133	0.31784496	0.32923359	0.33490107	0.33815371	0.346575 (14)
0.5	0.0	0.26740000	0.33317928	0.35057672	0.35745438	0.36088114	0.36284872	0.367950 (9)
0.5	0.5	0.35063553	0.37668215	0.38371283	0.38639599	0.38769210	0.38842126	0.390258 (3)
1.0	0.0	0.32923947	0.36764369	0.37752555	0.38137032	0.38327066	0.38435762	0.387166 (7)
1.0	0.5	0.40772622	0.41103439	0.41126990	0.41111188	0.41095017	0.41082815	0.410405 (2)
1.0	1.0	0.46298939	0.44576535	0.43960110	0.43671524	0.43513763	0.43418273	0.4315233 (4)
1.5	0.0	0.37601935	0.39599984	0.40063320	0.40233073	0.40314475	0.40360330	0.404771 (5)
1.5	0.5	0.45180855	0.43964968	0.43509788	0.43291844	0.43171625	0.43098591	0.4289459 (10)
1.5	1.0	0.50624745	0.47501911	0.46431698	0.45948057	0.45688890	0.45533728	0.4510742 (17)
1.5	1.5	0.54930614	0.50513652	0.49006459	0.48331974	0.47972832	0.47758588	0.471726 (2)
2.0	0.0	0.41389271	0.42018005	0.42097629	0.42111147	0.42113891	0.42114428	0.421145 (6)
2.0	0.5	0.48795109	0.46429984	0.45622604	0.45257417	0.45061901	0.44945033	0.4462607 (10)
2.0	1.0	0.54202495	0.50046092	0.48641918	0.48016010	0.47683419	0.47485271	0.4694505 (18)
2.0	1.5	0.58515036	0.53158535	0.51333087	0.50520022	0.50088581	0.49831761	0.491323 (3)
2.0	2.0	0.62122666	0.55918707	0.53795845	0.52850379	0.52348906	0.52050483	0.5123870 (19)

Table 14.6: Residual entropy s, obtained by extrapolating $f_0^{(0,0)}(L)$ to the infinite-system limit.

Here $z = 4$ is the coordination number of the square lattice. More recently, $\kappa \simeq 1.472$ was found from transfer matrix calculations [115] and $\kappa = 1.475(15)$ by exhaustive computer enumeration of short-chain configurations [97]. Both these results are very close to the mean-field value $\kappa_{\mathrm{MF}} = \frac{z}{e} = 1.4715\cdots$ [113], and it is tempting to conclude that conformations of compact polymers are in fact described by mean-field theory [97]. However, our result

$$\kappa = 1.472801(10) \tag{14.11}$$

demonstrates that this is not the case.

Resumo en Esperanto

La transfermatricoj donas nombrajn valorojn por la centra ŝargo c (Tab-o 14.1), la termika dimensio (Tab-o 14.2) kaj diversaj kord-dimensioj (Tab-oj 14.3–14.5) kiuj tre precize konfirmas la rezultojn de Ĉapitro 12.

Ĉar ni ekzakte konas c eblas komputi ege precizajn valorojn por la rezidua entropio (Tab-o 14.6). Por du reciproke ekskludaj kompaktaj polimeroj ni konjektas la rezulton (14.8), kaj en la kazo de unu kompakta polimero la konekta konstanto (14.11) malsamas ol ĝia valoro en la avaraĝkampa aproksimo.

Chapter XV

Discussion and outlook

As a by-product of the effective field theory constructed for the FPL^2 model some rather general conclusions regarding the scaling of compact polymers, and the relation between loop models and conformal field theory can be drawn. It also provides new insights into the three-state Potts antiferromagnet and the dimer loop model, which are identified with specific points in the phase diagram of the FPL^2 model. After a remark on the possible Bethe ansatz solvability of the model we conclude with a description of possible future developments.

15.1 Compact polymers

One of the main motivations for studying fully packed loop models is provided by compact polymers, their scaling properties in particular. Just like polymers in the dilute and dense phase, compact polymers form a critical geometrical system characterised by conformational exponents γ and ν. The exponent γ relates the number of conformations of the polymer to the number of monomers; see Chapter 9 for details. The other conformational exponent (ν) relates the linear size of the polymer to the number of monomers. For compact structures it has the trivial value 1/2 since these polymers are space filling.

Prior to our work, exact results have been obtained for compact polymers on the Manhattan [140, 141] and the honeycomb [91] lattice, and the mean-field value $\gamma = 1$ was found in both cases. This value of γ indicates that the two ends of the compact polymer are independent at large distances. This follows from the scaling relation $x_1 = 1 - \gamma = 0$, where the one-string dimension x_1 describes the probability $G_1(r) \sim r^{-2x_1}$ that the two chain ends are separated by a distance r. In this regard the scaling of compact polymers on the Manhattan and the honeycomb lattices is equivalent to that of ideal chains. Ideal chain configurations are described by simple random walks for which each step is independent of the previous one.

Here we have calculated the exact conformational exponent $\gamma = 117/112$ for compact polymers on the square lattice. The fact that $\gamma > 1$ is tantamount to an effective *repulsion* between the ends of the chain, indicating non-ideal

behaviour. Indeed, the fact that the connective constant κ in Eq. (14.11) is larger than its mean-field value indicates that the origin of this repulsion is *entropic*. Earlier numerical studies of this problem utilising direct enumerations of chain conformations have failed to see any deviation from the ideal chain result $\gamma_{\rm MF} = 1$ [97]; we can attribute this to the fact that the actual difference is indeed very small ($\gamma - \gamma_{\rm MF} = 5/112$) and below the numerical accuracy previously achieved. The same comment can be made for the connective constant.

Another interesting aspect of compact polymers is that their scaling properties are lattice dependent. This is in contrast to the dilute and dense case which are described by conformational exponents that do not depend on the lattice type (*e.g.*, honeycomb *versus* square). As remarked earlier this "lack of universality" is due to a kind of geometrical frustration that arises from the fully packing constraint imposed on the loop models which are employed in studies of compact polymers.

Finally, the field theory solution of the FPL^2 model uncovered a property of compact polymers that, to our knowledge, was not previously anticipated. The fact that there is a whole line of critical points in this loop model in the Hamiltonian walk limit ($n_{\rm b} \to 0$) indicates a continuum of universality classes described by compact polymers on the square lattice. In particular the exponent γ can be varied continuously between 43/48 and 9/8 by adjusting the fugacity of the loops uncovered by the polymer. The loop weight of the uncovered (grey) loops can be thought of as an effective interaction amongst the monomers, albeit a non-local one. A similar effect of interactions on directed self-avoiding walks was discovered by Cardy [123] from a field-theoretical analysis of the problem. The existence of a continuously varying γ in this case was recently challenged by numerical results [124].

15.2 Relation to other models

The FPL^2 model is a loop model which exhibits a two-dimensional manifold of fixed points in its phase diagram. Certain points in the critical region map to previously studied lattice models and here we comment on the relevance of our results for these models.

15.2.1 Dimer loop model

The dimer loop model studied by Raghavan *et al.* [107] is the $n_{\rm b} = 2$, $n_{\rm g} = 1$ FPL^2 model; see Fig. 8.1. The dimer loop model is defined by placing black and white dimers on the square lattice so that every vertex of the lattice is covered by exactly one black and one white dimer. Every such configuration is given equal weight. The mapping to the FPL^2 model is achieved by identifying the bonds covered by dimers as making up the black loops, whilst the uncovered bonds form the grey loops. The original motivation for studying this dimer problem is that it leads to a height model with a two-component height; *cfr.* the traditional dimer model which is described by a single component height.

Performing Monte Carlo simulations of the dimer loop model Raghavan *et al.* reached the conclusion that one of the two height components is rough whilst the other one is "anomalously smooth", *i.e.*, its structure function decays at small wave-vectors $\mathbf{q}$ slower then $1/\mathbf{q}^2$; a $1/\mathbf{q}^2$ dependence is to be expected in a Gaussian field theory.

In light of our results we would conclude that the dimer loop model is critical with a central charge $c = 2$. This follows from Eq. (12.4) for $n_\mathrm{b} = 2$ and $n_\mathrm{g} = 1$. The two components of the height found by Raghavan *et al.* should therefore both be rough, each contributing one to the central charge ($c = 1 + 1$). Furthermore, we believe that the observed anomalous behaviour of one of the heights can be attributed to the fact that this model is exactly at the boundary of the critical region of the FPL^2 model. We observe a similar effect in our numerical transfer matrix results which show largest deviations from the proposed exact formulae for loop fugacities at the critical-region boundary. The culprit might be logarithmic corrections due to the presence of marginal operators. To check this hypothesis and reconcile it with the fact that no such effects are seen in Monte Carlo simulations of the four-colouring model [93] ($n_\mathrm{b} = n_\mathrm{g} = 2$), which is also at the boundary of the critical region, simulations of the dimer-loop model for larger system sizes would be welcome.

15.2.2 Three-state Potts antiferromagnet

The critical ground state of three-state Potts antiferromagnet maps to the equal-weighted six-vertex model [106] which is the $n_\mathrm{b} = n_\mathrm{g} = 1$ point in the critical region of the FPL^2 model; see Fig. 8.1. Along the line $n_\mathrm{b} = n_\mathrm{g}$ the colouring representation of the FPL^2 model has the additional symmetry with respect to cyclic permutations of the four colours; see Sec. 11.3.2. This explains the origins of the $\mathbf{Z}_4$ symmetry found by Saleur for the *three*-state Potts antiferromagnet [142].

15.2.3 Folding model

The folding model of the square-diagonal lattice recently investigated by Di Francesco [143] maps onto a constrained version of the $(n_\mathrm{b}, n_\mathrm{g}) = (2, 2)$ FPL^2 model. The constraint consists in allowing only the vertices 1, 3, 5 and 6 of Fig. 10.1 for sites on the even sublattice, and similarly vertices 2, 4, 5 and 6 on the odd sublattice.

We have modified our transfer matrices to take this constraint into account. Our result for the folding entropy, $s = 0.4604(4)$, is in complete agreement with Ref. [143].[1] Interestingly enough the finite-size scaling of the gaps in the eigenvalue spectrum seems to indicate that the model is not critical for general values of the loop fugacities. From the field theory of the FPL^2 model we should be able to to understand why the constraint imposed by the folding model

[1] Our normalisation is "per vertex" whilst that of Di Francesco is "per triangle". Accordingly we find twice his result.

leads to a relevant perturbation which takes the system away from criticality. This we leave as an interesting open question. Incidentally, the situation is very reminiscent of the reformulation of the Q-state Potts model in terms of a staggered vertex model. Only at the critical point are the vertex weights on the even and odd sublattices identical, thus allowing for an exact solution of the model [125].

15.3 Bethe Ansatz solvability

Finally, we end with a speculatory note concerning the prospects of solving the FPL^2 model via Bethe Ansatz. Namely, all loop models to date have been solved by this method after mapping them to a vertex model, following a procedure analogous to the one outlined in Sect. 10.2. This does not seem to work for the FPL^2 model, at least not along the $n_{\mathrm{g}} = 1$ line [88]. Why this is so is an interesting open question.

One possibility is that the *full* FPL^2 model needs to be considered as opposed to the FPL model studied by Batchelor *et al.* for which $n_{\mathrm{g}} = 1$ is fixed. A more intriguing possibility is that a Bethe Ansatz solution might be hindered (or made more difficult) by the non-trivial elasticity displayed by the FPL^2 model in its interface representation. This statement we base solely on the observation that all previously solved loop models are simple as interface models in the sense that the height fluctuations are described by a single elastic constant. For the FPL^2 model, as described in Chapter 11, the stiffness tensor consists of three independent components. Whether indeed the interface representation of the loop model has any bearing on its Bethe Ansatz solvability remains to be seen.

15.4 Open questions

15.4.1 Universality

As mentioned in the introduction to Chapter 9 it is rather intriguing from the viewpoint of universality that dense polymers exhibit density- and lattice-independent critical exponents, whereas precisely at filling fraction $f = 1$ one encounters the problem of compact polymers with critical exponents that depend on the lattice. At the time when the dense polymer problem was solved [112] this possibility was not taken into account, and it appeared natural to identify compact and dense polymers.

Thus, although the conformational exponent γ places compact polymers on the honeycomb ($\gamma = 1$) and the square ($\gamma = 117/112$) lattices into different universality classes, it is an interesting question whether there exists other critical exponents that are identical for dense polymers and the two solved cases of compact polymers. A trivial example, of course, is $\nu = 1/2$ which just reflects the fact that both dense and compact polymers have Hausdorff dimension two.

More interestingly, from the exact results for the string dimensions one can compute an infinity of *contact exponents* [112] $\theta_{\mathcal{G}}$ giving the asymptotic decay of the probability $P_{\mathcal{G}}(y) \sim y^{\theta_{\mathcal{G}}}$ that a certain number of points on the polymer simultaneously have a spatial distance on the order of y, in the limit $y \to 0$. Since $P_{\mathcal{G}}$ is the ratio of two partition functions the resulting exponent $\theta_{\mathcal{G}}$ is independent of the boundary conditions. The most commonly encountered contact exponents pertain to the spatial approaching of the two endpoints of the polymer (θ_0), an endpoint and an interior point (θ_1), and two interior points (θ_2).

Now, θ_0 turns out to be simply related to γ, and so it is 'non-universal'. On the other hand, θ_1 and θ_2 are functions of the string dimensions only through the combinations $x_1 + x_3$ and x_4 respectively,[2] and these are 'universal'. Therefore

$$\theta_1 = \frac{1}{2}, \qquad \theta_2 = \frac{3}{4} \tag{15.1}$$

for both known cases of compact polymers and for dense polymers on arbitrary networks.

Contact exponents should be relevant to protein folding, since hydrophobic interactions take place at contacts. It would therefore be most interesting to examine whether the above equalities are purely coincidental or whether they hold true in general. In particular, this question gives motivation for investigating the compact polymer problem on the *triangular* lattice (see below).

15.4.2 Compact polymers on the triangular lattice

The six-colouring problem on the triangular lattice gives rise to a fully-packed loop model with three different flavours of loops. Extrapolating what is known for the honeycomb and the square lattices one would expect this FPL^3 model to have a five-dimensional interfacial representation embedded in seven-dimensional space and to possess a three-dimensional manifold of critical fixed points. The six-colouring problem would be critical with central charge $c = 5$ and have a scaling limit described by a $SU(6)_{k=1}$ symmetry.

If all of these statements were just obvious, of course, the FPL^3 model would be of rather limited interest from the point of view of learning new physics. But since closed loops on the triangular lattice can have *any* length ≥ 3 (and not just even ones), it poses a serious challenge to the mapping between the colouring and the interfacial representations used this far. In particular, the loops cannot be simply defined as alternating sequences of two different colours.

15.4.3 More about protein folding

The usefulness of the FPL^2 model for calculating conformational properties of proteins leaves one wondering whether further quantities of relevance to protein folding can be calculated. In particular we are currently looking into the

[2] For the FPL^2 model we have defined $x_1 = x_{1,1}$, the grey being assigned unit weight anyway. Similarly $x_3 = x_{3,1}$ and $x_4 = x_{4,0}$.

possibility of making quantitative predictions on the formation of secondary structure, *i.e.*, the two-dimensional restrictions of α-helices and β-sheets, in the compact states.

Even more interesting, one could imagine incorporating specific sequential (amino acid) information through the imposition of quenched randomness. A suitable model for this investigation is the FPL model on the honeycomb lattice. If we assign quenched random numbers $w_i \in [0, w]$ to each of the lattice edges and augment the conventional loop Hamiltonian with a term assigning energy penalties w_i to each of the $N/3$ edges which are not covered by a loop in any given configuration, one would end up with a model in which certain loop configurations are preferred to others. Evidently this would bring down the conformational entropy, and the question is whether it would go to *zero*, reflecting the fact that polymers fold to a unique native state.

On a more theoretical level this model would contain *both frustration and disorder* as two competing effects. Since the disorder just described couples to the local energy density the techniques of Part I would be applicable, in particular the perturbative renormalisation group. As such, this disordered loop model can be considered to be the missing 'Part III' of the present thesis, providing a unifying view of frustration and disorder in discrete lattice models.

Resumo en Esperanto

La devio inter niaj ekzaktaj rezultoj por kompaktaj polimeroj sur la kvadrata reto kaj la averaĝkampa aproksimo estas komentata. $\gamma > 1$ signifas ke la du ekstremaĵoj de la polimero sin reciproke forpuŝas pro entropiaj kialoj.

Por $(n_\mathrm{b}, n_\mathrm{g}) = (2, 1)$ ni retrovas la dimer-buklomodelon [107]. Kontraŭe al la konkludo de Ref-o [107] ni trovis ke tiu modelo estas krita kun centra ŝargo $c = 2$. Kun aldonaj kondiĉoj sur la permesataj verticoj nia buklomodelo ankaŭ povas priskribi la faldadon de membrano [143].

Finfine, ni listigas diversajn proponojn por estonta reserĉo. Aparte interesas nin buklomodelo kun malordo, esperante ke ĝi adekvate modelumas la specifan sekvencan informon en la proteinfalda problemo.

Appendix A

Dimensions of electric and magnetic operators

We calculate the scaling dimensions of electric and magnetic operators in the Coulomb gas theory described by the action

$$S_{\rm CG} = \frac{1}{2}\int d^2\mathbf{x}\; g_\alpha(\boldsymbol{\partial} H^\alpha)^2 + \frac{\mathrm{i}}{4\pi}\int d^2\mathbf{x}\; (\mathbf{E}_0\cdot\mathbf{H})\mathcal{R}\ , \tag{A.1}$$

where $\mathcal{R}$ is the scalar curvature. We are interested in the situation when the height field is defined on a flat surface, in which case $\mathcal{R}$ is zero everywhere except at the boundaries.

A.1 Electric charges

The scaling dimension $x(\mathbf{E})$, of the electric-type operator $\exp(\mathrm{i}\mathbf{E}\cdot\mathbf{H}(\mathbf{x}))$, follows from the two-point function

$$\left\langle \mathrm{e}^{i\mathbf{E}\cdot\mathbf{H}(\mathbf{x})}\mathrm{e}^{-i(\mathbf{E}-2\mathbf{E}_0)\cdot\mathbf{H}(\mathbf{y})}\right\rangle \sim |\mathbf{x}-\mathbf{y}|^{-2x(\mathbf{E})}, \tag{A.2}$$

where the expectation value is with respect to the measure defined by the action $S_{\rm CG}$. The extra electric charge $2\mathbf{E}_0$ appears due to the charged boundary conditions enforced by the curvature term in the Coulomb gas action, Eq. (A.1).

We break up the calculation into two parts. First we calculate the two-point function, Eq. (A.2), in the absence of the background charge ($\mathbf{E}_0 = 0$). We make use of the property of Gaussian integrals,

$$\left\langle \mathrm{e}^{i\mathbf{E}\cdot\mathbf{H}(\mathbf{x})}\mathrm{e}^{-i\mathbf{E}\cdot\mathbf{H}(\mathbf{y})}\right\rangle = \tag{A.3}$$
$$\exp\left(-\frac{1}{2}(E_\alpha)^2\left\langle (H^\alpha(\mathbf{x})-H^\alpha(\mathbf{y}))^2\right\rangle\right),$$

and of the known propagator for the massless scalar field in two dimensions (where we have dropped the regulators at large and small distances),

$$\left\langle (H^\alpha(\mathbf{x})-H^\alpha(\mathbf{y}))^2\right\rangle = \frac{1}{\pi g_\alpha}\ln|\mathbf{x}-\mathbf{y}|\ . \tag{A.4}$$

Combining the above two equations and comparing the result to Eq. (A.2), we find

$$2x_{\rm e}^{(0)}(\mathbf{E}) = \frac{1}{2\pi g_\alpha}(E_\alpha)^2 \ ; \tag{A.5}$$

the superscript (0) is there to remind us that this formula is valid only for $\mathbf{E}_0 = 0$.

This result for the two-point function can be rewritten as

$$\left\langle {\rm e}^{i\mathbf{E}\cdot\mathbf{H}(\mathbf{x})}{\rm e}^{-i\mathbf{E}\cdot\mathbf{H}(\mathbf{y})}\right\rangle = \exp[\mathcal{E}_{\mathbf{E}}^{(0)}(\mathbf{x},\mathbf{y})], \tag{A.6}$$

where

$$\mathcal{E}_{\mathbf{E}}^{(0)}(\mathbf{x},\mathbf{y}) = -\frac{1}{2\pi g_\alpha}(E_\alpha)^2 \ln|\mathbf{x}-\mathbf{y}| \tag{A.7}$$

is the energy for two (vector) electric charges interacting via the two-dimensional Coulomb force; in this language $S_{\rm E}$ is the energy of the electrostatic field set up by the electric charges $\pm\mathbf{E}$, expressed in terms of the electrostatic potential $\mathbf{h}$. This seemingly trivial rewriting makes the calculation of $x(\mathbf{E})$, the electric dimension in the presence of a background charge, physically transparent.

To properly take into account the curvature term we define the height field over a disc of radius R, instead of the infinite plane, keeping in mind that at the end of the calculation we need to take the limit $R\to\infty$. In the case of the disc $\mathcal{R} = 8\pi\delta(R)$, and the curvature term introduces a charge $-2\mathbf{E}_0$ at the disc boundary. Therefore, the vacuum of the modified Coulomb gas must contain a *floating charge* $+2\mathbf{E}_0$ in the disc interior, and the electrostatic energy of this charged vacuum is $\mathcal{E}_{2\mathbf{E}_0}^{(0)}(0,R) = -4{\rm E}_{0\alpha}^2\ln(R)/2\pi g_\alpha$. Now, to find the scaling dimension of a vertex operator of charge $\mathbf{E}$, we imagine placing charges $+\mathbf{E}$ and $-\mathbf{E}$ at points $\mathbf{x}$ and $\mathbf{y}$ in the disc interior, and we calculate the total electrostatic energy with respect to the charged vacuum. The floating charge being positive will coalesce with the negative charge $-\mathbf{E}$. Using Coulombs law, Eq. (A.7), we then calculate the interaction energy of charges $+\mathbf{E}$ at $\mathbf{x}$, $-\mathbf{E}+2\mathbf{E}_0$ at $\mathbf{y}$, and $-2\mathbf{E}_0$ at R, keeping in mind $R\gg|\mathbf{x}-\mathbf{y}|$. The final result

$$\mathcal{E}_{\mathbf{E}}(\mathbf{x},\mathbf{y}) = -\frac{1}{2\pi g_\alpha}E_\alpha(E_\alpha - 2E_{0\alpha})\ln|\mathbf{x}-\mathbf{y}| \tag{A.8}$$

is obtained after the energy of the charged vacuum is subtracted. Now it is a simple matter to read off the scaling dimension as the negative coefficient in front of the logarithm,

$$2x(\mathbf{E}) = \frac{1}{2\pi g_\alpha}E_\alpha(E_\alpha - 2E_{0\alpha}) \ . \tag{A.9}$$

This result can be derived in a more rigorous fashion by constructing the stress-energy tensor for the field theory $S_{\rm CG}$ and calculating its operator product with the vertex operator $\exp(i\mathbf{E}\cdot\mathbf{H})$ [62, 129].

A.2 Magnetic charges

To calculate the magnetic dimension $x(\mathbf{M})$ we consider the ratio of partition functions,

$$Z_{>\mathbf{M}}(\mathbf{x},\mathbf{y})/Z_{>} \sim |\mathbf{x}-\mathbf{y}|^{-2x(\mathbf{M})} \ . \tag{A.10}$$

$Z_{>\mathbf{M}}(\mathbf{x},\mathbf{y})$ is the sum (path integral) over height configurations where a vortex and an antivortex, of topological charge $\pm\mathbf{M}$, are placed at positions $\mathbf{x}$ and $\mathbf{y}$ of the basal plane, whilst $Z_{>}$ is the unconstrained sum:

$$Z_{>} = \int \mathcal{D}\mathbf{H} \ \exp\left(-\frac{1}{2}\int d^2\mathbf{x}\ g_\alpha(\boldsymbol{\partial}H^\alpha)^2\right) \ . \tag{A.11}$$

Here we have dropped the curvature term since it does not affect correlation functions of magnetic operators.

We can use the electrostatic analogy once again. Namely, we consider the interaction energy between two topological defects, $\mathcal{E}_{\mathbf{M}}(\mathbf{x},\mathbf{y})$. Since $Z_{>}$ is a Gaussian path integral, it follows that

$$Z_{>\mathbf{M}}(\mathbf{x},\mathbf{y})/Z_{>} = \exp[\mathcal{E}_{\mathbf{M}}(\mathbf{x},\mathbf{y})] \ , \tag{A.12}$$

where

$$-\mathcal{E}_{\mathbf{M}}(\mathbf{x},\mathbf{y}) = \frac{g_\alpha}{2\pi}(M^\alpha)^2 \ln|\mathbf{x}-\mathbf{y}| \ . \tag{A.13}$$

The above interaction energy is calculated as the Gaussian action of the the classical configuration of the height field, $\mathbf{h}_{\mathrm{c}}$. $\mathbf{h}_{\mathrm{c}}$ solves the classical equations of motion (Laplace's equation) with boundary conditions dictated by the presence of topological defects at $\mathbf{x}$ and $\mathbf{y}$ [127]. The scaling dimension of a magnetic-type operator is then the coefficient in front of the logarithm in Eq. (A.13),

$$2x(\mathbf{M}) = \frac{g_\alpha}{2\pi}(M^\alpha)^2 \ . \tag{A.14}$$

Bibliography

[1] S. Galam. *Rational group decision making: A random field Ising model at $T = 0$.* Physica A **238**, 66 (1997).

[2] N. W. Ashcroft and N. D. Mermin. *Solid state physics.* Saunders, Philadelphia (1976).

[3] L. Onsager. *Crystal statistics. I: A two-dimensional model with an order-disorder transition.* Phys. Rev **65**, 117 (1944).

[4] J. L. Cardy. *Conformal invariance.* In *Phase transitions and critical phenomena*, edited by C. Domb and J. L. Lebowitz, vol. 11, p. 54. Academic Press, London (1987).

[5] J. Cardy. *Scaling and renormalization in statistical physics.* Cambridge University Press, Cambridge (1996).

[6] T. Niemeijer and J. M. J. van Leeuwen. *Renormalization theory for Ising-like spin systems.* In *Phase transitions and critical phenomena*, edited by C. Domb and M. S. Green, vol. 6, p. 425. Academic Press, London (1977).

[7] B. Derrida. *Can disorder induce several phase transitions?* Phys. Rep. **103**, 29 (1984).

[8] A. B. Harris. *Effect of random defects on the critical behaviour of Ising models.* J. Phys. C **7**, 1671 (1974).

[9] A. Aharony. *Tricritical points in systems with random fields.* Phys. Rev. B **18**, 3318 (1978).

[10] Y. Imry and S.-K. Ma. *Random-field instability of the ordered state of continuous symmetry.* Phys. Rev. Lett. **35**, 1399 (1975).

[11] A. J. Bray and M. A. Moore. *Scaling theory of the random-field Ising model.* J. Phys. C **18**, L927 (1985).

[12] Y. Imry and M. Wortis. *Influence of quenched impurities on first-order phase transitions.* Phys. Rev. B **19**, 3581 (1979).

[13] K. Hui and A. N. Berker. *Random-field mechanism in random-bond multicritical systems.* Phys. Rev. Lett. **62**, 2507 (1989). *Ibid.* **63**, 2433 (1989) (erratum).

[14] M. Aizenman and J. Wehr. *Rounding of first-order phase transitions in systems with quenched disorder.* Phys. Rev. Lett. **62**, 2503 (1989).

[15] J. Cardy. *Effect of random impurities on fluctuation-driven first-order transitions.* J. Phys. A **29**, 1897 (1996).

[16] P. Pujol. *Effect of randomness in many coupled Potts models.* Europhys. Lett. **35**, 283 (1996).

[17] R. B. Potts. *Some generalized order-disorder transformations.* Proc. Camb. Phil. Soc. **48**, 106 (1952).

[18] F. Y. Wu. *The Potts model.* Rev. Mod. Phys. **54**, 235 (1982).

[19] R. J. Baxter. *Potts model at the critical temperature.* J. Phys. C **6**, L445 (1973).

[20] S. Chen, A. M. Ferrenberg, and D. P. Landau. *Monte Carlo simulation of phase transitions in a two-dimensional random-bond Potts model.* Phys. Rev. E **52**, 1377 (1995).

[21] S. Wiseman and E. Domany. *Critical behavior of the random-bond Ashkin-Teller model: A Monte Carlo study.* Phys. Rev. E **51**, 3074 (1995).

[22] M. Kardar, A. L. Stella, G. Sartoni, and B. Derrida. *Unusual universality of branching interfaces in random media.* Phys. Rev. E **52**, R1269 (1995).

[23] A. W. W. Ludwig and J. L. Cardy. *Perturbative evaluation of the conformal anomaly at new critical points with applications to random systems.* Nucl. Phys. B **285 [FS19]**, 687 (1987).

[24] A. W. W. Ludwig. *Critical behavior of the two-dimensional random q-state Potts model by expansion in* $(q-2)$. Nucl. Phys. B **285**, 97 (1987).

[25] A. W. W. Ludwig. *Infinite hierarchies of exponents in a diluted ferromagnet and their interpretation.* Nucl. Phys. B **330**, 639 (1990).

[26] V. Dotsenko, M. Picco, and P. Pujol. *Renormalisation-group calculation of correlation functions for the 2D random bond Ising and Potts models.* Nucl. Phys. B **455**, 701 (1995).

[27] M. Picco. *A study of cross-over effects for the 2D random bond Potts Model.* To appear (1998). Preprint cond-mat/9802092.

[28] C. Chatelain and B. Berche. *Finite-size scaling study of the surface and bulk critical behavior in the random-bond eight-state Potts model.* Phys. Rev. Lett. **80**, 1670 (1998).

[29] D. Stauffer. *Cross-over scaling, correlation function and connectivity in dilute low temperature ferromagnets.* Z. Phys. B **22**, 161 (1975).

[30] W. Kinzel and E. Domany. *Critical properties of random Potts models.* Phys. Rev. B **23**, 3421 (1981).

[31] P. W. Kasteleyn and C. M. Fortuin. *Phase transitions in lattice systems with random local properties.* J. Phys. Soc. Jap. (suppl.) **26**, 11 (1969).

[32] T. Bhattacharya, R. Lacaze, and A. Morel. *Large-q expansion of the 2D q-states Potts model.* J. Phys. I France **7**, 81 (1997).

[33] K. Binder. *Random-field induced interface widths in Ising systems.* Z. Phys. B **50**, 343 (1983).

[34] H. A. Kramers and G. H. Wannier. *Statistics of the two-dimensional ferromagnet.* Phys. Rev **60**, 252 (1941).

[35] F. Y. Wu. *Duality relations for Potts correlation functions.* Phys. Lett. A **228**, 43 (1997).

[36] J. L. Jacobsen. *Comment on "Duality relations for Potts correlation functions" by Wu.* Phys. Lett. A **233**, 489 (1997).

[37] F. Y. Wu and H. Y. Huang. *Sum rule identities and duality relation for the Potts n-point boundary correlation function.* Phys. Rev. Lett. **79**, 4954 (1997).

[38] W. T. Lu and F. Y. Wu. *On the duality relation for correlation functions of the Potts model.* J. Phys. A **31**, 2823 (1998).

[39] C.-N. Chen, C.-K. Hu, and F. Y. Wu. *Partition function zeros of the square lattice Potts model.* Phys. Rev. Lett. **76**, 169 (1996).

[40] V. Matveev and R. Shrock. *Some new results on complex-temperature singularities in Potts models on the square lattice.* Phys. Rev. E **54**, 6174 (1996).

[41] J. L. Jacobsen. In preparation.

[42] P. DiFrancesco, P. Mathieu, and D. Sénéchal. *Conformal field theory.* Springer Verlag, New York (1997).

[43] F. Y. Wu and Y. K. Wang. *Duality transformation in a many-component spin model.* J. Math. Phys. **17**, 439 (1976).

[44] H. W. J. Blöte and M. P. Nightingale. *Critical behaviour of the two-dimensional Potts model with a continuous number of states: A finite size scaling analysis.* Physica A **112**, 405 (1982).

[45] G. D. Birkhoff. Ann. Math. **14**, 42 (1912).

[46] J. Cardy and J. L. Jacobsen. *Critical behavior of random bond Potts models.* Phys. Rev. Lett. **79**, 4063 (1997).

[47] J. L. Jacobsen and J. Cardy. *Critical behavior of random bond Potts models: A transfer matrix study.* Nucl. Phys. B **515**, 701 (1998).

[48] J. L. Cardy, editor. *Finite-size scaling.* North-Holland, Amsterdam (1988).

[49] J. L. Cardy. *Critical exponents of directed self-avoiding walks.* J. Phys. A **16**, L355 (1983).

[50] H. W. Blöte, J. L. Cardy, and M. P. Nightingale. *Conformal invariance, the central charge, and universal finite-size amplitudes at criticality.* Phys. Rev. Lett. **56**, 742 (1986).

[51] I. Affleck. *Universal term in the free energy at a critical point and the conformal anomaly.* Phys. Rev. Lett. **56**, 746 (1986).

[52] H. Furstenberg. *Noncommuting random products.* Trans. Am. Math. Soc. **108**, 377 (1963).

[53] G. Benettin, L. Galgani, A. Giorgilli, and J.-M. Strelcyn. *Lyapunov characteristic exponents for smooth dynamical systems and for hamiltonian systems: A method for computing all of them.* Meccanica **15**, 9 (1980).

[54] H. W. J. Blöte and B. Nienhuis. *The phase diagram of the O(n) model.* J. Phys. A **22**, 1415 (1989).

[55] S. L. A. de Queiroz and R. B. Stinchcombe. *Connectivity-dependent properties of diluted sytems in a transfer-matrix description.* Submitted to Phys. Rev. E (1998). Preprint cond-mat/9804286.

[56] E. Buffenoir and S. Wallon. *The correlation length of the Potts model at the first-order transition point.* J. Phys. A **26**, 3045 (1993).

[57] S. L. A. de Queiroz. *Correlation decay and conformal anomaly in the two-dimensional random-bond Ising ferromagnet.* Phys. Rev. E **51**, 1030 (1995).

[58] F. D. A. Aarão Reis, S. L. A. de Queiroz, and R. R. dos Santos. *Universality and logarithmic corrections in two-dimensional random Ising ferromagnets.* Phys. Rev. B **56**, 6013 (1997).

[59] J. L. Cardy. *Logarithmic corrections to finite-size scaling in strips.* J. Phys. A **19**, L1093 (1986).

[60] A. B. Zamolodchikov. *"Irreversability" of the flux of the renormalization group in a 2D field theory.* Pis'ma Zh. Eksp. Teor. Fiz. **43**, 565 (1986). English translation in JETP Lett. **43**, 730 (1986).

[61] M. Picco. *Weak randomness for large q-state Potts models in two dimensions.* Phys. Rev. Lett. **79**, 2998 (1997).

[62] V. S. Dotsenko and V. A. Fateev. *Convormal algebra and multipoint correlation functions in 2D statistical models.* Nucl. Phys. B **240**, 312 (1984).

[63] Lederman, editor. *Handbook of applicable mathematics*, vol. 2: *Probability*, chap. 12.5.3. John Wiley and Sons, London (1980).

[64] M. Picco. *Numerical results for the two-dimensional random-bond three-state Potts model.* Phys. Rev. B **54**, 14930 (1996).

[65] M. Picco. Private communication.

[66] M.-A. Lewis. *Higher moments of spin-spin correlation functions for the ferromagnetic random bond Potts model.* Submitted to Europhys. Lett. (1998). Preprint cond-mat/9710312, version 3.

[67] J. T. Chayes, L. Chayes, D. S. Fisher, and T. Spencer. *Finite-size scaling and correlation lengths for disordered systems.* Phys. Rev. Lett. **57**, 2999 (1986).

[68] J. T. Chayes, L. Chayes, D. S. Fisher, and T. Spencer. *Correlation length bounds for disordered Ising ferromagnets.* Commun. Math. Phys. **120**, 501 (1989).

[69] J. L. Cardy. *Operator content of two-dimensional conformally invariant theories.* Nucl. Phys. B **270 [FS16]**, 186 (1986).

[70] J. L. Cardy. *Operator content and modular properties of higher-dimensional conformal field theories.* Nucl. Phys. B **366**, 403 (1991).

[71] M. P. Nightingale. *Scaling theory and finite systems.* Physica A **83**, 561 (1976).

[72] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical recipes in C: The art of scientific computing*, chap. 5.7. Cambridge University Press, Cambridge, 2 edn. (1992).

[73] Vi. Dotsenko, Vl. Dotsenko, M. Picco, and P. Pujol. *Renormalisation-group solution for the 2-dimensional random bond Potts-model with broken replica symmetry.* Europhys. Lett. **32**, 425 (1996).

[74] F. Pázmándi, R. T. Scalettar, and G. T. Zimányi. *Revisiting the theory of finite size scaling in disordered systems: ν can be less than $2/d$.* Phys. Rev. Lett. **79**, 5130 (1997).

[75] W. Janke and R. Villanova. *Two-dimensional eight-state Potts model on random lattices: a Monte Carlo study.* Phys. Lett. A **209**, 179 (1995).

[76] W. Janke and R. Villanova. *Monte Carlo study of 8-state Potts model on 2D random lattices.* Nucl. Phys. B (Proc. Suppl.) **47**, 641 (1996).

[77] A. Okabe, B. Boots, and K. Sugihara. *Spatial tessellations: Concepts and applications of Voronoi diagrams.* John Wiley and Sons, London (1992).

[78] C. F. Baillie, W. Janke, and D. A. Johnston. *Softening of first-order phase transition on quenched random gravity graphs.* Phys. Lett. B **388**, 14 (1996).

[79] G. H. Wannier. *Antiferromagnetism. The triangular Ising net.* Phys. Rev **79**, 357 (1950).

[80] S. T. Bramwell, S. G. Carling, C. J. Harding, K. D. M. Harris, *et al. The anhydrous alums as model triangular-lattice magnets.* J. Phys. Cond. Matt. **8**, L123 (1996).

[81] J. Stephenson. *Ising-model spin correlations on the triangular lattice. III.* J. Math. Phys. **11**, 413 (1970).

[82] H. W. J. Blöte and H. J. Hilhorst. *Roughening transitions and the zero-temperature triangular Ising antiferromagnet.* J. Phys. A **15**, L631 (1982).

[83] B. Nienhuis, H. J. Hilhorst, and H. W. J. Blöte. *Triangular SOS models and cubic-crystal shapes.* J. Phys. A **17**, 3559 (1984).

[84] E. H. Lieb and F. Y. Wu. *Two-dimensional ferroelectric models.* In *Phase transitions and critical phenomena*, edited by C. Domb and M. S. Green, vol. 1, p. 332. Academic Press, London (1972).

[85] E. H. Lieb. *Residual entropy of square ice.* Phys. Rev **162**, 162 (1967).

[86] H. van Beijeren. *Exactly solvable model for the roughening transition of a crystal surface.* Phys. Rev. Lett. **38**, 993 (1977).

[87] J. Kondev and C. L. Henley. *Kac-Moody symmetries of critical ground states.* Nucl. Phys. B **464 [FS]**, 540 (1996).

[88] M. T. Batchelor, H. W. J. Blöte, B. Nienhuis, and C. M. Yang. *Critical behavior of the fully packed loop model on the square lattice.* J. Phys. A **29**, L399 (1996).

[89] J. Kondev. *Liouville field theory of fluctuating loops.* Phys. Rev. Lett. **78**, 4320 (1997).

[90] J. L. Jacobsen and J. Kondev. *Field theory of compact polymers on the square lattice.* Submitted to Nucl. Phys. B (1998). Preprint cond-mat/9804048.

[91] M. T. Batchelor, J. Suzuki, and C. M. Yung. *Exact results for Hamiltonian walks from the solution of the fully packed loop model on the honeycomb lattice.* Phys. Rev. Lett. **73**, 2646 (1994).

[92] J. Kondev, J. deGier, and B. Nienhuis. *Operator spectrum and exact exponents of the fully packed loop model.* J. Phys. A **29**, 6489 (1996).

[93] J. Kondev and C. L. Henley. *Four-coloring model on the square lattice: A critical ground state.* Phys. Rev. B **52**, 6628 (1995).

[94] P.-G. de Gennes. *Scaling concepts in polymer physics.* Cornell University Press, Ithaca (1979).

[95] B. Nienhuis. *Exact critical point and critical exponents of O(n) models in two dimensions.* Phys. Rev. Lett. **49**, 1062 (1982).

[96] B. Duplantier and H. Saleur. *Exact tricritical exponents for polymers at the theta-point in two dimensions.* Phys. Rev. Lett. **59**, 539 (1987).

[97] C. J. Camacho and D. Thirumalai. *Minimum energy compact structures of random sequences of heteropolymers.* Phys. Rev. Lett. **71**, 2505 (1993).

[98] H. S. Chan and K. A. Dill. *Compact polymers.* Macromolecules **22**, 4559 (1989).

[99] B. Nienhuis. *Coulomb gas formulation of two-dimensional phase transitions.* In *Phase transitions and critical phenomena*, edited by C. Domb and J. L. Lebowitz, vol. 11, p. 1. Academic Press, London (1987).

[100] A. A. Belavin, A. M. Polyakov, and A. B. Zamolodchikov. *Infinite conformal symmetry in two-dimensional quantum-field theory.* Nucl. Phys. B **241**, 333 (1984).

[101] C. Itzykson, H. Saleur, and J.-B. Zuber, editors. *Conformal invariance and applications to statistical mechanics.* World Scientific, Singapore (1988).

[102] H. G. Evertz. *The loop algorithm.* To appear (1997). Preprint cond-mat/9707221.

[103] M. Aizenman and B. Nachtergaele. *Geometric aspects of quantum spin states.* Commun. Math. Phys. **164**, 17 (1994).

[104] J. T. Chalker and P. D. Coddington. *Percolation, quantum tunnelling and the integer Hall-effect.* J. Phys. C **21**, 2665 (1988).

[105] D.-H. Lee. *Network models of quantum percolation and their field-theory representations.* Phys. Rev. B **50**, 10788 (1994).

[106] H. Park and M. Widom. *Exact results on the antiferromagnetic three-state Potts model.* Phys. Rev. Lett. **63**, 1193 (1989).

[107] R. Raghavan, C. L. Henley, and S. L. Arouh. *New two-color dimer models with critical ground states.* J. Stat. Phys. **86**, 517 (1997).

[108] J. Bascle, T. Garel, and H. Orland. *Mean-field theory of polymer melting.* J. Phys. A **25**, L1323 (1992).

[109] B. Duplantier. In *Fundamental problems in statistical mechanics*, edited by H. van Beijeren, vol. 7. North-Holland, Amsterdam (1990).

[110] H. W. J. Blöte and B. Nienhuis. *Fully packed loop model on the honeycomb lattice.* Phys. Rev. Lett. **72**, 1372 (1994).

[111] D. A. Huse and A. D. Rutenberg. *Classical antiferromagnets on the Kagomé lattice.* Phys. Rev. B **45**, 7536 (1992).

[112] B. Duplantier and H. Saleur. *Exact critical properties of two-dimensional dense self-avoiding walks.* Nucl. Phys. B **290**, 291 (1987).

[113] H. Orland, C. Itzykson, and C. de Dominicis. *An evaluation of the number of Hamiltonian paths.* J. Phys. Lett. (Paris) **46**, L353 (1985).

[114] A. L. Owczarek, T. Prellberg, and R. Brak. *New scaling form for the collapsed polymer phase.* Phys. Rev. Lett. **70**, 951 (1993).

[115] T. G. Schmalz, G. E. Hite, and D. J. Klein. *Compact self-avoiding circuits on two-dimensional lattices.* J. Phys. A **17**, 445 (1984).

[116] S. Higuchi. *Field theoretic approach to the counting problem of Hamiltonian cycles of graphs.* To appear in Phys. Rev. E (1998). Preprint cond-mat/9711152.

[117] K. A. Dill, S. Bromberg, K. Yue, K. M. Fiebig, *et al. Principles of protein folding—A perspective from simple exact models.* Protein Science **4**, 561 (1995).

[118] J. Kondev and J. L. Jacobsen. *Conformational entropy of compact polymers.* Submitted to Phys. Rev. Lett. (1998). Preprint cond-mat/9805178.

[119] T. Garel, H. Orland, and E. Pitard. *Protein folding and heteropolymers.* In *Spin glasses and random fields*, edited by A. P. Young. World Scientific, Singapore (1997).

[120] A. E. Mirsky and L. Pauling. *On the structure of native, denatured, and coagulated proteins.* Proc. Natl. Acad. Sci. USA **22**, 439 (1936).

[121] W. Kauzmann. *Denaturation of proteins and enzymes.* In *The mechanism of enzyme action*, edited by W. D. McElroy and B. Glass, p. 70. Johns Hopkins Press, Baltimore (1954).

[122] N. Read. In *Proceedings of the Kagomé Workshop*, edited by P. Chandra. NEC Laboratories, Princeton (1992).

[123] J. L. Cardy. *Continuously varying exponents for oriented self-avoiding walks.* Nucl. Phys. B **419**, 411 (1994).

[124] A. Trovato and F. Seno. *Universality for interacting oriented self-avoiding walks: A transfer matrix calculation.* Phys. Rev. E **56**, 131 (1997).

[125] R. J. Baxter. *Exactly solved models in statistical mechanics.* Academic Press, New York (1982).

[126] J. Cardy. *Conformal invariance and statistical mechanics.* In *Fields, strings and critical phenomena*, edited by E. Brézin and J. Zinn-Justin. North-Holland, Amsterdam (1989).

[127] P. M. Chaikin and T. C. Lubensky. *Principles of condensed matter physics*, chap. 9. Cambridge University Press, Cambridge (1995).

[128] M. Aizenman. *On the number of incipient spanning clusters.* Nucl. Phys. B **485**, 551 (1997).

[129] V. S. Dotsenko and V. A. Fateev. *Four-point correlation functions and the operator algebra in 2D conformal invariant theories with central charge* $c \leq 1$. Nucl. Phys. B **251**, 691 (1985).

[130] J. Fuchs. *Affine Lie algebras and quantum groups.* Cambridge University Press, Cambridge (1992).

[131] J. Kondev and C. L. Henley. *Geometrical exponents of contour loops on random Gaussian surfaces.* Phys. Rev. Lett. **74**, 4580 (1995).

[132] H. Saleur and B. Duplantier. *Exact determination of the percolation hull exponent in two dimensions.* Phys. Rev. Lett. **58**, 2325 (1987).

[133] P. Ginsparg. *Applied conformal field theory.* In *Fields, strings and critical phenomena*, edited by E. Brézin and J. Zinn-Justin. North-Holland, Amsterdam (1989).

[134] A. Kast. *Correlation length and average loop length of the fully packed loop model.* J. Phys. A **29**, 7041 (1996).

[135] S. T. Chui and J. D. Weeks. *Phase transition in the two-dimensional Coulomb gas, and the interfacial roughening transition.* Phys. Rev. B **14**, 4978 (1976).

[136] J. V. José, L. P. Kadanoff, S. Kirkpatrick, and D. R. Nelson. *Renormalization, vortices, and symmetry-breaking perturbations in the two-dimensional planar model.* Phys. Rev. B **16**, 1217 (1977).

[137] R. J. Baxter. *Colouring of a hexagonal lattice.* J. Math. Phys. **11**, 784 (1970).

[138] P. J. Flory. *Thermodynamics of high polymer solutions.* J. Chem. Phys. **10**, 51 (1942).

[139] M. L. Huggins. *Some properties of long-chain compounds.* J. Phys. Chem. **46**, 151 (1942).

[140] P. W. Kasteleyn. *A soluble self-avoiding walk problem.* Physica **29**, 1329 (1963).

[141] B. Duplantier. *Critical exponents of Manhattan Hamiltonian walks in two dimensions, from Potts and O(n) models.* J. Stat. Phys. **49**, 411 (1987).

[142] H. Saleur. *The antiferromagnetic Potts model in two dimensions—Berker Kadanoff phase, antiferromagnetic transition, and the role of Beraha numbers.* Nucl. Phys. B **360**, 219 (1991).

[143] P. DiFrancesco. *Folding the square-diagonal lattice.* Cond-mat/9803051 (1998).

Index