Real-Space Renormalization from a Geometric Fractal Perspective

Marcel den Nijs

Department of Physics, University of Washington, P.O. Box 351560, Seattle, Washington 98195-1560

Abstract

An introduction to real-space renormalization transformation (RT) is presented with as focus geometric fractal aspects. Critical exponents are equivalent to fractal dimensions. Details of the construction rule of deterministic fractals and the definition of their fractal dimensions illustrate differences in approach between field theory and condensed matter physics. The appearance of fractal structures is typically less obvious within the context of partition functions (equilibrium statistical mechanics and field theory) and master equations (dynamic processes). A RT amounts to a reformulation of those construction rules into a form similar to the conventional ones for deterministic fractals, but on an algebraic level without ever explicitly referring to the geometric structure. This is illustrated explicitly in the context of the Sierpinsky Gasket. Its construction rule can be reformulated as a growth process and also as a partition function.

1 Introduction

In 1973 Hans van Leeuwen and Theodorus Niemeijer (vL&N) [1] transformed Kadanoff's [2] qualitative block spin explanation for the thermodynamic scaling laws of Widom [3] and Domb and Hunter [4], into a quantitative tool for calculating critical exponents at equilibrium phase transitions. They introduced a systematic cluster approximation scheme to evaluate critical exponents up to arbitrary accuracy and illustrated the method by applying it to the two dimensional (2D) Ising model on a triangular lattice numerically. This achievement triggered a flurry of activity during the following decade, when the real-space RT technique was refined, applied to numerous models, and used to map out many equilibrium universality classes [5,6]. The importance of real-space RT as a working tool to calculate scaling properties has waned in recent years. Other methods, such as Monte Carlo and transfer matrix based finite size scaling techniques are somewhat less elegant, but can

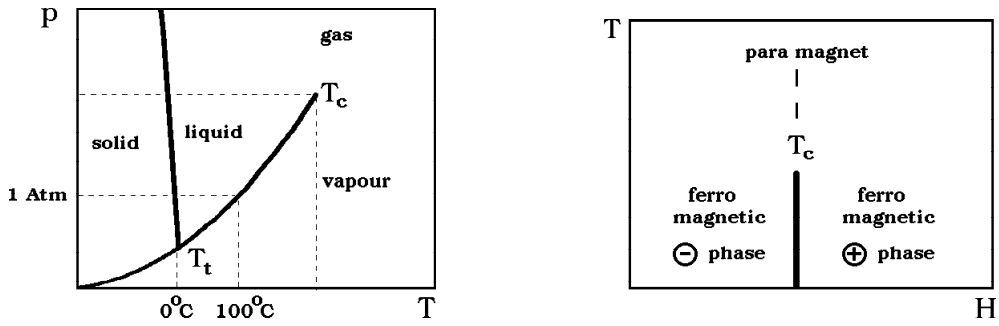

Fig. 1. Phase diagrams of gas-liquids and ferro magnets

be applied more directly to systems, e.g., with less pre-knowledge about their order parameters. However, the core RT concepts remain essential in the interpretation of such results, in particular the notion of fixed points, relevant versus irrelevant scaling fields, and their role in determining corrections to scaling.

Scale invariance has emerged as a unifying and common phenomena in all branches of physics and real-space RT has found its way into the standard statistical mechanics text books [7–9]. These texts present the conventional algebraic approach, following very closely the review by vL&N [5]. In my own graduate statistical mechanics course at the University of Washington I take a more geometric route. Scale invariant structures are not limited to partition function type construction rules. A geometric approach is more general and therefore more appropriate for the next generation of physicists. On a geometric level scale invariance implies the presence of fractal structures. Most senior physicists are aware of the connections between fractals and real-space RT, but a detailed presentation seems unavailable, in particular one suitable for beginning graduate students. This special issue of Physica dedicated to Hans van Leeuwen is an appropriate place to present such a discussion.

The outline of this paper is as follows: Section 2 contains a short introduction to scale invariance at equilibrium phase transitions. In Section 3 it is shown how fractality and scale invariance appear in equilibrium critical phenomena in the context of finite size scaling (FSS) using the 2D Ising model as example. Section 4 contains an overview of the properties of deterministic fractals. Their conventional recursive construction rule is nothing else than a real-space RT. The various methods to determine their fractal dimensions are closely related to techniques used to determine scaling properties in field theory and condensed matter physics. The construction rules of most physical processes are not in such a recursive form, and the presence of a fractal scale invariant structure is typically obscured. A RT is a reformulation of such rules into a recursive form reminiscent of those for deterministic fractals. The Sierpinsky gasket (SP-gasket) (one of the deterministic fractals) will be used to illustrate this. In section 5 the SP-gasket is generated in the context of a dynamic

process, and its scaling properties are determined from that growth rule by means of a real-space RT. In section 6 the SP-gasket is generated in terms of an Ising type partition function and it fractal dimensions are obtained by a classic vL&N type real-space RT.

2 SCALING IN CRITICAL PHENOMENA

The critical point in phase diagrams of gas-liquids is the end point of the boiling line where the liquid and the vapour phases become indistinguishable, see Fig. 1. Their density difference, the order parameter $\Delta\rho = \rho_l - \rho_v$, vanishes. Cagniard de la Tour ($\simeq$ 1820) is credited with having been the first to observe the disappearance of the meniscus between the gas and liquid. Andrews ($\simeq$ 1870) performed the first systematic studies of the liquification of CO_2 [10].

The Curie point, T_c, in easy-axis ferro magnets and Ising models, is the point where the distinction between the two ferro magnetic phases (with the majority of spins pointing up or down) vanishes. The ferro magnet becomes a para magnet. At temperatures high compared to the interaction strength, the spins behave independently. Below T_c they collectively prefer to point in the up or down direction. This is called spontaneous symmetry breaking. To refer to a symmetry is quite misleading however. In gas-liquids there is no symmetry between the vapour and the liquid phase. The essential point is the appearance in phase-space of two non-communicating sectors; the vapour phase and the liquid phase. Along the boiling line these two sectors switch stability. At those values of μ and T their pressures $p(\mu, T)$ become equal, and therefore the vapour and liquid coexist. Similarly, the free energies $f(h, T)$ of the up and down phase become equal along the first-order transition line in the magnet. We will use the magnetic language from now on (the canonical instead of the grand canonical ensemble).

The inter-molecular interactions are responsible for the spontaneous appearance of non-communicating sectors in phase space. Liquids are high-density collectively self-trapped clouds of molecules. They are adequately described in terms of molecular field approximations such as the van der Waals equation of state. The following argument, attributed to Peierls, explains why and when these two sectors fail to communicate. Along the boiling line the two phases coexist. Vapour bubbles in the liquid cost only surface free energy. Their surface tension is roughly of the form $\eta = E_{int} - TS_{wave}$, with E_{int} the interface energy per unit length, and S_{wave} the entropy of capillary waves in the interface. Imagine a thermodynamic fluctuation which evaporates the entire liquid. It would require an intermediate state where one of the bubbles grows and spans the entire system. The free energy barrier between the two universes is therefore of order $\Delta F = \eta L^{D-1}$. This barrier becomes infinitely

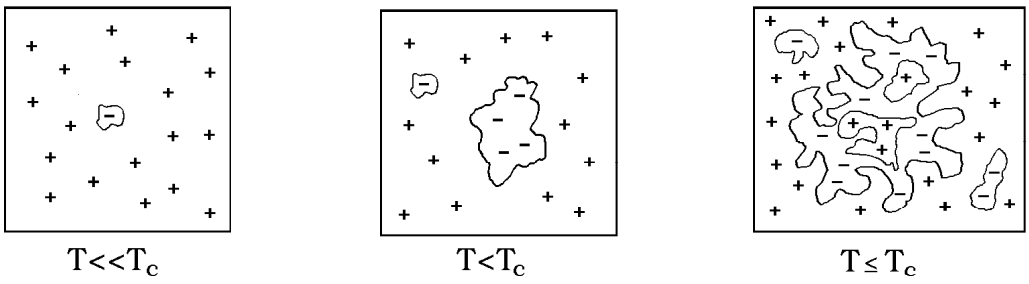

Fig. 2. Droplet excitation shape development with temperature

high in the thermodynamic limit, except below the so-called lower critical dimension $D_L = 1$. Therefore, phase transitions between discrete sets of coexisting phases are impossible in 1D systems with short range interactions [11]. The same argument yields $D_L = 2$ for systems with continuous degrees of freedom (so-called Goldstone modes). Elastic deformations in solids are an example of this. 2D solids of adsorbates on smooth substrates have powerlaw peaks instead of Bragg peaks [12].

At the critical point the two universes start to communicate. The free energy barrier vanishes because the surface tension goes to zero. It vanishes as $\eta \sim |T - T_c|^{\nu}$. The correlation length diverges as $\xi \sim |T - T_c|^{-\nu}$. The coefficient ν is one of the critical exponents. Two other examples are α and β. The order parameter vanishes at T_c as $\Delta\rho \sim |T - T_c|^{\beta}$, and the specific heat diverges as $C_v \sim |T - T_c|^{-\alpha}$. The values of these critical exponents disagree with the ones predicted by molecular field theory ($\alpha = 0$(disc), $\beta = 1/2$, and $\nu = 1/2$). Early observations of such differences where ignored; e.g., those by Verschaffelt in Leiden around 1900 [13]. This issue was not investigated seriously until Onsager solved the 2D Ising model in 1944 ($\alpha = 0$(log), $\beta = 1/8$, and $\nu = 1$)[14]. He proved that non-mean field values for the critical exponents are possible within the framework of statistical mechanics.

The critical exponents are universal. The same values, $\alpha = 0.11$, $\beta = 0.33$, and $\nu = 0.63$, apply to all 3D gas-liquids, easy-axis ferro magnets, binary mixtures, and the Ising model. Univerality emerged as an empirical fact between 1945-1965. The critical exponents turned out to be independent of most details of microscopic interactions. Only the number of coexisting phases, their symmetries, and the dimension D matter. An example of a different universality class is the so-called 2D 3-state Potts class. It applies to systems with three coexisting phases with full permutational symmetry ($\alpha = 1/3$, $\beta = 1/9$, and $\nu = 5/6$) [15,16]. Monolayers of He or Kr adsorbed on graphite are a realization of this [17,18].

Qualitatively te origin of universality can be explained as follows. Consider the droplet excitations, see Fig. 2. The typical droplet size ξ and the surface tension η are linked to each other by the thermodynamic uncertainty relation

$\Delta f \simeq k_B T$. The free energy of a droplet is proportional to $\Delta f \sim \eta \xi^{D-1}$. Therefore, $\xi \sim \eta^{1/(D-1)}$, if we measure η in units of $k_B T$. Initially, capillary wave excitations dominate the reduction of the surface tension and therefore the increase in droplet size. But close to T_c the droplets start to entangle. The topological rules of how these interfaces (domain walls) can cross and merge and their symmetry properties become the limiting factors. They determine how the entropy increases in the direct vicinity of T_c and therefore the value of the critical exponents. Most details of the interactions at the microscopic level do not matter.

In 1965 Widom [3] and Domb and Hunter [4] formulated the thermodynamic scaling postulate. They proposed that all experimental and theoretical evidence could be explained if the the singular part of the free energy takes the form of a generalized homogeneous function

$$f_{sing}(u_T, u_H) = b^{-D} f_{sing}(b^{y_T} u_T, b^{y_H} u_H). \tag{1}$$

b is an arbitrary (scaling) parameter; $u_T \sim T - T_c$ and $u_H \sim H$ are scaling fields defined locally at T_c. $u_H = 0$ coincides with boiling line in gas-liquids. y_T and y_H are the fundamental critical exponents. For the Ising univerality class they take the values: $y_T = 1.59$ and $y_H = 2.48$ in 3D, and $y_T = 1$ (log) and $y_H = 15/8$ in 2D. All thermodynamic exponents can be expressed in these fundamental ones; e.g., $\alpha = (2y_T - D)/y_T$ and $\beta = (D - y_H)/y_T$. The basic set of exponents is larger for other universality classes. For example, for the 3-state Potts universality class in 2D: $y_T = 6/5$, $y_{H,1} = 28/15$, and $y_{H,2} = 2/3$.

Kadanoff [2] explained the empirical thermodynamic scaling law in terms of the appearance of "droplets inside droplets on all length scales" (see Fig. 2). and introduced a block-spin transformation to describe this. This was the first (but only qualitative) real-space RT. The block spin picture is stronger than the thermodynamic scaling postulate, Eq. (1). It implies Fisher's finite size scaling postulate

$$f_s(u_T, u_H; L) = b^{-D} f_s(b^{y_T} u_T, b^{y_H} u_H; b^{-1} L) \tag{2}$$

and also that correlation functions, such as the spin-spin correlation function g_H, scale as

$$g_H(u_T, u_H; r) = b^{-2(D-y_H)} g_H(b^{y_T} u_T, b^{y_H} u_H; b^{-1} r). \tag{3}$$

Both lead to the so-called hyperscaling relation, $\nu = 1/y_T$. How to interpret such scaling laws is the topic of the next sections. The review papers by Fisher [19] and Kadanoff *et. al.* [20], and also the first edition of the text book by Stanley [21] summarize the status of the field in the late sixties at the threshold of the renormalization era.

Renormalization traces it origins also within quantum field theory (QFT).

There it arose independently as a cure to divergencies in perturbation theory. Path integrals in d+1 dimensional QFT are equivalent to partition functions in D=d+1 Statistical Mechanics. The two fields merged in the early 1970-ties when Ken Wilson for the first time explicitly formulated the concept of renormalization [22] and momentum-space type renormalization was introduced in the ϕ^4-theory formulation of Ising type critical behaviour [23–25]. Following this, van Leeuwen and Niemeijer took-up Kadanoff's block-spin description of scale invariance, transformed it into a quantitative tool, and thereby introduced the real-space RT method.

3 Fractal geometry and finite size scaling at T_c.

The founders of scaling did not have the luxury of computer graphics. Nowadays it is easy to generate typical configurations, e.g., by Monte-Carlo (MC) simulations. Fig. 3 is an example of an Ising spin configuration just below T_c [26]. Kadanoff's "droplets inside droplets" are clearly visible. Moreover, also their coast lines "wiggle inside wiggles". Mandelbrot [27] popularized the word fractal for such structures in the late 1970-ties. We will see that each critical exponent in the basic set $\{y_T, y_H, ...\}$ of Eq. (2) represents a fractal dimension.

Consider a square lattice with on each site a spin, $S_n = \pm 1$. These Ising spins are subject to a magnetic field and nearest neighbour spin-spin interactions

$$\mathcal{E} = -J \sum_{<i,j>} S_n S_{n+1} - h \sum_i S_n \tag{4}$$

such that nearest neighbours prefer to be alligned. Define the magnetiation and energy density as derivatives of the free energy in terms of the dimensionless coupling constants $K = J/k_BT$ and $H = h/k_BT$

$$m\ (K,H) = -\left(\frac{\partial f}{\partial H}\right)_K \qquad e\ (K,H) = -\left(\frac{\partial f}{\partial K}\right)_H \tag{5}$$

with

$$\mathcal{Z} = \exp[-Nf(K,H)] = \sum_{\{S_n\}} \exp[-\mathcal{E}/k_BT]. \tag{6}$$

The total magnetization measures the area of the "land mass" of up-spins versus that of the "sea" of down spins,

$$M(K,H) = m(K,H)L^D = \sum_i \langle s_i \rangle = \langle N_+ - N_- \rangle. \tag{7}$$

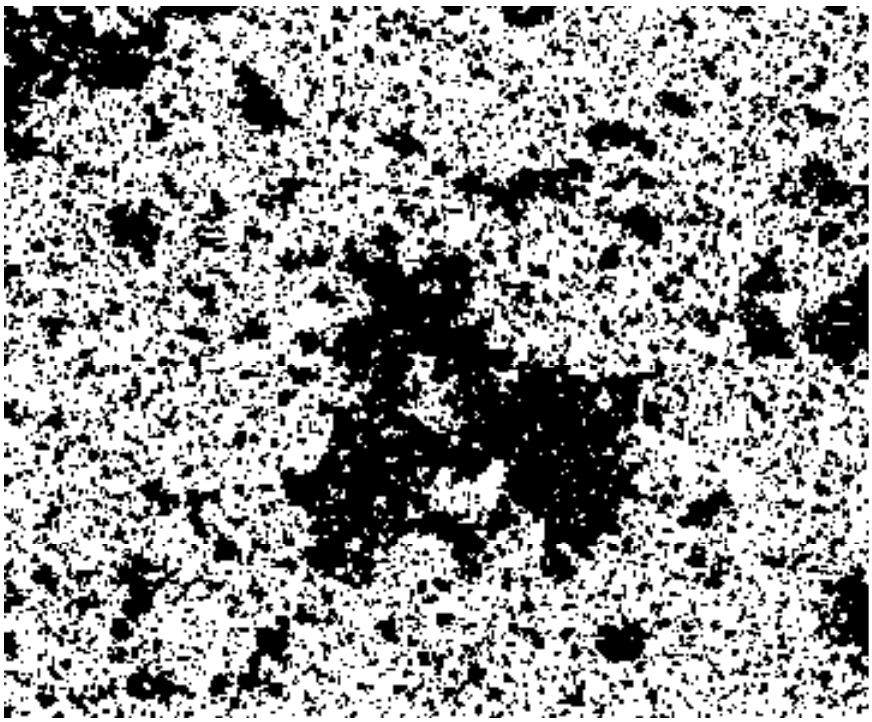

Fig. 3. Typical Ising configuration just below T_c

The free energy FSS postulate implies the following scaling form for the magnetization. The derivative of Eq. (2) with respect to $u_H \sim H$ yields

$$m(u_T, u_H; L) = b^{-D+y_H} m(b^{y_T} u_T, b^{y_H} u_H; b^{-1} L), \tag{8}$$

i.e., that the total magnetization scales as

$$M(u_T, u_H; L) = b^{y_H} M(b^{y_T} u_T, b^{y_H} u_H; b^{-1} L). \tag{9}$$

This means that $M \sim L^{y_H}$ at T_c (set $u_T = u_H = 0$ and choose $b = L$). We say that M scales with dimension y_H. Consider a typical configuration from a large MC run. Define a window with a tunable diameter L. Center it at an up-spin, more or less in the center of a large droplet. The total magnetization inside this window scales as $M \sim L^{y_H}$. In the fractal literature, y_H is known as the Haussdorfer dimension [28].

There are more fractal dimensions hiding in critical configurations. The energy

$$E = \sum_{\langle i,j \rangle} \langle s_i s_j \rangle = \langle N_{++} + N_{--} - N_{+-} - N_{-+} \rangle \tag{10}$$

measures the length of the interface "coastline". The derivative of Eq. (2) with respect to $u_T \sim K - K_c$ implies the following scaling form for the energy

$$E_s(u_T, u_H; L) = b^{y_T} E_s(b^{y_T} u_T, b^{y_H} u_H; b^{-1} L). \tag{11}$$

The singular part of the energy scales at T_c with dimension y_T, as $E_s \sim L^{y_T}$. The total energy includes also an analytic contribution and therefore scales as $E \simeq AL^D + BL^{y_T}$. y_T appears only in a so-called correction to scaling term (the analytic contribution is absent in the magnetization because of spin up-down symmetry).

We just identified the scaling dimensions of three geometric aspects: the unity operator (the total area inside the window, the analytic part of the free energy), the magnetization (the difference in area covered by up versus down spins), and the energy operator (the length of the interface coastline). Do other geometric aspects scale with their own fractal dimension as well? The set of so-called relevant scaling dimensions, those with positive values $y_{rel} > 0$, is finite. This number is set by the phase-rule. There are only two of them for critical behaviour in the Ising universality class. The coexistence surface must be of co-dimension 1 in any generic phase diagram since only two phases coexist. The critical surface has co-dimension 2 since it forms its edge. y_H controls the thermodynamic singularities in the direction across and y_T in the direction along the coexistence surface.

The critical surface is spanned by the set of so-called irrelevant scaling fields, u_{ir}, those with negative fractal dimensions $y_{ir} < 0$. These irrelevant scaling fields can be incorporated into the FSS scaling postulate

$$f_s(u_T, u_H, u_{ir}; L) = b^{-D} f_s(b^{y_T} u_T, b^{y_H} u_H, b^{y_{ir}} u_{ir}; b^{-1} L). \tag{12}$$

It is straight forward to show that the leading thermodynamic singularities along any path across the critical plane $u_H = u_T = 0$ are identical to those at the "fixed point" where all $u_\alpha = 0$. This is the explanation of universality within the context of the scaling postulate. The irrelevant scaling fields are associated with all other possible geometric aspects that can be represented in terms of local operators. For example the number of times the coastline touches itself can be counted by the next nearest neighbour spin-spin operator. In this sense the number of fractal dimensions in critical configurations is infinite.

Something special happens in 2D equilibrium critical phenomena. They are conformal invariant and described by 2D conformal field theory (CFT) [29]. In geometric language, conformal invariance implies that on average critical configurations, like Fig. 3, look the same under any complex function distortion of space, $z = x+iy$. For many 2D universality classes "the CFT algebra closes", and the set of relevant scaling dimensions $\{y_T, y_H, ...\}$ locks-in the values of all irrelevant exponents as well. They all differ from the relevant ones by integer amounts, e.g., as $y_T - n$ or $y_H - m$.

Even in 2D the finite set $\{y_T, y_H, ...\}$ does not fully characterize the fractal. It is possible to ask geometrical questions that cannot be expressed in terms of local Ising spin operators (nor in terms of dual so-called disorder type string

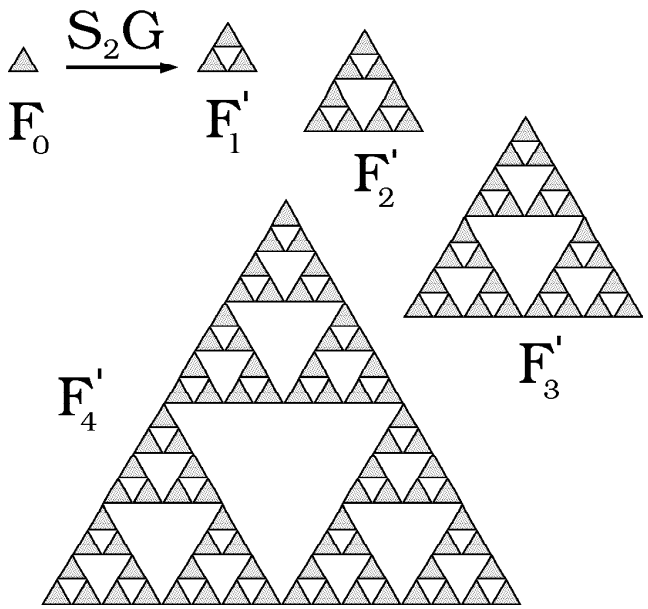

Fig. 4. Sierpinsky Gasket, $y_A = \log(3)/\log(2)$

operators). A typical example is the area of the backbone of the droplets, where all singly connected branches do not count.

In ϕ^4-theory the irrelevant scaling fields of the Ising universality class are described by the higher moments and gradients of the magnetization, like ϕ^n with $n > 4$. This suggests a connection between irrelevant scaling fields and multi-fractality [28].

Conformal invariance must be applicable beyond partition function type construction rules (statistical mechanics, QFT, and master equations). Suppose we convince ourselves that a particular growth mechanism for 2D trees, cracks, or whatever, creates structures that are (asymptotically) conformal invariant. CFT then provides us with a complete list of possible sets $\{y_h, y_T, ...\}$ of fractal dimensions [29].

4 Deterministic fractals

The equilibrium fractal structures discussed above are statistical in nature, defined in terms of an average over a set of configurations. The following elementary fractals are deterministic [27]. Fig. 4 shows the Sierpinsky Gasket (SP-gasket). The area of the up-triangles scales with $y_A = \log(3)/\log(2)$. Fig. 5 shows the Peano curve. Its coast line is space filling, i.e., scales with $y_L = \log(9)/\log(3) = 2$. The Peano curve is a special case of the so-called

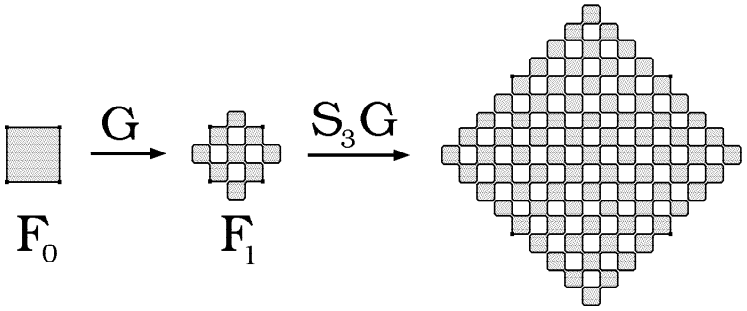

Fig. 5. Peano's space filling curve, $y_L = \log(9)/\log(3) = 2$

quadratic Koch islands. Fig. 6 shows an example with coastal dimension $y_L = \log(8)/\log(4) = 1.5$.

The conventional construction rule for these objects is iterative in nature. Define an initial figure, $\mathcal{F}_0$ (the initiator) and a modification rule $\mathcal{G}$ (the generator). For the SP-gasket the conventional initiator is a single up-triangle, and $\mathcal{G}$ is the rule that each up-triangle be divided into three up-triangles, see Fig. 4. $\mathcal{F}_n = \mathcal{G}^n \mathcal{F}_0$ is the n-th stage figure. The limiting process

$$\lim_{N\to\infty} \mathcal{F}_n = \lim_{N\to\infty} \mathcal{G}^n \mathcal{F}_0 \tag{13}$$

defines the fractal. Each application of $\mathcal{G}$ enhances the resolution by a scale factor b. The lattice cutoff shrinks as $a_n = a_0/b^n$, while the overall size of the object ξ remains constant. I call this the "field-theory construction rule". ξ is like the size of a particle in QFT. Lattice cutoffs are introduced in gauge theories to regulate divergencies, and are removed at the end (after solving the lattice model) by a procedure analogous to Eq.(13). In condensed matter physics the lattice cutoff is typically a given constant. Instead, the "droplet size" ξ diverges. Define a scale transformation $\mathcal{S}_b$, which blows-up the figure by a factor b. The alternative rule

$$\lim_{N\to\infty} \mathcal{F}'_n = \lim_{N\to\infty} (\mathcal{S}_b \mathcal{G})^n \mathcal{F}_0 \tag{14}$$

creates the same fractal structure, but with $a_n = a_0$ and $\xi_n = b^n \xi$. I call this the "condensed matter construction rule". Later, we will identify the inverse process with a renormalization transformation.

$$\mathcal{RT} = (\mathcal{S}_b \mathcal{G})^{-1}. \tag{15}$$

RT's involve two steps: (i) Removal of all details at length scales $a \le l < ba$, $\mathcal{G}^{-1}$. (ii) Restoration of the cutoff to its original value, $\mathcal{S}_b^{-1}$.

The fractal dimenions can be calculates in several ways. Imagine a fully developed fractal in the field-theory representation, $\mathcal{F}_\infty$. Cover it with an infinite

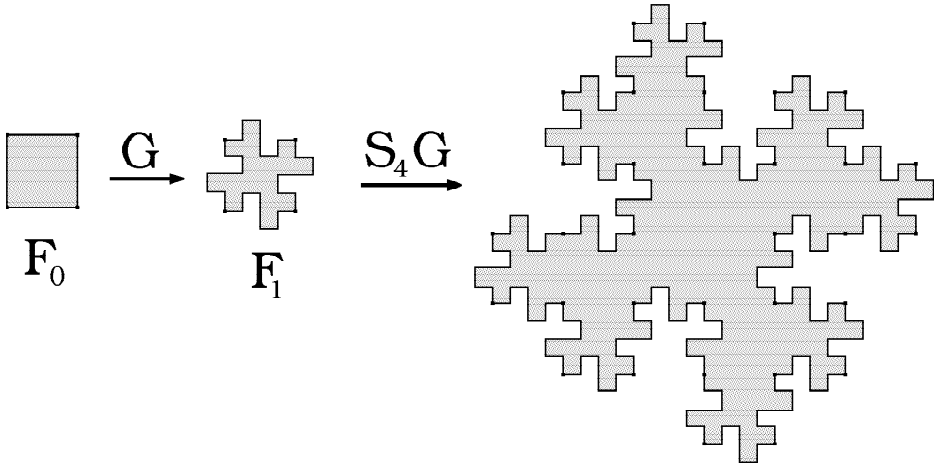

Fig. 6. Koch Island, $y_L = \log(8)/\log(4) = 1.5$

sequence of nested windows $W_n = b^D W_{n+1}$ with decreasing area's. $A(W_n)$ is the amount of aspect of type A (area, length,...) inside window W_n. A fractal dimension can be defined as

$$\lim_{n\to\infty} \frac{A(W_n)}{A(W_{n+1})} = b^{y_A}. \tag{16}$$

This definition reproduces the usual values of y_A for conventional objects. For example, center this set of windows on top of an arbitrary analytic curve in D dimensional space. Measure the total length $L(W_n)$ of the curve inside each window. The curve must become a straight line at small enough length scales (by definition, since it is analytic), therefore

$$\lim_{n\to\infty} \frac{L(W_n)}{L(W_{n+1})} = b^{y_L} \tag{17}$$

with $y_L = 1$, as it should be for any one dimensional object.

For the SP-gasket the natural choice for the scale factor is $b = 2$. Choose rectangular shaped windows with aspect ratio $\sqrt{3}$. Center them such that they share their left lower corner and are placed on top of the left corner of an up-triangle. "Obviously" the "number" of up-triangles inside each $W(n)$ is a factor 3 larger than in each $W(n+1)$. Therefore, the fractal dimension of the area of up-triangles is equal to $y_A = \log(3)/\log(2)$. Non of these $A(W_n)$ is countable however. Each of them is "zero". Our Euclidean yard sticks can deal only with objects of integer dimensionality. Since we measure area we get stricktly speaking

$$\frac{A(W_n)}{A(W_{n+1})} = \frac{3 \times \text{“}zero\text{”}}{\text{“}zero\text{”}} = 3. \tag{18}$$

Renormalization emerged in QFT to regulate precisely these types of diver-

gencies. Measuring lengths does not help. The total length of all line segments inside the windows yields

$$\frac{L(W_n)}{L(W_{n+1})} = \frac{3 \times \text{“}\infty\text{”}}{\text{“}\infty\text{”}} = 3. \tag{19}$$

The SP-gasket has the peculiar property that the area of its up-triangles scales with the same dimension as their coast lines, $y_A = y_L = \log(3)/\log(2)$.

Imagine a fully developed condensed matter type fractal ${\mathcal{F}'}_\infty$. This requires a piece of paper of infinite size, i.e., that we take the so-called thermodynamic limit. Cover ${\mathcal{F}'}_\infty$ with a set of nested windows $W_n = b^{-D} W_{n+1}$ with increasing area's . We are zooming-out instead of zooming-in. The fractal dimension is now defined as

$$\lim_{n\to\infty} \frac{A(W_n)}{A(W_{n+1})} = b^{-y_A}. \tag{20}$$

This definition is equivalent to the FSS approach described in the previous section. The cut-off removes all counting and measure problems.

The field theoretical zooming-in type definition focusses on local features. For example, it is possible to center the nested windows onto the edge of the fractal. Most fractal structures (but not the SP-gasket) have different scaling dimensions for the same aspect at their edge than inside their bulk. With the condensed matter type construction rule (in its naive form) the boundary vanishes to infinity, and the FSS definition of y_A naturally reproduces the bulk value of y_A.

The proper definition of fractal dimension is in terms of a limiting process applied to partially developed fractals ${\mathcal{F}'}_n$. Let $A({\mathcal{F}'}_n)$ be the total amount of aspect A inside object ${\mathcal{F}'}_n$. The following procedure to calculate y_A

$$\lim_{n\to\infty} \frac{A({\mathcal{F}'}_{n+1})}{A({\mathcal{F}'}_n)} = b^{y_A} \tag{21}$$

avoids the measure problem of the field-theory rule and also the thermodynamic limit problem of the condensed-matter rule. Let's apply this to Koch Islands, like those in Fig. 5 and Fig. 6. Their coast lines have fractal dimension $y_L = \log(p)/\log(b)$ because during each $\mathcal{S}_b\mathcal{G}$-step all line segments become a factor p longer.

The set of figures ${\mathcal{F}'}_n$ reminds us of how droplets evolve along the coexistence line towards T_c. Compare Fig. 2 to Figs. 4-6. In the next sections we will find out how Eq.(21) leads naturally to the scaling postulates of section 2 and to renormalization flow through a phase diagram.

5 The Sierpinsky Gasket as a growth process

The conventional construction rules for deterministic fractals presented in the previous section lead naturally to scale invariant structures. Ising model critical configurations are constructed according to the partition function rule. It also leads to scale invariance, but this is far from obvious from simply looking at this rule. In this and the next section we are going to play a game. We will define a dynamic Markov process and a partition function which obviously give rise to the SP-gasket. However, we must pretend not to be able to draw nor visualize this geometric structure (Fig. 4). It is like not being able to generate Fig. 3 for the Ising model. Instead, we are required to demonstrate scale invariance algebraically directly at the level of these two alternative construction rules. It takes some self-discipline (to distinguish honestly between what you know and what you are not supposed to know) but this is worth it.

Consider the following Markov process on a triangular lattice, see Fig. 7. Associate with each site an Ising spin. The site is occupied (empty) if $S_{i,\tau} = +1(-1)$. Interpret the vertical (down) axis as time τ. At time $\tau = 1$ only one site is occupied. The time evolution is deterministic. The value of each spin $S_{i,\tau}$ is determined by the values of the two just above it as

$$S_{i,\tau} = -S_{i-\frac{1}{2},\tau-1} S_{i+\frac{1}{2},\tau-1}. \tag{22}$$

This rule must be satisfied in every down-triangle of the triangular lattice.

The "MC-simulation" is rather trivial. By inspection it becomes clear very quickly that this rule produces the SP-gasket up to level $\mathcal{F}'_n$ at time $\tau = 2^N$. (Each up-spin represents an up-triangle in Fig. 4). But we are not supposed to know this in our game.

We are allowed to record the time evolution of observables like the total mass. This is analogous to a measurement of thermodynamic quantities at gas-liquid critical points. The total mass M of the space-time cluster does not grow uniformly. Mass extinctions at all time scales create a very irregular curve $M(\tau)$. We will discover however empirically the scaling law

$$M(\tau) \sim \tau^{y_A} \quad with \quad y_A = \log(3)/\log(2). \tag{23}$$

if we plot M only at times $\tau = 2^n$. How can we derive this algebraically without being aware of the SP-gasket configuration explicitly? We will acchieve it by applying the inverse of the "condensed matter construction rule", Eq.(14), directly onto the Markov process.

The empirical fluctuations in $M(\tau)$ suggest we try scale factor $b = 2$, and reformulate the local update rule of the spins at time τ in terms of those at

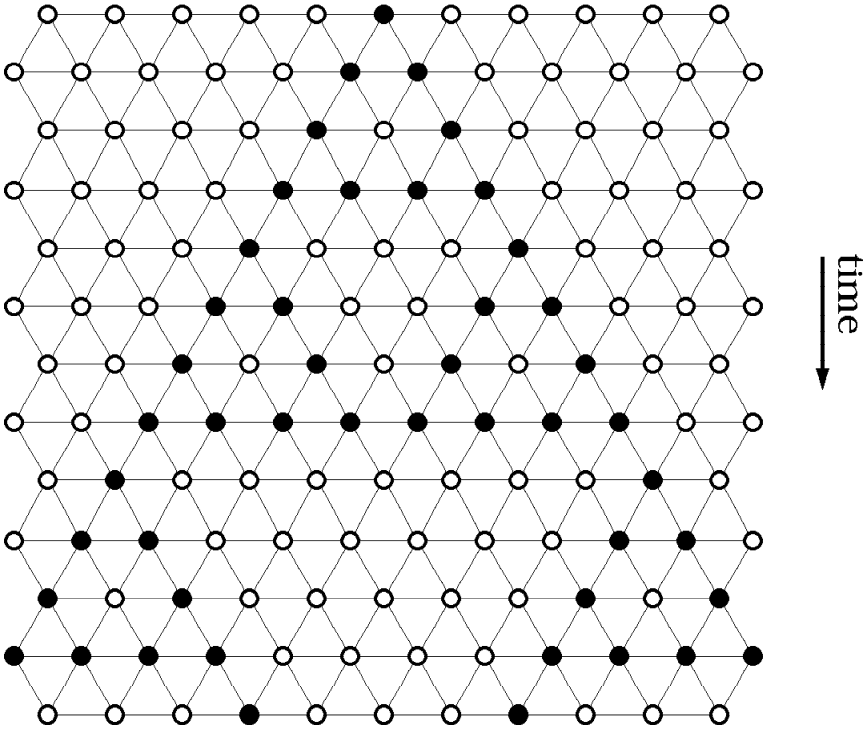

Fig. 7. Sierpinsky gasket as a growth pattern

time $\tau-2$ instead of $\tau-1$. The original Markov process will be replaced by a new one for spins at even times and every other site only. The temporal and spatial resolution will be decreased by a factor 2. In general, such a new rule is more complex than the original one, but we are "lucky". The spins on the triangular space-time lattice can be grouped into four sublattices. The rule reproduces itself exactly within each sublattice separately.

$$\begin{aligned} S_{i,\tau} &= -S_{i-\frac{1}{2},\tau-1}S_{i+\frac{1}{2},\tau-1} = -[-S_{i-1,\tau-2}S_{i,\tau-2}][-S_{i,\tau-2}S_{i+\frac{1}{2},\tau-2}] \\ &= -S_{i-1,\tau-2}S_{i+1,\tau-2}. \end{aligned} \tag{24}$$

The value of spin $S_{i,\tau}$ depends only on $S_{i+1,\tau-2}$ and $S_{i-1,\tau-2}$ which are part of the same sublattice. As far as the dynamics is concerned, we never have to look at the spins on the other 3 sublattices. They are linked only by the initial condition at time $\tau = 1, 2$. When you start the growth rule on one sublattice you must start it on two of the three others as well according to the original rule. This must give rise to three identical structures. Therefore, the total mass of the entire structure at time τ is equal to three times the mass on one sublattice, $M(\tau) = \frac{1}{3}M_s(\tau)$. Moreover, the growth rule on each sublattice is identical to that on the entire lattice. Therefore, the total mass of entire structure at time τ/b is equal to the mass on one sublattice at time τ. These two facts together imply that

$$M(\frac{\tau}{b}) = \frac{1}{3}M(\tau) \tag{25}$$

and confirms the empirical scaling relation Eq.(23). This is an example of a renormalization transformation. We implemented the inverse of the "condensed matter construction rule", Eq.(15), directly onto the growth rule, and thus derived Eq.(21) algebraically.

6 A real-space RT for the Sierpinsky Gasket

An important aspect of RT's is their representation as flows through phase diagrams. To illustrate this, we need to create the SP-gasket in terms of a generating function and a tunable parameter. Consider the set of finite resolution SP-Gaskets, $\mathcal{F}'_n$, see Fig. 4. Assign a chemical potential μ to every up-triangle. Each $\mathcal{F}'_n$ has a Boltzmann weight $\exp[-\mu M_n]$, with M_n its total mass. For each value of μ the thermodynamic average is centered around a finite resolution n. The fully developed fractal appears only at $\mu = 0$. By reducing μ we improve the resolution of the fractal, in complete analogy with droplet size evolution along the boiling line in gas liquids.

The following partition function generates this ensemble average. It looks somewhat complicated, since we insist on using the Ising model formalism, such that we can implement the vL&N real-space RT scheme literally. Consider the same triangular lattice with Ising spins $S(i,j) = \pm 1$ as in the previous section and Fig. 7, but time τ is now a second spatial coordinate j.

$$\exp[-F(\mu)] = \mathcal{Z} = \sum_{\{S_{i,j}\},\star} \exp[-\frac{1}{2}\mu \sum_r (S_r + 1)]. \tag{26}$$

$r = (i,j)$ denotes the position of each spin and $\star$ represents the following constraints on the allowed spin configurations. In each down-triangle the spins must satisfy the relation

$$S_{i,j+1} S_{i-\frac{1}{2},j} S_{i+\frac{1}{2},j} = -1. \tag{27}$$

Only two mistakes are allowed, both in the same but unspecified row $j = 2^n$, with n an integer. In that row all spins are down $S_r = -1$, such that the SG-gasket can terminate. Moreover, in the first row $j = 1$ all spins must be down except for one, such that only one SP-gasket exists. These constraints ensure that this Ising model reproduces an ensemble of finite resolution SP-gaskets.

Let's construct a vL&N type real-space RT. Define diamond shaped cells on the triangular lattice, see Fig. 8. The cell-grid can be laid down in 4 ways. Choose the one where the up-spin in the first row, at $j = 1$, coincides with the top-spin inside one of the cells in the first row. Associate with each cell a cell-spin variable $\sigma_{\mathcal{R}} = \pm 1$. with $\mathcal{R}$ the coordinate of the center of the cell. Relabel all original spin variables in terms of the cells $\mathcal{R}$ and their local

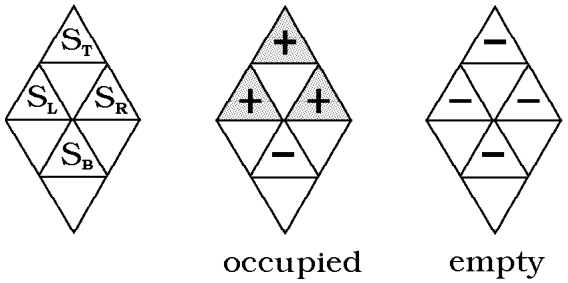

Fig. 8. diamond cells for the Sierpinsky gasket real-space RT

position, $k = T, R, L, B$, inside each cell. Define also a weight function $\mathcal{P}$ for each cell, normalized as

$$\sum_{\sigma_{\mathcal{R}}=\pm 1} \mathcal{P}(\sigma_{\mathcal{R}}; \{S_{\mathcal{R},k}\}) = 1. \tag{28}$$

Eq.(28) must be satisfied for all possible configurations of the original spins $\{S_{\mathcal{R},k}\}$. One common type choice is

$$\mathcal{P}(\sigma; \{S_k\}) = \frac{\exp[p\ \sigma\ (S_T + S_R + S_L + S_B)]}{\cosh[p(S_T + S_R + S_L + S_B)]} \tag{29}$$

with p a free parameter. This form reduces to the so-called majority rule in the limit $p \to \infty$. We are following vL&N [5] exactly. Next, the left hand sides of Eqs.(28), the weight functions, are inserted into $\mathcal{Z}$. This is harmless, since they are all normalized to one. It introduces summations over the cell spin variables. Finally, we switch the order of the summations and perform the summation over the original spin variables

$$\begin{aligned} \mathcal{Z} &= \sum_{\{S_r\},\star} e^{-\frac{1}{2}\mu \sum_r (S_r+1)} = \sum_{\{S_r\},\star} [\prod_{\mathcal{R}} \sum_{\sigma_{\mathcal{R}}} \mathcal{P}(\sigma_{\mathcal{R}}; \{S_{\mathcal{R},k}\})] e^{-\frac{1}{2}\mu \sum_r (S_r+1)} \\ &= \sum_{\{\sigma_{\mathcal{R}}\},\star} e^{-\frac{1}{2}\mu' \sum_{\mathcal{R}} (\sigma_{\mathcal{R}}+1)}. \end{aligned} \tag{30}$$

The resulting cell-spin model has the exact same structure as the original spin model, but the lattice constant has increased by a factor two ($b^2 = 4$ is the number of spins in each cell).

Eq.(30) defines the relation $\mu'(p, \mu)$. This is the so-called RT equation. It defines a flow through parameter space. Fixed points, $\mu' = \mu$, must have infinite or zero correlation length since $\xi' = \xi/b$. All points that flow to the same critical fixed point (those with $\xi = \infty$) must have infinite correlation length. All of them are critical points and belong to the same universality class.

The RT does typically not close in the sub space with only one type of inter-

action. Each RT step generates longer and longer ranged interactions. In practice, the strengths of those interactions remain small, and decay fast enough with range that the model remains intrinsically short-ranged. However, the non-closure means that we need to apply approximations to the fundamental trace:

$$e^{-\frac{1}{2}\mu'\sum_{\mathcal{R}}(\sigma_{\mathcal{R}}+1)} = \sum_{\{S_r\},\star} [\prod_{\mathcal{R}} \mathcal{P}(\sigma_{\mathcal{R}};\{S_{\mathcal{R},k}\})]\ e^{-\frac{1}{2}\mu\sum_r(S_r+1)}. \tag{31}$$

The basic premise is that the RT'eqs are analytic at T_c, and therefore that any type of conventional perturbation theory can be applied to evaluate Eq.(31). The thermodynamic singularities originate from fluctuations at large length scales. The cell spins are meant to carry these. The original spins fluctuate, but the coupling with the cell spins (which are frozen in Eq.(31)) removes the long range correlations between them. Not all choices of the cells and $\mathcal{P}$ will have this property. In particular, the RT eqs. are not analytic if the symmetries of the cells do not conserve the symmetries of the coexisting phases [5,6]. In our example we can avoid such complications. Our cells are already commensurate with the intrinsic scale factor 2 of the SP-gasket and we will find that for the majority rule (the $p \to \infty$ limit of Eq.(29) the RT closes with μ as only parameter.

Each diamond shaped cell could be in 2^4 spin-configurations. Only 2 of them occur, see Fig. 8: the "occupied" state $S_T = S_R = S_L = -S_B = 1$, and the "empty" state $S_T = S_R = S_L = S_B = -1$. Algebraically this follows from the fact that the boundary condition only allows these two cell configurations in the first row, and that the constraint Eq.(27) can reproduce only those two. For the majority rule the "occupied" and "empty" states relate one-to-one to the value of the cell spin $\sigma_{\mathcal{R}}$. All other aspects of the constraint $\star$ carry-over exactly onto the cell spins as well.

At finite p the cell spins would have a finite probability to be "down" ($\sigma_{\mathcal{R}} = -1$) in the "occupied" state and to be "up"($\sigma_{\mathcal{R}} = 1$) in the "empty-state, thus creating vacancies inside the SP-gasket or starting new SP-gaskets. Such excitations can be described by replacing the $S_i, S_j S_k = -1$ contraint in each down triangle by an three-body interaction $BS_iS_jS_k$. The coupling constant B is infinitely strong in Eq.(26). At finite p the RT equation does not close anymore in the one parameter space μ and requires the approximations schemes mentioned above.

Assume the majority rule. The trace over the original spins factorizes over the cells:

$$e^{-\frac{1}{2}\mu'(\sigma_{\mathcal{R}}+1)} = \sum_{\{S_k\}} \mathcal{P}(\sigma_{\mathcal{R}};\{S_k\})\ e^{-\frac{1}{2}\mu(S_T+S_R+S_L+S_B)} \tag{32}$$

and yields

$$\mu' = 3\mu. \tag{33}$$

This is the RT flow equation. It is a linear equation. In general, RT eqs. are non-linear. The precise form of Eq.(33) is hardly a surprise, and easy to anticipate due to the simplicity of the problem. However, it is instructive to see it appear in detail within the vL&N formalism.

Let's follow the conventional real-space RT type analysis of the renormalization equations. First we search for the fixed points of the RT flow. This is a non-trivial issue when the RT equations are non-linear. In our case $\mu = 0$ is the critical fixed point.

The RT trace preserves the total free energy

$$F(\mu, L) = F(\mu', \frac{L}{b}) \tag{34}$$

with L the size of the lattice. This is not an accident. The RT is designed to reproduce the FSS scaling postulate Eq.(2), it's baked into it. The free energy density (per site) transforms as

$$f(\mu, L) = b^{-D} f(\mu', \frac{L}{b}). \tag{35}$$

In our case this reduces to

$$f(\mu, L) = b^{-D} f(b^{y_A}\mu, \frac{L}{b}) \tag{36}$$

with $y_A = \log(3)/\log(2)$ by inserting our explicit form of the RT flow equation. In general, for non-linear RT equations, the RT will be linerized at the fixed point, and the eigenvalues, $\lambda_i = b^{y_i}$, of that linear transformation matrix yield the values of all fractal dimensions.

The free energy is the generating function of local observables, see Eq.(5). For our case, its derivative with respect to μ, see Eq.(26), gives the expectation value of the total mass of the up-triangles of the SP-gasket. Combined with the above free energy scaling relation this yields

$$M(\mu, L) = \left(\frac{\partial \mu'}{\partial \mu}\right) M(\mu', \frac{L}{b}). \tag{37}$$

Compare this with the definition of fractal dimension in section 4, Eq.(21). We obtained onces more the fractal dimension of the SP gasket but now the RT equation $\mu'(\mu)$ is doing the counting for us. We represented the "condensed matter construction rule" in terms of μ and a flow through parameter space. We evaluated the fractal dimension without any explicit reference to the actual geometric structure.

7 Closing Remarks

The real-space RT's of section 6 and 7 for the SP-gasket are quite transparant and simple. They serve the main purpose of this paper, to illustrate the basic concepts by pointing out the close relationship between geometric fractal properties and renormalization theory. These RT's are exact. In real life, construction rules are typically not "exactly soluble" and exact RT's are very rare. The advantage of the RT approach is that approximations to the RT rule automatically preserve scale invariance. They avoid the molecular field values of the critical exponents. Renormalization is the appropriate context to perform perturbation theory for scale invariant phenomena. Examples of such approximation schemes are presented elsewhere, e.g., in the review by Niemeijer and van Leeuwen [5] and also in the standard text books [7–9]. In his first paper on RT, Wilson [22] used the analogy of a ball rolling down a hill. The problem we face is to describe the properties of a singular function (the free energy at T_c; the hill; the fractal). It is advantageous to define a differential equation (the RT; the construction rule for determinsitic fractals), one that has the singular points automatically built into its solutions. Instead of straight perturbation theory (low or high temperature expansions analyzed with padé approximants) applied to the singular function itself, perturbation theory is now being applied onto the coefficients of that differential equation, which are much better behaved because they are analytic at those singular (critical) points.

Aknowledgements

I like to thank Hans van Leeuwen for his guidance during my early years in theoretical physics. In 1973 I was an undergraduate student in Delft. The first task Hans gave me for my senior independent study project was to study the "brand-new" real-space RT method from the proofs of that by now famous paper [1]. Many students contributed to this manuscript by their remarks, during classes at the University of Washington, and lectures at the 1997 Dutch AIO Cursus for graduate students. I like to thank Henk Blöte for providing Fig. 3, and Hubert Knops and the other members of the Instituut voor Theoretische Fysica for their hospitality during my stay at the University of Nijmegen in the Spring of 1997 where this final version was completed. This work is also supported by NSF grant DMR-9700430.

References

[1] Th. Niemeijer and J.M.J. van Leeuwen, Phys. Rev. Lett. **31**, 1411 (1973), and Physica **71**, 17 (1974).

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

[3] B. Widom, J. Chem. Phys. **433**, 3898 (1965).

[4] C. Domb and D.L. Hunter, Proc. Phys. Soc. **86**, 1147 (1965).

[5] Th. Niemeijer and J.M.J. van Leeuwen, in *Phase Transitions and Critical Phenomena* Vol. 6, eds. C. Domb and M.S. Green (Academic Press, New York, 1976).

[6] *Real-Space Renormalization*, eds. T.W. Burkhard and J.M.J. van Leeuwen (Springer, New York, 1982).

[7] K. Huang, *Statistical Mechanics* (Wiley, New York, 1987).

[8] M. Plischke and B. Bergersen, *Equilibrium Statistical Mechanics* (World scientific, Singapore, 1994).

[9] R. Balian, *From Microphysics to Macrophysics*, Vols.1-2, (Springer, New York, 1991).

[10] see e.g. *Mechanics, Molecular Physics, Heat, and Sound* by R.A. Millikan, D. Roller, and E.C. Watson (MIT press, Cambridge, 1965).

[11] L. van Hove, Physica **16**,137 (1950); reprinted in *Mathematical Physics in one Dimension*, *Exactly Soluble Models of Interacting Particles*, eds. E.H. Lieb and D.C. Mattis (Academic Press, New York, 1966).

[12] See e.g., M. Kosterlitz and D.J. Thouless, J. Phys. C **6**, 1181 (1973).

[13] A. Levelt-Sengers, Physica **82A**, 319 (1976).

[14] L. Onsager, Phys. Rev. **65**, 117 (1944).

[15] F.Y. Wu, Rev. Mod. Phys. **54**, 235 (1982).

[16] R.B. Stinchcombe, in *Phase Transitions and Critical Phenomena*, Vol. 7, eds. C. Domb and J.L. Lebowitz (Academic Press, London, 1977).

[17] M. Schick, Prog. Surf. Sci. **11**, 245 (1981).

[18] M. den Nijs, in *Phase Transitions and Critical Phenomena*, Vol. 12, eds. C. Domb and J.L. Lebowitz (Academic Press, London, 1988).

[19] M. Fisher, Rep. Prog. Phys. **30**, 615 (1967).

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

[21] E. Stanley, *Phase Transitions and Critical Phenomena* (Oxford, 1972).

[22] K.G. Wilson, Phys. Rev. B **4**, 3174, 3184 (1971), and Rev. Mod. Phys. **47**, 773 (1975).

[23] K.G. Wilson and J. Kogut, Phys. Rep. **12C**, 75 (1974).

[24] M.E. Fisher, Rev. Mod. Phys. **46**, 597 (1974).

[25] Vol. 6 of *Phase Transitions and Critical Phenomena* eds. C. Domb and M.S. Green (Academic Press, 1976).

[26] This critical configuration of the two-dimensional Ising model was created by the Delft Ising System Processor (DISP), a special purpose computer built by A. Hoogland in 1982 at the Technical University Delft.

[27] B. Mandelbrot, *The Fractal Geometry of Nature* (Freeman, 1983).

[28] See e.g., J. Feder, *Fractals* (Plenum Press, New York 1988).

[29] For a review see, J.L. Cardy, in *Phase Transitions and Critical Phenomena*, Vol. 11, eds. C. Domb and J.L. Lebowitz (Academic Press, London, 1987).