

Conformal Invariance in Percolation, Self-Avoiding Walks, and Related Problems

John Cardy

Abstract. Over the years, problems like percolation and self-avoiding walks have provided important testing grounds for our understanding of the nature of the critical state. I describe some very recent ideas, as well as some older ones, which cast light both on these problems themselves and on the quantum field theories to which they correspond. These ideas come from conformal field theory, Coulomb gas mappings, and stochastic Loewner evolution.

This talk is about ‘geometric’ critical phenomena. These are random spatial processes, where either (1) the probability distribution is determined by equilibrium statistical mechanics, and we ask questions about geometrical properties, or (2) the probability distribution is itself geometrical in nature. The simplest example of (1) is clustering in percolation (see Fig. 1), in which the probability distribution is trivial (in this case sites of a triangular lattice are independently coloured black or white with equal probability,) but we ask questions like whether there exists a path on (say) the black sites connecting opposite edges of a large rectangle. The paradigm example of (2) is the ensemble self-avoiding walks (SAWs) of a fixed (large) length, all weighted equally (see Fig. 2). In this case one might ask, for example, questions about the distribution of the distance between the ends. Other examples abound: for example the clusters formed by the spins in a critical Ising model, or the boundaries of the Fortuin-Kasteleyn clusters in the Potts model.

Note that although percolation is relevant to disordered media and SAWs relevant but to polymer physics, the emphasis in this talk is on understanding the nature of their fractal geometry. The outline is as follows: first, we shall discuss the various expectations for ‘geometric’ critical behaviour from conventional critical behaviour, largely based on well-known mappings between these types of problem. Since, in two dimensions, critical behaviour is described by conformal field theory, we may deduce all sorts of (non-rigorous) results, and we shall describe some of these. A related set of non-rigorous techniques is based on a mapping of cluster boundaries to a height models and then to a Coulomb gas, and we shall mention these.

But the main point of my talk will be to bring to the attention of theoretical physicists relatively new ideas which, it so happens, were developed by mathematicians, based on the direct construction of continuum limit of cluster boundaries, and known as SLE. I will describe how these provide rigorous and new results for percolation, as an example. I finish with some (slightly) provocative conclu-

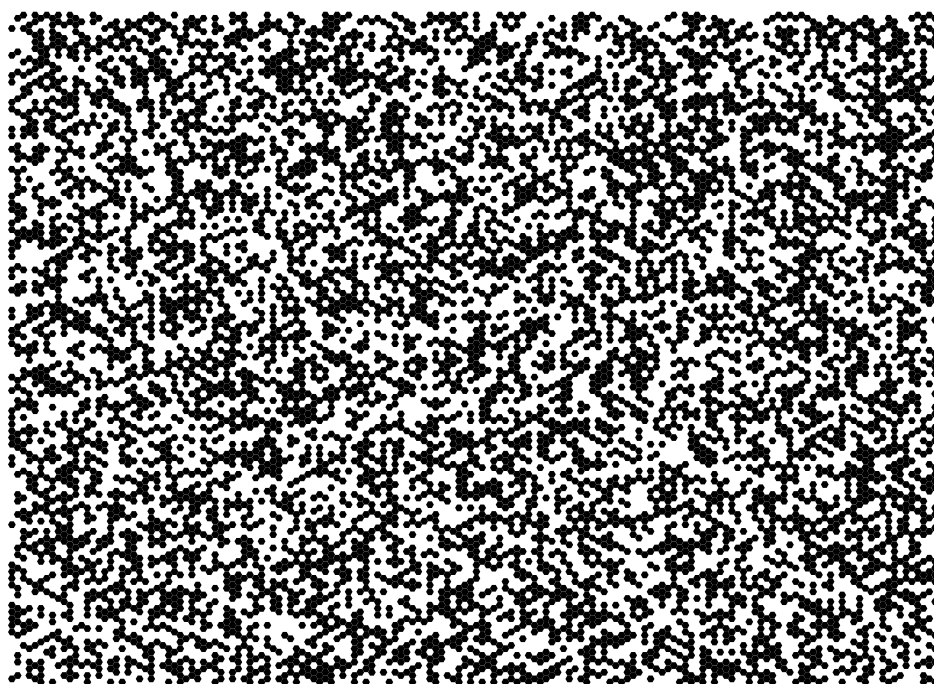


Figure 1: Critical site percolation on the triangular lattice. Each hexagon is independently coloured black or white with probability $\frac{1}{2}$. Is there a path on neighbouring black hexagons which connects the left and right sides of the rectangle?

sions about the usefulness of rigorous methods and of the traditional approach to quantum field theory, in the study of critical phenomena in general.

‘Geometric’ vs. ‘conventional’ critical behaviour

Much of our intuition about geometric critical behaviour is based on two well-known mappings:

1. (Fortuin-Kasteleyn [1]) Q -state Potts model \Leftrightarrow Random Cluster Model: the Potts model is a generalisation of the Ising model in which ‘spins’ at the sites of a lattice can each take one of Q values. Initially, Q must be an integer ≥ 2 , but the partition function may also be written

$$Z = \langle Q^{|\text{clusters}|} \rangle_{\text{percolation}}$$

where the clusters are weighted as in percolation, with the parameter p being simply related to the temperature in the Potts model. Evidently, the limit $Q \rightarrow 1$ reproduces percolation. Although in this limit the partition function

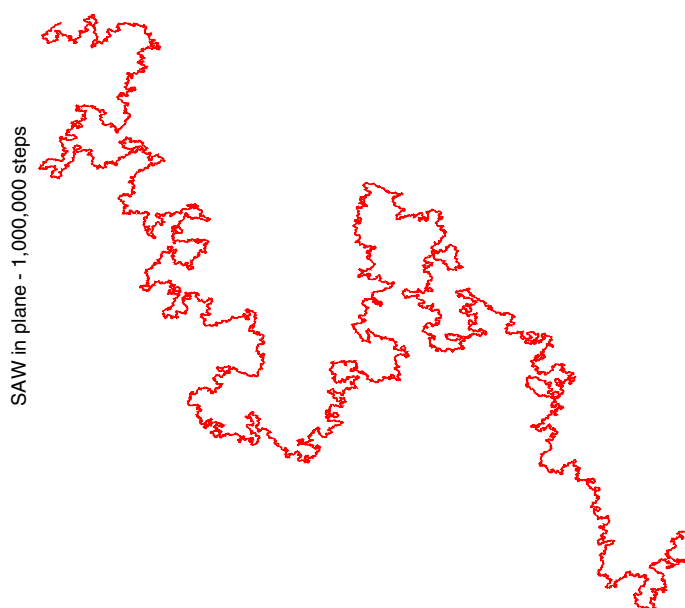


Figure 2: A typical self-avoiding walk. In this case the lattice is too fine to be visible.

(with free boundary conditions) is trivially equal to 1, the correlations are nontrivial.

2. (de Gennes [2]) $O(n)$ model \Leftrightarrow Self-Avoiding Loops: This model similarly generalises the Ising model to n -component spins and a hamiltonian which is invariant under $O(n)$ rotations. The partition function may be expressed as that of a gas of non-intersection loops:

$$Z = \langle n^{|\text{loops}|} \rangle_{\text{loop gas}}$$

Evidently the case $n = 1$ corresponds to the Ising model, while the limit $n \rightarrow 0$ leaves a single self-avoiding loop. In fact the $O(n)$ model has two types of critical behaviour for $n < 2$: one corresponding to the dilute phase, and a *dense* critical phase. In the dense phase, the loops are in the same universality class as the hulls of F-K clusters with $Q = n^2$.

From these correspondences, we can build a dictionary which relates geometrical properties to more conventional thermodynamic quantities. For example

$$\begin{aligned} \text{Cluster size} &\Leftrightarrow \text{susceptibility} \propto (p - p_c)^{-\gamma(Q)} \\ \text{Radius} &\Leftrightarrow \text{correlation length} \sim \text{mass}^{\nu(n)} \end{aligned}$$

Critical Behaviour & Euclidean Field Theory

It has been realised since the late 1960s that the scaling limit of an isotropic system near a continuous phase transition is a euclidean quantum field theory. If we take a near-critical lattice model, such that the correlation length $\xi \gg$ the lattice spacing a , the following limit exists:

$$\langle \phi(r_1) \dots \phi(r_N) \rangle_{\text{QFT}} = \lim_{a \rightarrow 0, \xi \text{ fixed}} a^{-n x_\phi} \langle S(r_1) \dots S(r_N) \rangle_{\text{lattice}}$$

where $\phi(r)$ is a local quantum field, and $S(r)$ is the corresponding lattice quantity. The non-trivial power x_ϕ is the scaling dimension of ϕ . This correspondence is rooted in an emphasis on correlation functions of *local* (or quasi-local) operators and their algebra encoded in the operator product expansion (and is therefore not always the best tool to investigate other quantities.) It has been *proved* in very few examples, but if *assumed* it has many powerful consequences: the Renormalisation Group, universality, and, in particular *scaling*: the property that the critical exponents describing off-critical behaviour of thermodynamic quantities are simply related to those describing decay of correlation functions at the critical point. This means that, for many purposes, we may restrict ourselves to studying the behaviour at the critical point. This means that the corresponding quantum field theory is massless: a *conformal field theory*.

Conformal Field Theory

Conformal field theory provides a very powerful tool to study critical behaviour, especially in two dimensions:

- in local *classical* field theories, scale invariance implies that the trace of the stress tensor T^μ_μ , vanishes, and this by itself implies conformal invariance.
- CFT assumes this holds (up to conformal anomaly c) in the full theory including fluctuation effects.
- In CFT, unlike normal QFT, there is a 1-1 correspondence between local operators and states of the Hilbert space, and the spectrum of these is usually discrete rather than continuous.
- For $d = 2$, these transform according to irreducible representations of an infinite-dimensional Virasoro algebra. Classifying these essentially classifies all possible CFTs, that is, all universality classes.
- There is, however, a problem: given some critical lattice model, to which CFT does it correspond?
- this was answered in part in the work of Friedan, Qiu and Shenker [3], who showed that in theories having reflection positivity (eg. $Q = 2, 3, 4$ Potts

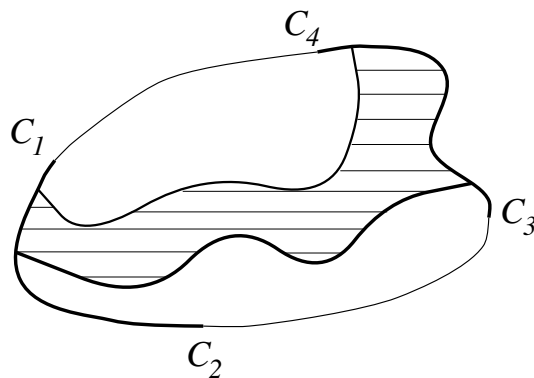


Figure 3: Schematic representation of a crossing cluster which connects the arcs C_1C_2 and C_3C_4 . Of course in reality this is a fractal object with a fractal boundary.

models, or $n = 1, 2$ in the $O(n)$ model), we should look for *unitary* representations, and at least when $c < 1$ this leads to a discrete series of possibilities. Moreover, the necessary decoupling of null states in these theories leads to linear differential equations for correlation functions (Belavin, Polyakov, Zamolodchikov [4].)

But percolation, SAWs, and related models are not unitary: in fact they have partition function $Z = 1$ ($c = 0$) even though their correlations are nontrivial. In fact non-unitary $c = 0$ CFTs are very poorly understood. Nevertheless they are important not just for percolation and SAWs, but for all critical problems with quenched disorder (e.g. the quantum Hall plateau transition, a major unsolved problem.)

The Crossing Formula in Percolation

A recurring theme in this talk will be the following problem (see Fig. 3):

- Given a simply connected region D of the plane, with suitably smooth boundary ∂D with 4 marked points C_j , what is the probability of a spanning cluster connecting C_1C_2 with C_3C_4 (in the limit lattice spacing $\rightarrow 0$)?

The following conjecture was made (Cardy [5]), based on ideas of CFT and the mapping to the $Q \rightarrow 1$ limit of the Potts model: the Riemann mapping theorem allows us to conformally map interior of D into the unit disc, with the marked points $C_j \rightarrow z_j$. Then the crossing probability depends only the anharmonic ratio $\eta = z_{12}z_{34}/z_{13}z_{24}$ and is

$$\frac{\Gamma(\frac{2}{3})}{\Gamma(\frac{4}{3})\Gamma(\frac{1}{3})} \eta^{\frac{1}{3}} {}_2F_1(\frac{1}{3}, \frac{2}{3}, \frac{4}{3}; \eta)$$

The argument depends on:

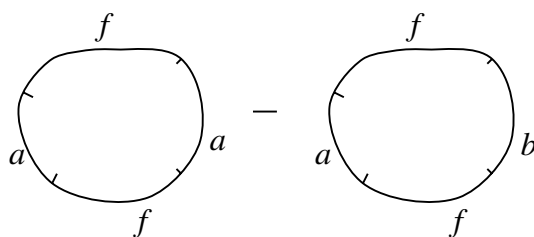


Figure 4: The crossing probability is a difference of Potts model partition functions, in which the Potts spins are fixed into the same (a), or different (a and b), states along the two arcs C_1C_2 and C_3C_4 .

- assuming that the scaling continuum limit exists and is given by a CFT with $c = 0$;
- realising that crossing probability is related to a difference of partition functions of the $Q \rightarrow 1$ Potts model, with different boundary conditions (Fig. 4);
- realising that states of the CFT induced by changes in the boundary condition ('boundary condition changing operators') also should correspond to Virasoro representations;
- guessing the right representation, and hence deducing the appropriate differential equation.

The formula has been numerically verified to high precision, but it is hard to see how to make arguments rigorous, or to go beyond them.

Cluster Boundary Approach

Instead of thinking about clusters, in the Potts language, it is sometimes easier to think about cluster boundaries, or hulls, using the $O(n)$ language. There are really two different approaches here: the older Coulomb gas arguments, and the more recent ideas of SLE.

'Coulomb gas' method

(den Nijs [6], Nienhuis [7], Duplantier & Saleur [8], Kondev [9], ...) The elements of these arguments are:

- thinking of cluster boundaries in random cluster model (or loops in $O(n)$ model) as a gas of (unoriented) closed loops.

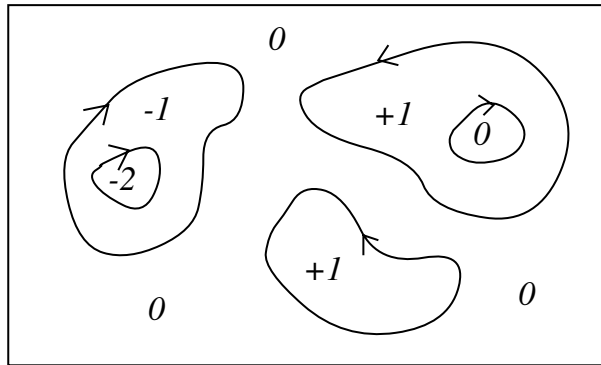


Figure 5: Oriented loops and the mapping to configurations of the height model.

- randomly orientate each loop: each configuration of oriented loop than maps onto one of a *height model*, with degrees of freedom $h(r) \in \text{integers}$, on the dual lattice (see Fig. 5):
- the factors of Q (resp. n) can be associated with local (but in general complex!) Boltzmann weights;
- *assume* in the continuum limit that $h(r)$ takes values in the real numbers, and that the measure converges to a gaussian $\exp(-g \int (\partial h)^2 d^2r)$;
- this gives a $c = 1$ CFT, with, however, various bells and whistles like a charge at ∞ , screening charges, etc., which make it nontrivial;
- within this formulation critical exponents are calculable, eg $\nu_{\text{SAW}} = \frac{3}{4}$ and $\nu_{\text{perc}} = \frac{4}{3}$;
- *but* correlation functions are ambiguous, and, more seriously, it has proven very hard to make this approach rigorous.
- nevertheless new results are still emerging from this method: eg distribution of *internal areas* of loops: density $n(A)$ of large loops with area $> A$

$$n(A) \sim C/A \quad C \text{ universal}$$

where [10]

$$\begin{aligned} C_{\text{perc}} = 1/8\sqrt{3\pi} &= 0.0229720 \quad \text{predicted} \\ &= 0.022972(1) \quad \text{measured} \end{aligned}$$

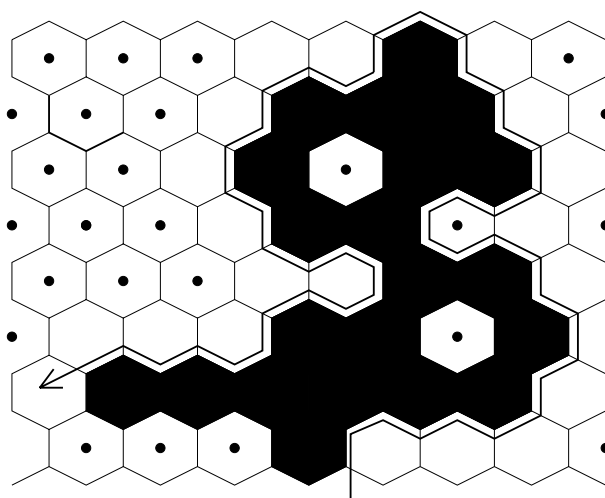


Figure 6: The process in which a random walker lays down the randomness as it progresses.

Dynamical description of cluster boundaries (SLE)

Although we usually think of percolation in terms of first laying down a particular random configuration of black or white sites, and then identifying the clusters and the boundaries of the clusters, it is statistically equivalent to construct these boundaries as certain random walks, where the walker lays down the random configuration as it goes. More specifically, this ‘exploration’ process involves random walker laying down the configuration as it moves: black sites to the left, white to the right (see Fig. 6). Notice that the path automatically reflects from itself, and also from the boundary if we choose the correct boundary conditions (black sites to the left of the starting point, white to the right.)

How should one characterise the continuum limit of these paths (assuming this exists)? In what sense is it conformally invariant? These were the questions addressed by Schramm [11] and Lawler, Schramm and Werner [12], and the answer is:

Stochastic Loewner Evolution(SLE)

For definiteness, consider the half-plane with black sites along $x < 0$ and white sites along $x > 0$. The walker starts at the origin, as shown in Fig. 7:

Rather than trying to write an equation for the path, consider the conformal mapping $z \rightarrow g(z; t)$ which sends {region of the half-plane which has not been excluded by the path} \rightarrow {upper half-plane}. Then, instead of the dynamics of the path, we may think about dynamics on conformal mappings. In particular,

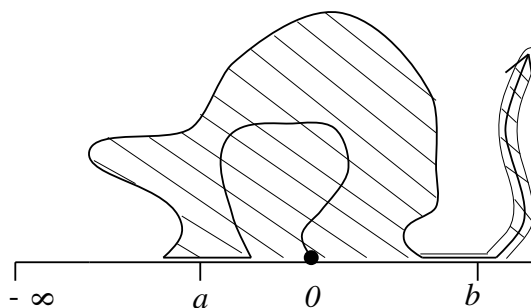


Figure 7: Schematic representation of SLE. At any given time, the path, and the region between it and the real axis, form an excluded part of the upper half plane. The complement of this is conformally mapped by the function $g_t(z)$ into the whole upper half plane. This function obeys the stochastic Loewner equation.

Schramm [11] conjectured that the continuum limit of the percolation exploration process corresponds to the Loewner equation

$$\frac{\partial g(z; t)}{\partial t} = \frac{2}{g(z; t) - a(t)}$$

This has several important properties:

- if $a(t)$ is a real continuous function, the excluded region of the half-plane grows with increasing t ;
- from requirements of scaling and locality, we conclude that $a(t)$ must be Brownian motion, ie $\dot{a} = \zeta(t)$ with $\zeta(t)\zeta(t') = \kappa\delta(t-t')$. (Rohde & Schramm [13])

If $0 \leq \kappa \leq 4$ path is simple
 $4 < \kappa < 8$ it touches itself
 $8 < \kappa$ it fills space

- the fractal dimension of the path is $= 1 + \kappa/8$ (for $\kappa < 8$).
- *only* for $\kappa = 6$ does the SLE path not ‘feel’ where the boundary of the domain is as long as it does not hit it, as one expects for percolation with uncorrelated site probabilities. This led to the conjecture:

**SLE₆ is the conformally invariant continuum limit
of percolation cluster boundaries**

- one can actually *compute* with SLE: it involves arguments quite familiar to theoretical physicists working in a different area, namely stochastic processes. In this way, one obtains all previously conjectured critical exponents at p_c (and with the help of rigorous scaling relations (Kesten [14]), exponents away from p_c), including multifractal irrational but algebraic exponents (related to 2d quantum gravity (Duplantier [15]), and some new ones, eg the
- *backbone exponent* (Lawler, Schramm, Werner [16]), the fractal dimension of the part of the infinite cluster which would carry electric current. This is given by the lowest eigenvalue of a 2d Dirichlet problem (and is probably not a rational or even an algebraic number, which makes its derivation by CFT or Coulomb gas methods a real challenge.)
- once SLE is assumed to describe the limit of percolation hulls, one easily gets and the crossing formula. As illustrated in Fig. 7, if we map the region into the upper half plane so that the arcs C_1C_2 and C_3C_4 map to $(-\infty, a)$ and $(0, b)$ respectively, then

$$\begin{aligned} Pr(\text{white crossing } (-\infty, a) \leftrightarrow (0, b)) \\ = Pr(a \text{ gets excluded before } b) \end{aligned}$$

- this gives the same ${}_2F_1$ formula as conjectured from CFT.

However, all this is so far dependent on Schramm's conjecture.

The missing link: Smirnov's proof of the crossing formula

Smirnov [17] proved that the crossing formula holds for the continuum limit of site percolation on a triangular lattice, and thereby that SLE_6 is the continuum limit of percolation cluster boundaries. Therefore all the results derived from SLE_6 are *rigorous*.

- First, it was observed by Carleson that the crossing formula is simple in an equilateral triangle (see Fig. 8).
- Note that the formula as proposed is the boundary value of an analytic function $P(z)$: so what is its interpretation for z not on the boundary? Smirnov proposed studying $P(z) = Pr(z \text{ separated from } C_1C_4 \text{ by at least one cluster spanning from } C_1C_2 \text{ to } C_2C_4)$. Then:
- on the triangular lattice, $P(z)$ satisfies linear relations, which mean that its continuum limit exists, and in fact just is the real part of a harmonic function, ie, a solution of Laplace's equation. It is the rather unusual boundary conditions which then determine that $P(z) \propto$ distance from C_1C_4 , which leads directly to the crossing formula when z is taken to lie on the boundary.

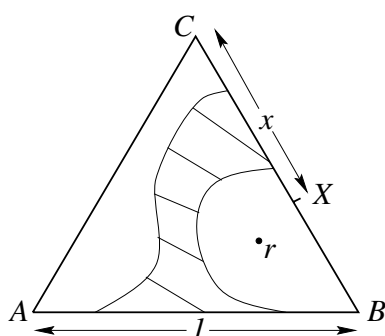


Figure 8: Carleson's version of the crossing formula in an equilateral triangle, and Smirnov's generalisation: what is the probability that there is a cluster which connects AB with BC , at the same time disconnecting the point z from AC ?

Other values of κ

It turns out that other values of κ in SLE correspond to different values of n in the $O(n)$ model, for example

- Self-Avoiding Walks: if we use the principle that uniform measure on set of simple paths must remain uniform when restricted to a subset, and assume the continuum limit is SLE_κ , then we are led to $\kappa = \frac{8}{3}$. This reproduces all the conjectured results for SAWs (and more) (Lawler, Schramm, Werner [18]).
- the critical Ising model is conjectured to be described by SLE_3 . This leads to all the standard results and some new ones: eg, if we take an Ising model inside a simply connected region of the plane, with boundary conditions that the spins are up on one segment of the boundary, and down on its complement (so that there exists a domain wall crossing the region, see Fig. 9), what is the probability that the domain passes above a given point z ? The result is conformally invariant, and an explicit formula is provided by SLE (Schramm [19]). Of course these results still need the analogue of Smirnov's proof to make them completely rigorous, but they appear to be completely beyond the reach of traditional approaches to the Ising model, which focus on spin correlation functions.

Final Remarks

I shall finish with a couple of possibly provocative remarks. The first concerns progress in rigorous results for critical behaviour. The first such nontrivial results were of course obtained for the Ising model in 1944 by Onsager, followed by the work of Yang, and this gave rise to a whole industry of computations of correlation

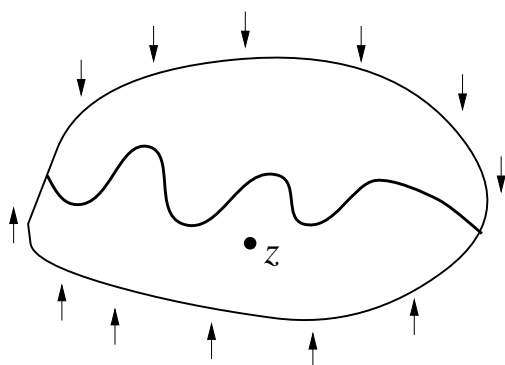


Figure 9: The Ising model in a domain with + boundary conditions on one segment of the boundary, and - boundary conditions on the complement: what is the probability that the consequent domain wall passes above the point z ?

functions and other quantities in this and variant models. The next step was perhaps Lieb's 1967 solution of 6-vertex type models, then Baxter's solution of the 8-vertex model in 1971. The methods used by Baxter and others (Bethe ansatz, commuting transfer matrices, corner transfer matrices) are powerful and elegant, and they produce 'exact' results, but they are not, in general, fully rigorous.

Now we finally have, thanks to the work of Kesten, Smirnov, Lawler, Schramm, Werner, and others *rigorous* results for a different sort of two-dimensional critical behaviour, namely percolation. The methods used are completely different, relying on identifying the continuum limit explicitly, and performing computations directly in that limit, rather than taking the limit of lattice results. As a result, only universal quantities are calculated. The methods focus on the geometrical aspects of the measure, rather than the correlation functions of local operators. It is quite likely that in the next few years we shall see similar rigorous results for a whole set of such models, at least in two dimensions.

Of course, particularly at a meeting on Theoretical Physics, one might question the necessity of mathematical rigour. Are we not only interested in the answers, and in the physical picture which they convey? Of course this is true, but the historical fact remains that theoretical physicists in this subject have been perhaps too happy for many years with their heuristic and incomplete arguments, which had once been very fruitful but perhaps had reached a dead end. It took mathematicians to question the true content of these arguments, to find them wanting, and then to develop a completely independent approach, that of SLE. The result is that we now are beginning to have a new physical way of understanding the origins of conformal invariance and of the emergence of non-trivial critical exponents, quite separate from that of the renormalisation group. The methods of SLE could have been developed by physicists – after all they are based on Brownian motion

– but the fact is that they were not. This affords a strong example for continuing to make theoretical physics as mathematically rigorous as possible.

The second point concerns progress in quantum field theory, especially from the point of view of critical behaviour. Historically, the development of the renormalisation group point of view in 1969 by Wilson and others gave answers to many outstanding questions at the time. It gave a non-rigorous framework in which to understand many important features of critical phenomena, such as universality and scaling, and to do approximate computations. From 1984 the methods of conformal field theory, originally developed for string theory, were brought in, and they provided a plethora of exact but non-rigorous results in two dimensions. But even today field theory is still tied to particle physics ideas. We see that both in the way it is taught and in the continuing focus on the objects of interest to particle theory: namely correlation functions of local fields. While this has been successful in the past, it is ill-suited to study other objects (eg the crossing formula, which required the introduction of ‘boundary-condition changing’ operators to be expressed in the old language.) One might ask whether QFT has outlived its usefulness to generate new results, or whether (rather like string theory) it can reinvent itself in the 21st century, in a different and more powerful form – perhaps

QFT as fractal geometry?

Note added. Very recently Bauer and Bernard [20] have, among other things, exposed the relationship between Virasoro null vectors and SLE, thus establishing the connection with the CFT approach to the crossing formula.

This work was supported in part by EPSRC Grant GR/J78327. I thank R. Ziff for help in producing Fig. 1, and T. Kennedy for permission to reproduce Fig. 2.

References

- [1] C. M. Fortuin and P. W. Kasteleyn, *Physica* **57**, 536, 1972.
- [2] P.-G. de Gennes, *Phys. Lett. A* **38**, 339, 1972.
- [3] D. Friedan, Z. Qiu and S. Shenker, *Phys. Rev. Lett.* **52**, 1575, 1984.
- [4] A. A. Belavin, A. M. Polyakov and A. B. Zamolodchikov, *J. Stat. Phys.* **34**, 763, 1984.
- [5] J. L. Cardy, *J. Phys. A* **25**, L201, 1992.
- [6] M. P. M. den Nijs, *Phys. Rev. B* **27**, 1674, 1983.

- [7] B. Nienhuis in *Phase Transitions and Critical Phenomena*, v. 11, C. Domb and J. L. Lebowitz eds. (Academic, NY, 1987.)
- [8] H. Saleur and B. Duplantier, *Phys. Rev. Lett.* **38**, 2325, 1987.
- [9] J. Kondev, *Phys. Rev. Lett.* **78**, 4320, 1997.
- [10] J. Cardy and R. M. Ziff, *J. Stat. Phys.*, to appear; cond-mat/0205404.
- [11] O. Schramm, *Israel J. Math.* **118**, 221, 2000.
- [12] G. F. Lawler, O. Schramm and W. Werner, *Acta Math.* **187**, 237, 2001; *Acta Math.* **187**, 275, 2001; *Ann. Inst. Henri Poincaré PR* **38**, 109, 2002; arXiv:math.PR/0108211.
- [13] S. Rohde and O. Schramm, arXiv:math.PR/0107096.
- [14] H. Kesten, *Comm. Math. Phys.* **109**, 109, 1987.
- [15] B. Duplantier, this volume.
- [16] G. F. Lawler, O. Schramm and W. Werner, arXiv:math.PR/0108211.
- [17] S. Smirnov, *C. R. Acad. Sci. Paris Sr. I Math.* **333**, 239, 2001.
- [18] G. F. Lawler, O. Schramm and W. Werner, arXiv:math.PR/0204277.
- [19] O. Schramm, arXiv:math.PR/0107096.
- [20] M. Bauer and D. Bernard, arXiv:math-ph/0206028; arXiv:hep-th/0210015.

John Cardy
Theoretical Physics
1 Keble Road
Oxford OX1 3NP
United Kingdom
and
All Souls College, Oxford