Simulations of lattice animals and trees

2005 J. Phys. A: Math. Gen. 38 775

(http://iopscience.iop.org/0305-4470/38/4/001)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **38** (2005) 775–806 doi:10.1088/0305-4470/38/4/001

Simulations of lattice animals and trees

Hsiao-Ping Hsu, Walter Nadler and Peter Grassberger

John-von-Neumann Institute for Computing, Forschungszentrum Jülich, D-52425 Jülich, Germany

Received 3 August 2004, in final form 8 November 2004
Published 12 January 2005
Online at stacks.iop.org/JPhysA/38/775

Abstract

The scaling behaviour of randomly branched polymers in a good solvent is studied in two to nine dimensions, using as microscopic models lattice animals and lattice trees on simple hypercubic lattices. As a stochastic sampling method we use a biased sequential sampling algorithm with re-sampling, similar to the pruned-enriched Rosenbluth method (PERM) used extensively for linear polymers. Essentially we start simulating percolation clusters (either site or bond), re-weigh them according to the animal (tree) ensemble, and prune or branch the further growth according to a heuristic fitness function. In contrast to previous applications of PERM, this fitness function is *not* the weight with which the actual configuration would contribute to the partition sum, but is closely related to it. We obtain high statistics of animals with up to several thousand sites in all dimension $2 \leqslant d \leqslant 9$. In addition to the partition sum (number of different animals) we estimate gyration radii and numbers of perimeter sites. In all dimensions we verify the Parisi–Sourlas prediction, and we verify all exactly known critical exponents in dimensions 2, 3, 4 and $\geqslant 8$. In addition, we present the hitherto most precise estimates for growth constants in $d \geqslant 3$. For clusters with one site attached to an attractive surface, we verify for $d \geqslant 3$ the superuniversality of the cross-over exponent ϕ at the adsorption transition predicted by Janssen and Lyssy, but not for $d = 2$. There, we find $\phi = 0.480(4)$ instead of the conjectured $\phi = 1/2$. Finally, we discuss the collapse of animals and trees, arguing that our present version of the algorithm is also efficient for some of the models studied in this context, but showing that it is *not* very efficient for the 'classical' model for collapsing animals.

PACS numbers: 02.70.Uu, 05.10.Ln, 05.50.+q, 82.35.Lr

1. Introduction

Lattice animals (or polyominoes, as they are sometimes called in mathematics [1]) are just clusters of connected sites on a regular lattice. Such clusters play an important role in many models of statistical physics, as e.g. percolation [2], the Ising model (Fortuin–Kastleyn

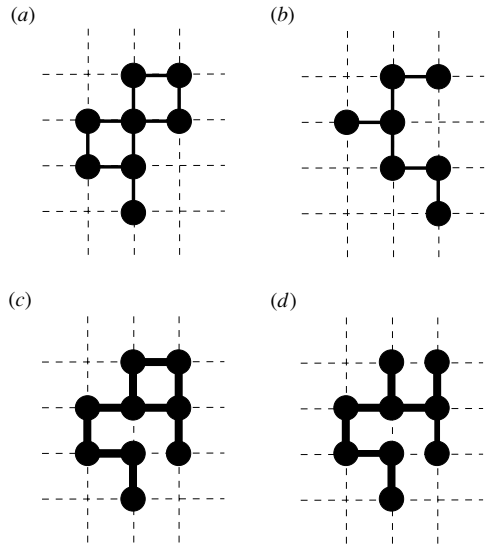

Figure 1. (*a*) A site animal with eight sites. (*b*) A site tree ('strongly embeddable tree'). (*c*) A bond animal which is not a tree. (*d*) A bond tree ('weakly embeddable tree').

clusters, Swendsen–Wang algorithm [3, 4]), and even lattice gauge theories [5]. The basic combinatorial problem associated with them is to count the number Z_N of different animals of N sites. Two animals are considered as identical, if they differ just by a translation (i.e., we deal with *fixed* animals in the notation of [6]), but are considered as different, if a rotation or reflection is needed to make them coincide. Thus there are, e.g., d animals of $N = 2$ sites on a simple hypercubic lattice of dimension d, and $d(2d - 1)$ animals with $N = 3$.

The animal problem can be turned into a statistical problem by giving a statistical weight to every cluster. In contrast to percolation, where different shapes acquire different weights, all clusters with the same number N of sites are given the same weight. This is similar to self-avoiding walks (SAW). A SAW on a lattice is a connected cluster of N sites with equal weight on all clusters, but with a restriction on its shape: each SAW has to be topologically linear, i.e. each site is connected by bonds to at most two neighbours. No such constraint holds for animals, thus animals are the natural model for randomly branched polymers [7].

In addition to animals (or *site animals*, to be more precise) one can also consider *bond animals* and *lattice trees*. A bond animal is a cluster where bonds can be established between neighbouring sites (just as in SAWs), and connectivity is defined via these bonds: if there is no path between any two sites consisting entirely of established bonds, these sites are considered as not connected, even if they are nearest neighbours. Different configurations of bonds are considered as different clusters, and clusters with the same number of bonds (irrespective of their number of sites) have the same weight [8]. *Weakly embeddable trees* are bond animals with tree topology, i.e. the set of weakly embeddable trees is a subset of bond animals, each with the same statistical weight. *Strongly embeddable trees* are, in contrast, the subset of site animals with tree-like structure. All these definitions are illustrated in figure 1.

Like many other statistical models, animals are characterized by scaling laws in the limit of large N. It is believed that all the above statistics (site and bond animals, weakly and strongly embeddable trees) are in the same universality class (same exponents, same scaling functions)

which is that of randomly branched polymers. The number of animals (i.e. the microcanonical partition sum) should scale as [7]

$$Z_N \sim \mu^N N^{-\theta}, \tag{1}$$

and the gyration radius as

$$R_N \sim N^{\nu}. \tag{2}$$

Here μ is the *growth constant* or inverse critical fugacity, and is not universal. In contrast, the Flory exponent ν and the entropic exponent θ should be universal.

In spite of the obvious similarity to the SAW and percolation problems, there are a number of features in which the animal problem is unusual:

- The upper critical dimension is $d = 8$. There, $\nu = 1/4$ and $\theta = 5/2$ [9].
- The model is not conformally invariant [10], and thus the Flory exponent ν is not known exactly in $d = 2$.
- Using supersymmetry, it has been argued by Parisi and Sourlas [11] that the animal problem in d dimensions is related to the Yang–Lee problem (Ising model in an imaginary external field) in $D = d - 2$ dimensions. Based on this relationship (which is now proven rigorously [12], using a mapping onto the hypercubes problem at negative fugacity [13]) they argued that θ and ν should not be independent, but

  $$\theta = (d - 2)\nu + 1. \tag{3}$$

  This implies in particular that $\theta = 1$ in $d = 2$. In addition, they showed that $\nu = 1/2$ in $d = 3$.
- Assuming universality so that scaling of the hard squares model at negative fugacity can be inferred from Baxter's solution of the hard hexagon model, and mapping the hard squares model onto lattice animals in four dimensions, Dhar [14] obtained $\theta = 11/6$ for $d = 4$. Thus one knows the exact values of ν for $d = 1$, 3, 4 and 8, but not for $d = 2$, 5, 6 and 7.
- In a series of papers, Janssen and Lyssy [15–17] studied animals attached to an adsorbing plane surface. For weak adsorption (high temperature) the animals have basically the same structure as in the bulk, and the partition sum has the same scaling, $Z'_N \sim \mu'^N N^{-\theta'}$ with $\mu' = \mu$ and $\theta' = \theta$ [18, 19][1]. For strong adsorption (low temperature) there is an adsorbed phase. Janssen and Lyssy argued that the cross-over exponent between these two phases should be super-universal, $\phi = 1/2$ for all dimensions $d \geqslant 2$.

In the present paper we address all these points by means of a novel Monte Carlo algorithm which follows essentially the strategy used in the *pruned-enriched Rosenbluth method* (PERM) [20]. This is a recursively (depth first) implemented sequential sampling method with importance sampling (bias) and re-sampling ('population control'). It seems that PERM in the present implementation is much more efficient than previous sampling methods for animals and trees. Indeed we shall present numerous new estimates for critical exponents and growth constants which had previously been measured only with much larger error bars or not at all.

All this holds for athermal animals and trees, i.e. when there are no attractive forces between monomers. When such forces become strong, a number of different collapse phase transitions are claimed to occur, depending on the detail of the model [21–33]. The simplest one of these involves site animals and a simple contact energy for each occupied nearest neighbour

[1] DeBell *et al* write $\theta' = \theta + 1$, since they obviously counted shifted animals in the bulk as different, while we consider the ensemble of 'fixed animals' where translated shapes are identified.

pair [30–33] and is undisputed. But another transition, between two collapsed phases with different densities of bonds [21–23, 27, 29], is still controversial [24–26]. At present all versions of our algorithm become inefficient for the first model, when the collapse point is approached. This is a bit disappointing in view of the fact that PERM for linear polymers is dramatically *more* efficient at the coil–globule transition than for athermal SAWs [20]. Obviously this leaves much room for further improvements. On the other hand, our method should work well for the other transition in large parts of the phase diagram.

Details of the algorithm for site animals will be given in section 2. Detailed studies of site animals in the bulk and in contact with a wall will be presented in sections 3 and 4. Bond animals and trees will be discussed in section 5. Finally, in section 6 we will study animal (and tree) collapse due to attractive forces between monomers. The paper ends with conclusions in section 7.

2. Numerical methods

2.1. Previous methods

2.1.1. ϵ-Expansions. Field theoretic ϵ-expansions (where ϵ is the distance from the critical dimension) were applied already very early to animals [7] and to the Yang–Lee problem [34]. When the relationship between both problems was established, the latter gave the most precise predictions for critical animal behaviour in high dimensions ($d \geqslant 5$).

2.1.2. Exact enumeration. Exact enumeration of animals and trees is surprisingly non-trivial [35–38]. Nevertheless, very extensive enumerations have recently been performed by Jensen [6, 39, 40] for site animals and site trees in $d = 2$. At present they give the best numerical verification of the prediction $\theta = 1$, and they give the most precise estimates for the Flory exponent ($\nu = 0.641\,15 \pm 0.000\,05$) and for the growth constants: $\mu = 4.062\,5696 \pm 0.000\,0005$ for animals [6], and $3.795\,254 \pm 0.000\,008$ for trees. These values are more precise than old estimates obtained by finite size scaling using strip geometry [41]. There are also enumerations of various animals and trees in higher dimensions [9, 21, 42–48], but they are much less complete and in general they do not at present give the best estimates for critical parameters.

2.1.3. Markov chain Monte Carlo methods. The latter are obtained nowadays by Markov chain Monte Carlo (MCMC) methods. Such algorithms have been used for animals since at least 20 years [31, 49–51]. At present, the most efficient MCMC algorithm for lattice trees is a version of the pivot algorithm [8, 52–55]. These simulations showed that $\nu = 1/2$ in $d = 3$, as predicted by [11]. Simulations of animals attached to an attractive wall verified that indeed $\phi = 1/2$ in $d = 2$ [53] (as also verified with transfer matrix and similar methods [56, 57]), although simulations in $d = 3$ gave $\phi \approx 0.714$ [58], in gross violation of the Janssen–Lyssy prediction.

When applied to SAWs, the pivot algorithm works by choosing a pivot point and proposing a rotation of the shorter arm around the pivot, and accepting it when this leads to no violation of self-avoidance [59]. When adapted to trees, one again chooses a random pivot point, but now the entire branch hinging on this pivot is cut and glued somewhere else. Again this move is accepted only if this leads to no violation of self-avoidance and if it would not lead to wrong cluster topology.

This method also allows us to estimate growth constants, if it is used together with the atmosphere method [54]. In the latter, it is counted how many possible ways there exist to

grow the cluster by one further step, giving in this way an estimate of Z_{N+1}/Z_N. Basically the same method had been used in [60] to obtain very precise estimates for the critical percolation thresholds in high-dimensional lattices.

2.1.4. Cluster growth ('sequential sampling', 'static') methods. The first stochastic growth algorithm for trees seems to have been devised by Redner [61]. Similar methods were then used by Meirovitch [62] and Lam [46] for animals. But already Leath seems to have realized that his well-known algorithm for growing percolation clusters [63] could be used also for the study of animals, simply by reweighing the clusters. Recently this was taken up systematically in [64].

In the following we shall discuss the latter in some detail, and we shall restrict our discussion to site animals. The authors of [64] basically use a standard growth algorithm for percolation clusters [63, 65, 66], and then reweigh the cluster so that they obtain the correct weight for the animal ensemble. In a percolation cluster growth algorithm for site percolation, one starts with a single seed site and writes it into an otherwise empty list of 'growth sites'. Then one recursively picks one item in the list of growth sites, removes it from the list and adds it with probability p to the cluster, and adds all its wettable neighbours to the list. This is repeated until either the cluster size exceeds some fixed limit (in which case the cluster is discarded), or until the list is empty. A cluster with N sites, b boundary sites, and with fixed shape is obviously obtained with probability

$$P_{Nb} = p^N(1-p)^b, \tag{4}$$

i.e. with the correct probability so that any unweighed average is just an average over the percolation ensemble. Repeating this many times, the animal partition sum is then

$$Z_N = \langle 1/P_{Nb}\rangle = p^{-N}\langle(1-p)^{-b}\rangle. \tag{5}$$

The authors of [64] called their method a Rosenbluth method, in view of the obvious analogy with the Rosenbluth–Rosenbluth method [67] for SAWs. In the latter one also samples from a biased ensemble and then reweighs each configuration with the inverse of the sampling probability to obtain the correct partition sum.

2.2. PERM

Like the original Rosenbluth–Rosenbluth method for SAWs, the method of [64] has the disadvantage that the weights have a very wide distribution for large N. Thus even a very large sample will finally, when N gets too big, be dominated by a single configuration, and the method becomes inefficient even though it is easy to generate huge samples.

PERM (or any other strategy with resampling) tries to avoid this by trimming the width of the distribution of weights, by pruning configurations with very low weight and making clones of high weight configurations which then share the weight among themselves. In many situations this has proven to be extremely efficient [68–70]. But we cannot yet apply it to animals, since we have to be able to estimate the weight of a cluster *while it is still growing*, and up to now we have only discussed the relationship between animals and percolation clusters after they had stopped growing.

In the following we shall again discuss only site animals, bond animals and trees being discussed in section 5.

To obtain the relationship between still growing percolation clusters and animals, let us consider a cluster with N sites, g growth sites, and b sites which definitely belong to the

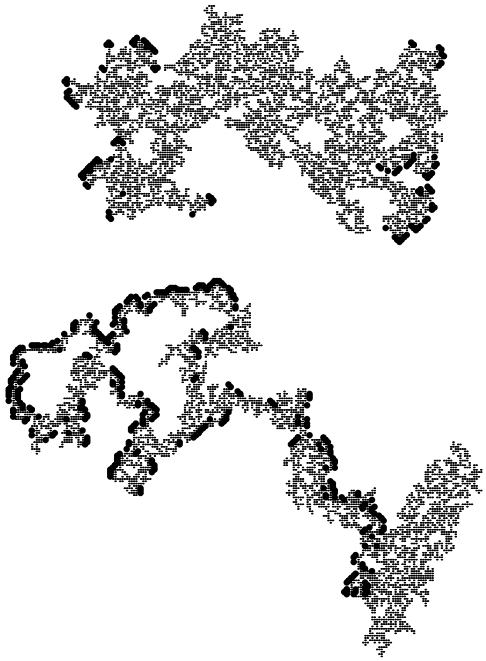

Figure 2. Growing clusters generated breadth first (top) and depth first (bottom). In both cases we used $p = p_c = 0.5927$, and in both cases $N = 4000$. Occupied sites are depicted by small points, growth sites by heavy dots. Both figures are plotted with the same scale.

boundary. At each of the growth sites the cluster can grow further, or it can stop growing with probability $1 - p$. Thus this cluster will contribute with weight $(1 - p)^g$ to the sample of percolation clusters with N sites and $b + g$ boundary sites. Its contribution to the animal ensemble is smaller by a factor $[p^N(1 - p)^{b+g}]^{-1}$, and we have thus

$$Z_N = p^{-N}\langle(1 - p)^{-b}\rangle. \tag{6}$$

This is exactly the same formula as above, but now the average is taken over all growing clusters, while before we had averaged over clusters which had stopped growing.

Before we can implement these ideas, we have to discuss two problems:

(a) How are the clusters to be grown precisely?
(b) On what basis should we decide when to prune and when to clone a cluster?

As we shall see, both questions are not trivial.

(a) Percolation cluster growth algorithms can be depth first or breadth first. In the former, growth sites are written into a first-in last-out list (a *stack*). A typical code for this is given in [66]. For a breadth first implementation we use, instead, a first-in first-out list ('fifo-list' or *queue*). Two 2d clusters growing according to these two schemes, with $p = p_c = 0.5927$, are shown in figure 2. Both have $N = 4000$. But the cluster grown using a stack has a completely different shape and has $\approx$3 times as many growth sites as that grown with a queue! Most of these growth sites are nearly dead (their descendents will die after a few generations), but this is not realized because they are never tested. Since also the fluctuations in the number of growth sites are much bigger in a depth first implementation, the weights in equation (6) will also fluctuate much more, and we expect much worse behaviour. This is indeed what we

found numerically: results obtained when using a stack for the growth sites were dramatically much worse than results obtained with a queue.

Note that this is independent of the way how pruning and cloning is done. Indeed we implemented this 'population control' recursively as a depth first strategy, as was done for all previous applications of PERM.

In addition, there are also some minor ambiguities in percolation cluster growth algorithms, such as the order in which one searches the neighbours of a growth site and writes them into the list. In two dimensions one can e.g., use the preferences east–south–west–north, or east–west–north–south, or a different random sequence at every point. We found no big differences in efficiency.

(b) In most previous applications of PERM, the best strategy was to base the decision whether to prune or branch directly on the weight with which the configuration contributes to the partition sum[2]. This would mean in the present case that we clone, if $W_n \equiv p^{-n}(1-p)^{-b} > c_+\hat{Z}_n$ where c_+ is a constant of the order $1, \ldots, 10$ and $\hat{Z}_n$ is the current estimate of Z_n. Similarly, a cluster would be killed (with probability $1/2$ [20]), if $W_n < c_-\hat{Z}_n$ with c_- slightly smaller than 1.

In the present case this would not be optimal, since it would mean that mostly clusters with few growth sites are preferred (they tend to have larger values of b, for the same n), and these clusters would die soon and would contribute little to the growth of much larger clusters. Thus we defined a *fitness function*

$$f_n = W_n/(1-p)^{\alpha g} = p^{-n}(1-p)^{-b-\alpha g} \tag{7}$$

with a parameter α to be determined empirically, and used

$$f_n > c_+\langle f_n\rangle, \qquad f_n < c_-\langle f_n\rangle \tag{8}$$

as criteria for cloning and pruning. We checked in quite extensive simulations that best results were obtained with $\alpha = 1$ (except when N is small), and in the following we shall use only this choice.

Finally we have to discuss the optimal values of p. It is clear that we should not use $p > p_c$, where p_c is the critical percolation threshold. Since minimal reweighing is needed for small p (subcritical percolation is in the animal universality class), one might expect $p \ll p_c$ to be optimal. This is indeed true for small values of N (which we are not primarily interested in), but not for large N. For the latter it is more important that clusters grown with $p \ll p_c$ have to be cloned excessively, since they otherwise would die rapidly in view of their few growth sites.

To decide this problem empirically, we show in figure 3 the errors of the estimated free energies $F_N = -\ln Z_N$ for $d = 2$. More precisely, we show there one standard deviation multiplied by the square root of the CPU time (measured in seconds), for different values of p. Each simulation was done on a Pentium with 3 GHz using the gcc compiler under Linux, and each simulation was done for $N_{\max} = 4000$ (although we plotted some curves only up to smaller N, omitting data which might not have been converged). We see clearly that small values of p are good only for small N. As N increases, the best results were obtained for $p \to p_c$. The same behaviour was observed also in all other dimensions, and also for animals on the bcc and fcc lattices in three dimensions (data not shown). As an example we show in figure 4 the analogous results for $d = 8$. There we used a 64-bit machine (a 600 MHz Compaq ALPHA), because this simplified hashing (for large d we used hashing as described e.g. in [60]).

[2] The only previous exception was in ground-state predictions for the HP protein model, where it was found useful to include a bias against *H–P* contacts in the population control [71, 72].

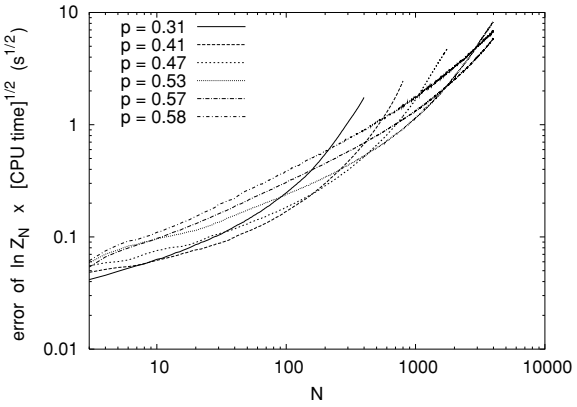

Figure 3. Statistical errors of ln Z for lattice animals in $d = 2$ for various values of p. To make the different runs comparable, errors are multiplied by the square root of the CPU time measured in seconds. In two dimensions $p_c = 0.5927$.

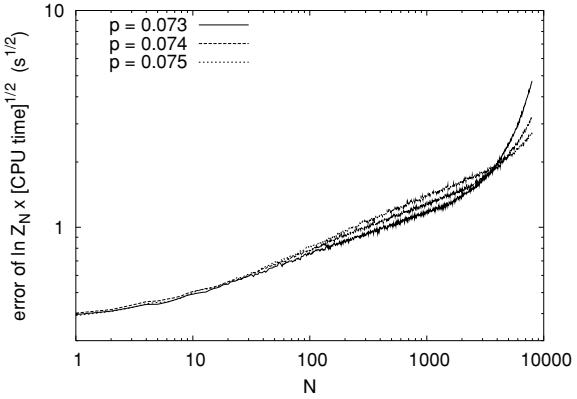

Figure 4. Same as figure 3, but for $d = 8$ where $p_c = 0.0752$. The straightest curve corresponds to the largest p.

Note that the errors shown in figure 4 are much smaller than those in figure 3, although the machine was slower and the animals were twice as large ($N_{max} = 8000$). Indeed, the errors decreased monotonically with d, being largest for $d = 2$. Using p slightly smaller than p_c we can obtain easily very high statistics samples of animals with several thousand sites for dimensions $\geqslant 2$. A typical 2d animal with 12 000 sites is shown in figure 5, and a 3d animal on the bcc lattice with 16 000 sites is shown in figure 6.

To check the reliability of our error bars we looked at distributions of *tour weights* as described in [73]. A tour is the set of all configurations generated by cloning from one common start and therefore possibly being strongly correlated. If the distribution P of tour weights W is very broad, we are back to the problem of the Rosenbluth method that averages might be dominated by a single tour. To check for this, we plot $P(\ln W)$ against $\ln W$, and compare its right-hand tail to the function $1/W$. If the tail decays much faster, we are presumably on the safe side, because then the product $W P(\ln W)$ has its maximum where the distribution is well sampled. If not, then the results can still be correct, but we have no guarantee for it.

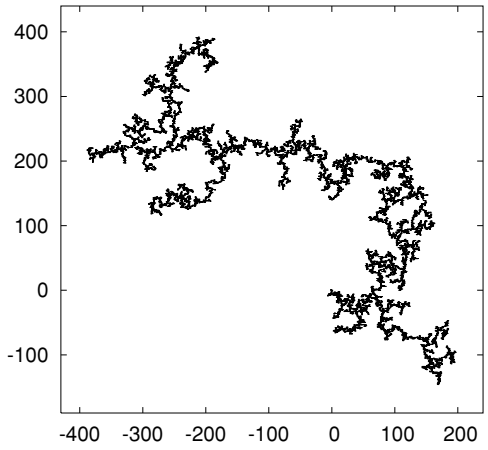

Figure 5. A typical lattice animal with 12 000 sites on the square lattice.

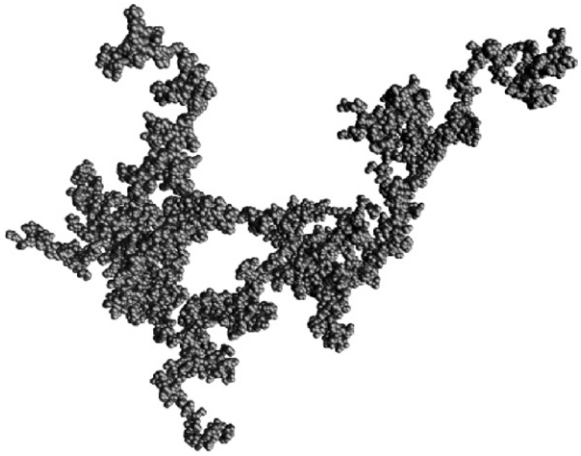

Figure 6. A 3d lattice animal with 16 000 sites on the bcc lattice.

In figure 7, we show these tour weight distributions for two-dimensional animals with 4000 sites, for $p = 0.57$ and for $p = 0.47$.[3] We see that the simulation with $p = 0.57$ is on the safe side, but not that for $p = 0.47$. Similar plots for other simulations described in this paper showed that all results reported below are converged and reliable.

Error bars quoted in the following on raw data (partition sums, gyration radii and average numbers of perimeter sites or bonds) are straightforwardy obtained single standard deviations. Their estimate is easy since clusters generated in different tours are independent, and therefore errors can be obtained from the fluctuations of the contributions of entire tours (note that clusters within one tour are *not* independent, and estimating errors from their individual values would be wrong).

On the other hand, errors on critical exponents and on growth constants are obtained by extrapolation. This is an ill-posed problem, and therefore any error obtained this way is to some degree subjective. All such errors quoted in the following are based on plotting the data

[3] Note the very large values of $\ln W$. Much effort was spent to write the codes such that overflows were avoided.

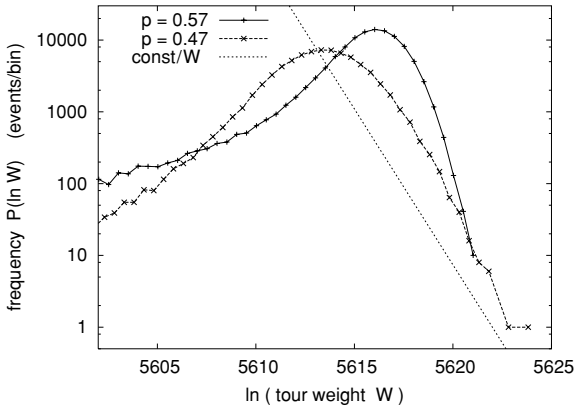

Figure 7. Log–log plot of distributions of tour weights of 2d animals with $N = 4000$, for $p = 0.57$ and for $p = 0.47$, together with a straight line representing the function $y = \text{const}/W$.

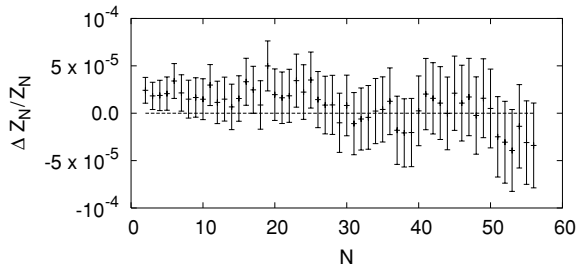

Figure 8. Deviations of the logarithm of the number of 2d site animals from exact enumerations of Jensen [40].

in different ways, plotting effective exponents against different powers of $1/N$, trying different ansatze for higher order correction to scaling terms, etc. They are *not* based on simply making least square fits over fixed intervals of N, as this could lead to very large underestimations of corrections to scaling. All quoted numbers are such that we believe, to the best of our knowledge, that the true value is most likely within one quoted error bar.

The total CPU time spent on the simulations reported in this paper is ≈25 000 h on fast PCs and work stations.

3. Site animals in two to nine dimensions

3.1. $d = 2$

Before we report our final results, we show one more test where we compare our raw data for $d = 2$ with the exact enumerations of [6]. In figure 8 we show the true relative errors of our estimates of the partition sum. Although there is some systematic trend visible, this is still within two standard deviations and thus not significant (note that our values for different N are not independent). Relative errors of the squared gyration radii are shown in figure 9. These data show that our estimates are basically correct, including the error bars.

Plotting directly our values of Z_N would not be very informative, neither would be a plot of $\ln Z_N - aN$, where $a = \ln \mu$. Both ways of plotting would hide any statistical errors.

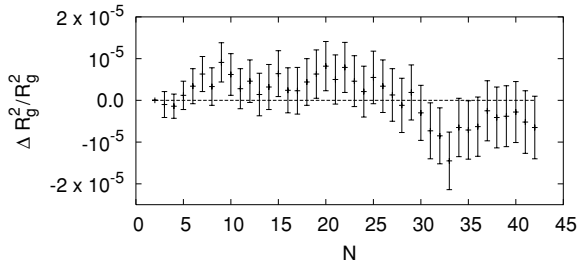

Figure 9. Relative deviations of the squared radii of gyration from exact values of [40].

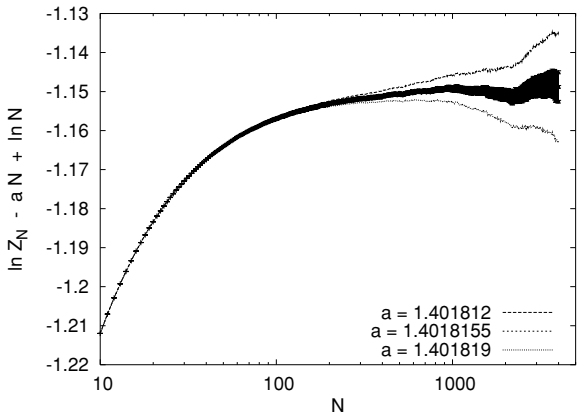

Figure 10. Plot of $\ln Z_N - aN + \theta \ln N$ against $\ln N$ for $d = 2$, with $a = 1.401\,812$, $1.401\,8155$ and $1.401\,819$ (top to bottom). Error bars are plotted only for the central curve.

A more meaningful way of plotting our full data for Z_N is used in figure 10, where we plot $\ln Z_N - aN + \ln N$ against N for three values of a. Error bars are shown only for the central curve, although all three curves have of course the same errors. In view of equation (1), and accepting the prediction that $\theta = 1$, we would expect a curve which becomes horizontal for large N. This is indeed seen for the central curve, but the obvious corrections to scaling make a precise estimate of μ difficult. The same is true for the gyration radii. In figure 11 we show $R_N^2/N^{2\nu}$ against N for three candidate values of ν. Again strong corrections to scaling are seen.

For these corrections one expects

$$Z_N \sim \mu^N N^{-\theta}(1 + b_Z N^{-\Delta} + \cdots) \tag{9}$$

and

$$R_N^2 \sim N^{2\nu}(1 + b_R N^{-\Delta} + \cdots), \tag{10}$$

where Δ is the correction exponent [9], b_Z and b_R are non-universal amplitudes, and the dots stand for higher order terms in $1/N$. Note that Δ is universal, and is the same in both equations. There are several methods discussed in the literature for estimating Δ. We estimated it by plotting $\ln Z_N - aN + \ln N$ and $R_N^2/N^{2\nu}$ against $x \equiv 1/N^{\delta}$. Straight lines are expected near $x = 0$ if and only if $\delta = \Delta$. We could not find a value of δ where these lines were absolutely straight for all x, but the straightest behaviour near $x \approx 0$ was obtained with $\delta \approx 0.9$, see figures 12 and 13.

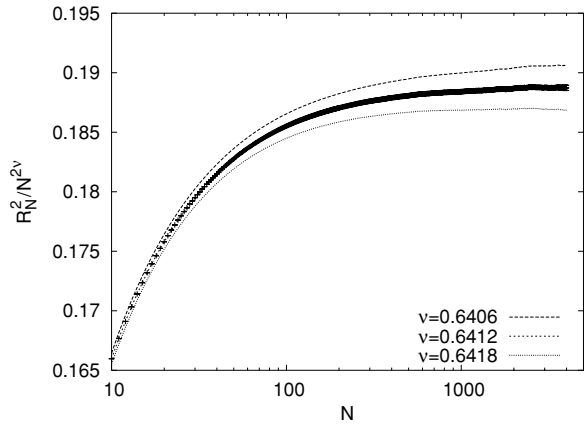

Figure 11. Plot of $R_N^2/N^{2\nu}$ against $\ln N$ for $\nu = 0.6406, 0.6412$ and 0.6418. Again, error bars are plotted only for the central curve.

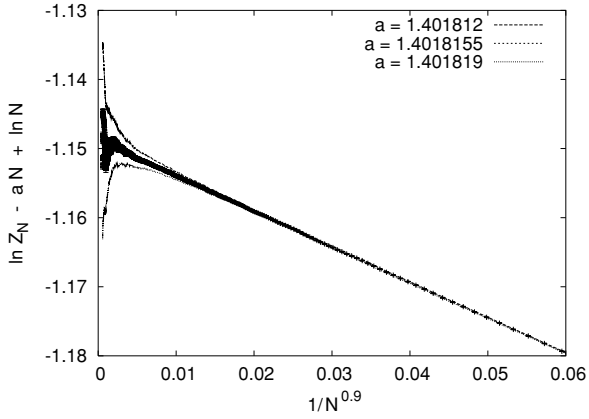

Figure 12. Same data as in figure 10, but plotted against $1/N^{0.9}$.

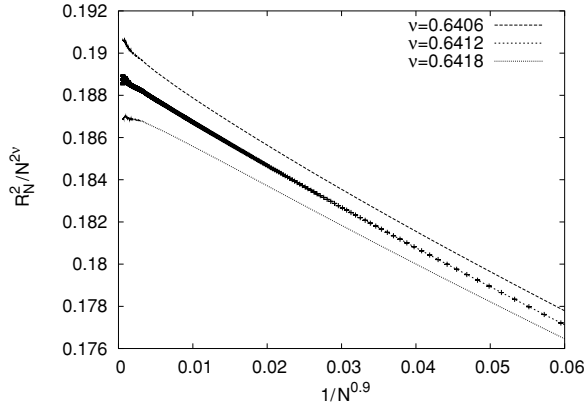

Figure 13. Same data as in figure 11, but plotted against $1/N^{0.9}$.

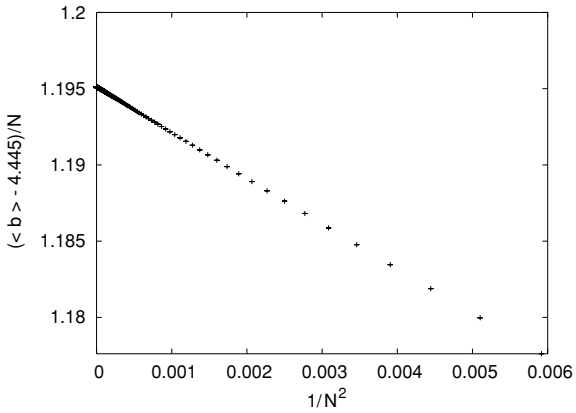

Figure 14. Average number of boundary sites per cluster site for 2d animals, plotted against $1/N^2$. To reduce finite size effects we actually subtracted 4.445 units from $\langle b \rangle$, before dividing it by N. Statistical errors are much smaller than symbol sizes.

We thus conclude that $\Delta = 0.9 \pm 0.1$ which suggests that the leading corrections to scaling are analytic ($\Delta = 1$ exactly). This is in agreement with the exact enumerations of [6, 39, 40] and with the exactly known correction exponent for the Lee–Yang problem [9, 74], but disagrees with the Monte Carlo estimate 0.65 ± 0.20 of [55]. Note that originally [11] the connection between the Lee–Yang and animal problems was established only for the leading terms, and therefore the authors of [9] suggested not to use the Lee–Yang correction to scaling exponents for animals. But the recent proof of [12] gives an exact mapping between two models in the respective universality classes, and therefore we should use the mapping also for the corrections to scaling.

The critical exponent ν and the growth constant μ can be read off figures 12 and 13, and are reported in table 1. The latter contains also our main results for all other dimensions. We see that our estimates for μ and ν are still much worse than the results obtained by the extremely long exact enumerations of Jensen, but they are more precise than all other previous estimates.

We have also made 'unbiased' fits where we did not assume the theoretical values $\theta = 1$. We do not show details, we just mention that our data would seem to exclude $|\theta - 1| > 0.002$.

Finally, we show in figure 14 the average numbers of boundary sites. More precisely, with $\langle b \rangle$ being this average, we plot $(\langle b \rangle - 4.445)/N$ against $1/N^2$. Subtracting 4.445 units was done in order to reduce finite size effects. Without the very large correction $4.445/N$, this term would dominate any other correction term, and would mask in particular any possible non-analytic correction. The linear shape of the curve suggests that there are no non-analytic corrections, and that the next to leading term is $\sim 1/N^2$, but the data are too poor to allow a firm conclusion.

3.2. $d > 2$

For $d = 3$ we show the data for Z_N and R_N^2 in figures 15 and 16, plotted in the same way as in figures 12 and 13. Now the straightest line is clearly obtained for $\delta < 1$, i.e. there definitely seem non-analytic corrections to scaling. The best fit was obtained with $\Delta = 0.75 \pm 0.08$ (upper panels in figures 15 and 16). But from the Lee–Yang problem we know [74] that we have $\Delta = 1$ also in $d = 3$. As seen from the lower panels in figures 15 and 16, where these data are plotted against $1/N$, this is clearly not supported by our data. But we cannot, of course,

Table 1. Main results for site animals. For convenience we also give in the second column the critical p values for site percolation.

d	p_c	$a = \ln \mu$	θ	ν	$\frac{\theta}{(d-2)\nu+1}$	Δ	Method
2	0.5927	1.401 815 5(30)	1[a,c]	0.6412(5)	–	0.9(1)	Present work
		1.401 815 696(5)		0.641 15(5)	–	1.0	Series [6, 39]
				0.642(10)	–	0.65(20)	MC [55]
3	0.311 6	2.121 858 8(25)	3/2[a,c]	1/2[a,c]	1[a,c]	0.75(8)	Present work (partially constrained)
		2.121 859 2(20)	3/2[a,c]	1/2[a,c]	1[a,c]	1[a,c]	Present work (constrained)
		2.120(2)					Series [75]
			1.502(3)			1[a,c]	Series [74]
				0.498(10)		0.54(12)	MC [55]
4	0.1968	2.587 858(6)	1.835(6)	0.416 3(30)	1.001(7)	0.57(8)	Present work (unrestricted)
		2.587 858 3(40)	1.833(5)	0.418 1(25)	0.998(4)	5/6[a,c]	Present work (partially constrained)
		2.587 858 3(30)	1.834(4)	0.417(2)	1[a,c]	5/6[a,c]	Present work (partially constrained)
		2.587 848 3(30)	11/6[a,c]	5/12[a,c]	1[a,c]	5/6[a,c]	Present work (constrained)
			1.839(8)			5/6[a,c]	Series [74]
		2.601 2(15)					MC [46]
				0.415(11)		0.46(11)	MC [55]
5	0.140 7	2.922 319 4(60)	2.080(7)	0.359(4)	1.001(9)	0.47(7)	Present work (unrestricted)
		2.922 320 5(30)	2.081 5(60)	0.360 5(20)	1[a,c]	0.622[b,c]	Present work (constrained)
			2.087 7(25)			0.622(12)	Series [74]
			2.080 7				ϵ-expansion [74]
			2.10(3)	0.367(11)	1[a,c]	0.65(15)	Series [9]
		2.899(9)					Expansion in $1/(2d-1)$ [76]
		2.940(15)					MC [46]
				0.359(11)		0.40(14)	MC [55]
6	0.109 0	3.178 524 5(40)	2.261(12)	0.315(4)	1.000(12)	0.39(6)	Present work (unrestricted)
		3.178 521(3)	2.256(8)	0.314(2)	1[a,c]	0.412[b,c]	Present work (constrained)
			2.264 8(15)			0.412(8)	Series [74]
			2.264 9				ϵ-expansion [74]
			2.30(4)	0.325(10)	1[a,c]	0.5(2)	Series [9]
		3.172(3)					Expansion in $1/(2d-1)$ [76]
		3.20(2)					MC [46]
				0.321(19)		0.34(13)	MC [55]
7	0.088 9	3.384 080(5)	2.40(2)	0.282(5)	0.996(20)	0.26(6)	Present work (unrestricted)
		3.384 079(3)	2.390(9)	0.278(2)	1[a,c]	0.205[b,c]	Present work (constrained)
			2.402(5)			0.205(5)	Series [74]
			2.499 9				ϵ-expansion [74]
			2.41(3)	0.282(6)	1[a,c]	0.4(2)	Series [9]
		3.382(1)					Expansion in $1/(2d-1)$ [76]
		3.41(1)					MC [46]
				0.291(11)		0.35(7)	MC [55]
8	0.075 2	3.554 827(4)	5/2[a,c]	1/4[a,c]	1[a,c]	0 (+logs)	Present work
		3.554 4(7)					Expansion in $1/(2d-1)$ [76]
9	0.065 2	3.700 523(10)	5/2[a,c]	1/4[a,c]	1[a,c]	0.25(5)	Present work

[a] Exact value.
[b] From [74].
[c] Used as constraint in the fit.

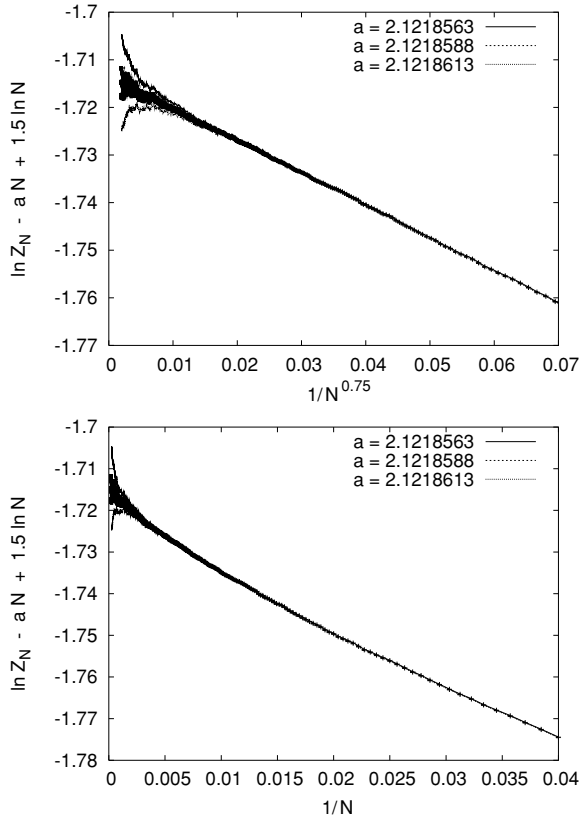

Figure 15. Similar to figure 12, but for $d = 3$. The straightest curve was now obtained by plotting the data against $1/N^{0.75}$ (upper panel). In the lower panel, the same data are plotted against $1/N$.

exclude the possibility that this is due to very large higher order corrections to scaling. In view of this we show in table 1 two fits, one unrestricted where Δ is fitted from the present data and one constrained fit where $\Delta = 1$ is imposed. In both fits the Parisi–Sourlas condition $\theta = 1 + (d-2)\nu$ and the exact values $\nu = 1/2$ and $\theta = 3/2$ were also used as constraints. Fits without imposing $\nu = 1/2$ and $\theta = 3/2$, and without assuming the Parisi–Sourlas relation, gave bigger errors for the growth constant, but gave exponents in full agreement with the predictions: $\nu = 0.500 \pm 0.002$ and $\theta = 1.500 \pm 0.001$.

The problem with the correction to scaling exponent is obviously due to large subleading corrections. It persists also in higher dimensions. For $d = 4$, e.g., we estimated $\Delta = 0.57 \pm 0.08$, while the exact value obtained from the Lee–Yang problem is $\Delta = 5/6$. We present therefore in table 1 four different fits with various constraints: One completely unrestricted, another with $\Delta = 5/6$ imposed, a third with the Parisi–Sourlas relation imposed in addition, and a final one with even the values $\theta = 11/6$ and $\nu = 5/12$ fixed. Note that the growth constant can be obtained without knowing θ, if the Parisi–Sourlas relation is assumed: in this case $N Z_N R_N^{d-2} \sim \mu^N (1 + \text{const}/N^{\Delta} + \cdots)$. From the values listed in table 1 we see that all four fits are mutually consistent.

Similar fits were also made for $d = 5, 6$ and 7, but we list in table 1 only the results of the unrestricted and of the completely restricted fits. In all the cases the agreement between the fits is very good, showing the consistency of the data.

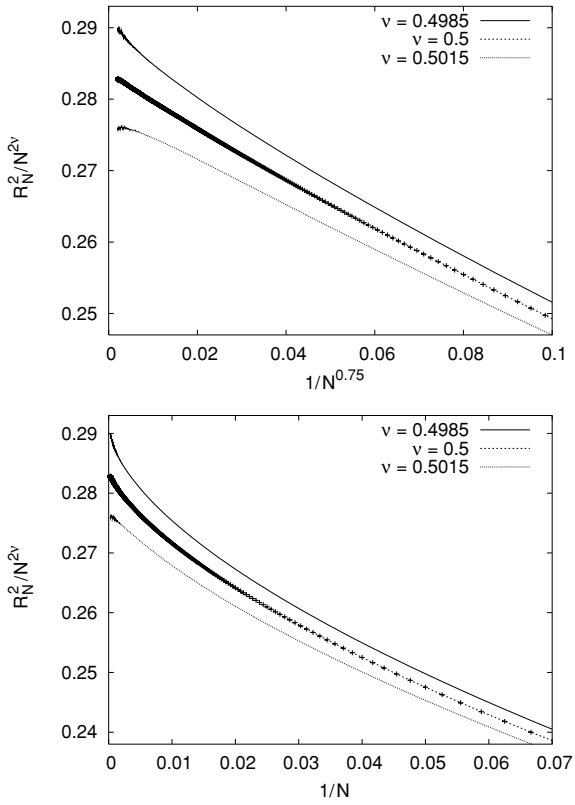

Figure 16. Similar to figure 13, but for $d = 3$. In the upper panel, $R_N^2/N^{2\nu}$ is plotted against $1/N^{0.75}$, in the lower panel it is plotted against $1/N$.

For $d = 8$ we show in figures 17 and 18 the data for free energies and for gyration radii, plotted against $\ln N$. From these plots one sees clearly that the data agree with the predicted exponents $\theta = 5/2$ and $\nu = 1/4$. But there are very large (presumably logarithmic) corrections, compatible with the fact that $d = 8$ is the upper critical dimension. We have not tried to make a detailed fit to these corrections, since we are not aware of any theoretical prediction beyond the leading order [9], and because verifying logarithmic corrections is notoriously difficult.

In figure 19, the average number of boundary sites are plotted in a way similar to figure 14. This time an even bigger term $6.5/N$ had to be subtracted, in order to see any possible non-analytic term. The fact that the curve is reasonably straight when plotted against $1/N^2$ suggests again (as for $d = 2$) that there is no non-analytic correction term. Our estimates for the critical exponents and for the growth constant are given in table 1.

The estimates for ν and for $(\theta - 1)/(d - 2)$ obtained by the unrestricted fits are also shown in figure 20. According to Parisi and Sourlas, they should coincide. The agreement is practically perfect. Moreover, both estimates have roughly the same errors, and estimating ν indirectly, using equation (3), seems to give slightly smaller errors for $d \geqslant 5$ than the direct measurement. For $d = 6$ and $d = 7$ our results are in very good agreement with the $\epsilon(= 8 - d)$-expansion results of [34]. Our results are also in very good agreement ($\leqslant 1\sigma$) with

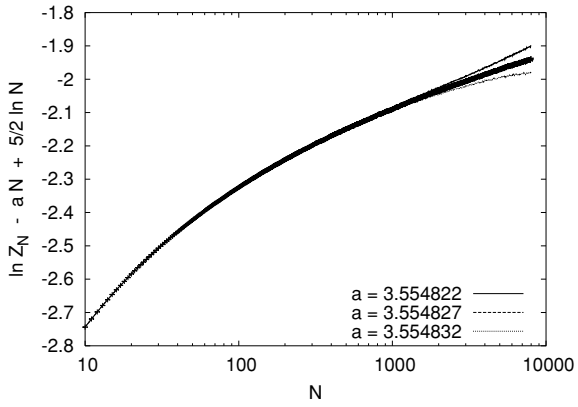

Figure 17. $\ln Z_N - aN + 2.5 \ln N$ for $d = 8$, plotted against $\ln N$, with three different values of a: 3.554 822, 3.554 8327, and 3.554 832 (top to bottom). Error bars are again plotted only for the central curve.

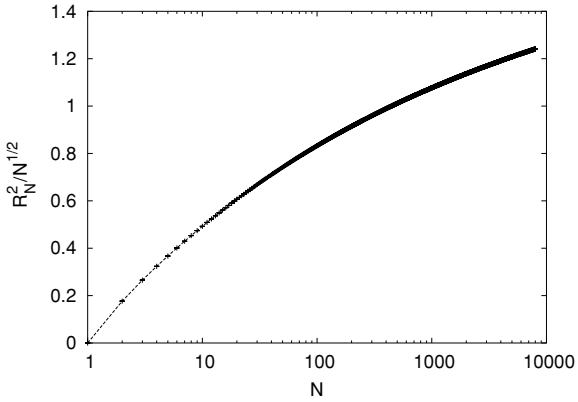

Figure 18. Plot of $R_N^2/\sqrt{N}$ against $\ln N$ for animals in $d = 8$.

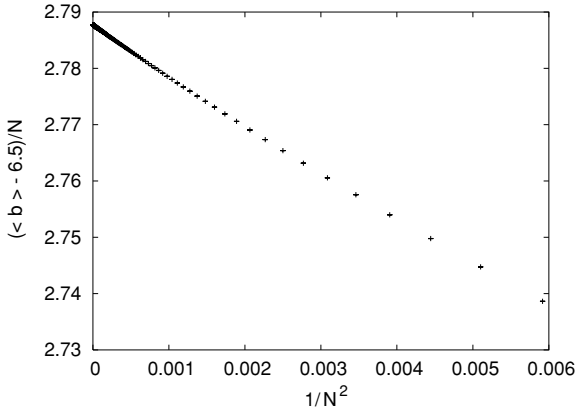

Figure 19. $(\langle b \rangle - 6.5)/N$, where $\langle b \rangle$ is the average number of boundary sites per cluster site for 3d animals, plotted against $1/N^2$. Statistical errors are much smaller than symbol sizes.

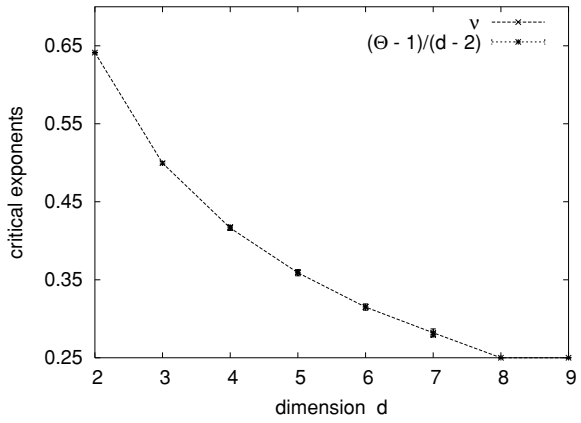

Figure 20. The critical exponents ν and $(\theta - 1)/(d - 2)$ against d.

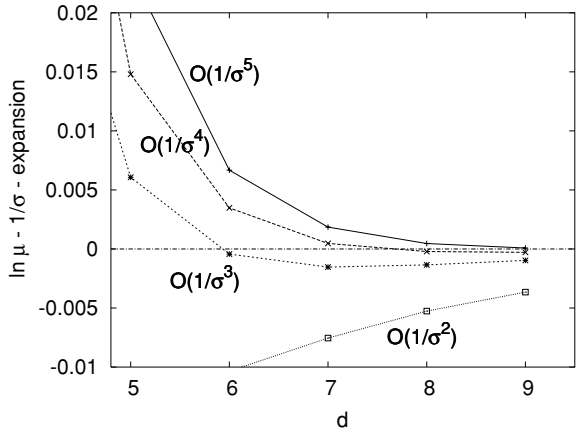

Figure 21. Difference between the measured values of $\ln \mu(d)$ and successive truncations of the expansion equation (12), plotted against d. Error bars are smaller than the sizes of the symbols.

the series expansion of the *binary Gaussian molecule mixture* of [74], which gives for high dimensions the most precise previous exponents for the Lee–Yang problem.

The growth constants seem to grow linearly with dimension,

$$\mu(d) \approx 5.49d - 8.94 \tag{11}$$

for large d, although there are small but statistically significant deviations. More precise expressions for the large-d behaviour are obtained from an expansion of $\ln \mu$ in powers of $1/\sigma$ with $\sigma = 2d - 1$ [76]:

$$\ln \mu(d) = \ln \sigma + 1 - \frac{2}{\sigma} - \frac{79}{24\sigma^2} - \frac{317}{24\sigma^3} - \frac{18\,321}{320\sigma^4} - \frac{123\,307}{240\sigma^5} + O(\sigma^{-6}). \tag{12}$$

A comparison of our data with different truncations of this expansion is shown in figure 21. This comparison suggests strongly that the expansion is only asymptotic: for any fixed d, there is an order beyond which the expansion gives values smaller than the true value, and continues to decrease with increasing order.

Our estimates for the asymptotic number of boundary sites per cluster site are given in table 2. The latter were all obtained by assuming no non-analytic corrections to scaling, since

Table 2. Asymptotic ratios between boundary and cluster sites, $\lim_{N\to\infty}\langle b\rangle/N$ (column #2); ratios between partition sums of bulk and wall-grafted animals (column #3); and average number of contact of wall-grafted animals with that wall (column #4).

Dimension	$\lim_{N\to\infty}\langle b\rangle/N$	$\lim_{N\to\infty} Z'_N/Z_N$	$\langle m\rangle$
2	1.1951(1)	1.987(8)	2.892(2)
3	2.7877(1)	2.97(3)	5.07(1)
4	4.5859(2)	3.98(5)	7.50(5)
5	6.4909(2)	4.91(4)	10.12(6)
6	8.4503(1)	–	–
7	10.4363(2)	–	–
8	12.4346(1)	–	–
9	14.4378(2)	–	–

our data can be fitted for all d to $\langle b\rangle/N = \beta + \beta_1/N + \beta_2/N^{\Delta'}$ with $\Delta' \approx 2$. For large d, our data seem to scale as

$$\beta(d) = \lim_{N\to\infty}\langle b\rangle/N \sim 2d - \text{const.} \tag{13}$$

4. Animals attached to a wall

4.1. Athermal walls

In this section we will consider d-dimensional animals grafted with one monomer to an impenetrable planar wall modelled by a hyperplane $x_d = 0$. For this case, it was shown in [18] that the partition sum, written now Z'_N instead of Z_N, scales as

$$Z'_N \sim \mu^N N^{-\theta'} \tag{14}$$

with the same μ as in the bulk, and with

$$\theta' = \theta. \tag{15}$$

The last equation has a very simple heuristic explanation. Let us first map lattice points $\mathbf{x} = (x_1, \ldots, x_d)$ on a lattice of size L^d onto integers

$$i_{\mathbf{x}} = x_1 + x_2 L, \ldots, x_d L^{d-1} \tag{16}$$

(we actually used this in our codes to index points by a single integer, as this simplifies programming and makes memory access faster). Consider now the problem of counting all animals restricted to the half space $I_+ = \{\mathbf{x} : i_{\mathbf{x}} \geqslant i_0 \equiv i_{\mathbf{x}_0}\}$ and positioned such that the site $\mathbf{x}_0$ belongs to the animal. On the one hand, this means just that we consider animals with fixed positions: for each shape we consider only that animal whose 'smallest' point is $\mathbf{x}_0$. Since we had counted only once all cluster shapes related by translations, this means that the partition sum obtained now is exactly equal to Z_N. On the other hand this model is equivalent to the animal being grafted to an impenetrable wall located at $x_d \approx x_{d,0}$ which is however not quite flat: x_d jumps from $x_{d,0} + 1$ to $x_{d,0}$ when any one of the other coordinates x_j $(j = 1, \ldots, d-1)$ goes through $x_{j,0}$. The proof of [18] then just shows that the scaling behaviour is independent of these steps, and is the same as for a flat surface. In addition, this argument shows that $Z'_N > Z_N$ for all N [18]. Indeed, one easily sees that the ratio $\langle m_N\rangle_0 = Z'_N/Z_N$ is just the average number of contacts a *free* animal in the bulk would have with a flat imaginary wall placed just below it. This is *not* equal to the average number of contacts $\langle m_N\rangle$ of a *grafted* animal with its wall, because the latter animals are counted m times

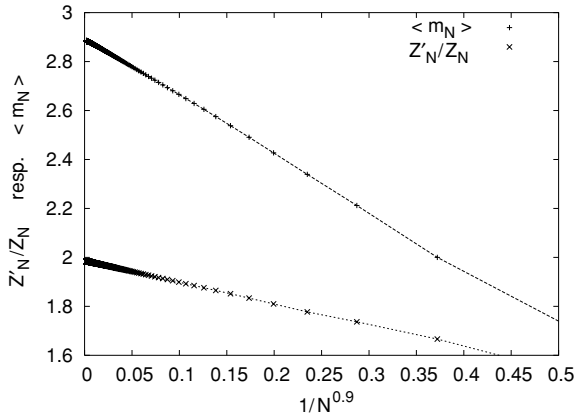

Figure 22. Ratio Z'_N/Z_N for 2d site animals, plotted against $1/N^{0.9}$ (lower curve) and average number of contacts of grafted 2d animals (upper curve). Statistical errors are smaller than the sizes of the points.

if they have m contacts: if we denote by $Z_{N,m}$ the number of configurations with m sites in the bottom hyperplane and N sites total (so that $Z_N = \sum_m Z_{N,m}$), and by $Z'_{N,m}$ the analogous quantity for grafted animals, then $Z'_{N,m} = mZ_{N,m}$. Therefore, the average number of contacts of a grafted animal is given by the *second* moment of m in the bulk ensemble divided by the first moment, $\langle m_N\rangle \equiv \sum_m mZ'_{N,m} \big/ \sum_m Z'_{N,m} = \sum_m m^2 Z_{N,m} \big/ \sum_m mZ_{N,m} \equiv \langle m_N^2\rangle_0 \big/ \langle m_N\rangle_0$.

In figure 22 we show the ratio Z'_N/Z_N against $1/N^{0.9}$, together with the average number of contacts $\langle m_N\rangle$ of grafted animals, for $d = 2$. We see again straight lines, showing that

$$Z'_N/Z_N, \quad \langle m_N\rangle \sim \text{const} - \frac{\text{const}}{N^{\Delta'}} \tag{17}$$

with $\Delta' = 0.9 \pm 0.05$. Similar results were found for larger dimensions. The values of Δ' are close to those of Δ but somewhat larger, and we see no theoretical reason why they should be the same. We do not quote numbers since they are rather poorly determined. The asymptotic values of Z'_N/Z_N and $\langle m_N\rangle$ are given in table 2.

4.2. Animals attached to an attractive surface

The partition sum now is written as

$$Z_N^{(1)}(q) = \sum_{m=1}^{N} A_N(m) q^m \tag{18}$$

where $A_N(m)$ is the number of configurations of lattice animals with N site having m sites on the walls, and $q = \mathrm{e}^{\epsilon/KT}$ is the Boltzmann factor, $\epsilon > 0$ is the attractive energy between the monomer and the wall.

As $q \to 1$, there is no attraction between the monomer and the wall, i.e. $Z_N^{(1)}(1) = Z'_N$. On the other hand it becomes clear that any cluster will collapse onto the wall, if q becomes sufficiently large. Therefore we expect a phase transition from a grafted but otherwise detached to an adsorbed phase, similar to the transition observed also for linear polymers.

Exactly at the transition point $q = q_c$ we expect the usual scaling laws

$$Z_N^{(1)}(q_c) \sim \mu^N N^{-\theta_s} \tag{19}$$

and

$$R_N \sim N^{\nu}. \tag{20}$$

In analogy to critical surface phenomena where this transition would correspond to the 'special' point [77], we expect ν to be the same Flory exponent as in the bulk, while θ_s should be a new and independent exponent. The growth constant μ, although being not universal, should be the same as in the bulk.

Away from the critical point we expect a scaling ansatz

$$Z_N^{(1)}(q) \sim \mu^N N^{-\theta_s} \Psi[(q - q_c)N^{\phi}], \tag{21}$$

with the crossover exponent ϕ being a second new exponent. Taking the derivative of $\ln Z_N^{(1)}(q)$ with respect to q and setting $q = q_c$ thereafter, we obtain for the average energy

$$E_N(q_c) = \langle \epsilon m \rangle \sim N^{\phi}. \tag{22}$$

Taking two derivatives we obtain for the specific heat per monomer near (but not exactly at) the critical point

$$C_N(q) = \frac{1}{NKT^2}(\langle (\epsilon m)^2 \rangle - \langle \epsilon m \rangle^2) \sim (q - q_c)^{-\alpha} \tag{23}$$

with

$$\alpha = 2 - 1/\phi, \tag{24}$$

while

$$C_N(q_c) \sim N^{2\phi - 1}. \tag{25}$$

In principle, all four scaling laws can be used to locate the critical value q_c. With conventional (Metropolis type) Monte Carlo simulations one cannot use easily equation (19), since precise estimates of the partition sum are difficult to obtain. In this case it is usually equation (23) which is used. With PERM we do have very precise estimates of $Z_N^{(1)}(q)$, and therefore we can use equation (19), but we shall see that it is indeed equation (22) which gives—together with the two others—the most precise estimate. This is very similar as for adsorption of linear polymers [78].

In the following we shall assume $\epsilon = 1$ without loss of generality. In order to compare with previous analyses we want to have specific heats for discrete values of N, but for a continuous range of q. They are most easily obtained from histograms

$$P(m; q) = \sum_{\text{Config.}} q^{m'} \delta_{m,m'} \tag{26}$$

which are normalized such that $Z_N^{(1)}(q) = \sum_m P(m; q)$. Note that we obtain from the simulations not only the shape of the histogram, but also its absolute normalization, which makes it easy to combine two histograms obtained in runs with different nominal values of q. All we have to know are rough values of their relative statistical errors. These we can estimate from the number of tours which contribute to a particular value of m, $\Delta P(m; q) \propto 1/\sqrt{\#\text{tours}}$. Although this estimate is not very precise, it is fully sufficient to obtain smooth global histograms by joining histograms which cover narrow regions in m.

Specific heats for 2d animals with lengths up to $N = 1200$ are shown in figure 23. These data are very similar to the results of [53], although the latter are for trees. According to Janssen and Lyssy [15–17] we expect $\alpha = 0$, i.e. the specific heat curves for different N should intersect exactly at the critical point. This gives roughly $q_c = 2.27$. But a close look at the inset in figure 23 reveals that these intersections slightly shift to larger q as N increases. Thus

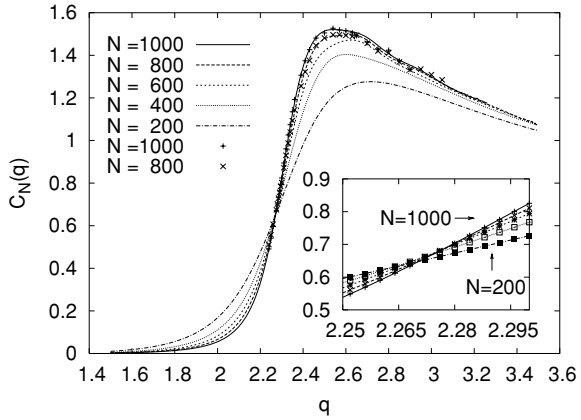

Figure 23. Specific heats $C_N(q)$ per monomer for 2d site animals, plotted against q for various values of N. Smooth curves show results obtained by histogram reweighing, points indicate results of single runs. The inset shows the region around $q \approx 2.27$ where the curves intersect.

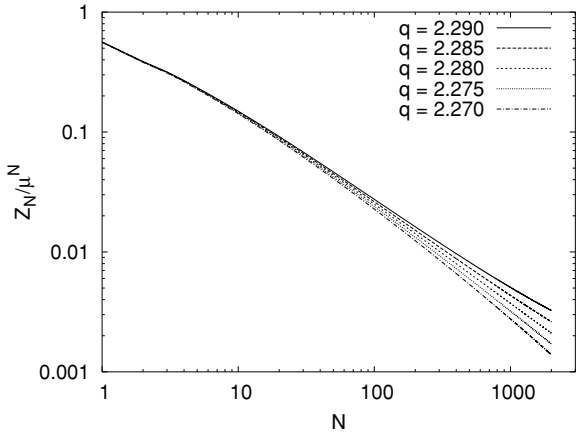

Figure 24. $Z_N^{(1)}(q)/\mu^N$ against N for various values of q, for site animals on the square lattice grafted to a wall. Error bars here and in the following figures are comparable to the thickness of the lines.

we have considerable corrections to scaling, preventing us from attributing error bars to this estimate.

Alternatively, we turn towards the partition sum itself. In figure 24 we show log–log plots of $Z_N^{(1)}(q)/\mu^N$, for various values of q close to q_c. We see the expected power law, but determining the critical point from this figure is difficult because of the substantial corrections to scaling. We thus multiply with an estimated power N^{θ_s} and plot the data against $1/N^{0.8}$. The result is shown in figure 25 (where we actually plot the logarithm on the y-axis). Note that we did not have to make a new estimate of μ; rather, we could take the old and very precise estimate. The value $\theta_s = 0.87$ was chosen so as to give the best straight line for a suitably chosen q_c. Indeed, from this plot we would conclude that $q_c \approx 2.278$.

We now turn towards equation (22). In figure 26 we plotted $E/N^{0.48}$ instead of $E/\sqrt{N}$. This was chosen because it suggests that $q_c \approx 2.278$, in agreement with the value obtained above from Z_N. Assuming $\phi = 1/2$, in contrast, would have given $q_c \approx 2.283$ which would be incompatible with the data for Z_N.

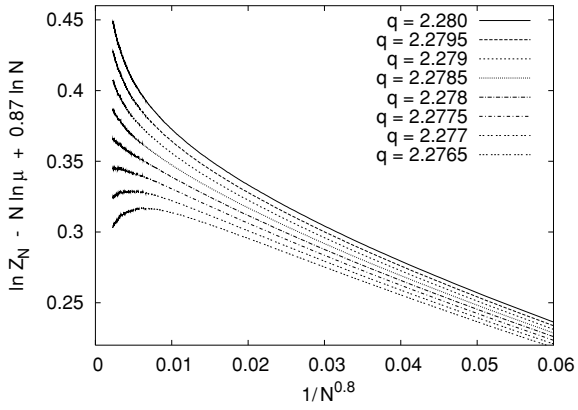

Figure 25. Part of the data shown in figure 24, after interpolating them to a finer set of q-values, multiplying them with $N^{0.87}$, and taking their logarithm, plotted against $1/N^{0.8}$.

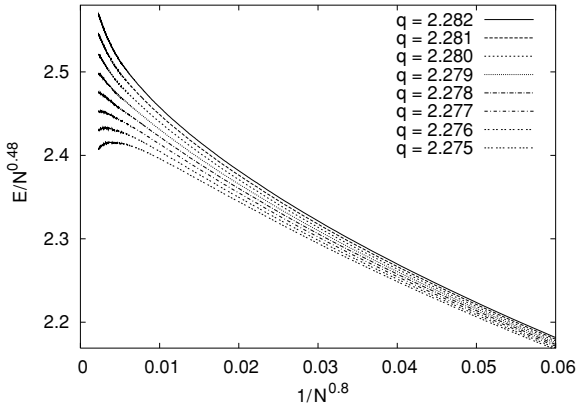

Figure 26. $E/N^{0.48}$ against $1/N^{0.8}$ for various values of q, for site animals on the square lattice grafted to a wall.

Finally, we plot $N^{1-2\phi}C_N(q)$ against $1/N^{0.8}$, in order to compare with equation (25). As seen from figure 27 this is fully consistent with $q_c \approx 2.278$. If we had taken $\phi = 1/2$, we would again get a too large estimate $q_c \approx 2.285$.

Summarizing, we obtain as our best estimates:

$$q_c = 2.2778 \pm 0.0008, \qquad \phi = 0.480 \pm 0.04 \tag{27}$$

together with

$$\theta_s = 0.870 \pm 0.009, \qquad \Delta_s = 0.8 \pm 0.2. \tag{28}$$

The large error of Δ_s reflects the fact that the best estimates obtained from the different observables would be quite different, suggesting again large non-leading corrections. But this seems to have little effect on the estimates of the other quantities. Since we believe that we have taken into account all systematic errors, we claim that the Janssen–Lyssy conjecture $\phi = 1/2$ is slightly but significantly violated in $d = 2$. The previous estimates $\phi = 0.505(15)$ [56] and $\phi = 0.503(3)$ [57] most likely suffer from such systematic errors. On the other hand, our result is in agreement with the MC estimate $\phi = 0.50(3)$ of [53]. Surprisingly, our estimate

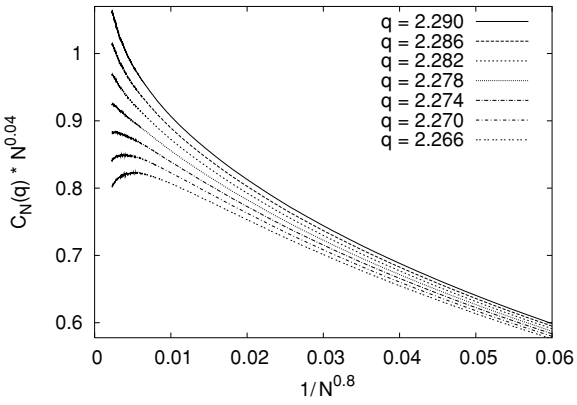

Figure 27. Part of the data shown in figure 23, but multiplied by $N^{1-2\phi}$ and plotted as curves with fixed q against $1/N^{0.8}$.

for ϕ agrees within the error bars with the most recent estimate of the cross-over exponent for unbranched polymers attached to an attractive wall in three dimensions, $\phi_{\mathrm{SAW}} = 0.484(2)$ [79], while $\phi_{\mathrm{SAW}} = 0.5$ for unbranched 2d SAWs [80].

Before leaving this problem, we should discuss the gyration radii. Their behaviour near the critical adsorption point is somewhat more complicated. For $q > q_c$ the gyration radius scales as $R_N \sim N$. At $q \leqslant q_c$ we expect it to scale as $R_N \sim N^{\nu}$ with the same value of ν as in the bulk, as in other surface critical phenomena [77]. The effect of the wall is only seen then in the amplitude $A = \lim_{N\to\infty} \langle R_N^2 \rangle / N^{2\nu}$. For $q < q_c$ it should be larger than for free animals, and for $q = q_c$ we should expect it to be even larger, $A_{\mathrm{bulk}} < A_{\mathrm{wall}} < A_c$. The reason is that the main effect of the wall is to squeeze the animal in the direction perpendicular to the wall, which by the excluded volume effect makes it more extended in the direction parallel to the wall.

This is somewhat analogous to the case of an unbranched polymer between two athermal walls [81]: if the distance D between the walls is decreased, at first the shrinking of the perpendicular extension dominates any increase parallel to the plates. However, if D is much smaller than the Flory radius, the stretching parallel to the walls dominates, and R_N increases in comparison to a free polymer [81].

Our data (figure 28) indicate that R_N/N^{ν} is larger than for animals in the bulk (where $A_{\mathrm{bulk}} \approx 0.189$), and that it increases with q. But at $q \approx q_c$ it is not monotonic in N: it increases with N until $N \approx 300$, and then decreases sharply. This strange behaviour might have been expected from the analogy with unbranched self-avoiding walks between two athermal walls. For small N and $q = q_c$ the effect of the wall is strong, and the increase of the size parallel to the wall dominates. But for $N \to \infty$ the effect of the wall becomes increasingly weaker, and the stretching along the wall becomes less important. We verified that it is indeed the slower increase of the parallel component which lets R_N/N^{ν} decrease for large N, but we found no similar effect in simulations (unpublished) of unbranched polymers at the critical adsorption point. Thus we have at the moment no good explanation for this effect.

For higher dimensions, the same kind of analysis as in $d = 2$ gave the estimates given in table 3. The main problem in these analyses is again that the best estimate for the leading correction to scaling exponents for the different observables did not quite agree with each other. This hints at the presence of more than one important term in the scaling corrections, and it dominates the error estimates. The most remarkable result seen in table 3 is the perfect

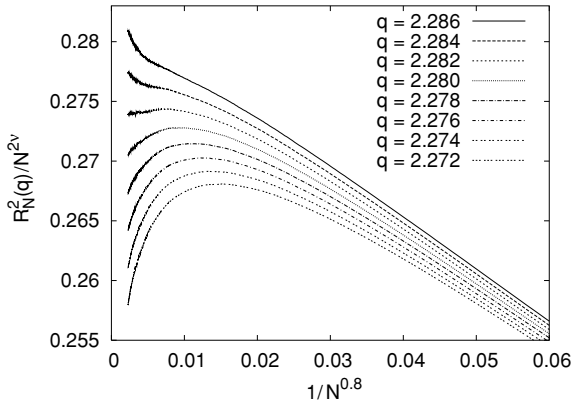

Figure 28. Ratios $\langle R_N^2 \rangle / N^{2\nu}$, where R_N is the gyration radius of grafted 2d animals, plotted against $1/N^{0.8}$.

Table 3. Critical Boltzmann factors, crossover exponents and critical exponents θ_s at the adsorption transition for site animals on simple (hyper-)cubic lattices grafted to a flat attractive wall.

Dimension	q_c	ϕ	θ_s
2	2.2778(8)	0.480(4)	0.870(9)
3	1.4747(6)	0.50(1)	1.476(7)
4	1.2674(6)	0.50(2)	1.91(1)
5	1.1786(5)	0.51(3)	2.18(4)

agreement with the Janssen–Lyssy prediction $\phi = 1/2$ in all dimensions $\geqslant 3$. In particular, it seems that the strong violation seen in [58] for $d = 3$ was due to an underestimation of finite size effects. Actually, Janssen and Lyssy had derived $\phi = 1/2$ only for $d = 3, 4$ and $\geqslant 8$, but not for $d = 2$.

Finally, we show in figure 29 the monomer density profile $\rho(z)$ for 3d animals, where z is the distance from the wall. Animal sizes for this figure were $N = 4000$. All curves must of course go to zero for $z \to \infty$. At temperatures far above T_c, i.e. at $q \ll q_c$, the entropic repulsion from the wall dominates and $\rho(z)$ has a maximum at a finite value of z. In the adsorbed phase $\rho(z)$ has its maximum at $z = 0$ and decreases monotonically with z. Note that the transition from non-monotonic to monotonic behaviour does not happen exactly at q_c, but for q slightly smaller than q_c. Presumably this is a finite size effect, and the transition would happen at q_c for much larger animals.

5. Trees and bond animals

5.1. Site trees

The simplest modification of the codes presented so far is needed for simulating site trees. As we pointed out in the introduction, site trees are site animals without loops. Thus the number of nearest neighbour pairs is just $N - 1$ for a tree of N sites. It is easy to count the number of occupied nearest neighbour pairs as the cluster grows. We have just to prune the growth as soon as this number is equal or larger than N. Apart from that, pruning and branching is done exactly as before, and all comments made in section 2 about the efficiency of the algorithm apply also to site trees.

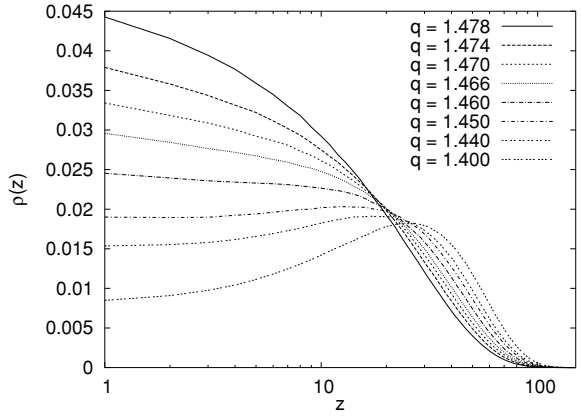

Figure 29. Monomer density profiles $\rho(z)$ (with z being the distance from the wall) for a 3d animal grafted to an attracting wall.

We made simulations only for 2d site trees, and only with rather modest statistics. Our results were fully in agreement with those of Jensen [40]. In particular we obtained $\mu = 3.795\,27(4)$ after three days of CPU time on a 3 GHz Pentium, to be compared to the estimate 3.795 254(8) obtained in [40].

5.2. Bond animals and bond trees

In order to simulate bond animals and bond trees, one has to grow bond instead of site percolation clusters. Cluster growth algorithms for bond percolation are very similar to those for site percolation and about as easy. One just has to remember that in bond percolation one often does not distinguish between clusters with the same configurations of sites, but with different bond configurations. However, for animals it is essential to make this distinction.

Let us denote by k the number of non-bonded nearest neighbour pairs (often called 'contacts' in this context), by b the number of surface bonds, and by m the number of established bonds between nearest neighbours. In the bond percolation ensemble, a cluster with these 'quantum numbers' has a weight (cf equation (4))

$$P_{Nkbm} = p^m(1-p)^{b+k}. \tag{29}$$

This is slightly more complicated than in the site percolation case, but one can follow the same strategy when using this to simulate (bond) animals. We just have to replace the number of perimeter sites in the weight factor by $b + k$, and if we want to simulate trees, we have of course to prune all configurations which are not tree-like. Growth sites have to be replaced by growth bonds. Moreover, it is a bit more natural to consider ensembles with fixed m, i.e. with fixed number of established bonds, rather than sites [8].

The heuristics worked out in section 2 remain valid: one obtains much better results when the trees are grown breadth first instead of depth first; one should use a fitness function $f(\mathcal{C}) = (1-p)^{-g}W(\mathcal{C})$, where g is now the number of growth *bonds*; and one should simulate at a slightly subcritical value of p which approaches p_c as the trees to be simulated become larger and larger. For the same values of N, the optimal values of p/p_c were however somewhat smaller. For $d = 2$ and $N = 1000$, e.g., best results were obtained with $p \approx 0.45$ (with $p_c = 1/2$).

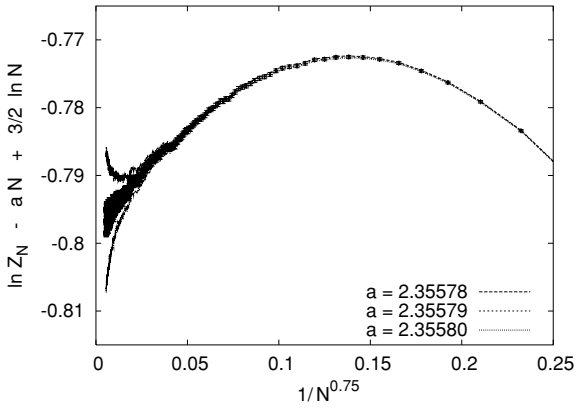

Figure 30. Plot of $\ln Z_N - aN + 3/2 \ln N$ against $1/N^{0.75}$, for bond trees on the simple cubic lattice.

Let us first discuss trees. In two dimensions, we obtained $\mu = 5.142\,76 \pm 0.000\,02$. This is compatible with the best previous estimate, $\mu = 5.143\,39 \pm 0.000\,72$ [54], but more than an order of magnitude more precise. Our estimate is based on rather small clusters ($N_{\max} = 500$), but very high statistics: The error of $\ln Z_N$ is $\Delta \ln Z_{N=500} = 0.000\,451$. These simulations confirmed also that the leading correction to scaling exponent is between 0.9 and 1.0. In three dimensions we obtained $\mu = 10.546\,46 \pm 0.000\,10$. This case is a bit special, since it has strange looking corrections to scaling, as seen from figure 30. Apart from figure 28, this is our clearest and most striking example showing that the corrections cannot be described by a single power, and illustrates the pitfalls in estimating correction to scaling exponents from poor data. In other cases, the existence of more than one power in corrections to scaling have often to be inferred less directly, e.g. by comparing different observables (see section 4.2) or by invoking universality [82].

General animals were simulated only with somewhat lower statistics, since we are not aware of any other recent high statistics simulations. On the square lattice we obtained $\mu = 5.207\,89 \pm 0.000\,04$, to be compared with the previous series expansions estimate 5.208 ± 0.004 of [75]. On the simple cubic lattice we got $\mu = 10.615\,39 \pm 0.000\,06$, to be compared with 10.63 ± 0.05 [34] and 10.62 ± 0.08 [83]. Mean-square gyration radii were e.g. $R^2_{N=499} = 347.974 \pm 0.036$ in $d = 2$ and $R^2_{N=999} = 165.669 \pm 0.022$ in $d = 3$, to be compared to the best previous MC estimates, 348.32 ± 0.88 and 166.03 ± 0.34 [8] (the quantity n displayed in the first lines of tables A1 to A7 of [8] is not the number of bonds, as stated there, but one plus this number).

Finally we point out that we can also use bond percolation as a starting point for the simulation of site animals. We just have to use the fact that site animals are isomorphic to the subset of bond animals with maximal number of bonds for a given configuration of occupied sites. Using this we obtained for $d = 2$ results in agreement with those of section 3, but the algorithm was somewhat less efficient than that based on site percolation.

6. Animal collapse

6.1. Collapse of site animals

In order to describe collapsing animals and/or trees, one has to introduce attractive monomer–monomer interactions. Historically the first model of this type [30, 31] starts from site animals

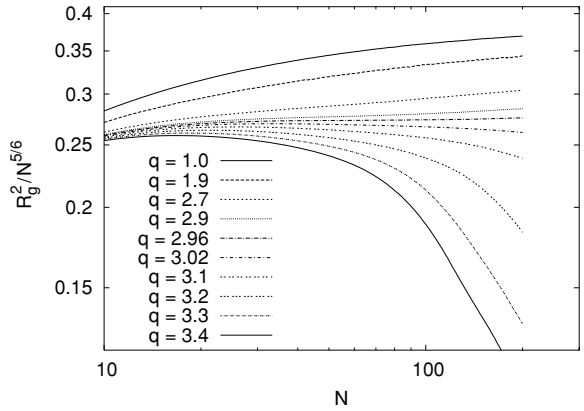

Figure 31. Squared gyration radii of collapsing site animals in four dimensions divided by $N^{2\nu}$, where $\nu = 5/12$ is the exact Flory exponent for non-collapsing animals.

and introduces a contact energy for each 'contact', where a contact is a pair of occupied nearest neighbour sites (note that the definition of contacts used here differs slightly from that used for bond animals). Let us denote this energy as $-\epsilon$, and the corresponding Boltzmann factor as $q = \exp(\beta\epsilon) > 1$. Let us furthermore denote by m the number of contacts. The partition sum is then written as

$$Z_N(q) = \sum_m C_{N,m} q^m \tag{30}$$

where $C_{N,m}$ is the number of different clusters with N sites and m contacts. If the clusters are embedded in some solvent, repulsive monomer–solvent interactions need not be included explicitly, since the number s of solvent contacts satisfies $\mathcal{N}N = s + 2m$, where $\mathcal{N}$ is the coordination number of the lattice ($\mathcal{N} = 2d$ on a simple hypercubic lattice).

Simulations with PERM are straight forward. We just have to modify the weight factors W_N to $W_{N,m} = W_N e^{\beta m}$. We found that again, as for the previous case $q = 1$, it was better to simulate clusters breadth first than depth first. We also verified that it was advantageous to include a factor $\approx (1 - p)^{-g}$ in the fitness functions, just as for athermal clusters. But the results were rather disappointing, at least for low dimensions. This might seem at first surprising, given the fact that PERM works extremely well at the collapse transition of linear polymers in three dimensions [20]. But it is easy to see the reason for this difference. For linear polymers, $d = 3$ is the upper critical dimension, and θ polymers form essentially random walks with very small logarithmic corrections. Thus starting off with random walks, PERM can do with very few resampling steps. There is only one pruning or cloning needed for every 2000 simple forward steps [20]. Collapsing site trees, at least in low dimensions, are however very different from site percolation clusters. A convenient observable to see this difference is the average number $\langle m \rangle$ of contacts. For both models, $\langle m \rangle$ is roughly proportional to N. But for site percolation on the simple cubic lattice one finds $\langle m \rangle / N \approx 0.15$, while the same number for collapsing animals (at $q \approx 3.22$ [33]) is ≈ 0.40. Thus there is still a huge amount of re-sampling needed, even more than for athermal animals where $\langle m \rangle / N \approx 0.065$.

For this reason we do not even show any data for the collapse in $d = 2$ and $d = 3$. The situation improves when d is increased. Results for $d = 4$ are shown in figure 31. We see a rather sharp collapse transition at $q = q_c = 2.98 \pm 0.02$. It is hard to give precise estimates of any critical exponents from these data. But the Flory exponent ν seems to be the same as

for athermal animals, within rather small errors. This would not be easy to understand, if it were exactly true. In any case, we are not aware of any theoretical prediction to compare this with. We are not even aware of any prediction of the upper critical dimension for this collapse transition.

6.2. Collapse of bond animals and bond trees

Let us now switch attention to collapse models based on bond animals. We have now two different possible interactions. Instead of the single parameter ϵ for the interaction strength in site animals, we can now introduce different parameters ϵ_1 for bonded and ϵ_2 for non-bonded neighbour contacts, and define $y = e^{\epsilon_1}, \tau = e^{\epsilon_2}$. Note that as before we do *not* have to introduce also an additional interaction with the solvent, because the number b of surface bonds is not independent of m, k and N:

$$2m + 2k + b = \mathcal{N}N. \tag{31}$$

We then define

$$Z_N(y, \tau) = \sum_{m,k} C_{N,m,k} y^m \tau^k. \tag{32}$$

The model discussed in the last subsection is obtained by taking the limit $\tau = 0$ [25]. In this limit only bond configurations with the maximal number of bonds (for a given configuration of sites) contribute to the partition sum. On the other hand, bond percolation without reweighing corresponds, due to equations (29) and (31), to the curve

$$y = p/(1-p)^2, \qquad \tau = 1/(1-p), \qquad 0 \leqslant p \leqslant 1, \tag{33}$$

or, explicitly,

$$y = \tau(\tau - 1), \qquad \tau \geqslant 1. \tag{34}$$

Critical bond percolation ($p_c = 1/2$) corresponds to $y = \tau = 2$. Simulations should be very easy in the neighbourhood of this point, but they also should be not too difficult in the neighbourhood of the entire bond percolation line. The reason is simply that along the entire line one should not need much resampling. This should be enough to obtain precise estimates for large parts of the phase diagram, and in particular to clarify the existence of two different collapsed phases. Results will be given in a forthcoming paper.

7. Conclusion

We have shown that the basic idea of PERM, namely the recursive implementation of biased sequential sampling with re-sampling, can be applied also to lattice animals and lattice trees. These are two models for randomly branched polymers. Our algorithm is extremely efficient (except for the collapse of site animals), obviously much more efficient than previous Monte Carlo algorithms. We applied it to simple (hyper-)cubic lattices in up to nine dimensions, but made also less complete simulations on bcc and fcc lattices. Our algorithm works indeed better in higher dimensions, nevertheless we obtained high statistics results also for large animals in two dimensions.

We verified a number of theoretical predictions. In particular, we verified the Parisi–Sourlas connection between entropic and Flory exponents, and we verified the values of these exponents whenever they are exactly known. We also verified that the cross-over exponents for branched polymer adsorption on plane walls is super-universal for $d > 2$, as predicted some

time ago (but not for $d = 2$!), and we gave precise estimates of the other critical exponents at this adsorption transition.

There are a number of problems we have not yet studied, although our algorithm seems ideally suited for them, and which we plan to address in forthcoming papers. One is that of collapsing animals where we hope to be able to verify or disprove the existence of two different collapsed phases. Another is the dependence on the wedge angle, of the entropic exponent of a 2d animal grafted at the tip of this wedge. In conformally invariant 2d theories this angle dependence can be predicted, but lattice animals are not conformally invariant.

Apart from these specific problems we believe that the present simulations have demonstrated again the power of sophisticated sequential sampling methods, and of PERM in particular. Although there are certainly many problems where other MC strategies are more efficient, there are by now many examples where PERM seems unchallenged by any other known method. Unfortunately (or, rather, fortunately for the livelihood of the subject) it is hard to predict when PERM or a similar strategy will be the method of choice. But we are confident that lattice animals will not be the last such problem.

Acknowledgments

We thank Michael Fisher for bringing [74] to our attention, and him and Buks van Rensburg for very useful correspondence.

References

[1] Golomb S 1994 *Polyominoes: Puzzles, Patterns, Problems and Packings* (Princeton, NJ: Princeton University Press)
[2] Stauffer D and Aharony A 1992 *Introduction to Percolation Theory* 2nd edn (London: Taylor and Francis)
[3] Fortuin C M and Kasteleyn P W 1972 *Physica* **57** 536
[4] Swendsen R H and Wang J-S 1986 *Phys. Rev. Lett.* **57** 2607
[5] Drouffe J M, Parisi G and Sourlas N 1979 *Nucl. Phys.* B **161** 397
[6] Jensen I 2003 *Counting Polyominoes: A Parallel Implementation for Cluster Counting ICCS 2003* ed P M A Sloot *et al* (Berlin: Springer) pp 203–12
[7] Lubensky T C and Isaacson J 1978 *Phys. Rev. Lett.* **41** 829
Lubensky T C and Isaacson J 1979 *Phys. Rev. Lett.* **42** 410(E)
Lubensky T C and Isaacson J 1979 *Phys. Rev.* A **20** 2130
[8] Janse van Rensburg E J and Madras N 1997 *J. Phys. A: Math. Gen.* **30** 8035
[9] Adler J, Meir Y, Harris A B and Aharony A 1988 *Phys. Rev.* B **38** 4941
[10] Miller J D and De'Bell K 1993 *J. Phys.* I **3** 1717
[11] Parisi G and Sourlas N 1981 *Phys. Rev. Lett.* **46** 871
[12] Imbrie J Z 2003 *Preprint* arXiv.org/abs/math-ph/0303015
[13] Baxter R J, Enting I G and Tsang S K 1980 *J. Stat. Phys.* **22** 465
[14] Dhar D 1983 *Phys. Rev. Lett.* **51** 853
[15] Jassen H K and Lyssy A 1992 *J. Phys. A: Math. Gen.* **25** L679
[16] Janssen H K and Lyssy A 1994 *Phys. Rev.* E **50** 3784
[17] Janssen H K and Lyssy A 1995 *Europhys. Lett.* **29** 25
[18] De'Bell K, Lookman T and Zhao D 1991 *Phys. Rev.* A **44** 1390
Lookman T, Zhao D and De'Bell K 1991 *Phys. Rev.* **44** 4814
[19] De'Bell K and Lookman T 1992 *Rev. Mod. Phys.* **65** 87
[20] Grassberger P 1997 *Phys. Rev.* E **56** 3682
[21] Madras N, Soteros C E, Whittington S G, Martin J L, Sykes M F, Flesia S and Gaunt D S 1990 *J. Phys. A: Math. Gen.* **23** 5327
[22] Flesia S, Gaunt D S, Soteros C E and Whittington S G 1992 *J. Phys. A: Math. Gen.* **25** L1169
[23] Flesia S, Gaunt D S, Soteros C E and Whittington S G 1994 *J. Phys. A: Math. Gen.* **27** 5831

[24] Seno F and Vanderzande C 1994 *J. Phys. A: Math. Gen.* **27** 5813
[25] Henkel M and Seno F 1996 *Phys. Rev.* E **53** 3662
[26] Stratychuk L M and Soteros C E 1996 *J. Phys. A: Math. Gen.* **29** 7067
[27] Madras N and Janse van Rensburg E J 1997 *J. Stat. Phys.* **86** 1
[28] Janse van Rensburg E J, Orlandini E and Tesi M C 1999 *J. Phys. A: Math. Gen.* **32** 1567
[29] Janse van Rensburg E J 2000 *J. Phys. A: Math. Gen.* **33** 3653
[30] Derrida B and Herrmann H J 1983 *J. Phys.* **44** 1365
[31] Dickman R and Schieve W C 1984 *J. Phys.* **45** 1727
[32] Lam P M 1987 *Phys. Rev.* B **36** 6988
[33] Lam P M 1988 *Phys. Rev.* B **38** 2813
[34] de Alcantara Bonfim D F, Kirkham J E and McKane A J 1981 *J. Phys. A: Math. Gen.* **14** 2391
[35] Martin J L 1974 Computer enumerations *Phase Transitions and Critical Phenomena* vol 3 ed C Domb and M S Green (London: Academic)
[36] Redner S 1982 *J. Stat. Phys.* **29** 309
[37] Redelmeier D H 1981 *Discrete Math.* **36** 191
[38] Mertens S 1990 *J. Stat. Phys.* **58** 1095
[39] Jensen I and Guttmann A J 2000 *J. Phys. A: Math. Gen.* **33** L257
[40] Jensen I 2001 *J. Stat. Phys.* **102** 865
[41] Derrida B and De Seze L 1981 *J. Phys.* **43** 475
[42] Peters H P, Stauffer D, Hölters H P and Loewenich K 1979 *Z. Phys. B: Condens. Matter* **34** 399
[43] Sykes M F and Glen M 1976 *J. Phys. A: Math. Gen.* **9** 87
[44] Gaunt D S, Sykes M F, Torrie G M and Whittington S G 1982 *J. Phys. A: Math. Gen.* **15** 3209
[45] Whittington S G, Torrie G M and Gaunt D S 1983 *J. Phys. A: Math. Gen.* **16** 1695
[46] Lam P M 1986 *Phys. Rev.* A **34** 2339
[47] Edwards B F, Gyure M F and Ferer M 1992 *Phys. Rev.* A **46** 6252
[48] Foster D P and Seno F 1993 *J. Phys. A: Math. Gen.* **26** 1299
[49] Stauffer D 1978 *Phys. Rev. Lett.* **41** 1333
[50] Seitz W A and Klein D J 1981 *J. Chem. Phys.* **75** 5190
[51] Glaus U 1985 *J. Phys. A: Math. Gen.* **18** L609
[52] Janse van Rensburg E J and Madras N 1992 *J. Phys. A: Math. Gen.* **25** 303
[53] You S and Janse van Rensburg E J 2001 *Phys. Rev.* E **64** 046101
[54] Janse van Rensburg E J and Rechnitzer A 2003 *Phys. Rev.* E **67** 036116
[55] You S and Janse van Rensburg E J 1998 *Phys. Rev.* E **58** 3971
[56] de Queiroz S L A 1995 *J. Phys. A: Math. Gen.* **28** 6515
[57] Vujić D 1999 *J. Stat. Phys.* **95** 767
[58] Lam P M and Binder K 1988 *J. Phys. A: Math. Gen.* **21** L405
[59] Madras N and Sokal A D 1988 *J. Stat. Phys.* **50** 109
[60] Grassberger P 2003 *Phys. Rev.* E **67** 036101
[61] Redner S 1979 *J. Phys. A: Math. Gen.* **12** L239
[62] Meirovitch H 1987 *J. Phys. A: Math. Gen.* **20** 6059
[63] Leath P 1976 *Phys. Rev.* B **14** 5046
[64] Care C M 1997 *Phys. Rev.* E **56** 1181
Care C M and Ettelaie R 2000 *Phys. Rev.* E **62** 1397
[65] Grassberger P 1983 *Math. Biosci.* **62** 157
[66] Swendsen R H and Wang J-S 1990 *Physica* A **167** 565
[67] Rosenbluth M N and Rosenbluth A W 1955 *J. Chem. Phys.* **23** 356
[68] Grassberger P 2002 *Comp. Phys. Commun.* **147** 64
[69] Hsu H-P, Nadler W and Grassberger P 2004 *Macromolecules* **37** 4658
[70] Hsu H-P and Grassberger P 2004 *Europhys. Lett.* **66** 874
[71] Hsu H-P, Mehra V, Nadler W and Grassberger P 2003 *Phys. Rev.* E **68** 021113
[72] Grassberger P 2004 *Preprint* cond-mat/0408571
[73] Grassberger P 1999 *J. Chem. Phys.* **111** 440
[74] Lai S-N and Fisher M E 1995 *J. Chem. Phys.* **103** 8144
[75] Guttmann A J and Gaunt D S 1978 *J. Phys. A: Math. Gen.* **11** 949
[76] Peard P J and Gaunt D S 1995 *J. Phys. A: Math. Gen.* **28** 6109
[77] Diehl H-W 1986 *Phase Transitions and Critical Phenomena* vol 10 ed C Domb and J L Lebowitz (New York: Academic)
[78] Hegger R and Grassberger P 1994 *J. Phys. A: Math. Gen.* **27** 4069

[79] Grassberger P 2004 *Preprint* cond-mat/0410055
[80] Burkhardt T W, Eisenriegler E and Guim I 1989 *Nucl. Phys.* B **316** 559
[81] van Vliet J H and Brinke G 1990 *J. Chem. Phys.* **93** 1436
van Vliet J H, Luiken M C and Brinke G 1992 *Macromolecules* **25** 3802
[82] Grassberger P 1999 *Physica* A **262** 251
[83] Gaunt D S and Ruskin H 1978 *J. Phys. A: Math. Gen.* **11** 1369