

# Counting lattice animals in high dimensions

Sebastian Luther<sup>1</sup> and Stephan Mertens<sup>1,2</sup>

<sup>1</sup> Institut für Theoretische Physik, Otto-von-Guericke Universität, PF 4120, D-39016 Magdeburg, Germany

<sup>2</sup> Santa Fe Institute, 1399 Hyde Park Rd, Santa Fe, NM 87501, USA

E-mail: [SebastianLuther@gmx.de](mailto:SebastianLuther@gmx.de) and [mertens@ovgu.de](mailto:mertens@ovgu.de)

Received 7 June 2011

Accepted 29 August 2011

Published 22 September 2011

Online at [stacks.iop.org/JSTAT/2011/P09026](http://stacks.iop.org/JSTAT/2011/P09026)

[doi:10.1088/1742-5468/2011/09/P09026](https://doi.org/10.1088/1742-5468/2011/09/P09026)

**Abstract.** We present an implementation of Redelemeier's algorithm for the enumeration of lattice animals in high-dimensional lattices. The implementation is lean and fast enough to allow us to extend the existing tables of animal counts, perimeter polynomials and series expansion coefficients in  $d$ -dimensional hypercubic lattices for  $3 \leq d \leq 10$ . From the data we compute formulae for perimeter polynomials for lattice animals of size  $n \leq 11$  in arbitrary dimension  $d$ . When amended by combinatorial arguments, the new data suffice to yield explicit formulae for the number of lattice animals of size  $n \leq 14$  and arbitrary  $d$ . We also use the enumeration data to compute numerical estimates for growth rates and exponents in high dimensions that agree very well with Monte Carlo simulations and recent predictions from field theory.

**Keywords:** critical exponents and amplitudes (theory), percolation problems (theory), disordered systems (theory)

**ArXiv ePrint:** [1106.1078](https://arxiv.org/abs/1106.1078)

---

**Contents**

<b>1. Introduction</b>	<b>2</b>
<b>2. The algorithm</b>	<b>4</b>
<b>3. Performance</b>	<b>6</b>
<b>4. Proper animals</b>	<b>8</b>
<b>5. Mean cluster size</b>	<b>12</b>
<b>6. Growth rates and exponents</b>	<b>12</b>
<b>7. Conclusions</b>	<b>14</b>
<b>Appendix A. Formulae for <math>DX(n, n - k)</math>: structure</b>	<b>15</b>
<b>Appendix B. Formulae for <math>DX(n, n - k)</math>: coefficients</b>	<b>17</b>
<b>References</b>	<b>19</b>

---

**1. Introduction**

A polyomino of size  $n$  is an edge-connected set of  $n$  squares on the square lattice and a polycube of size  $n$  is a face-connected set of  $n$  cubes in the cubic lattice. Polyominoes and polycubes are a classical topic in recreational mathematics and combinatorics [1]. In statistical physics and percolation theory, polyominoes and polycubes are called lattice animals [2, 3]. Lattice animals are not restricted to dimension 2 or 3: a  $d$ -dimensional lattice animal of size  $n$  is a set of  $n$  face-connected hypercubes on  $\mathbb{Z}^d$ .

In this paper we will address the problem of counting the number of fixed animals of size  $n$  in dimension  $d > 2$ . Fixed animals are considered distinct if they have different shapes or orientations. Free animals, on the other hand, are distinguished only by shape, not by orientation. Figure 1 shows all fixed polyominoes of size 3 and all free polycubes of size 4.

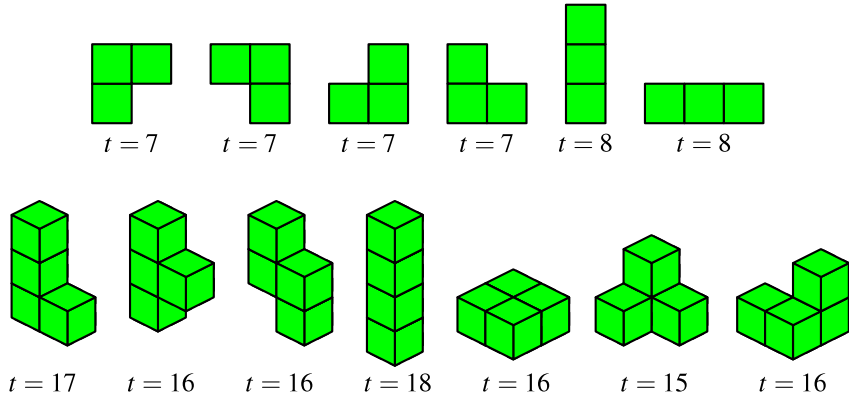
We denote the number of  $d$ -dimensional fixed animals of size  $n$  by  $A_d(n)$ . There is no formula for  $A_d(n)$ , but we know that  $A_d(n)$  grows exponentially with  $n$ . Using subadditivity and concatenation arguments [4], one can show that there are constants  $1 < \lambda_d < \infty$  such that

$$\lim_{n \rightarrow \infty} \sqrt[n]{A_d(n)} = \lambda_d. \quad (1)$$

The constant  $\lambda_2$  is known as Klarner's constant.

A slightly stronger result due to Madras [5] asserts that

$$\lim_{n \rightarrow \infty} \frac{A_d(n+1)}{A_d(n)} = \lambda_d. \quad (2)$$



**Figure 1.** All fixed polyominoes of size  $n = 3$  (top) and all free polycubes of size  $n = 4$  (bottom) and their perimeters.

Intuitively, the growth rate  $\lambda_d$  should grow with the coordination number  $2d$  of the lattice. In fact, in [6] it is shown that

$$\lambda_d = 2de - o(d), \tag{3}$$

and in the same paper it is conjectured that  $\lambda_d = (2d - 3)e + O(1/d)$ . For finite  $d$ , however, we know only lower and upper bounds for  $\lambda_d$ . Numerical estimates for  $\lambda_d$  can be derived from extrapolating  $A_d(n + 1)/A_d(n)$ , which is one motivation to compute  $A_d(n)$  for  $n$  as large as possible. We will try our hands at that in section 6.

In percolation theory one is interested in counting lattice animals of a given size  $n$  according to their perimeter  $t$ , i.e. to the number of adjacent cells that are empty (see figure 1). If each cell of the lattice is occupied independently with probability  $p$ , the average number of clusters of size  $n$  per lattice site is

$$\sum_t g_{n,t}^{(d)} p^n (1 - p)^t, \tag{4}$$

where  $g_{n,t}^{(d)}$  denotes the number of fixed  $d$ -dimensional lattice animals of size  $n$  and perimeter  $t$ . The  $g$ 's define the perimeter polynomials

$$P_d(n, q) = \sum_t g_{n,t}^{(d)} q^t. \tag{5}$$

We can easily compute  $A_d(n)$  from the perimeter polynomial  $P_d(n, q)$  through

$$A_d(n) = P_d(n, 1) = \sum_t g_{n,t}^{(d)}. \tag{6}$$

In fact we can compute  $A_d(n + 1)$  from the perimeter polynomials up to size  $n$ :

$$A_d(n + 1) = \frac{1}{n + 1} \sum_{m \leq n} m \sum_t g_{m,t}^{(d)} \binom{t}{n + 1 - m} (-1)^{n-m}. \tag{7}$$

This equation follows from the observation that below the percolation threshold  $p_c$ , each occupied lattice site belongs to some finite size lattice animal:

$$p = \sum_{n=1}^{\infty} np^n P_d(n, 1-p) \quad (8)$$

for  $p < p_c$ . The right-hand side is a power series in  $p$  and equation (7) follows from the fact that the coefficient of  $p^{n+1}$  must be zero.

As we will see in section 2, the algorithm for counting lattice animals keeps track of the perimeter anyway. Hence it is reasonable to use the algorithm to compute the perimeter polynomials and to apply (7) to get an extra value of  $A_d$ .

## 2. The algorithm

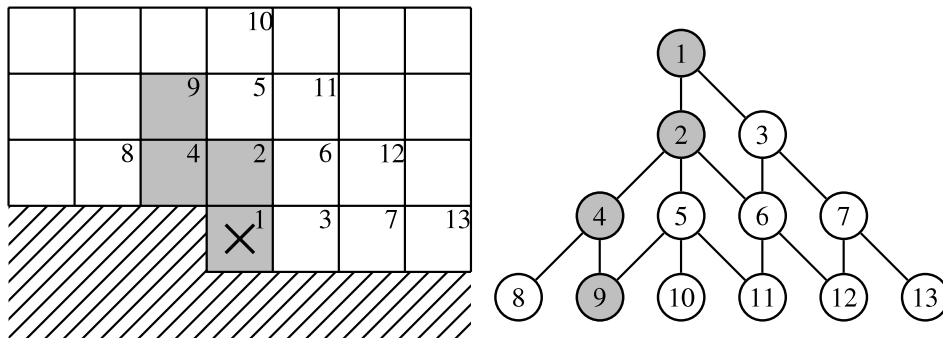
The classical algorithm for counting lattice animals is due to Redelmeier [7]. Originally developed for the square lattice, Redelmeier's algorithm was later shown to work on arbitrary lattices and in higher dimensions [8] and to be efficiently parallelizable [9]. For two-dimensional lattices there is a much faster counting method based on transfer matrices [10], but for  $d \geq 3$  Redelmeier's algorithm is still the most efficient known way to count lattice animals.

The algorithm works by recursively generating all lattice animals up to a given size  $n_{\max}$ . Given an animal of size  $n$ , the algorithm generates animals of size  $n+1$  by adding a new cell in the perimeter of the given animal. The lattice sites that are available for extending the current animal are stored in a set  $U$  called the untried set. To avoid generating the same fixed animal more than once, lattice sites that have previously been added to the untried set are marked on the lattice.

In order to break the translational symmetry we demand that the initial site, which is contained in all animals, is an extremal site with respect to the lexicographic order of lattice coordinates. Figure 2 illustrates how this can be achieved in the square lattice. We simply block all lattice sites from further consideration that are in a row below the initial site (marked with a cross) or in the same row and to the left of the initial site. The generalization to  $d > 2$  is straightforward. These blocked sites are never added to the untried set, but they need to be taken into account when we compute the perimeter  $t$ .

We start with all lattice sites being marked 'free' or 'blocked', except for the initial site, which is marked 'counted'. Furthermore  $n = 1$ ,  $t = 1$  and the initial site being the only element of the untried set  $U$ . Redelmeier's algorithm works by invoking the following routine with this initial setting:

- Iterate until  $U$  is empty:
  - (1) Remove a site  $s$  from  $U$ .
  - (2)  $F :=$  set of 'free' neighbors,  $B :=$  set of 'blocked' neighbors of  $s$ .  $N := |U| + |B|$ .
  - (3) Count new cluster: increase  $g_{n,t+N-1}$  by one.
  - (4) If  $n < n_{\max}$ :



**Figure 2.** Part of the square lattice that can be reached from animals up to size 4 (left). The number of lattice animals of size  $n \leq 4$  equals the number of subgraphs in the neighborhood graph (right) that contain vertex 1.

- (a) Mark all sites in  $F$  and  $B$  as ‘counted’.
  - (b) Call this routine recursively with  $U' = U \cup F$ ,  $n' = n + 1$  and  $t' = t + N - 1$ .
  - (c) Relabel sites in  $F$  as ‘free’ and sites in  $B$  as ‘blocked’.
- Return.

Since the algorithm generates each lattice animal explicitly, its running time scales like  $A_d(n)$ . This exponential complexity implies hard limits for the accessible animal sizes. All we can do is to keep the prefactor in the time complexity function small, i.e. to implement each step of Redelmeier’s algorithm as efficiently as possible by using an appropriate data structure for the untried set and for the lattice, see [8].

Another crucial element of tuning Redelmeier’s algorithm is the computation of the neighborhood of a lattice cell. In a recent paper [11], Aleksandrowicz and Barequet observed that Redelmeier’s algorithm can be interpreted as the counting of subgraphs in a graph that represents the neighborhood relation of the lattice. Figure 2 illustrates this for the square lattice.

For any lattice, the neighborhood graph can be precomputed and be represented as an adjacency list which is then fed to the actual subgraph counting algorithm. That way the computation of the neighbors of a lattice cell is taken out of the counting loop, and the prefactor in the exponential scaling is reduced.

The size of the neighborhood graph or, equivalently, the number of lattice points required to host lattice animals of size  $n$  scales like  $\Theta(n^d)$ . Aleksandrowicz and Barequet [11] claimed that this exponential growth of memory with  $d$  represents a serious bottleneck for Redelmeier’s algorithm in high dimensions. In a subsequent paper [12], they therefore present a variation of the algorithm that avoids the storage of the full graph by computing the relevant parts of the graph on demand. This cuts down the space complexity to a low-order polynomial in  $d$ , but it forfeits the gain in speed that can be obtained by precomputing the complete neighborhood graph.

We claim that in practice the space complexity of Redelmeier’s algorithm is no bottleneck. The reason is that the prefactor in the  $\Theta(n^d)$  scaling can be made small enough to hold the complete graph in memory for all values of  $n$  and  $d$  for which  $A_d(n)$  is computable in reasonable time.

The key observation is that in [11] Aleksandrowicz and Barequet used a hypercube of side length  $2n$  of the lattice to host the lattice animals, whereas a hypersphere of radius  $n$  suffices. This is a significant difference, as can be seen from the analogous situation in  $\mathbb{R}^d$ . Here the volume of a cube is much larger than the volume of the inscribed Euclidean sphere:

$$\frac{\text{volume hypercube}}{\text{volume inscribed hypersphere}} = \frac{2^d \Gamma(d/2 + 1)}{\pi^{d/2}} \simeq \left(\frac{2d}{\pi e}\right)^{d/2} \frac{1}{\sqrt{\pi d}}. \quad (9)$$

In the lattice  $\mathbb{Z}^d$ , the number of lattice points in a cube of side length  $2n + 1$  is

$$C(d, n) = (2n + 1)^d. \quad (10)$$

Let  $B(d, n)$  denote the number of lattice sites that are  $n$  steps or less away from the origin. This ‘volume of the crystal ball’ can be computed recursively via

$$B(1, n) = 2n + 1 \quad B(d, n) = B(d - 1, n) + 2 \sum_{k=0}^{n-1} B(d - 1, k), \quad (11)$$

which reflects the fact that the crystal ball in dimension  $d$  can be decomposed into  $(d - 1)$ -dimensional slices whose diameter decreases with increasing distance from the central slice. The recursion (11) tells us that  $B(d, n)$  is a polynomial in  $n$  of degree  $d$  which can easily be computed, see [A001845](#) to [A001848](#) on [oeis.org](#) for the polynomials for  $d = 3$ – $6$ . Note that  $B(d, n)$  can also be computed through the generating function [13]

$$\frac{(1 + x)^d}{(1 - x)^{d+1}} = \sum_{n=0}^{\infty} B(d, n) x^n. \quad (12)$$

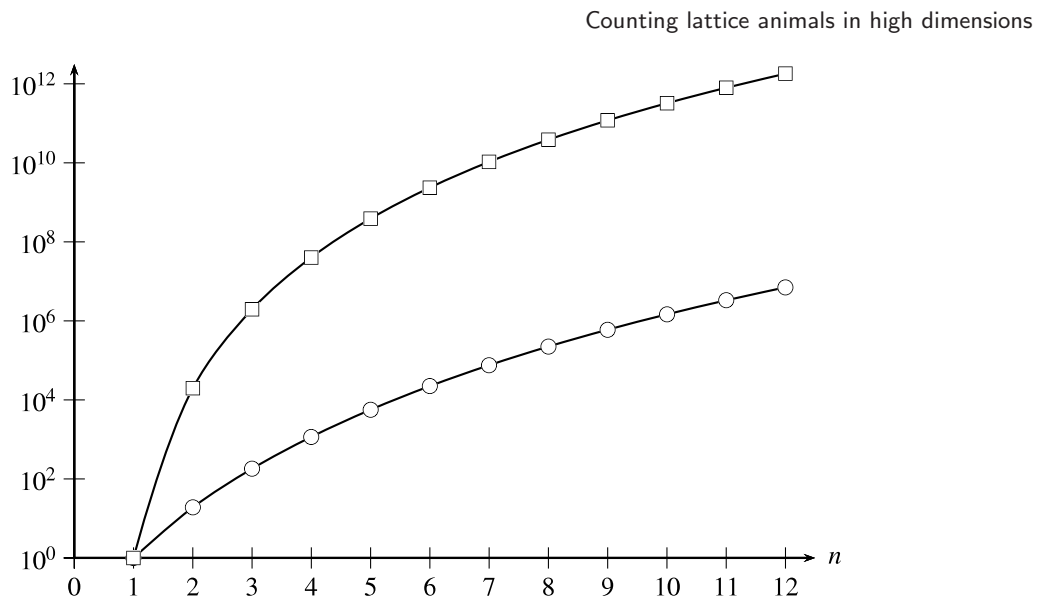
The number of lattice sites needed to cage lattice animals of size  $n$  is very close to  $B(d, n - 1)/2$  for spherical cages and  $C(d, n - 1)/2$  for cubical cages. Figure 3 shows both numbers for the case  $d = 9$ . As you can see,  $d = 9$  and  $n = 12$  require memory on the terabyte scale if one uses cubic cages, but only a few megabytes for spherical cages.

### 3. Performance

Our implementation of the Redelmeier algorithm consists of two programs. The first program computes the neighborhood graph of a specified lattice and writes this graph as an adjacency list into a file. The second program reads this file and computes the corresponding perimeter polynomials. The programs are written in C++ and can be downloaded from the project webpage [15].

When run on a laptop with an Intel<sup>®</sup> Core<sup>™</sup> 2 Duo CPU at 2 GHz, the program enumerates perimeter polynomials at a rate of roughly  $2 \times 10^7$  animals  $\text{s}^{-1}$ . This means that generating and counting one lattice animal and measuring its perimeter takes about 100 clock cycles, which is reasonable for a program compiled from C++.

At this rate, our laptop needs 35 days to enumerate the perimeter polynomials for  $d = 9$  and  $n \leq 11$  (see table 2). Computing the next perimeter polynomial ( $n = 12$ ) would take more than three years. Note that, according to figure 3, the neighborhood graph for  $d = 9$  and  $n = 12$  easily fits into the memory of a run-of-the-mill laptop. These numbers



**Figure 3.** Volume needed to cage animals of size  $n$  on the hypercubic lattice of dimension  $d = 9$ . Cubic ( $\square$ ) versus spherical cages ( $\circ$ ).

**Table 1.** Range of perimeter polynomials and animal numbers in dimensions  $d \geq 3$  that have been found by exhaustive enumerations. The old perimeter polynomials are from [8] and [14], while the old values of  $A_d(n)$  are from [11] and [12].

$d$	Perimeter polynomial		$A_d(n)$	
	Old $n_{\max}$	New $n_{\max}$	Old $n_{\max}$	New $n_{\max}$
3	15	18	18	19
4	10	15	15	16
5	9	14	13	15
6	8	14	10	15
7	8	13	10	14
8		11	8	12
9		11	4	12
10		11		12

illustrate that, for all practical purposes, the bottleneck of Redelemeier’s algorithm is time, not memory.

Using a parallel implementation [9] that we ran on a Linux cluster with 128 Intel<sup>®</sup>Xeon<sup>®</sup> 3.2 GHz CPUs, or for the most demanding computations, on a SciCortex SC5832 with 972 MIPS64 6-core nodes, we could extend the table of known perimeter polynomials and animal counts considerably, see table 1. The new values for  $d \leq 5$  are

$$\begin{aligned}
 A_3(19) &= \mathbf{651\ 459\ 315\ 795\ 897}, & A_4(16) &= \mathbf{692\ 095\ 652\ 493\ 483}, \\
 A_5(14) &= \mathbf{227\ 093\ 585\ 071\ 305}, & A_5(15) &= \mathbf{3689\ 707\ 621\ 144\ 614}.
 \end{aligned}$$

The numbers for  $6 \leq d \leq 9$  are given in table 2, while the corresponding perimeter polynomials can be found on the project webpage [15]. Before we evaluate the results, we will discuss a combinatorial argument that allows us to extend the enumeration data considerably.

**Table 2.** Number of lattice animals in the hypercubic lattice for  $d = 6 \dots 9$  obtained by direct enumeration. New results in boldface; the numbers of smaller animals are from [11, 12] and references therein. Note that in [16],  $A_8(n)$  and  $A_9(n)$  for  $n \leq 9$  were computed rather than enumerated by the same method that we will use in section 4 to extend this table to  $n \leq 14$  and all values of  $d$ .

$n$	$A_6(n)$	$A_7(n)$	$A_8(n)$	$A_9(n)$
1	1	1	1	1
2	6	7	8	9
3	66	91	120	153
4	901	1484	2276	3309
5	13881	27468	49204	<b>81837</b>
6	231008	551313	1156688	<b>2205489</b>
7	4057660	11710328	28831384	<b>63113061</b>
8	74174927	259379101	750455268	<b>1887993993</b>
9	1398295989	5933702467	<b>20196669078</b>	<b>58441956579</b>
10	27012396022	139272913892	<b>558157620384</b>	<b>1858846428437</b>
11	<b>532327974882</b>	<b>3338026689018</b>	<b>15762232227968</b>	<b>60445700665383</b>
12	<b>10665521789203</b>	<b>81406063278113</b>	<b>453181069339660</b>	<b>2001985304489169</b>
13	<b>227093585071305</b>	<b>2014611366114053</b>		
14	<b>4455636282185802</b>	<b>50486299825273271</b>		
15	<b>92567760074841818</b>			

Note that the most demanding computation in this paper was the enumeration of the perimeter polynomial for  $n = 14$  in  $d = 6$ . On a single core of a MIPS64, this enumeration would have taken 77 CPU years, and on our laptop from above it would still have taken about 7 CPU years. In practice we used a parallel implementation that ran on many cores (and several different machines) such that no computation took longer than two weeks' wall clock time.

#### 4. Proper animals

A lattice animal of size  $n$  cannot span more than  $n - 1$  dimensions. This simple observation allows us to derive explicit formulae for  $A_d(n)$  for fixed  $n$ . Obviously  $A_d(1) = 1$  and  $A_d(2) = d$ . A lattice animal of size  $n = 3$  is either a one-dimensional 'stick' with  $d$  possible orientations or 'L-shaped' and spanning two out of  $d$  dimensions. Within these two dimensions there are four possible orientations for the L-shaped animal (see figure 1), hence

$$A_d(3) = d + 4 \binom{d}{2} = 2d^2 - d.$$

For  $n = 4$ , we have again the 'stick' that lives in one dimension, 17 animals that span two dimensions and 32 animals that span three dimensions:

$$A_d(4) = d + 17 \binom{d}{2} + 32 \binom{d}{3} = \frac{16}{3}d^3 - \frac{15}{2}d^2 + \frac{19}{6}d.$$

**Table 3.** Number of lattice animals of sizes 2–14 in hypercubic lattices of dimension  $d$ . The polynomials for  $n \leq 11$  have been obtained by direct enumeration and confirm those listed in [6]. Polynomials for  $n > 11$  have been computed from enumeration data and known values of  $\text{DX}(n, n - k)$ .

---


$$\begin{aligned}
 A_d(2) &= d \\
 A_d(3) &= 2d^2 - d \\
 A_d(4) &= \frac{16}{3}d^3 - \frac{15}{2}d^2 + \frac{19}{6}d \\
 A_d(5) &= \frac{50}{3}d^4 - 42d^3 + \frac{239}{6}d^2 - \frac{27}{2}d \\
 A_d(6) &= \frac{288}{5}d^5 - 216d^4 + \frac{986}{3}d^3 - 231d^2 + \frac{926}{15}d \\
 A_d(7) &= \frac{9604}{45}d^6 - 1078d^5 + \frac{20651}{9}d^4 - \frac{14927}{6}d^3 + \frac{120107}{90}d^2 - \frac{827}{3}d \\
 A_d(8) &= \frac{262144}{315}d^7 - \frac{26624}{5}d^6 + \frac{132320}{9}d^5 - \frac{65491}{3}d^4 + \frac{1615991}{90}d^3 - \frac{113788}{15}d^2 + \frac{52589}{42}d \\
 A_d(9) &= \frac{118098}{35}d^8 - 26244d^7 + \frac{447903}{5}d^6 - \frac{511082}{3}d^5 + \frac{23014949}{120}d^4 - \frac{1522261}{12}d^3 \\
 &\quad + \frac{38839021}{840}d^2 - \frac{30089}{4}d \\
 A_d(10) &= \frac{800000}{567}d^9 - \frac{272000}{21}d^8 + \frac{14272000}{27}d^7 - \frac{11092360}{9}d^6 + \frac{239850598}{135}d^5 \\
 &\quad - \frac{14606026}{9}d^4 + \frac{1067389643}{1134}d^3 - \frac{42595493}{126}d^2 + \frac{2804704}{45}d \\
 A_d(11) &= \frac{857435524}{14175}d^{10} - \frac{67319318}{105}d^9 + \frac{2884481974}{945}d^8 - \frac{380707987}{45}d^7 + \frac{40341440233}{2700}d^6 \\
 &\quad - \frac{1260803635}{72}d^5 + \frac{79118446751}{5670}d^4 - \frac{19252021283}{2520}d^3 + \frac{17126616179}{6300}d^2 - \frac{7115086}{15}d \\
 A_d(12) &= \frac{509607936}{1925}d^{11} - \frac{15925248}{5}d^{10} + \frac{607592448}{35}d^9 - \frac{1956324864}{35}d^8 + \frac{2930444704}{25}d^7 \\
 &\quad - \frac{2522387284}{15}d^6 + \frac{17894522696}{105}d^5 - \frac{1242881121}{10}d^4 + \frac{22272055467}{350}d^3 - \frac{4225468993}{210}d^2 \\
 &\quad + \frac{181356011}{66}d \\
 A_d(13) &= \frac{551433967396}{467775}d^{12} - \frac{75047226332}{4725}d^{11} + \frac{166095324499}{1701}d^{10} - \frac{48436628461}{135}d^9 \\
 &\quad + \frac{49499551181119}{56700}d^8 - \frac{1335959158369}{900}d^7 + \frac{248648897740349}{136080}d^6 - \frac{25156285613453}{15120}d^5 \\
 &\quad + \frac{757565736903221}{680400}d^4 - \frac{5607318230581}{10800}d^3 + \frac{12648671104037}{83160}d^2 - \frac{135165335}{6}d \\
 A_d(14) &= \frac{4628074479616}{868725}d^{13} - \frac{23612624896}{297}d^{12} + \frac{3309261190144}{6075}d^{11} - \frac{304034058496}{135}d^{10} \\
 &\quad + \frac{12648090831712}{2025}d^9 - \frac{553376997376}{45}d^8 + \frac{758347226205724}{42525}d^7 - \frac{2633038200122}{135}d^6 \\
 &\quad + \frac{98388569956577}{6075}d^5 - \frac{2734657007119}{270}d^4 + \frac{11824147558382}{2475}d^3 - \frac{560344373791}{330}d^2 \\
 &\quad + \frac{97500388612}{273}d
 \end{aligned}$$


---

In general we can write

$$A_d(n) = \sum_{i=0}^d \binom{d}{i} \text{DX}(n, i), \tag{13}$$

where  $\text{DX}(n, i)$  denotes the number of fixed proper animals of size  $n$  in dimension  $i$ . An animal is called ‘proper’ in dimension  $d$  if it spans all  $d$  dimensions. Equation (13) is due to Lunnon [17]. If we know  $A_d(n)$  for a given  $n$  and  $d \leq d_{\max}$ , we can use (13) to compute  $\text{DX}(n, d)$  for the same value of  $n$  and all  $d \leq d_{\max}$ , and vice versa.

Since  $\text{DX}(n, i) = 0$  for  $i \geq n$ , Lunnon’s equation tells us that  $A_d(n)$  is a polynomial of degree  $n - 1$  in  $d$ , and since  $A_0(n) = 0$  for  $n > 1$ , it suffices to know the values  $A_1(n), A_2(n), \dots, A_{n-1}(n)$  to compute the polynomial  $A_d(n)$ . From our enumeration data (table 1), we can compute these polynomials up to  $A_d(11)$ , see table 3.

In order to compute  $A_d(12)$ , we need to know  $A_{11}(12)$  or, equivalently,  $\text{DX}(12, 11)$ . The latter can actually be computed with pencil and paper. That is because an animal of size 12 in 11 dimensions has to span a new dimension with each of its cells to be proper.

In particular, its cells cannot form loops. Hence computing  $\text{DX}(12, 11)$  is an exercise in counting trees. This is true for  $\text{DX}(n, n - 1)$  in general, so let us compute this function.

The adjacency graph of a lattice animal of size  $n$  is an edge-labeled graph with  $n$  vertices, in which each vertex represents a cell of the animal and two vertices are connected if the corresponding cells are neighbors in the animal. Every edge of the adjacency graph is labeled with the dimension along which the two cells touch each other.

In the case  $\text{DX}(n, n - 1)$ , every pair of adjacent cells must span a new dimension. Therefore the corresponding adjacency graph contains exactly  $n - 1$  edges, i.e. it is a tree, and each edge has a unique label. There are two directions for each dimension that we represent by the orientation of the edge in the tree. Hence  $\text{DX}(n, n - 1)$  equals the number of directed, edge-labeled trees of size  $n$ , where in our context ‘directed’ means that each edge has an arbitrary orientation in addition to its label.

The number of vertex labeled trees of size  $n$  is given by  $n^{n-2}$ , the famous formula published by Cayley in 1889 [18]. The number of edge-labeled trees seems to be much less known; at least it is proven afresh in recent papers like [19]. The following nice derivation is from [6]. Start with a vertex-labeled tree of size  $n$  and mark the vertex with label  $n$  as the root. Then shift every label smaller than  $n$  from its vertex to the incident edge towards the root. This gives an edge-labeled tree with a single vertex marked (the root). Since the mark can be on any vertex, the number of edge-labeled trees equals the number of vertex-labeled trees divided by the number of vertices. According to Cayley’s formula, this number is  $n^{n-3}$ . And since each directed edge can have two directions, we get

$$\text{DX}(n, n - 1) = 2^{n-1} n^{n-3}. \quad (14)$$

This formula has been known in the statistical physics community for a long time [20]. We used it to compute  $\text{DX}(12, 11)$  and then  $A_d(12)$  (table 3).

We can proceed further along this line. To compute  $A_d(13)$  we have to extend our enumeration data by  $A_{12}(13), \dots, A_8(13)$  or equivalently by  $\text{DX}(13, 12), \dots, \text{DX}(13, 8)$ . What we need are formulae  $\text{DX}(n, n - k)$  for  $k > 1$ .

For  $k > 1$ , there is no longer a simple correspondence between edge-labeled trees and proper animals. We need to take into account that there are edge labels with the same value, that the adjacency graph may contain loops, and that some labeled trees represent a self-overlapping and therefore illegal lattice animal. A careful consideration of these issues yields

$$\text{DX}(n, n - 2) = 2^{n-3} n^{n-5} (n - 2)(9 - 6n + 2n^2), \quad (15)$$

see [6] for the derivation of (15).

For  $k > 2$ , the computation of  $\text{DX}(n, n - k)$  gets very complicated and is better left to a computer. In appendix A we show that

$$\text{DX}(n, n - k) = 2^{n-2k+1} n^{n-2k-1} g_k(n), \quad (16)$$

where  $g_k(n)$  is a polynomial of degree  $3k - 3$ . Hence we can compute  $g_k$  from  $3k - 2$  data points, like the values of  $\text{DX}(n, n - k)$  for  $n = k, \dots, 4k - 3$ . Our enumeration data suffice to compute  $g_2$  and  $g_3$  with this method, but not  $g_4$ .

However, there is a trick that allows us to compute  $g_k$  from many fewer data points. The free energy

$$f_n = \frac{1}{n} \log A_d(n)$$

**Table 4.** Polynomials  $g_k(n)$  that appear in  $\text{DX}(n, n - k)$  (16). The polynomials  $g_2, \dots, g_6$  can be found as  $g_{k,0}$  in appendix 2 of [21]. As far as we know, the polynomial  $g_7$  has not been published before. See the appendix of this paper for the method on how to compute the  $g_k$ .

---


$$\begin{aligned}
 g_2(n) &= (n - 2)(9 - 6n + 2n^2) \\
 g_3(n) &= \frac{n-3}{6}(-1560 + 1122n - 679n^2 + 360n^3 - 104n^4 + 12n^5) \\
 g_4(n) &= \frac{n-4}{6}(204\,960 - 114\,302n + 41\,527n^2 - 17\,523n^3 + 7404n^4 - 2930n^5 \\
 &\quad + 828n^6 - 128n^7 + 8n^8) \\
 g_5(n) &= \frac{n-5}{360}(-3731\,495\,040 + 1923\,269\,040n - 535\,510\,740n^2 + 150\,403\,080n^3 - 42\,322\,743n^4 \\
 &\quad + 12\,397\,445n^5 - 4062\,240n^6 + 1335\,320n^7 - 356\,232n^8 + 62\,240n^9 - 6000n^{10} + 240n^{11}) \\
 g_6(n) &= \frac{n-6}{360}(1785\,362\,705\,280 - 939\,451\,308\,048n + 248\,868\,418\,932n^2 - 56\,265\,094\,748n^3 \\
 &\quad + 11\,984\,445\,891n^4 - 2448\,081\,038n^5 + 535\,284\,255n^6 - 127\,651\,774n^7 + 33\,940\,138n^8 \\
 &\quad - 9580\,440n^9 + 2398\,912n^{10} - 440\,688n^{11} + 51\,856n^{12} - 3424n^{13} + 96n^{14}) \\
 g_7(n) &= \frac{n-7}{45\,360}(-156\,017\,752\,081\,551\,360 + 85\,163\,968\,967\,728\,896n - 22\,517\,704\,978\,919\,136n^2 \\
 &\quad + 4585\,470\,174\,542\,376n^3 - 851\,686\,123\,590\,540n^4 + 146\,137\,469\,433\,102n^5 \\
 &\quad - 24\,441\,080\,660\,523n^6 + 4148\,836\,864\,606n^7 - 747\,463\,726\,205n^8 \\
 &\quad + 149\,724\,735\,468n^9 - 33\,793\,043\,592n^{10} + 8322\,494\,124n^{11} - 1946\,680\,944n^{12} \\
 &\quad + 363\,148\,352n^{13} - 47\,679\,184n^{14} + 4019\,904n^{15} - 193\,536n^{16} + 4032n^{17})
 \end{aligned}$$


---

has a well-defined  $1/d$  expansion whose coefficients depend on  $n$ . If we assume that these coefficients are bounded in the limit  $n \rightarrow \infty$ , most of the coefficients in  $g_k$  are fixed, and we only need to know  $k + 1$  data points to fully determine  $g_k$ . See appendix B for the details of this argument. In our case this enables us to compute  $g_k$  up to  $k = 7$ , see table 4, and consequently  $A_d(13)$  and  $A_d(14)$ , see table 3.

We actually know all data to compute  $A_d(15)$  with the exception of the number  $A_7(15)$ . On our laptop, the enumeration of the missing number  $A_7(15)$  would take about 80 years. On a parallel system with a few hundred CPUs this would still take several months, which is not out of reach. Computing the formula for  $A_d(16)$ , however, is definitely beyond the power of our machinery.

Before we turn our attention to the analysis of the enumeration data we note that Lunnon’s equation (13) has a corresponding equation for perimeter polynomials:

$$g_{n,t}^{(d)} = \sum_i \binom{d}{i} G_{n,t-2(d-i)n}^{(i)}. \tag{17}$$

$G_{n,t}^{(d)}$  denotes the number of proper  $d$ -dimensional animals of size  $n$  and perimeter  $t$ . Since  $G_{n,t}^{(d)} = 0$  for  $d > n - 1$ , we can write

$$P_d(n, q) = q^{2dn-2(n-1)} \sum_{i=1}^{n-1} \binom{d}{i} \sum_t G_{n,t}^{(i)} q^{t-2-2n(i-1)}. \tag{18}$$

For a given value of  $n$ , (18) represents the perimeter polynomial for general dimension  $d$ . Our enumeration data allowed us to compute the  $G_{n,t}^{(d)}$  and hence the formulae (18) for  $n \leq 11$  (see [15] for the data), extending the previously known formulae for  $n \leq 7$  [14]. A computation of the next formula  $P_d(12, q)$  requires the knowledge of the perimeter polynomials for  $d \leq 11$  and  $n \leq 12$ . The enumeration for  $d = 11$  and  $n = 12$  alone would take approx. 38 years on our laptop.

### 5. Mean cluster size

From the perimeter polynomials we can compute moments of the cluster statistics like the mean cluster size:

$$S(p) = \frac{1}{p} \sum_{n=1}^{\infty} n^2 p^n P_d(n, 1-p) = \sum_r b_d(r) p^r. \tag{19}$$

The coefficients of the series expansion are

$$\begin{aligned} b_d(r) &= \sum_{n=1}^{r+1} n^2 \sum_t g_{n,t}^{(d)} \binom{t}{r+1-n} (-1)^{r+1-n} \\ &= (r+1)^2 A_d(r+1) + \sum_{n=1}^r n^2 \sum_t g_{n,t}^{(d)} \binom{t}{r+1-n} (-1)^{r+1-n}. \end{aligned} \tag{20}$$

Since we can compute  $A_d(r+1)$  from the perimeter polynomials  $P_d(n, q)$  for  $n \leq r$  via (7), we can also compute the series coefficients  $b_d(r)$  from this set of perimeter polynomials. If we happen to know  $A_d(r+2)$ , we can get an extra coefficient through

$$b_d(r+1) = (r+2)A_d(r+2) + \sum_{n=1}^r n(n-r-1) \sum_t g_{n,t}^{(d)} \binom{t}{r+2-n} (-1)^{r-n}. \tag{21}$$

This formula can be derived by solving (7) for  $\sum_t g_{n,t} t$  and plugging the result into (20).

We used (21) to extend the table of coefficients (table 5) for  $d \geq 8$ , since here we know perimeter polynomials only up to  $n = 11$ , but the cluster numbers  $A_d$  up to  $n = 14$ .

### 6. Growth rates and exponents

The cluster numbers  $A_d(n)$  are expected to grow asymptotically as

$$A_d(n) \sim C \lambda_d^n n^{-\Theta_d} \left( 1 + \frac{b}{n^\Delta} + \text{corrections} \right), \tag{22}$$

where the exponents  $\Theta_d$  and  $\Delta$  are universal constants, i.e. their value depends on the dimension  $d$ , but not on the underlying lattice, while  $C$  and  $b$  are nonuniversal, lattice-dependent quantities [22]. The universality facilitates the computation of  $\Theta_d$  for some values of  $d$  using field theoretic arguments. We know  $\Theta_3 = 3/2$  [23, 24],  $\Theta_4 = 11/6$  [25] and  $\Theta_d = 5/2$  (the value for the Bethe lattice) for  $d \geq d_c = 8$ , the critical dimension for animal growth [26].

The enumeration data for  $A_d(n)$  can be used to estimate both  $\lambda_d$  and  $\Theta_d$ . For that we compute  $\lambda_d(n)$  and  $\Theta_d(n)$  as the solutions of the system

$$\log A_d(n-k) = \log C + (n-k) \log \lambda_d(n) - \Theta_d(n) \log(n-k) \tag{23}$$

for  $k = 0-2$ . We need three equations to eliminate the constant  $\log C$ . Growth rate  $\lambda_d$  and exponent  $\Theta_d$  are obtained by extrapolating the numbers  $\lambda_d(n)$  and  $\Theta_d(n)$  to  $n \rightarrow \infty$ . From (22) we expect that

$$\log \lambda_d(n) \sim \log \lambda_d + \frac{b}{n^{\Delta+1}} \tag{24}$$

**Table 5.** Series coefficients of the mean cluster size  $S(p) = \sum_r b_r p^r$  in hypercubic lattices of dimension  $d$ . New values in boldface, while older values from [8] ( $d = 3$ ) and [14] ( $d = 4-7$ ) and references therein.

$r$	$d = 3$	$d = 4$	$d = 5$	$d = 6$
1	6	8	10	12
2	30	56	90	132
3	114	320	690	1272
4	438	1832	5290	12252
5	1542	9944	39210	115332
6	5754	55184	293570	1091472
7	19574	290104	2135370	10159252
8	71958	1596952	15839690	95435172
9	233574	8237616	113998170	<b>883192392</b>
10	870666	45100208	<b>840643170</b>	<b>8258076192</b>
11	2696274	<b>229502616</b>	<b>6017266290</b>	<b>76196541732</b>
12	10375770	<b>1254330128</b>	<b>44178511010</b>	<b>710151162432</b>
13	30198116	<b>6307973352</b>	<b>315024296150</b>	<b>6540805549192</b>
14	<b>122634404</b>	<b>34574952952</b>	<b>2307462163110</b>	<b>60831844077672</b>
15	<b>327024444</b>	<b>171364602736</b>		
16	<b>1460721616</b>			
17	<b>3347244554</b>			
18	<b>17795165832</b>			

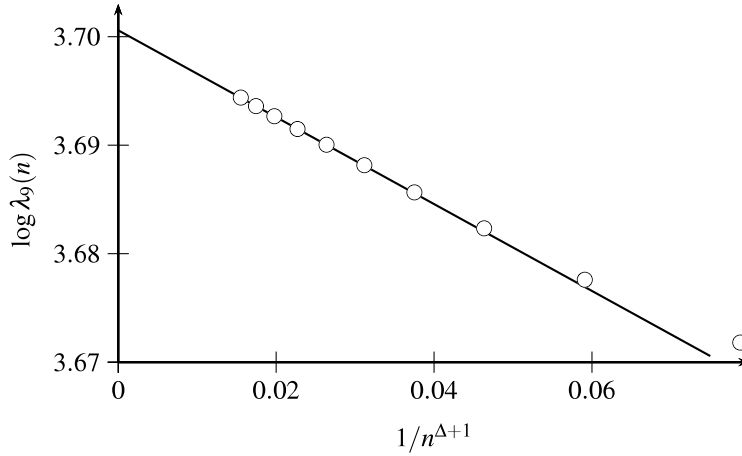
$r$	$d = 7$	$d = 8$	$d = 9$
1	14	<b>16</b>	<b>18</b>
2	182	<b>240</b>	<b>306</b>
3	2114	<b>3264</b>	<b>4770</b>
4	24542	<b>44368</b>	<b>74322</b>
5	280238	<b>595632</b>	<b>1146834</b>
6	3210074	<b>8012384</b>	<b>17720514</b>
7	36394302	<b>107053424</b>	<b>272530194</b>
8	414610014	<b>1434259248</b>	<b>4198328082</b>
9	<b>4685293438</b>	<b>19125485024</b>	<b>64487361906</b>
10	<b>53201681162</b>	<b>255662267296</b>	<b>991886672898</b>
11	<b>600207546946</b>	<b>3405928921264</b>	<b>1522687319670</b>
12	<b>6800785109594</b>	<b>45466350310880</b>	<b>233996383280898</b>
13	<b>76649757121000</b>		

J. Stat. Mech. (2011) P09026

for large values of  $n$ . We used the data points  $\lambda_d(n)$  for the three largest values of  $n$  to fit the parameters  $\log \lambda_d$ ,  $b$  and  $\Delta$  in (24). A plot of  $\log \lambda_d(n)$  versus  $n^{-\Delta-1}$  (figure 4) then shows that the data points in fact scale like (24). The resulting estimates for  $\log \lambda_d$  are listed in table 6. They agree very well with the high precision values from large scale Monte Carlo simulations [27, 28].

The same approach can be used to compute the exponent  $\Theta_d$ . Here we expect

$$\Theta_d(n) \sim \Theta_d + \frac{b}{n^\Delta}. \tag{25}$$



**Figure 4.** Growth rate  $\lambda_d$  for  $d = 9$ . The symbols are  $\lambda_9(n)$  computed from equation (23). The correction exponent  $\Delta = 0.58$  and the line are the result of a numerical fit to the three leftmost data points.

**Table 6.** Growth rates  $\lambda_d$  and exponents  $\theta_d$  obtained from extrapolating the enumeration data. The columns marked MC contain values from large scale Monte Carlo simulations [27, 28].

$d$	$\log \lambda_d$		$\Theta_d$	
	Enum.	MC	Enum.	Exact, MC
3	2.121 69	2.121 8588(25)	1.489	3/2
4	2.587 50	2.587 858(6)	1.796	11/6
5	2.922 54	2.922 318(6)	2.113	2.080(7)
6	3.178 38	3.178 520(4)	2.232	2.261(12)
7	3.384 03	3.384 080(5)	2.357	2.40(2)
8	3.554 84	3.554 827(4)	2.441	5/2
9	3.700 57	3.700 523(10)	2.489	5/2

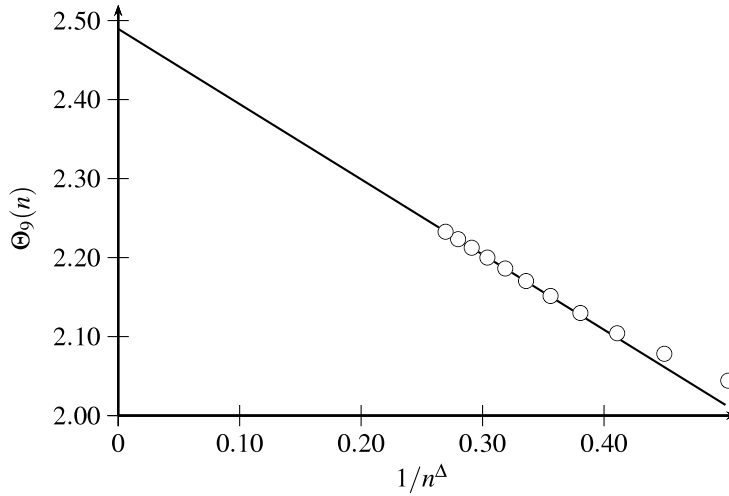
Again we used the data points  $\Theta_d(n)$  for the three largest values of  $n$  to fit the parameters  $\Theta_d$ ,  $b$  and  $\Delta$ . Figure 5 shows that  $\Theta_d(n)$  in fact scales like (25). The resulting estimates for  $\Theta_d$  (table 6) deviate from the Monte Carlo results and the exact values by no more than 3%.

The estimates for  $\lambda_d$  and  $\Theta_d$  based on the current known values of  $A_d(n)$  are much more precise than previous extrapolations based on shorter sequences  $A_d(n)$ , see [14, 16].

### 7. Conclusions

We have seen that the memory requirements of Redelmeier’s algorithm can be kept low by using hyperspherical regions of the lattice. Even in high dimensions, the limiting resource in Redelmeier’s algorithm is time, not space.

We used a lean and efficient implementation of Redelmeier’s algorithm to compute new perimeter polynomials in hypercubic lattices of dimensions  $d \leq 10$ . We have used these new perimeter polynomials together with combinatorial arguments based on proper



**Figure 5.** Exponent  $\Theta_d$  for  $d = 9$ . The symbols are  $\theta_9(n)$  computed from equation (23). The correction exponent  $\Delta = 0.50$  and the line are the result of a numerical fit to the three leftmost data points.

animals to compute new values of the cluster numbers  $A_d(n)$  and new formulae for  $A_d(n)$  for  $n \leq 14$  and arbitrary  $d$ . We have also used the new data to compute formulae for the perimeter polynomials  $P_d(n; q)$  for  $n \leq 11$  and arbitrary  $d$ . We have not shown these formulae here, but you can download them from the project webpage [15].

We have also used our data to compute the formula for  $\text{DX}(n, n - 7)$ , the number of proper animals of size  $n$  in dimension  $n - 7$  and new coefficients in the series expansion of the mean cluster size  $S(p)$ .

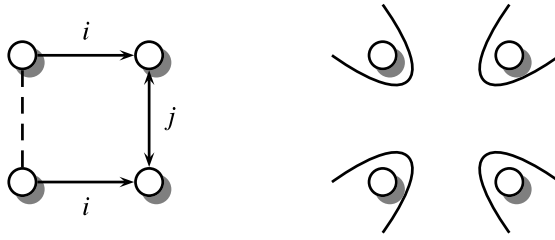
Based on the enumeration data, we have finally computed numerical values for the growth rates  $\lambda_d$  and the critical exponents  $\theta_d$  that agree very well with the results of Monte Carlo simulations and field theoretical predictions.

All in all we have explored the limits of computerized counting of lattice animals in dimensions  $d \geq 3$ . Any significant extension of the results presented here would require either a considerable amount of CPU time or an algorithmic breakthrough comparable to the transfer matrix methods for  $d = 2$ .

### Appendix A. Formulae for $\text{DX}(n, n - k)$ : structure

In the physics literature like [21], equation (16) is usually assumed to be true just because it is supported by the available enumeration data. But as a matter of fact, one can actually motivate (16) using the type of arguments that were used in [6] to prove the formula for  $\text{DX}(n, n - 2)$ . The idea is to show that the leading order of  $\text{DX}(n, n - k)$  is  $\sim 2^n n^{n+k-4}$  whereas the lowest order contributions are  $\sim 2^n n^{n-2k-1}$ . This is exactly the range of terms in (16) if  $g_k$  is a polynomial of degree  $3(k - 1)$ .

Equation (16) is obviously correct for  $k = 1$  (with  $g_1 = 1$ ). For  $k > 1$ , we can still represent lattice animals by trees, namely the spanning trees of their adjacency graph. Each spanning tree is again an edge-labeled tree, but this time there are only  $n - k$  labels for  $n - 1$  edges, i.e.  $k - 1$  edges will carry a label that is also used elsewhere in the tree.



**Figure A.1.** A labeled spanning tree that contains a part  $\circ \xrightarrow{i} \circ \xleftarrow{j} \circ \xleftarrow{i} \circ$  corresponds to a four-loop in the lattice animals, a quadrilateral that lives in the  $i$ - $j$  plane (left). To count the number of spanning trees with such a four-loop, the edges of the quadrilateral are removed and the vertices of the quadrilateral are considered the root vertices of disconnected trees (right). The number of the latter is given by (A.1) with  $\ell = 4$ .

Consider a tree whose  $n - 1$  edges are labeled with numbers  $1, \dots, n - 1$ . If we identify each of the high value labels  $n - k + 1, n - k + 2, \dots, n - 1$  with one of the low value labels  $1, \dots, n - k$ , we get the right set of labels. Since there are  $(n - k)^{k-1}$  ways to do this, the number of edge-labeled trees with  $n - k$  distinct labels scales like  $n^{k-1}n^{n-3} = n^{n+k-4}$  in leading order. With two directions for every edge we get  $2^n n^{n+k-4}$  for the leading order in  $\text{DX}(n, n - k)$ .

For  $k > 1$ , a proper animal can contain loops. Figure A.1 shows the simplest case: a loop that arises because one dimension ( $i$ ) is explored twice. The result is a quadrilateral that lives in the  $i$ - $j$  plane of the lattice. In the spanning tree, this corresponds to a part that is labeled  $\circ \xrightarrow{i} \circ \xleftarrow{j} \circ \xleftarrow{i} \circ$ . Since graphs with loops have several spanning trees, our count of edge-labeled trees overcounts the number of proper animals. We need to subtract some contributions from loopy animals. The idea is to break up the part of the spanning tree that corresponds to the loop and to separately count the number of trees that are attached to the vertices on the loopy part.

Consider an animal with a loop that contains  $\ell$  cells and one of its spanning trees. If we remove all the edges from the spanning tree that connect the vertices in the loop, the remaining graph is a forest, i.e. a collection of trees, where each tree is rooted in one of the  $\ell$  vertices. The forest has  $n - \ell$  vertices with a total of  $n - \ell$  edges, and it is independent from the way that the root vertices have been connected in the loop.

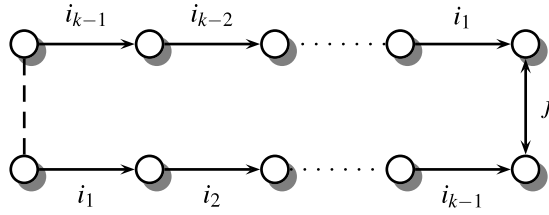
Now the number of ordered sequences of  $\ell \geq 1$  directed rooted trees with a total of  $n - \ell$  edges and  $n - \ell$  distinct edge labels is

$$2^{n-\ell} n^{n-\ell-1} \ell. \tag{A.1}$$

See [6] for a proof of (A.1). The lowest order corrections come from those animals for which the number  $\ell$  of cells in a loop is maximal. This is the case for a loop that joins all  $k - 1$  non-unique edge labels, see figure A.2. The number of vertices in these loops is  $\ell = 2k$ , hence the lowest order corrections are  $\sim 2^{n-2k} n^{n-2k-1}$ .

If we want the exact number of spanning trees for loopy animals, we need to count the number of ways to reconnect the roots of the forest to form a single tree. But since this number does depend on  $k$  but not on  $n$  and we are interested only in the scaling with  $n$ , we do not need to enter this discussion here. The same is true for animals that contain

J. Stat. Mech. (2011) P09026



**Figure A.2.** An animal that contributes to  $\text{DX}(n, n - k)$  uses at most  $k - 1$  dimensions more than once, and the longest loop arises when each of these  $k - 1$  dimensions is explored twice. Such a loop contains  $2k$  vertices, as shown here.

several small loops instead of a single loop of maximal length. If we apply the separation trick to one of the shorter loops, we get a scaling of order larger than  $\sim 2^{n-2k} n^{n-2k-1}$ , and the resulting forest is then labeled with fewer labels than edges, which increases the order even further. So the lowest order corrections from loops come in fact from single loops of maximum length, and these contributions are of order  $\sim 2^{n-2k} n^{n-2k-1}$ , as claimed in (16).

Besides the non-uniqueness of spanning trees for loopy graphs, there is another type of error that needs to be corrected: some edge-labeled trees correspond to animals with overlapping cells, i.e. to illegal animals. For instance, if a spanning tree contains the subtree

$$\bullet \xrightarrow{i} \circ \xleftarrow{i} \bullet \quad \text{or} \quad \bullet \xleftarrow{i} \circ \xrightarrow{i} \bullet,$$

the two  $\bullet$ 's represent the very same cell of the animal. But these 'colliding' configurations can be interpreted as three-loops or, more generally, as  $\ell$  loops, and counted in the same way as the legal loops above. Again the lowest order contributions come from the longest 'colliding' loops which are formed by  $k - 1$  labels assigned to two edges each and arranged like

$$\bullet \xrightarrow{i_1} \circ \xrightarrow{i_2} \circ \dots \circ \xrightarrow{i_{k-1}} \circ \xleftarrow{i_1} \circ \xleftarrow{i_2} \circ \dots \circ \xleftarrow{i_{k-1}} \bullet.$$

These longest collision loops contain  $2k - 1$  vertices of the tree (representing  $2k - 2$  cells of the animal). According to (A.1), their number scales like  $2^{n-2k+1} n^{n-2k}$ , one order above the lowest order of legal loops.

This concludes the motivation of (16). Note that a proof of (16) would require a thorough analysis to exclude contributions outside the range covered by (16).

**Appendix B. Formulae for  $\text{DX}(n, n - k)$ : coefficients**

Having established the fact that  $\text{DX}(n, n - k)$  is given by (16) we still have to determine the coefficients of the polynomials  $g_k(n)$ . Since  $g_k$  has degree  $3(k - 1)$ , it seems that we need to know the  $3k - 2$  values  $\text{DX}(k, 0), \text{DX}(k + 1, 1), \dots, \text{DX}(4k - 3, 3k - 3)$  to compute the coefficients. In terms of our enumeration data this means knowledge of  $A_d(n)$  for  $n \leq 4k - 3$  and  $d \leq n - k$ . The data in table 2 suffice to compute the coefficients of  $g_k$  for  $k = 2, 3$ , but not for  $k \geq 4$ . Nevertheless we can compute  $g_k$  for  $k \leq 7$  by assuming that the 'free energy'

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log A_d(n)$$

has a well-defined  $1/d$  series expansion. This approach has been used to compute  $g_k$  for  $k \leq 6$  from much less enumeration data in [21] and [29], and we used it to compute  $g_7$  from the new enumeration data. Since the method has not been described in detail elsewhere, we provide a description in this appendix.

Let us start with Lunnon's equation, which tells us that  $A_d(n)$  is a polynomial of degree  $n - 1$  in  $d$  with coefficients that depend on  $n$ . For  $d \geq n$  we have

$$A_d(n) = \sum_{k=1}^{n-1} \text{DX}(n, k) \binom{d}{k} = \sum_{j=1}^{n-1} a_j(n) d^j \tag{B.1}$$

with

$$a_j(n) = \sum_{k=1}^{n-1} \frac{\text{DX}(n, k)}{k!} \begin{bmatrix} k \\ j \end{bmatrix}, \tag{B.2}$$

where  $\begin{bmatrix} k \\ j \end{bmatrix}$  denotes the Stirling number of the first kind. In particular we get

$$\begin{aligned} a_{n-1}(n) &= \frac{\text{DX}(n, n-1)}{(n-1)!} & a_{n-2}(n) &= \frac{\text{DX}(n, n-2)}{(n-2)!} + \frac{\text{DX}(n, n-1)}{(n-1)!} \begin{bmatrix} n-1 \\ n-2 \end{bmatrix} \\ a_{n-3}(n) &= \frac{\text{DX}(n, n-3)}{(n-3)!} + \frac{\text{DX}(n, n-2)}{(n-2)!} \begin{bmatrix} n-2 \\ n-3 \end{bmatrix} + \frac{\text{DX}(n, n-1)}{(n-1)!} \begin{bmatrix} n-1 \\ n-3 \end{bmatrix} \end{aligned}$$

and so on. From (B.1) we get

$$A_d(n) = a_{n-1}(n) d^{n-1} \left( 1 + \sum_{j=1}^{n-2} \frac{a_{n-1-j}}{a_{n-1}} d^{-j} \right),$$

and with  $\ln(1+x) = x + x^2/2 + x^3/3 \dots$  this gives the  $1/d$  series for the 'free energy':

$$\frac{1}{n} \ln A_d(n) = \left( 1 - \frac{1}{n} \right) \ln d + \frac{1}{n} \ln a_{n-1}(n) + \frac{1}{n} \frac{a_{n-2}(n)}{a_{n-1}(n)} \frac{1}{d} + \mathcal{O} \left( \frac{1}{d^2} \right). \tag{B.3}$$

We assume that all coefficients in this series remain bounded in the limit  $n \rightarrow \infty$ . This is definitely true for the zeroth-order term:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \ln a_{n-1}(n) = 1 + \ln 2.$$

For the first-order coefficient we get

$$\begin{aligned} \frac{1}{n} \frac{a_{n-2}}{a_{n-1}} &= \frac{n-1}{n} \frac{\text{DX}(n, n-2)}{\text{DX}(n, n-1)} + \frac{1}{n} \begin{bmatrix} n-1 \\ n-2 \end{bmatrix} \\ &= \left( 1 - \frac{1}{n} \right) \left( \frac{g_2(n)}{4n^2} - \frac{n-2}{2} \right). \end{aligned}$$

This is only bounded if the  $g_2(n)$  term balances the second term, i.e. if the  $n^3$  coefficient of  $g_2$  equals 2. Using also the fact that  $\text{DX}(2, 0) = 0$ , we can write

$$g_2(n) = (n-2)(2n^2 + bn + c).$$

To compute the remaining coefficients, we only need to know  $\text{DX}(3, 1) = 1$  and

$$\text{DX}(4, 2) = A_2(4) - 2 = 17$$

to get

$$g_2(n) = (n - 2)(2n^2 - 6n + 9).$$

The postulation of bounded coefficients in the series (B.3) has saved us from knowing the value  $\text{DX}(5, 3)$  to compute  $g_2$ . How much does it help us to compute  $g_k$ ?

The polynomial  $g_k$  enters the series expansion (B.3) via the term

$$\frac{1}{n} \frac{\text{DX}(n, n - k)}{\text{DX}(n, n - 1)} \frac{(n - 1)!}{(n - k)!} = 2^{2-2k} n^{1-2k} g_k(n) \underbrace{(n - 1)(n - 2) \cdots (n - k + 1)}_{\Theta(n^{k-1})}$$

in the coefficient of  $d^{-(k-1)}$ . The leading order of this term is  $n^{-k} g_k(n)$ . All terms of degree larger than  $k$  in the polynomial  $g_k$  lead to unbounded contributions to the series coefficient that need to be counterbalanced by other terms. These balancing terms always exist, a fact that gives additional support for the claim of bounded coefficients. The coefficients of the terms of order larger than  $k$  in  $g_k$  are therefore computable from the known terms  $g_{k-1}(n), g_{k-2}(n), \dots$  that also enter the same coefficient. Only the  $k + 1$  low-order terms of  $g_k$  are not fixed by the postulate of bounded coefficients and we need  $k + 1$  data points  $\text{DX}(k, 0), \text{DX}(k + 1, 1), \dots, \text{DX}(2k, k)$  to complete  $g_k$ .

Our enumeration data suffices to compute  $g_7$  (see table 4). The computation of  $g_8$  requires knowledge of  $\text{DX}(15, 7)$  and  $\text{DX}(16, 8)$  or, in terms of  $A_d(n)$ ,

$$\text{DX}(15, 7) = A_7(15) - 572\,521\,427\,068\,702\,741$$

and

$$\text{DX}(16, 8) = A_8(16) + 48\,366\,334\,433\,679\,758 - 56A_5(16) + 28A_6(16) - 8A_7(16).$$

## References

- [1] Golomb S W, 1994 *Polyominoes* 2nd edn (Princeton, NJ: Princeton University Press)
- [2] Stauffer D and Aharony A, 1992 *Introduction to Percolation Theory* 2nd edn (London: Taylor and Francis)
- [3] Guttman A J (ed), 2009 *Polygons, Polyominoes and Polycubes (Springer Lecture Notes in Physics vol 775)* (Heidelberg: Springer)
- [4] Klarner D A, *Cell growth problems*, 1967 *Can. J. Math.* **19** 851
- [5] Madras N, *A pattern theorem for lattice clusters*, 1999 *Ann. Comb.* **3** 357
- [6] Barequet R, Barequet G and Rote G, *Formulae and growth rates of high-dimensional polycubes*, 2010 *Combinatorica* **30** 257
- [7] Redelmeier D H, *Counting polyominoes: yet another attack*, 1981 *Discrete Math.* **36** 191
- [8] Mertens S, *Lattice animals: a fast enumeration algorithm and new perimeter polynomials*, 1990 *J. Stat. Phys.* **58** 1095
- [9] Mertens S and Lautenbacher M E, *Counting lattice animals: a parallel attack*, 1992 *J. Stat. Phys.* **66** 669
- [10] Jensen I, *Enumerations of lattice animals and trees*, 2001 *J. Stat. Phys.* **102** 865
- [11] Aleksandrowicz G and Barequet G, *Counting d-dimensional polycubes and nonrectangular planar polyominoes*, 2009 *Int. J. Comput. Geometry Appl.* **19** 215
- [12] Aleksandrowicz G and Barequet G, *Counting polycubes without the dimensionality curse*, 2009 *Discrete Math.* **309** 4576
- [13] Conway J H and Sloane N J A, *Low-dimensional lattices. VII coordination sequences*, 1997 *Proc. R. Soc. A* **453** 2369
- [14] Gaunt D S, Sykes M F and Ruskin H, *Percolation processes in d-dimensions*, 1976 *J. Phys. A: Math. Gen.* **9** 1899
- [15] Mertens S, *Lattice animals*, 2011 [www.ovgu.de/mertens/research/animals](http://www.ovgu.de/mertens/research/animals)
- [16] Gaunt D S, *The critical dimension for lattice animals*, 1980 *J. Phys. A: Math. Gen.* **13** L97
- [17] Lunnon W F, *Counting multidimensional polyominoes*, 1975 *Comput. J.* **18** 366
- [18] Cayley A, *A theorem on trees*, 1889 *Q. J. Pure Appl. Math.* **23** 376

- [19] Cameron P J, *Counting two-graphs related to trees*, 1995 *Electron. J. Comb.* **2** R4
- [20] Fisher M E and Essam J W, *Some cluster size and percolation problems*, 1961 *J. Math. Phys.* **2** 609
- [21] Peard P J and Gaunt D S, *1/d-expansions for the free energy of lattice animal models of a self-interacting branched polymer*, 1995 *J. Phys. A: Math. Gen.* **28** 6109
- [22] Adler J, Meir Y, Harris A B, Aharony A and Duarte J A M S, *Series study of random animals in general dimensions*, 1988 *Phys. Rev. B* **38** 4941
- [23] Parisi G and Sourlas N, *Critical behavior of branched polymers and the Lee–Yang edge singularity*, 1981 *Phys. Rev. Lett.* **46** 871
- [24] Imbrie J Z, *Dimensional reduction and crossover to mean-field behavior for branched polymers*, 2003 *Annal. Henri Poincaré* **4** S445
- [25] Dhar D, *Exact solution of a directed-site animals-enumeration problem in three dimensions*, 1983 *Phys. Rev. Lett.* **51** 853
- [26] Lubensky T C and Isaacson Joel, *Statistics of lattice animals and dilute branched polymers*, 1979 *Phys. Rev. A* **20** 2130
- [27] Hsu H-P, Nadler W and Grassberger P, *Simulations of lattice animals and trees*, 2005 *J. Phys. A: Math. Gen.* **38** 775
- [28] Hsu H-P, Nadler W and Grassberger P, *Statistics of lattice animals*, 2005 *Comput. Phys. Commun.* **169** 114
- [29] Gaunt D S and Peard P J, *1/d-expansions for the free energy of weakly embedded site animal models of branched polymers*, 2000 *J. Phys. A: Math. Gen.* **33** 7515