IOPscience iopscience.iop.org

Home Search Collections Journals About Contact us My IOPscience

Area versus length distribution for closed random walks

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2003 J. Phys. A: Math. Gen. 36 1539

(http://iopscience.iop.org/0305-4470/36/6/301)

View the table of contents for this issue, or go to the journal homepage for more

Download details:
IP Address: 171.66.16.89
The article was downloaded on 02/06/2010 at 17:21

Please note that terms and conditions apply.

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

J. Phys. A: Math. Gen. **36** (2003) 1539–1551 PII: S0305-4470(03)53618-6

Area versus length distribution for closed random walks

Filippo Colomo

INFN, Sezione di Firenze and Dipartimento di Fisica, Università di Firenze, Via G Sansone 1, 50019 Sesto Fiorentino (FI), Italy

E-mail: colomo@fi.infn.it

Received 17 September 2002, in final form 20 November 2002
Published 29 January 2003
Online at stacks.iop.org/JPhysA/36/1539

Abstract
Using a connection between the q-oscillator algebra and the coefficients of the high temperature expansion of the frustrated Gaussian spin model, we derive an exact formula for the number of closed random walks of given length and area, on a hypercubic lattice, in the limit of infinite number of dimensions. The formula is investigated in detail, and asymptotic behaviour is evaluated. The area distribution in the limit of long loops is computed. As a byproduct, we also obtain an infinite set of new, nontrivial identities.

PACS numbers: 05.40.Fb, 05.50.+q

1. Introduction

Let us consider, on a hypercubic lattice of dimension D, all closed paths of length $2k$, starting from a given vertex. An interesting and still unanswered question concerns the counting of such paths according to the area they enclose. Besides being a challenging combinatorial problem in its own right, the question also has some relevance from a physical point of view: considering, for instance, the hopping of a charged particle around a two-dimensional lattice in a uniform magnetic field [1–3], its energy spectrum displays an intricate hierarchic structure [4], whose properties are closely related to the area versus length distribution of closed paths [5]. Such a distribution finds applications also in polymer physics [6, 7], and in explaining anomalous magnetoconductance [8]. A different but somehow related problem is the counting of random walks according to the number of distinct sites they visit [9, 10]: relatively compact loops are expected to enclose a small area and to visit relatively few sites, and vice versa. This counting problem is relevant when investigating random walks on lattices with static traps.

Coming back to the initial question of counting loops according to their enclosed area, in the case of a two-dimensional lattice, the answer was given only in the limit of asymptotically large k in [11], and the first subleading correction was provided in [12]. In higher dimension, the problem is in general even more difficult. Here we want to address the question in the limit of infinite dimensionality of the lattice, a very peculiar situation, in which however some simplifications occur, and an exact answer can be given. The latter is to be interpreted as a 'mean-field' approximation to the exact counting on a lattice of finite dimensionality D. Such a formulation of the problem was first considered by Parisi *et al* [13, 14]. They were investigating spin models with frustration but without any quenched disorder, in order to test the conjecture that such deterministic models could behave at low temperature as some suitably chosen spin-glass model with quenched disorder. They considered the frustrated spherical and XY spin models in the limit of large dimensionality D of the lattice, where the saddle-point approximation becomes exact. In their analysis of these models, they showed how the high temperature expansion (i.e. loop expansion) can be nicely rewritten by using the q-oscillator algebra [15], where q measures the frustration per plaquette. Similar q-deformed algebraic relations have also appeared in the Hofstadter problem of quantum particles hopping on a two-dimensional lattice in a uniform magnetic field [16]. This problem is closely related to the frustrated spin models, which can be considered as simplified models of hopping in the large D and classical limits. The use of somehow related non-commutative geometry techniques also appeared as a crucial ingredient in [12].

The problem we address and solve in this paper may be stated as follows: given an infinite-dimensional hypercubic lattice, what is the number $G_{k,l}$ of loops of length $2k$, starting from a given vertex, enclosing a minimal area of exactly l plaquettes[1]?

The paper is organized as follows. In the next section we summarize some results of [13, 14] which are then used in section 3 to obtain an exact expression for the counting numbers $G_{k,l}$. In section 4 a generating function for these counting numbers is presented. In section 5 a probability distribution is naturally associated with the $G_{k,l}$, and a systematic procedure is built to compute its moments of arbitrary order. The asymptotic behaviour of this probability distribution for large k is evaluated in section 6, while section 7 is devoted to the presentation of some new, nontrivial identities. The last section contains our conclusions, while some technical details are relegated to the appendix.

2. Frustrated lattice and q-oscillator algebra

In [13, 14], Parisi *et al* considered the frustrated Gaussian, spherical and XY models on a D-dimensional hypercubic lattice, in the large D limit. Among other things, they unveiled a remarkable connection between the coefficients of the high temperature expansion (i.e. loop expansion) and the q-oscillator algebra [15], where q measures the frustration per plaquette and varies continuously on the real interval $[-1, +1]$, between the fully-frustrated case ($q = -1$, fermionic algebra) and the ferromagnetic case ($q = 1$, bosonic algebra). For our purposes, the relevant part of the Hamiltonians describing the models considered in [13, 14] is

$$H = -\frac{1}{\sqrt{2D}} \sum_{\langle jk \rangle} \phi_j^\dagger U_{jk} \phi_k + \text{h.c.} \tag{1}$$

The complex field $\phi_j \in \mathbb{C}$ is defined on the sites (labelled by j) of a D-dimensional hypercubic lattice; frustration is induced on the lattice through the nearest-neighbour couplings U_{jk}, which are complex numbers of modulus 1 and satisfy the relation $U_{jk} = U_{kj}^*$; they are the link variables of a background Abelian lattice gauge field, chosen in such a way as to produce a static and constant magnetic field, suitably oriented to give the same magnetic flux $\pm B$ for any plaquette of the lattice (the product of the four U around the plaquette being $e^{\pm iB}$).

[1] When facing this problem of area versus length distribution of closed random walks, the area is most often considered in its algebraic sense, i.e. as oriented; in our case, however, an average over orientations is intrinsically performed in the approach (see section 2), and we are therefore referring to the absolute value of the area.

Such a magnetic field has the same projection over all the axes, modulo the sign; in [13, 14] these signs were chosen randomly in order to avoid the selection of a preferred direction. The usual unfrustrated ferromagnetic spin interaction is obtained for $B = 0$. Non-vanishing values of B induce a frustration around each plaquette, which is maximal for $B = \pi$, the fully-frustrated case.

In the framework of the high temperature expansion, the free energy of such models is expressed as a sum over the contribution of loops of increasing length $2k$:

$$\beta F = \sum_{k=0}^{\infty} \frac{\beta^{2k}}{2k} G_k. \tag{2}$$

Each loop encloses a number of plaquettes; in the case of the models considered in [13, 14], for each loop the magnetic field yields a weight proportional to $\exp(\mathrm{i}BA)$, where A is the sum of plaquettes with signs depending on the orientations. The total contribution of all loops of length $2k$ is given by G_k. Due to the average over orientations and loops, the quantity G_k is a polynomial in the variable $q = \cos B$, the coefficient $G_{k,l}$ of q^l being given by the number of loops of length $2k$ and area l.[2] The infinite dimensionality of the lattice ensures that for any given loop, no two plaquettes of the subtended minimal surface will lie on the same plane; this in turn accounts for such a simple averaging over orientations, performed independently for each plane. The order of polynomial $G_k(q)$ is given by $k(k-1)/2$, the maximal area can be enclosed by a loop of length $2k$. The coefficients $G_{k,l}$ can also be interpreted as the number of Feynman diagrams with $2k$ external points, which are joined pairwise by lines (propagators) intersecting l times. These diagrams also occur in the topological (large N) expansion of matrix models [17], where the planar limit corresponds to no intersections, i.e. to the $q = 0$ case. Equivalently, in simple graphical terms, they can just be seen as the number of ways of connecting pairwise $2k$ points on a circle with k segments intersecting exactly l times. In the following we shall refer to this last picture[3]. In [13], the enumeration of such diagrams was investigated. In particular, a recursion relation was found for the coefficients of the polynomial $G_k(q)$—a sort of Wick theorem—which can be nicely expressed in terms of the algebra of the q-oscillators $a_q, a_q^\dagger$:

$$a_q a_q^\dagger - q a_q^\dagger a_q = 1. \tag{3}$$

These operators [15] act on the Hilbert space spanned by the vectors: $|m\rangle, m = 0, 1, \ldots,$ as follows:

$$\begin{aligned} a_q|m\rangle &= \sqrt{[m]_q}\,|m-1\rangle \qquad & a_q|0\rangle &= 0 \\ a_q^\dagger|m\rangle &= \sqrt{[m+1]_q}\,|m+1\rangle \qquad & [m]_q &= \frac{1-q^m}{1-q}. \end{aligned} \tag{4}$$

Using the recursion relation, the weighted multiplicities of the diagrams of equation (2) were neatly written as an expectation value over the ground state of the q-oscillators [13, 14]:

$$G_k(q) = \langle 0|\left(a_q^\dagger + a_q\right)^{2k}|0\rangle. \tag{5}$$

The authors of [13, 14] did not exploit the consequences of this result, preferring to turn themselves to the numerical investigation of the models they were interested in. As a consequence, till now only the two limiting cases $G_k(0)$ and $G_k(1)$ were explicitly known; when $q = 1$ it is only a matter of counting the ways of connecting $2k$ points on a circle, with no restrictions, and this is simply the number of pairings of $2k$ objects: $(2k-1)!!$.

[2] The area of a loop is defined as the minimal area of a surface of lattice plaquettes having that loop as boundary.

[3] This formulation of the problem was considered, e.g., by Pauling [18] when trying to simplify the calculation of matrix elements involved in Slater's treatment of the electronic structure of molecules.

When $q = 0$ we are in fact evaluating the planar limit of the zero-dimensional $2k$-point Green function of a matrix model in the limit of vanishing interaction [17]. In simple graphical terms this corresponds to the number of ways of joining pairwise $2k$ points on a circle with no intersection. In other words, this is just one of the many possible definitions of Catalan numbers (see, e.g., [19]), given as

$$G_k(0) = \frac{(2k)!}{k!(k+1)!}. \tag{6}$$

But the numbers $G_{k,l}$, counting the ways of connecting pairwise $2k$ points on a circle, with exactly l interactions (i.e. the coefficients of q^l in $G_k(q)$), are not so easily accessible. In [13, 14], it was found by direct enumeration of the graphs on a computer. In [20], a generating function for coefficients $G_{k,l}$ was obtained, proposed, but then the explicit evaluation of such coefficients relied upon heavy symbolic manipulations.

3. Exact solution of the enumeration problem

We shall now present a simple, easy to evaluate, formula for a generic coefficient $G_{k,l}$. To this purpose let us first introduce the x_q coordinate representation, $x_q = a_q^\dagger + a_q$, $x_q|x\rangle = x|x\rangle$, which is given by the so-called continuous q-Hermite polynomials [22, 23]. These are defined by

$$H_n(x) = \langle x|n\rangle \mathcal{C}_n \qquad \mathcal{C}_n = ([n]_q!)^{1/2}\mathcal{C}_0 \tag{7}$$

where the normalization constant $\mathcal{C}_0$ is fixed by $H_0(x) = 1$ and the q-factorial is

$$[n]_q! = [n]_q[n-1]_q \cdots [1]_q \qquad [1]_q = [0]_q = 1. \tag{8}$$

These polynomials satisfy, of course, a three-term recursion relation in the index n:

$$xH_n(x) = H_{n+1}(x) + [n]_q H_{n-1}(x) \qquad n \geqslant 1. \tag{9}$$

x ranges over the interval $x \in [-2/\sqrt{1-q}, 2/\sqrt{1-q}]$, and a convenient parametrization is

$$x = \frac{2}{\sqrt{1-q}}\cos\theta \qquad \theta \in [0, \pi]. \tag{10}$$

More properties of these q-Hermite polynomials can be found in [22], where they are defined as $\mathcal{H}_n(\cos\theta) = (1-q)^{n/2}H_n(x)$. The most important property for us is the orthogonalizing measure $\nu_q(x)$ [21–23]:

$$\int_{-2/\sqrt{1-q}}^{2/\sqrt{1-q}} \nu_q(x)\,\mathrm{d}x\, H_n(x)H_m(x) = \delta_{n,m}[n]_q! \tag{11}$$

$$\begin{aligned}\nu_q(x) &= \frac{\sqrt{1-q}}{2\pi} q^{-1/8}\Theta_1\left(\frac{\theta}{\pi}, \tau\right) \\ &= \frac{\sqrt{1-q}}{\pi}\sum_{n=0}^{\infty}(-1)^n q^{n(n+1)/2}\sin[(2n+1)\theta]\end{aligned} \tag{12}$$

where $\Theta_1(z, \tau)$ is the first Jacobi theta function [24]:

$$\Theta_1(z, \tau) = \sum_{n\in\mathbb{Z}} \exp\left(\mathrm{i}\pi\tau\left(n+\tfrac{1}{2}\right)^2 + 2\pi\mathrm{i}\left(n+\tfrac{1}{2}\right)\left(z-\tfrac{1}{2}\right)\right) \qquad q = \mathrm{e}^{2\pi\mathrm{i}\tau}. \tag{13}$$

It is worth emphasizing that, interestingly enough, measure (12) interpolates continuously between the Wigner semi-circle law for Gaussian matrix models, and the Gaussian distribution, as q varies from 0 to 1^-; this reminds us of the behaviour of the density distribution of eigenvalues for a Gaussian ensemble of random matrices, as the dimension of the matrices, N, varies from ∞ to 1.

Polynomial $G_k(q)$ may now be rewritten as

$$G_k(q) = \int_{-2/\sqrt{1-q}}^{2/\sqrt{1-q}} \nu_q(x) x^{2k}\, \mathrm{d}x. \tag{14}$$

Indeed, using explicit form (12) for the integration measure, performing the integral and playing a little bit with indices, we may write

$$G_k(q) = \left(\frac{1}{1-q}\right)^k \sum_{l=0}^{k} (-1)^l \binom{2k+1}{k-l} \frac{2l+1}{2k+1} q^{l(l+1)/2}. \tag{15}$$

We now only need to perform explicitly the division of the polynomial of degree $k(k+1)/2$ defined by the sum in the previous formula. To this purpose, let us state the following:

Theorem[4]**.** *Let $P(q) = \sum_{l=0}^{N} p_l q^l$ be an integer coefficient polynomial of degree N in q, exactly divisible by $(1-q)^k$. Then $A(q) = P(q)/(1-q)^k$ is an integer coefficient polynomial of degree $N-k$ whose coefficients are simply expressed in terms of the p_l as follows:*

$$A(q) = \sum_{l=0}^{N-k} a_l q^l \qquad a_l = \sum_{i=0}^{l} \binom{k+l-1-i}{k-1} p_i. \tag{16}$$

We therefore readily get a simple closed expression for $G_{k,l}$:

$$G_{k,l} = \sum_{i=0}^{i_{\max}} (-1)^i \binom{k+l-1-i(i+1)/2}{k-1} \binom{2k+1}{k-i} \frac{2i+1}{2k+1} \qquad l \leqslant \frac{k(k-1)}{2} \tag{17}$$

where $i_{\max} = [(-1+\sqrt{1+8l})/2]_{i.p.}$ is the largest integer i satisfying $i(i+1)/2 \leqslant l$. Because of their definition as counting numbers, all integers $G_{k,l}$ should be positive. This fact is indeed not apparent from (17), and we are not able to present a rigorous proof of this statement, nevertheless we are convinced of its validity. Sensible arguments rely upon the intrinsic positivity of (5) and its derivatives with respect to q, for any value of q in the interval [0, 1).

We have computed explicitly with Mathematica [25], through formula (17), which is very efficient, all values $G_{k,l}$ for k ranging from 1 to 9, and verified that they indeed match the results of [14] which were found by direct enumeration of the graphs on a computer.

4. Generating function

Let us define a generating function for the combinatorial numbers $G_{k,l}$:

$$R(z,q) \equiv \sum_{k=0}^{\infty} G_k(q) z^k = \sum_{k,l=0}^{\infty} G_{k,l} q^l z^k. \tag{18}$$

The running of index l may or may not be restricted to the range $0 \leqslant l \leqslant k(k-1)/2$, without loss of generality: coefficients $G_{k,l}$ automatically vanish for $l > k(k-1)/2$. From expression (14), we may readily write generating functions for coefficients $G_k(q)$:

$$R(z,q) = \int_{-2/\sqrt{1-q}}^{2/\sqrt{1-q}} \nu_q(x) \frac{1}{1-zx^2}\, \mathrm{d}x. \tag{19}$$

[4] We are unfortunately unable to quote any reference. We are of course convinced that this theorem was enounced a long time ago, indeed it is just a corollary of Ruffini's rule. The only proof we have been able to give, absolutely inelegant and unworthy of being presented here, is by direct verification.

Inserting for the measure its explicit expression (12), and performing the integral, we obtain

$$R(z,q)=\sqrt{\frac{1-q}{z}}\sum_{n=0}^{\infty}(-1)^n q^{n(n+1)/2}\left[K\left(\frac{z}{1-q}\right)\right]^{2n+1} \tag{20}$$

where we introduced

$$K(\alpha)=\frac{1-\sqrt{1-4\alpha}}{2\sqrt{\alpha}}. \tag{21}$$

For $q=0$ function (20) reduces to the standard generating function of Catalan numbers. The singularities of $R(z,q)$ in the complex z-plane are completely determined by those of the simpler function $K(z/(1-q))$. Actually the series is very well convergent for $|q|<1$ as any Jacobi theta function. Moreover, for $|q|=1$, it is a geometric series, which is still convergent because $|K(\alpha)|<1$ for $|z|<z_c\equiv(1-q)/4$.

Generating function (20), as shown in [20] in a different context, can be interpreted physically as the internal energy at temperature $T=1/\sqrt{z}$ of a Gaussian spin model defined on the sites of the infinite-dimensional frustrated lattice introduced in [13, 14].

5. Moments of the distribution

The interpretation of coefficients $G_{k,l}$ naturally suggests the introduction, for each positive integer value of index k, of the following normalized (discrete) probability distribution:

$$P_k(l)=\frac{G_{k,l}}{\sum_{l=0}^{k(k-1)/2}G_{k,l}}=\frac{G_{k,l}}{G_k(1)}=\frac{G_{k,l}}{(2k-1)!!}\qquad\begin{cases}k=0,1,\ldots\\ l=0,1,\ldots,k(k-1)/2\end{cases} \tag{22}$$

whose interpretation is obvious: given an arbitrary diagram with $2k$ legs, the probability for it to have exactly l crossings is simply given by $P_k(l)$; equivalently, $P_k(l)$ measures the probability, for a randomly chosen closed path of length $2k$ on our infinite-dimensional hypercubic lattice, to have area equal to l. The moments of the previous probability distribution are defined as

$$M_s^{(k)}\equiv\langle l^s\rangle_k=\sum_{l=0}^{k(k-1)/2}l^s P_k(l) \tag{23}$$

and can be computed from the following generating function:

$$\tilde{P}_k(q)\equiv\langle q^l\rangle_k=\frac{G_k(q)}{G_k(1)}. \tag{24}$$

Indeed,

$$M_s^{(k)}=\left[\left(q\frac{\partial}{\partial q}\right)^s\tilde{P}_k(q)\right]_{q=1}=\left[\left(\frac{\partial}{\partial\log q}\right)^s\tilde{P}_k(q)\right]_{q=1}. \tag{25}$$

The moments of the deviations with respect to the mean $M_1^{(k)}$ (central moments) are given by a similar expression:

$$m_s^{(k)}\equiv\left\langle\left(l-M_1^{(k)}\right)^s\right\rangle_k=\left[\left(\frac{\partial}{\partial\log q}\right)^s\tilde{P}_k(q)q^{-M_1^{(k)}}\right]_{q=1}. \tag{26}$$

The computation of such moments, besides being interesting in its own right, is (at least for the first two) a necessary step towards the evaluation of the asymptotic form of probability distribution (22) in the limit $k\to\infty$. We therefore need a systematic procedure to calculate $G_k(q)$, together with its derivatives, at $q=1$. Indeed, from expression (14), we essentially need to evaluate the derivatives at $q=1$ of the integration measure (12). Unfortunately, such a limit, due to the presence of the Jacobi theta function, is highly singular: for fixed values of x, $\nu_q(x)$ develops an essential singularity at $q = 1$. A prescription should be given, in our situation the natural one being to choose to approach $q = 1$ from real, smaller than 1, values of q; in a more convenient, equivalent notation, we want to evaluate the behaviour of $\nu_q(x)$ in the limit $\epsilon \to 0^+$, ϵ real, defined by $q \equiv \mathrm{e}^{-\epsilon}$. With such a choice, as shown in the appendix, we can write any given moment $M_s^{(k)}$ as a simple combination of trivial Gaussian moments:

$$M_s^{(k)} = \frac{(-1)^s}{(2k-1)!!} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{d}x\, x^{2k} c_s(x)\, \mathrm{e}^{-\frac{x^2}{2}} \tag{27}$$

where $c_s(x)$ is an even polynomial of degree $4s$ in x. The only practical difficulty is the explicit computation of the sth coefficient $c_s(x)$ of the Taylor expansion around $\epsilon = 0^+$, of the function

$$\sqrt{1-\mathrm{e}^{-\epsilon}}\, \mathrm{e}^{\frac{\epsilon}{8}} \frac{1}{\sqrt{\epsilon}} \exp\left(-\frac{1}{2}\left[\frac{4}{\epsilon}\arcsin^2\left(\frac{x}{2}\sqrt{1-\mathrm{e}^{-\epsilon}}\right) - x^2\right]\right). \tag{28}$$

The expansion can be calculated manually up to the first few orders in ϵ, but soon becomes intractable. It is, however, a trivial task if we use some symbolic manipulation program such as Mathematica [25]. We shall give here explicit results for the first few moments:

$$\begin{aligned}
M_0^{(k)} &= 1\\
M_1^{(k)} &= k(k-1)/6\\
M_2^{(k)} &= k(k-1)(5k^2-k+12)/180\\
M_3^{(k)} &= k(k-1)(35k^4+14k^3+235k^2-188k+24)/7560\\
M_4^{(k)} &= k(k-1)(175k^6+315k^5+2341k^4-1959k^3+2056k^2-5664k-2160)/226\,800.
\end{aligned} \tag{29}$$

Analogously, the evaluation of the central moments $m_s^{(k)}$ is reduced to the explicit computation of the sth coefficient of the Taylor expansion around $\epsilon = 0^+$, of the function

$$\sqrt{1-\mathrm{e}^{-\epsilon}}\, \mathrm{e}^{\frac{\epsilon}{8}} \frac{1}{\sqrt{\epsilon}} \exp\left(-\frac{1}{2}\left[\frac{4}{\epsilon}\arcsin^2\left(\frac{x}{2}\sqrt{1-\mathrm{e}^{-\epsilon}}\right) - x^2\right]\right) \mathrm{e}^{\epsilon k(k-1)/6} \tag{30}$$

with the following results:

$$\begin{aligned}
m_0^{(k)} &= 1\\
m_1^{(k)} &= 0\\
m_2^{(k)} &= (k+3)k(k-1)/45\\
m_3^{(k)} &= (2k+3)(2k+1)k(k-1)/945\\
m_4^{(k)} &= k(k-1)(7k^4+37k^3+7k^2-108k-45)/4725.
\end{aligned} \tag{31}$$

By direct inspection of the formulae, it is possible to extract the large k behaviour of generic moments. Since $c_s(x)$ is a polynomial of degree $4s$ in x, $M_s^{(k)} \sim (2k+4s-1)!!/(2k-1)!! = O(k^{2s})$ as $k \to \infty$.[5] Following a similar line of thought, and taking into account some cancellations which may occur more or less severely, according to the parity of s, it is also possible to guess that $m_s^{(k)} = O\left(k^{[\frac{3}{2}s]_{i.p.}}\right)$ for asymptotically large k.

Let us further note that, on the infinite-dimensional hypercubic lattice, from $M_1^{(k)} \sim k^2/6$, the area of random loops increases on average as the square of their length. This should be contrasted with the two-dimensional lattice [12], where the average area increases of course as the square of the average linear extension of the loop, that is as the first power of the length itself.

[5] More precisely, it is easy to show that $M_s^{(k)} = (k^2/6)^s[1 + O(1/k)]$ for large k.

We shall conclude this section by considering coefficients $G_{k,l}$ for fixed values of l. We then have for each integer value l an infinite succession of integer numbers. From direct inspection of formula (17), for asymptotically large values of k we can write

$$G_{k,l} = G_{k,0}\frac{k^l}{l!}\left[1 + O\left(\frac{1}{k}\right)\right] \qquad l \text{ fixed.} \tag{32}$$

6. Asymptotic behaviour of distribution $P_k(l)$

Let us turn back to probability distribution $P_k(l)$, equation (22). We want to investigate its asymptotic behaviour as $k \to \infty$. The standard procedure is to rescale the discrete index l to a continuous variable t using the average and the standard deviation as computed in section 5 for generic k:

$$l \to \mu_k + \sigma_k t \qquad \begin{aligned} \mu_k &\equiv M_1^{(k)} = \frac{k(k-1)}{6} \\ \sigma_k^2 &\equiv m_2^{(k)} = \frac{(k+3)k(k-1)}{45}. \end{aligned} \tag{33}$$

It is now possible to define the asymptotic probability distribution (22) as

$$P(t) = \lim_{k\to\infty} \sigma_k P_k(\mu_k + \sigma_k t). \tag{34}$$

Unfortunately, the coefficients $G_{k,l}$ are defined in terms of alternate sums of very large numbers (about $\sqrt{2l}$ terms, each one of order $O(k^k)$, summing up to a much smaller number[6]), so that the evaluation of the asymptotic distribution $P(t)$ is not completely straightforward. Indeed, we are interested in the behaviour of the $G_{k,l}$ for large k and large l, $l \sim k^2/6$, and the simple estimate of equation (32) is now completely useless.

In principle, the asymptotic behaviour of coefficients $G_k(q)$ for large values of k is encoded in the behaviour of generating function $R(z, q)$, equation (20), in the vicinity of its closest (with respect to the origin) singularity in the variable z. The latter is situated at $z = z_c \equiv (1-q)/4$. But since we are interested in the behaviour of $G_{k,l}$ where l is also asymptotically large, but in a controlled way, see (33), we must inspect the generating function in the vicinity of its closest (again, with respect to the origin) singularity in variable q, that is near $q = 1$. But in this way z_c is collapsing towards the origin, conspiring to the construction of an essential singularity in $z = z_c = 0$ as $q \to 1^-$. To tackle this difficulty we shall not investigate the generating function $R(z, q)$, but we shall rather resort to the Laplace transform method, applied directly to probability distribution $P_k(l)$; let us first introduce the Laplace transform

$$\phi_k(s) = \sum_{l=0}^{k(k-1)/2} P_k(l)\, \mathrm{e}^{-s(\frac{l-\mu_k}{\sigma_k})} \tag{35}$$

of probability distribution $P_k(l)$, suitably shifted and rescaled according to equation (33). In the large k limit, $\phi_k(s)$ is expected to converge pointwise to the Laplace transform of distribution $P(t)$:

$$\int_{-\infty}^{+\infty} \mathrm{d}t\, P(t)\, \mathrm{e}^{-st}. \tag{36}$$

Since we may write

$$\phi_k(s) = \mathrm{e}^{s\frac{\mu_k}{\sigma_k}} \left\langle \mathrm{e}^{-s\frac{l}{\sigma_k}} \right\rangle = \mathrm{e}^{s\frac{\mu_k}{\sigma_k}} \frac{G_k\left(\mathrm{e}^{-\frac{s}{\sigma_k}}\right)}{G_k(1)} \tag{37}$$

[6] Just as an example, let us quote that $G_{100,1000}$ is expressed through formula (17) as an alternate sum of 45 numbers of order up to 10^{200}, being itself of order 10^{179}, i.e. 21 orders of magnitude smaller.

and $\sigma_k \sim k^{3/2}$ for large k, the asymptotic behaviour of $P_k(l)$ is ruled by the behaviour of $G_k(q)$ in the neighbourhood of unity. Indeed, after identification of ϵ with s/σ_k, we can make use of the machinery developed in section 5 and the appendix. We start by using the fact that, up to terms vanishing exponentially for small ϵ ($q = \mathrm{e}^{-\epsilon}$),

$$G_k(q) = \int_{-\infty}^{+\infty} \mathrm{d}x\, x^{2k} \mathcal{F}(x, \epsilon) \tag{38}$$

with $\mathcal{F}(x, \epsilon)$ given by equation (52). For $\epsilon = 0$ ($q = 1$), the evaluation of the integral reduces to a straightforward application of the saddle-point method: we have to sum up the contribution of the neighbourhoods of two stationary points situated in $x = x_\pm = \pm\sqrt{2k}$, respectively. We readily get

$$G_k(1) = \sqrt{2}(2k)^k\, \mathrm{e}^{-k} \left[1 + O\left(\frac{1}{k}\right)\right] \tag{39}$$

which is indeed the leading asymptotics for $(2k-1)!!$, as it should be. Now, for ϵ small but positive (we are therefore specializing to positive real values of s), for each of the two stationary points $x = x_\pm$, we just shift the integration variable x to $x_\pm + y$ and expand the integrand, including $\mathcal{F}(x, \epsilon)$, around $k = \infty$. Considering, e.g., the contribution coming from the positive stationary point x_+, and taking into account only terms up to $O(\log k)$, we get

$$I_+ \sim \frac{1}{\sqrt{2\pi}} \int \mathrm{d}y \exp\left(\left[k \log 2k - k - y^2 - \frac{s}{2}\sqrt{5k} - \sqrt{10}sy - 2s^2\right]\right) \qquad (s > 0). \tag{40}$$

Of course the situation in which $s < 0$ has to be computed, too, but the result turns out to be independent of the sign of s. Integrating and adding the analogous contribution coming from the region in the neighbourhood of x_-, we finally obtain

$$G_k\left(\mathrm{e}^{-\frac{s}{\sigma_k}}\right) = \sqrt{2}(2k)^k\, \mathrm{e}^{-k} \exp\left(\frac{s^2}{2} - \frac{s}{2}\sqrt{5k}\right)\left[1 + O\left(\frac{1}{\sqrt{k}}\right)\right] \tag{41}$$

and thus, for asymptotically large values of k

$$\phi_k(s) = \mathrm{e}^{\frac{s^2}{2}} \left[1 + O\left(\frac{1}{\sqrt{k}}\right)\right]. \tag{42}$$

Indeed, by comparison with equation (36), we have thus proved that $P(t)$ as defined in equation (34) is a standard Gaussian distribution:

$$P(t) = \frac{1}{\sqrt{2\pi}}\, \mathrm{e}^{-\frac{t^2}{2}}. \tag{43}$$

This result gives us further information about the asymptotic behaviour of the central moments $m_s^{(k)}$ of distribution $P_k(l)$: indeed, rescaling each central moment $m_s^{(k)}$ by a factor $\sigma_k^{-s} \sim k^{-\frac{3}{2}s}$ (correspondingly to the rescaling of the discrete variable l to the continuous one t), and performing the $k \to \infty$ limit, we must find the moments of a standard (unit variance) Gaussian distribution, namely 0 or $(s-1)!!$, for odd or even values of s, respectively. This on one hand confirms rigorously the estimate $m_s^{(k)} = \sigma_k^{-s} \cdot (s-1)!![1 + O(1/k)]$ (s even) and on the other hand matches the guessed behaviour $m_s^{(k)} \sim k^{[\frac{3}{2}s]_{i.p.}}$ (s odd).

As for the area versus length distribution, the result we have found, equation (43), does not agree with the behaviour proposed in [11, 12]; however, the role of $D \to \infty$ limit, inherent to our treatment, must be taken into proper account. In this limit, all loops with more than one plaquette of the subtended surface lying over the same plane are neglected; on the other hand, on the two-dimensional lattice considered in [11, 12] there is only one possible plane. Moreover, having averaged over orientations, we restrict ourselves to the absolute value of area, while in [11, 12], algebraic (oriented) area is considered.

When considering very compact loops, their distribution is ruled by the left tail of equation (43). In [10] the different but somehow related problem of the distribution $p(s,t)$ of random walks of t steps visiting exactly s different sites was addressed on a cubic lattice. In particular, an asymptotic formula in the regime of large t and relatively small s was proposed, which describes the distribution of compact walks. However, the two distributions are not comparable: indeed, in our situation of infinite dimensionality, the most compact loops of length $2k$ have zero enclosed area, but still visit a relatively large number of distinct sites, at least $k+1$; more precisely, when the dimensionality of the lattice is very large, the number of walks which take more than one step in the same direction is of order $1/D$ with respect to those stepping in k different directions, and therefore even the most compact walk, the one coming back to the origin every two steps, at leading order in D visits k different sites, in addition to the origin. In other words, the connection between the two notions of compactness, as small enclosed area, or as small number of visited sites, is lost in the limit of infinite dimensionality.

7. An infinite hierarchy of nontrivial identities

We would like to conclude with a simple consideration, which in our opinion is, however, rather striking and deserves further investigation. The exact evaluation of moments $M_s^{(k)}$ gives as a byproduct an infinite set of nontrivial and somehow intriguing identities, the simplest of them being

$$\sum_{l=0}^{k(k-1)/2} G_{k,l} = (2k-1)!! \qquad k = 0, 1, 2, \ldots \tag{44}$$

and in general

$$\sum_{l=0}^{k(k-1)/2} G_{k,l} l^s = M_s^{(k)} (2k-1)!! \qquad k = 0, 1, 2, \ldots \tag{45}$$

with $G_{k,l}$ and $M_s^{(k)}$ explicitly given by formulae (17) and (27), (28), respectively. These identities are of course encoded in the Jacobi theta function, essentially in its modular invariance properties, equation (48), but they come nevertheless rather unexpected.

Other sets of identities can be obtained if we are able to compute the $G_{k,l}$ in an independent and simple way. This is indeed very easy, e.g., $G_{k,k(k-1)/2}$, which counts the number of diagrams with $2k$ external points and maximal crossing number, and is thus obviously equal to 1 for any k (just connect the jth point with the $(j+k)$th one, $j = 1, \ldots, k$). From equation (17) we may therefore write

$$\sum_{i=0}^{k-1} (-1)^i \binom{k(k+1)/2 - 1 - i(i+1)/2}{k-1} \binom{2k+1}{k-i} \frac{2i+1}{2k+1} = 1 \qquad k = 0, 1, 2, \ldots. \tag{46}$$

Identities of this second sort are however more obvious: equation (46) is just the simplest example of a whole bunch of identities encoded in an obvious symmetry of polynomial $G_k(q)$: by applying the theorem of section 3 (equation (16)) to the new polynomial $\tilde{G}_k(q) = q^{k(k-1)/2} G_k(1/q)$, with $G_k(q)$ expressed in terms of equation (15), one obtains new coefficients $\tilde{G}_{k,l}$, which obviously satisfy $\tilde{G}_{k,l} = G_{k,(k(k-1)/2)-l}$; setting in particular $l = 0$, we readily get equation (46).

8. Conclusions

The problem of counting closed paths of given length on a lattice, according to the area they enclose, is a difficult one. In the present paper we have tackled the question on a hypercubic lattice, in the limit of infinite dimensionality. Moreover, an average over orientations is inherent to the method we have used. But these conditions have allowed us to give a closed answer, which is moreover exact under the restriction we have assumed. The main flaw of the whole method resides in our opinion in the fact that the computation of even only the first corrections in $1/D$ to coefficients $G_{k,l}$ is out of reach. On the other hand, we have presented an explicit solution to a combinatorial problem which could hardly be addressed directly; in this respect, the connection between the abstract combinatorial puzzle and the physical model proposed in [13] is crucial. The exact solution proposed for coefficients $G_{k,l}$ has been investigated in detail; their asymptotic behaviour has been evaluated; finally, as a byproduct, an infinite set of new nontrivial identities has been obtained.

Acknowledgments

We would like to thank Andrea Cappelli and Stefano Ruffo for fruitful discussions, Vladimir Rittenberg for suggesting to us the use of the Laplace transform method (section 6) and Giuseppe Santoro for informing us of [18] and its connection with our work.

Appendix

We want to build up a procedure to compute in a systematic way derivatives of arbitrary order of $G_k(q)$ with respect to $\log q$, at $q = 1$. From the explicit expression (14) of $G_k(q)$, we have to take into account the q dependence of both the integration interval and the integration measure $\nu_q(x)$, given by equation (12). As for the latter, the task is indeed not completely straightforward, since for fixed values of x, $\nu_q(x)$ develops an essential singularity in the limit $q \to 1$, and should be regularized. In the presence of an essential singularity, the result of the limiting procedure depends on the direction chosen to approach the singularity, and a regularization is a prescription on how the singularity is approached. Indeed, if we restrict q to the real axis and approach $q = 1$ from smaller values, it is possible to write the integration measure $\nu_q(x)$ $\left(x = \frac{2}{\sqrt{1-q}}\cos\theta\right)$ as a sum of a completely regular contribution, Taylor expandable in the real variable ϵ ($q = \mathrm{e}^{-\epsilon}$), plus a singular contribution vanishing with all its derivatives for $\epsilon \to 0^+$. Indeed, we shall show that

$$\nu_q(x) = \mathcal{F}(x,\epsilon)[1 + O(\mathrm{e}^{-1/\epsilon})] \tag{47}$$

$\mathcal{F}(x,\epsilon)$ being an analytic function of ϵ.

Using modular invariance, we can write

$$\Theta_1\left(\frac{\theta}{\pi}\middle|\tau\right) = \frac{\mathrm{i}}{\sqrt{-\mathrm{i}\tau}}\,\mathrm{e}^{-\mathrm{i}\theta^2/\pi\tau}\,\Theta_1\left(\frac{\theta}{\pi\tau}\middle| -\frac{1}{\tau}\right). \tag{48}$$

On the other hand, from definition (13), letting $\tau = \frac{\mathrm{i}\epsilon}{2\pi}$,

$$\Theta_1\left(\frac{\theta}{\pi\tau}\middle| -\frac{1}{\tau}\right) = \frac{1}{\mathrm{i}}\,\mathrm{e}^{-\frac{\pi^2}{2\epsilon}}\,\mathrm{e}^{\frac{2\pi\theta}{\epsilon}}\sum_{n\in\mathbb{Z}}(-1)^n \exp\left(-\frac{2\pi^2}{\epsilon}\left(n^2 + n - 2n\frac{\theta}{\pi}\right)\right). \tag{49}$$

For arbitrary small positive ϵ, and for $\frac{\theta}{\pi}$ in the open interval (0, 1), the leading contribution is given by the $n = 0$ term in the infinite sum, all other terms being exponentially small. We may therefore write, for $\tau = \frac{\mathrm{i}\epsilon}{2\pi}$, and $0 < \epsilon \ll 1$,

$$\Theta_1\left(\frac{\theta}{\pi\tau}\middle| -\frac{1}{\tau}\right) \sim \frac{1}{\mathrm{i}}\,\mathrm{e}^{-\frac{\pi^2}{2\epsilon}}\,\mathrm{e}^{\frac{2\pi\theta}{\epsilon}}[1 + O(\mathrm{e}^{-\frac{1}{\epsilon}})] \tag{50}$$

and for the measure $\nu_q(x)$ we readily get, in the limit $q \to 1^-$, or $q = \mathrm{e}^{-\epsilon}, \epsilon \to 0^+$, the approximate behaviour (47), with

$$\mathcal{F}(x,\epsilon) = \sqrt{1-\mathrm{e}^{-\epsilon}}\,\mathrm{e}^{\frac{\epsilon}{8}}\frac{1}{\sqrt{2\pi\epsilon}}\exp\left[-\frac{2}{\epsilon}\left(\theta - \frac{\pi}{2}\right)^2\right] \tag{51}$$

$$= \sqrt{1-\mathrm{e}^{-\epsilon}}\,\mathrm{e}^{\frac{\epsilon}{8}}\frac{1}{\sqrt{2\pi\epsilon}}\exp\left[-\frac{2}{\epsilon}\arcsin^2\left(\frac{x}{2}\sqrt{1-\mathrm{e}^{-\epsilon}}\right)\right] \tag{52}$$

$$= \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}\sum_{r=0}^{\infty}\frac{c_r(x)}{r!}\epsilon^r \tag{53}$$

where the coefficients $c_r(x)$ are given by the Taylor expansion in the neighbourhood of the origin $\epsilon = 0$ of the analytical function:

$$\sqrt{1-\mathrm{e}^{-\epsilon}}\,\mathrm{e}^{\frac{\epsilon}{8}}\frac{1}{\sqrt{\epsilon}}\exp\left(-\frac{1}{2}\left[\frac{4}{\epsilon}\arcsin^2\left(\frac{x}{2}\sqrt{1-\mathrm{e}^{-\epsilon}}\right) - x^2\right]\right). \tag{54}$$

$c_0(x) = 1$ trivially while higher coefficients $c_r(x)$ can be shown to be even polynomials of degree $4r$ in x. In the neighbourhood of $q = 1$ we may therefore write, up to terms vanishing exponentially for small ϵ,

$$G_k(q) = \frac{1}{\sqrt{2\pi}}\int_{-\frac{2}{\sqrt{1-q}}}^{\frac{2}{\sqrt{1-q}}}\mathrm{d}x\,x^{2k}\,\mathrm{e}^{-\frac{x^2}{2}}\sum_{r=0}^{\infty}\frac{c_r(x)}{r!}\epsilon^r. \tag{55}$$

Let us now observe that the integration limits diverge as $1/\sqrt{\epsilon}$, while the integrand falls as a quadratic exponential. Therefore, up to terms which vanish exponentially for small ϵ, the integration interval can be readily extended to the whole real axis, and the computation of moments (25) is finally reduced to the Taylor expansion of (54) and the evaluation of a sum of standard Gaussian integrals:

$$M_s^{(k)} = \frac{(-1)^s}{(2k-1)!!}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\mathrm{d}x\,\mathrm{e}^{-\frac{x^2}{2}}x^{2k}c_s(x). \tag{56}$$

As an example, let us compute explicitly the first moment, $M_1^{(k)}$. The first coefficient of the Taylor expansion of (54) is

$$c_1(x) = -\tfrac{1}{8} + \tfrac{1}{4}x^2 - \tfrac{1}{24}x^4 \tag{57}$$

and therefore, denoting by $\langle\langle\cdots\rangle\rangle$ the Gaussian average (with variance equal to 1), and exploiting the standard relation $\langle\langle x^{2k}\rangle\rangle = (2k-1)!!$

$$M_1(k) = \frac{1}{\langle\langle x^{2k}\rangle\rangle}\left[\frac{1}{8}\langle\langle x^{2k}\rangle\rangle - \frac{1}{4}\langle\langle x^{2k+2}\rangle\rangle + \frac{1}{24}\langle\langle x^{2k+4}\rangle\rangle\right] \tag{58}$$

$$= \frac{1}{8} - \frac{1}{4}(2k+1) + \frac{1}{24}(2k+3)(2k+1) \tag{59}$$

$$= \frac{k(k-1)}{6}. \tag{60}$$

References

[1] Harper P G 1955 *Proc. Phys. Soc.* A **68** 874
[2] Zak J 1964 *Phys. Rev.* **134** 1602
[3] Hofstadter D R 1976 *Phys. Rev.* B **14** 2239
[4] Abanov A G, Talstra J C and Wiegmann P B 1998 *Nucl. Phys.* B **525** 571
[5] Lipan O 2000 *J. Phys. A: Math. Gen.* **33** 6875
[6] Brereton M G and Butler C 1987 *J. Phys. A: Math. Gen.* **20** 3955
[7] Duplantier B 1989 *J. Phys. A: Math. Gen.* **22** 3033
[8] Minkov G M, Germanenko A V, Larionova V A, Negashev S A and Gornyi I V 2000 *Phys. Rev.* B **61** 13164
[9] Stanley H E, Kang K, Redner S and Blumberg R L 1983 *Phys. Rev. Lett.* **51** 1223
[10] Nieuwenhuizen Th M 1989 *Phys. Rev. Lett.* **62** 357
[11] Lévy P 1965 *Processus stochastiques et Mouvements Browniens* (Paris: Gauthier-Villars)
[12] Bellissard J, Camacho C J, Barelli A and Claro F 1997 *J. Phys. A: Math. Gen.* **30** L707
[13] Parisi G 1994 *J. Phys. A: Math. Gen.* **27** 7555
[14] Marinari E, Parisi G and Ritort F 1995 *J. Phys. A: Math. Gen.* **28** 4481
[15] Biedenharn L 1989 *J. Phys. A: Math. Gen.* **22** 4581
Macfarlane A 1989 *J. Phys. A: Math. Gen.* **22** L873
Chaichian M and Kulish P 1990 *Phys. Lett.* B **234** 72
[16] Wiegmann P B and Zabrodin A V 1994 *Nucl. Phys.* B **422** 495
[17] Brezin E, Itzykson C, Parisi G and Zuber J B 1978 *Commun. Math. Phys.* **59** 35
[18] Pauling L 1933 *J. Chem. Phys.* **1** 280
[19] Stanley R P 1997 *Enumerative Combinatorics* (Cambridge: Cambridge University Press)
[20] Cappelli A and Colomo F 1998 *J. Phys. A: Math. Gen.* **31** 3141
[21] van Leeuwen H and Maassen H 1995 *J. Math. Phys.* **36** 4743
Bozejko M, Kümmerer B and Speicher R 1997 *Commun. Math. Phys.* **185** 129
[22] Gasper G and Rahman M 1990 *Basic Hypergeometric Functions* (Cambridge: Cambridge University Press)
[23] Szëgo G 1926 *Sitz. Preuss. Akad. Wiss. Phys. Math.* **19** 242
[24] Mumford D 1983 *Tata Lectures on Theta* (Boston, MA: Birkhauser)
[25] Wolfram S 1991 *Mathematica* (New York: Addison-Wesley)