3

The Lace Expansion for the Self-Avoiding Walk

The lace expansion was derived by Brydges and Spencer in [45]. Their derivation, which is given below in Sects. 3.2–3.3, involves an expansion and resummation procedure closely related to the cluster expansions of statistical mechanics [40]. It was later noted that the lace expansion can also be seen as resulting from repeated application of the inclusion-exclusion relation [186]. For a more combinatorial description of the lace expansion, see [211]. We first discuss the inclusion-exclusion approach.

3.1 Inclusion-Exclusion

The inclusion-exclusion approach to the lace expansion is closely related to the method of proof of Theorem 2.3. In that proof, a single inclusion-exclusion was used to obtain upper and lower bounds. Here, we will derive an identity by using repeated inclusion-exclusion.

For simplicity, we restrict attention to the strictly self-avoiding walk ($\lambda = 1$). We consider a walk taking steps in a finite set Ω, so that $\omega(i+1)-\omega(i) \in \Omega$ for each i, but there is no need here for a symmetry assumption and Ω is an arbitrary finite set. As in (1.10), we write

$$D(x) = \frac{1}{|\Omega|} I[x \in \Omega]. \tag{3.1}$$

We rewrite $c_n(x)$ using the inclusion-exclusion relation. Namely, we first count all walks from 0 to x which are self-avoiding *after* the first step, and then subtract the contribution due to those which are not self-avoiding from the beginning, i.e., walks that return to the origin. Since $c_1(0, y) = 1$ for $y \in \Omega$, this gives

$$c_n(x) = (c_1 * c_{n-1})(x) - \sum_{y\in\Omega} \sum_{\omega^{(1)}\in\mathcal{S}_{n-1}(y,x)} I[0 \in \omega^{(1)}]. \tag{3.2}$$

Comparing with (1.5), it is the second term on the right hand side that makes the above equation interesting.

The inclusion-exclusion relation can now be applied to the last term of (3.2), as follows. Let s be the first (and only) time that $\omega^{(1)}(s) = 0$. Then for $y \in \Omega$,

$$\begin{aligned} &\sum_{\omega^{(1)} \in \mathcal{S}_{n-1}(y,x)} I[0 \in \omega^{(1)}] \\ &= \sum_{s=1}^{n-1} \sum_{\substack{\omega^{(2)} \in \mathcal{S}_s(y,0) \\ \omega^{(3)} \in \mathcal{S}_{n-1-s}(0,x)}} I[\omega^{(2)} \cap \omega^{(3)} = \{0\}] \qquad (3.3) \\ &= \sum_{s=1}^{n-1} \Big[c_s(y,0) c_{n-1-s}(0,x) - \sum_{\substack{\omega^{(2)} \in \mathcal{S}_s(y,0) \\ \omega^{(3)} \in \mathcal{S}_{n-1-s}(0,x)}} I[\omega^{(2)} \cap \omega^{(3)} \neq \{0\}] \Big]. \end{aligned}$$

We can interpret $c_s(y,0)$ as the number of $(s+1)$-step walks which step from the origin directly to y, then return to the origin in s steps, and which have distinct vertices apart from the fact that they return to their starting point. Let $\mathcal{U}_s$ denote the set of all s-step self-avoiding loops at the origin (s-step walks which begin and end at the origin but which otherwise have distinct vertices), and let u_s be the cardinality of $\mathcal{U}_s$. Then

$$\begin{aligned} &\sum_{y \in \Omega} \sum_{\omega^{(1)} \in \mathcal{S}_{n-1}(y,x)} I[0 \in \omega^{(1)}] \\ &= \sum_{s=2}^{n} u_s c_{n-s}(x) - \sum_{s=2}^{n} \sum_{\substack{\omega^{(2)} \in \mathcal{U}_s \\ \omega^{(3)} \in \mathcal{S}_{n-s}(0,x)}} I[\omega^{(2)} \cap \omega^{(3)} \neq \{0\}]. \qquad (3.4) \end{aligned}$$

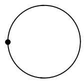
Continuing in this fashion, in the last term on the right hand side of the above equation, let $t \geq 1$ be the first time along $\omega^{(3)}$ that $\omega^{(3)}(t) \in \omega^{(2)}$, and let $v = \omega^{(3)}(t)$. Then the inclusion-exclusion relation can be applied again to remove the avoidance between the portions of $\omega^{(3)}$ before and after t, and correct for this removal by the subtraction of a term involving a further intersection. Repetition of this procedure leads to the convolution equation

$$c_n(0,x) = (|\Omega| D * c_{n-1})(x) + \sum_{m=2}^{n} (\pi_m * c_{n-m})(x), \qquad (3.5)$$

where we have used $c_1(x) = |\Omega| D(x)$, and where π_m is given by

$$\pi_m(v) = \sum_{N=1}^{\infty} (-1)^N \pi_m^{(N)}(v), \qquad (3.6)$$

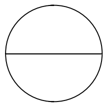
with the terms on the right hand side defined as follows. The $N = 1$ term is given by

$$\pi_m^{(1)}(v) = \delta_{0,v} u_m = \delta_{0,v} \; 0$$  ,

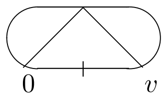
where the diagram represents u_m. The $N = 2$ term is

$$\pi_m^{(2)}(v) = \sum_{\substack{m_1, m_2, m_3 : \\ m_1 + m_2 + m_3 = m}} \sum_{\omega_1 \in \mathcal{S}_{m_1}(0,v)} \sum_{\omega_2 \in \mathcal{S}_{m_2}(v,0)} \sum_{\omega_3 \in \mathcal{S}_{m_3}(0,v)} I(\omega_1, \omega_2, \omega_3),$$

where $I(\omega_1, \omega_2, \omega_3)$ is equal to 1 if the ω_i are pairwise mutually avoiding apart from their common endpoints, and otherwise equals 0. Diagrammatically this can be represented by

$$\pi_m^{(2)}(v) = \; 0$$  v ,

where each line represents a sum over self-avoiding walks between the endpoints of the line, with mutual avoidance between the three pairs of lines in the diagram. Similarly

$$\pi_m^{(3)}(v) \quad =$$  ,

0 v

where now there is mutual avoidance between some but not all pairs of lines in the diagram; a precise description requires some care. The unlabelled vertex is summed over $\mathbb{Z}^d$. A slashed diagram line is used to indicate a walk which may have zero steps, i.e., be a single site, whereas lines without a slash correspond to walks of at least one step. All the higher order terms can be expressed as diagrams in this way, and with some care it is possible to keep track of the pattern of mutual avoidance between subwalks (individual lines in the diagram) which emerges. The algebraic derivation of the expansion, described next, keeps track of this mutual avoidance automatically. Equations (3.5)–(3.6) constitute the lace expansion. No laces have appeared yet, but they will come later.

Exercise 3.1. Determine a precise expression for $\pi_m^{(3)}(v)$. What is the picture for $\pi_m^{(4)}(v)$?

3.2 Expansion

In this and the following section, we give the original derivation of the lace expansion due to Brydges and Spencer [45]. The expansion applies in a more general context than we have considered so far, and we will give a quite general derivation.

Consider walks taking steps in a finite subset $\Omega \subset \mathbb{Z}^d$. Suppose that to each walk $\omega = (\omega(0), \omega(1), \dots, \omega(n))$ and each pair $s, t \in \{0, 1, \dots, n\}$, we are given a complex number $\mathcal{U}_{st}(\omega)$ (for example, (2.1)).

Definition 3.2. (i) *Given an interval $I = [a, b]$ of positive integers, we refer to a pair $\{s, t\}$ ($s < t$) of elements of I as an* edge. *To abbreviate the notation, we usually write st for $\{s, t\}$. A set of edges is called a* graph. *The set of all graphs on $[a, b]$ is denoted $\mathcal{B}[a, b]$.*
(ii) *A graph Γ is said to be* connected *if both a and b are endpoints of edges in Γ, and if in addition, for any $c \in (a, b)$, there are $s, t \in [a, b]$ such that $s < c < t$ and $st \in \Gamma$. In other words, Γ is connected if, as intervals, $\cup_{st \in \Gamma}(s, t) = (a, b)$. The set of all connected graphs on $[a, b]$ is denoted $\mathcal{G}[a, b]$.*

An apology is required for graph theorists. The above notion of connectivity is not the usual notion of path-connectivity in graph theory. Instead, the above notion relies heavily on the fact that the vertices of the graph are linearly ordered in time, and may be justified by the fact that connected graphs are those for which $\cup_{st \in \Gamma}(s, t)$ is equal to the connected interval (a, b). In any event, it is decidedly not path-connectivity. There are connected graphs that are not path-connected, and vice versa. It is convenient to have in mind the representation of graphs illustrated in Fig. 3.1.

We set $K[a, a] = 1$, and for $a < b$ we define

$$K[a, b] = \prod_{a \le s < t \le b} (1 + \mathcal{U}_{st}), \tag{3.7}$$

where the dependence on ω is left implicit. By expanding the product in (3.7), we obtain

$$K[a, b] = \sum_{\Gamma \in \mathcal{B}[a,b]} \prod_{st \in \Gamma} \mathcal{U}_{st}. \tag{3.8}$$

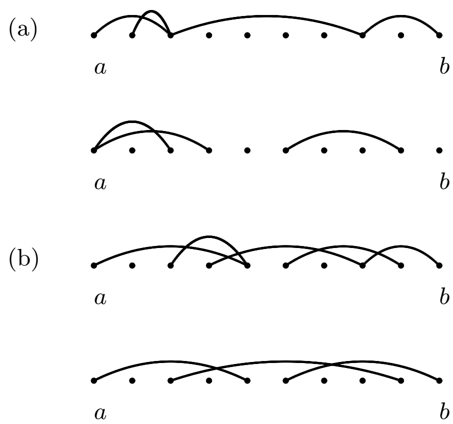


Fig. 3.1. Graphs in which an edge st is represented by an arc joining s and t. The graphs in (a) are not connected, whereas the graphs in (b) are connected.

Note that $\mathcal{B}[a,b]$ contains the graph with no edges, so our convention that $K[a,a]=1$ is consistent with the standard convention that an empty product is equal to 1.

Exercise 3.3. Prove (3.8).

We set $J[a,a]=1$, and for $a<b$ we define a quantity analogous to $K[a,b]$, but with the sum over graphs restricted to connected graphs:

$$J[a,b] = \sum_{\Gamma \in \mathcal{G}[a,b]} \prod_{st \in \Gamma} \mathcal{U}_{st}. \tag{3.9}$$

Lemma 3.4. *For any $a<b$,*

$$K[a,b] = K[a+1,b] + \sum_{j=a+1}^{b} J[a,j]K[j,b]. \tag{3.10}$$

Proof. The contribution to the sum on the right hand side of (3.8) due to all graphs Γ for which a is not in an edge is exactly $K[a+1,b]$. To resum the contribution due to the remaining graphs, we proceed as follows. If Γ does contain an edge containing a, let $j(\Gamma)$ be the largest value of j such that the set of edges in Γ with both ends in the interval $[a,j]$ forms a connected graph on $[a,j]$. Then the sum over Γ factorizes into sums over connected graphs on $[a,j]$ and arbitrary graphs on $[j,b]$, and resummation of the latter gives

$$K[a,b] = K[a+1,b] + \sum_{j=a+1}^{b} \sum_{\Gamma \in \mathcal{G}[a,j]} \prod_{st \in \Gamma} \mathcal{U}_{st}\, K[j,b], \tag{3.11}$$

which with (3.9) proves the lemma. ∎

Let

$$c_n(x) = \sum_{\omega \in \mathcal{W}_n(x)} K[0,n] = \sum_{\omega \in \mathcal{W}_n(x)} \prod_{0 \le s < t \le n} (1 + \mathcal{U}_{st}(\omega)), \tag{3.12}$$

a generalization of (2.2). It is simplest if we assume that $\mathcal{U}_{st}(\omega)$ is invariant under spatial translation of ω, and under an equal shift of each of s,t and the time parameter of ω, and we make this assumption. Note that (2.1) obeys the assumption. We substitute (3.10) into (3.12). A key point is that in the last term of (3.10) the portion of the walk from time j onwards is independent of the portion up to time j. Let

$$\pi_m(x) = \sum_{\omega \in \mathcal{W}_m(0,x)} J[0,m]. \tag{3.13}$$

Then for $n \ge 1$, we obtain

$$c_n(x) = (|\Omega| D * c_{n-1})(x) + \sum_{m=1}^{n} (\pi_m * c_{n-m})(x), \tag{3.14}$$

as in (3.5).[1] To obtain a more useful representation of π_m than (3.13), we perform a resummation of (3.13) using the notion of laces.

3.3 Laces and Resummation

Definition 3.5. *A* lace *is a minimally connected graph, i.e., a connected graph for which the removal of any edge would result in a disconnected graph. The set of laces on $[a,b]$ is denoted by $\mathcal{L}[a,b]$, and the set of laces on $[a,b]$ which consist of exactly N edges is denoted $\mathcal{L}^{(N)}[a,b]$.*

We write $L \in \mathcal{L}^{(N)}[a,b]$ as $L = \{s_1t_1, \ldots, s_Nt_N\}$, with $s_l < t_l$ for each l. The fact that L is a lace is equivalent to a certain ordering of the s_l and t_l. For $N = 1$, we simply have $a = s_1 < t_1 = b$. For $N \geq 2$, $L \in \mathcal{L}^{(N)}[a,b]$ if and only if

$$a = s_1 < s_2, \quad s_{l+1} < t_l \leq s_{l+2} \quad (l = 1, \ldots, N-2), \quad s_N < t_{N-1} < t_N = b \tag{3.15}$$

(for $N = 2$ the vacuous middle inequalities play no role); see Fig. 3.2. Thus L divides $[a,b]$ into $2N-1$ subintervals:

$$[s_1, s_2],\ [s_2, t_1],\ [t_1, s_3],\ [s_3, t_2],\ \ldots\ , [s_N, t_{N-1}],\ [t_{N-1}, t_N]. \tag{3.16}$$

Of these, intervals number 3, 5, ..., $(2N-3)$ can have zero length for $N \geq 3$, whereas all others have length at least 1.

Exercise 3.6. Prove that (3.15) characterizes laces.

Given a connected graph $\Gamma \in \mathcal{G}[a,b]$, the following prescription associates to Γ a unique lace $\mathsf{L}_\Gamma \subset \Gamma$: The lace L_Γ consists of edges $s_1t_1, s_2t_2, \ldots$, with $t_1, s_1, t_2, s_2, \ldots$ determined, in that order, by

$$t_1 = \max\{t : at \in \Gamma\}, \quad s_1 = a,$$

$$t_{i+1} = \max\{t : \exists s < t_i \text{ such that } st \in \Gamma\}, \quad s_{i+1} = \min\{s : st_{i+1} \in \Gamma\}.$$

The procedure terminates when $t_{i+1} = b$. Given a lace L, the set of all edges $st \notin L$ such that $\mathsf{L}_{L \cup \{st\}} = L$ is denoted $\mathcal{C}(L)$. Edges in $\mathcal{C}(L)$ are said to be *compatible* with L. Fig. 3.3 illustrates these definitions.

Exercise 3.7. Show that $\mathsf{L}_\Gamma = L$ if and only if L is a lace, $L \subset \Gamma$, and $\Gamma \setminus L \subset \mathcal{C}(L)$.

[1] For $m = 1$, there is a single connected graph $\{01\}$, and when $\mathcal{U}_{st}$ is given by (2.1) we have $\pi_1(x) = \sum_{\omega \in \mathcal{W}_1(0,x)} U_{01}(\omega) = 0$, since it is always the case that $\omega(0) \neq \omega(1)$. Thus the sum over m in (3.14) can be started at $m = 2$ in this case.

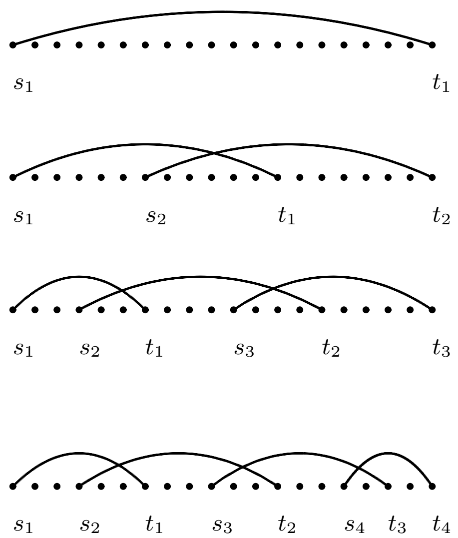


Fig. 3.2. Laces in $\mathcal{L}^{(N)}[a,b]$ for $N = 1, 2, 3, 4$, with $s_1 = a$ and $t_N = b$.

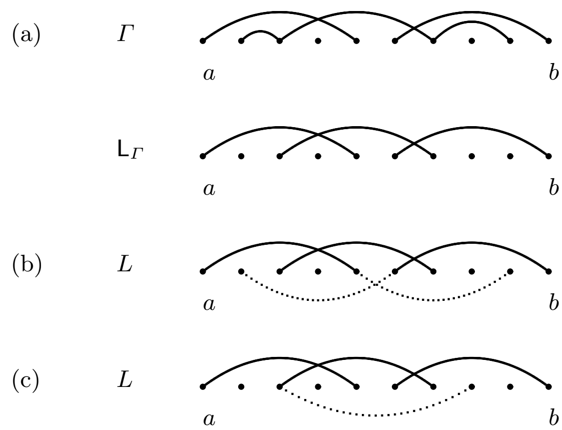


Fig. 3.3. (a) A connected graph Γ and its associated lace $L = \mathsf{L}_\Gamma$. (b) The dotted edges are compatible with the lace L. (c) The dotted edge is not compatible with the lace L.

The sum over connected graphs in (3.9) can be performed by first summing over all laces and then, given a lace, summing over all connected graphs associated to that lace by the above prescription. This gives

$$J[a,b] = \sum_{L \in \mathcal{L}[a,b]} \prod_{st \in L} \mathcal{U}_{st} \sum_{\Gamma : \mathsf{L}_\Gamma = L} \prod_{s't' \in \Gamma \setminus L} \mathcal{U}_{s't'} . \tag{3.17}$$

But, writing $\Gamma' = \Gamma \setminus L$, it follows from Exercise 3.7 that

$$\sum_{\Gamma : \mathsf{L}_\Gamma = L} \prod_{s't' \in \Gamma \setminus L} \mathcal{U}_{s't'} = \sum_{\Gamma' \subset \mathcal{C}(L)} \prod_{s't' \in \Gamma'} \mathcal{U}_{s't'} = \prod_{s't' \in \mathcal{C}(L)} (1 + \mathcal{U}_{s't'}). \tag{3.18}$$

Therefore,

$$J[a,b] = \sum_{L \in \mathcal{L}[a,b]} \prod_{st \in L} \mathcal{U}_{st} \prod_{s't' \in \mathcal{C}(L)} (1 + \mathcal{U}_{s't'}). \tag{3.19}$$

Inserting this in (3.13) gives

$$\pi_m(x) = \sum_{\omega \in \mathcal{W}_m(0,x)} \sum_{L \in \mathcal{L}[0,m]} \prod_{st \in L} \mathcal{U}_{st} \prod_{s't' \in \mathcal{C}(L)} (1 + \mathcal{U}_{s't'}). \tag{3.20}$$

For $a < b$ we define $J^{(N)}[a,b]$ to be the contribution to (3.17) from laces consisting of exactly N bonds:

$$J^{(N)}[a,b] = \sum_{L \in \mathcal{L}^{(N)}[a,b]} \prod_{st \in L} \mathcal{U}_{st} \prod_{s't' \in \mathcal{C}(L)} (1 + \mathcal{U}_{s't'}). \tag{3.21}$$

For the special case in which $\mathcal{U}_{st}$ is given by (2.1), each term in the above sum is either 0 or $(-1)^N$. By (3.17) and (3.21),

$$J[a,b] = \sum_{N=1}^{\infty} J^{(N)}[a,b]. \tag{3.22}$$

The sum over N in (3.22) is a finite sum, since the sum in (3.21) is empty for $N > b - a$ and hence $J^{(N)}[a,b] = 0$ if $N > b - a$.

Now we define

$$\begin{aligned} \pi_m^{(N)}(x) &= (-1)^N \sum_{\omega \in \mathcal{W}_m(x)} J^{(N)}[0,m] \\ &= \sum_{\omega \in \mathcal{W}_m(x)} \sum_{L \in \mathcal{L}^{(N)}[0,m]} \prod_{st \in L} (-\mathcal{U}_{st}) \prod_{s't' \in \mathcal{C}(L)} (1 + \mathcal{U}_{s't'}). \end{aligned} \tag{3.23}$$

The factor $(-1)^N$ on the right hand side of (3.23) has been inserted to arrange that

$$\pi_m^{(N)}(x) \geq 0 \quad \text{for all } N, m, x \tag{3.24}$$

when $\mathcal{U}_{st}$ is given by U_{st} of (2.1). By (3.13), (3.22) and (3.23),

$$\pi_m(x) = \sum_{N=1}^{\infty} (-1)^N \pi_m^{(N)}(x). \tag{3.25}$$

For the special case in which $\mathcal{U}_{st}$ is given by (2.1), walks making a nonzero contribution to (3.23) are constrained to have the topology indicated in Fig. 3.4. In the figure, for $\prod_{s't' \in \mathcal{C}(L)}(1+U_{s't'}) \neq 0$, each of the $2N-1$ subwalks

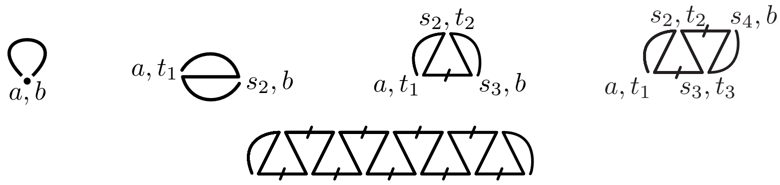


Fig. 3.4. Self-intersections required for a walk ω with $\prod_{st\in L} U_{st}(\omega) \neq 0$, with U_{st} given by (2.1), for the laces with $N = 1, 2, 3, 4$ bonds depicted in Fig. 3.2. The picture for $N = 11$ is also shown.

must be a self-avoiding walk, and in addition there must be mutual avoidance between some (but not all) of the subwalks. The number of loops (faces excluding the "outside" face) in a diagram is equal to the number of edges in the corresponding lace. The lines which are slashed correspond to subwalks which may consist of zero steps, but the others correspond to subwalks consisting of at least one step. This gives an interpretation of $\pi_m^{(N)}$ identical to that obtained in Sect. 3.1, but here there is the advantage that explicit formulas keep track of the mutual avoidance between subwalks.

It is sometimes convenient to modify the definitions of "connected graph" and "lace," and we will do so in Sect. 8.1. A more general theory of laces is developed and applied in [124, 126], for the analysis of networks of mutually-avoiding self-avoiding walks. See also [125] for an application of the more general theory to lattice trees.

3.4 Transformations

Equation (3.14) involves convolution in both space and time. It has been studied in this form in [29], via fixed point methods.

It is tempting to use transformations to eliminate one or both of these convolutions. We can eliminate the convolution in space if we take the Fourier transform (1.6). For $n \geq 1$, this gives

$$\hat{c}_n(k) = |\Omega| \hat{D}(k) \hat{c}_{n-1}(k) + \sum_{m=1}^{n} \hat{\pi}_m(k) \hat{c}_{n-m}(k). \tag{3.26}$$

Conditions are given in [120] which ensure that solutions of (3.26) have Gaussian asymptotics, via an analysis based on induction on n.

We may instead prefer to eliminate the convolution in time, by going to generating functions. Using (2.18) and (3.14), this gives

$$\begin{aligned} G_z(x) &= \delta_{0,x} + \sum_{n=1}^{\infty} c_n(x) z^n \\ &= \delta_{0,x} + z|\Omega|(D * G_z)(x) + (\Pi_z * G_z)(x), \end{aligned} \tag{3.27}$$

where

$$\Pi_z(x) = \sum_{m=1}^{\infty} \pi_m(x) z^m. \tag{3.28}$$

Equation (3.27) has been studied in [90, 91].

Finally, we may prefer to eliminate both convolutions by using both the Fourier transform and generating functions. Taking the Fourier transform of (3.27) gives

$$\hat{G}_z(k) = 1 + z|\Omega|\hat{D}(k)\hat{G}_z(k) + \hat{\Pi}_z(k)\hat{G}_z(k), \tag{3.29}$$

which can be solved to give

$$\hat{G}_z(k) = \frac{1}{1 - z|\Omega|\hat{D}(k) - \hat{\Pi}_z(k)}. \tag{3.30}$$

Equation (3.30) has been the point of departure for several studies of the self-avoiding walk, and we will work with (3.30) in Chap. 5.

Exercise 3.8. The memory-2 walk is the walk with $\mathcal{U}_{st} = U_{st}$ if $t - s \leq 2$, and otherwise $\mathcal{U}_{st} = 0$. This is a random walk with no immediate reversals. Suppose that $0 \notin \Omega \subset \mathbb{Z}^d$ is finite and invariant under the symmetries of the lattice.

(a) What is the value of $\hat{c}_n(0)$, the number of n-step memory-2 walks? (Calculation is not required.)

(b) Prove that for the memory-2 walk, for $m \geq 2$,

$$\pi_m(x) = \begin{cases} -|\Omega|\delta_{x,0} & \text{if } m \text{ is even} \\ I[x \in \Omega] & \text{if } m \text{ is odd.} \end{cases}$$

(c) Suppose that $|\Omega| > 2$. Show that the mean-square displacement for the memory-two walk is given by

$$\sigma^2 \left[\left(\frac{1+\delta}{1-\delta} \right) n - \frac{2\delta(1-\delta^n)}{(1-\delta)^2} \right] \sim \left(\frac{\sigma^2 |\Omega|}{|\Omega| - 2} \right) n,$$

where $\sigma^2 = \sum_x |x|^2 D(x)$ is the variance of D and $\delta = (|\Omega| - 1)^{-1}$. One approach[2] is to use (3.26) to compute $\nabla^2 \hat{c}_n(0)$. This problem goes back a long way [18, 63, 72].

(d) Show that for the memory-two walk,

$$\hat{G}_z(k) = \frac{1 - z^2}{1 + (|\Omega| - 1)z^2 - z|\Omega|\hat{D}(k)}$$

[2] Verification of the formula by induction seems an unsatisfactory solution, since it requires prior knowledge of the formula.

(compare Exercise 2.2 for $d = 1$). This formula was used to compute the mean-square displacement via contour integration in [158, Sect. 5.3].

The memory-τ walk is the walk with $\mathcal{U}_{st} = U_{st}$ if $t - s \leq \tau$, and otherwise $\mathcal{U}_{st} = 0$. Finite-memory walks played an important role in the original analysis of the lace expansion in [45], but will not concern us further here. For a study of generating functions of the number of memory-τ walks, for $\tau \leq 8$, see [171].