Acta Math., 185 (2000), 239–286

The asymptotic determinant of the discrete Laplacian

by

RICHARD KENYON

Universit 'ede Paris-Sud
Orsay, France

Contents

1. Introduction

The determinant of the Laplacian on a graph arises in two related statistical mechanical models, the uniform spanning tree model and the 2-dimensional lattice dimer model. For the spanning tree model, Kirchhoff [Ki] is attributed with showing that the product of the non-zero eigenvalues of the Laplacian on a finite graph is the same as the number of spanning trees on that graph. For the 2-dimensional dimer model, Temperley [T], based on work of Kasteleyn [Kas1], showed that the number of dimer coverings of certain subgraphs of $\mathbf{Z}^2$ can be computed by the determinant of the Laplacian on related graphs. Precise estimates on these determinants provide important information about these models, in particular allowing one to compute certain critical exponents and correlation functions [DD], [Ken1].

In this paper we compute the asymptotic expansion of the determinant of the Laplacian on a special family of graphs: subgraphs of $\mathbf{Z}^2$ which are approximating rectilinear polygons (a polygon is *rectilinear* if its sides are parallel to the axes). Our main motivation is not to study the Laplacian in itself but rather to study both the dimer (domino tiling) model and the uniform spanning tree model. For this reason we use the language of domino tilings. (Domino tilings are tilings with (1×2)- and (2×1)-rectangles.)

Temperley [T] gave a bijection between the number of spanning trees of a subgraph H of $\mathbf{Z}^2$ and domino tilings of a polyomino $P=P(H)$ constructed from the superposition of H and its dual. We call a polyomino *Temperleyan* if it arises from a graph H by Temperley's construction. Such polyominos have a simple description: let Q be a polyomino such that each side between a concave and convex corner has even length, and each side between two concave or two convex corners has odd length. Such a polyomino has odd area; let P be obtained from Q by removing one lattice square at some convex corner. Then P is Temperleyan.

THEOREM 1. *Let $U\subset\mathbf{R}^2$ be a rectilinear polygon with V vertices. For each $\varepsilon>0$, let P_ε be a Temperleyan polyomino in $\varepsilon\mathbf{Z}^2$ approximating U in the natural sense (the corners of P_ε are converging to the corners of U). Let A_ε be the area and* $\mathrm{Perim}_\varepsilon$ *be the perimeter of P_ε. Then the* log *of the number of domino tilings of P_ε is*

$$\frac{c_0 A_\varepsilon}{\varepsilon^2}+\frac{c_1\,\mathrm{Perim}_\varepsilon}{\varepsilon}-\frac{\pi}{48}\left(c_2(\varepsilon)\log\frac{1}{\varepsilon}+c_3(U)\right)+c_4+o(1), \tag{1}$$

where $c_0=G/\pi$, $G=1-1/3^2+1/5^2-\dots$ is Catalan's constant, $c_1=G/2\pi+\frac14\log(\sqrt{2}-1)$, c_4 is a constant independent of U, and $c_2(\varepsilon)\log(1/\varepsilon)+c_3(U)$ is the ε-normalized Dirichlet energy of the limiting average height function on U (see definitions below). The term

$-\frac{1}{48}\pi c_2(\varepsilon)$ *is of the form*

$$-\frac{1}{2}-\frac{V-4}{36}(1+\mathrm{ERR}(\varepsilon))$$

where $\mathrm{ERR}(\varepsilon)$ *is* $o(1)$.

COROLLARY 2. *Under the above hypotheses on* U, *for each* $\varepsilon>0$ *let* H_ε *be the subgraph of* $\varepsilon\mathbf{Z}^2$ *whose vertices are in* U. *Let* $N(H_\varepsilon)$ *be the number of vertices in* H_ε *and* $B(H_\varepsilon)$ *the number of edges of* $\varepsilon\mathbf{Z}^2$ *on the boundary of* H_ε. *Then the* log *of the determinant of the Laplacian on* H_ε *is*

$$\frac{4G}{\pi}N(H_\varepsilon)+\frac{\log(\sqrt{2}-1)}{2}B(H_\varepsilon)-\frac{\pi}{48}\left(c_2(\varepsilon)\log\frac{1}{\varepsilon}+c_3(U)\right)+c_5+o(1),$$

where c_2, c_3 *are as in Theorem* 1 *and* c_5 *is another constant independent of* U.

The limiting average height function has the following description. Let $b_0\in\partial U$ be a base point. For $x\in\partial U$ define $u_0(x)$ to be the total turning (in radians) of the boundary tangent on the boundary path counterclockwise from b_0 to x (the function u_0 has jump discontinuities at each corner of U and at b_0). The limiting average height function is the harmonic function on U whose boundary values are $2u_0/\pi$. The ε-normalized Dirichlet energy is by definition the Dirichlet energy contained in the complement of the ε-neighborhoods of the jump discontinuities.

Remarks. (1) In the case U is a rectangle, formula (1) follows from the formula for the exact number of tilings computed by Kasteleyn [Kas1] and Temperley and Fisher [TF] (the asymptotic expansion of which was computed by Duplantier and David [DD]); see Proposition 16 below.

(2) The leading term in the above formula, involving the constant c_0, essentially follows from work of Burton and Pemantle [BP]: they constructed a measure μ of entropy c_0 on the space X of domino tilings of the plane and proved that it was the unique translation-invariant measure of maximal entropy on X. Furthermore they proved that for regions of the type used in the theorem, the entropy (the coefficient of ε^{-2} in (1)) is c_0A.

(3) Our boundary conditions give rise to a correction to the number of tilings which is only exponential in the length of the boundary (the 'perimeter' term in the theorem). In [CKP], on the contrary, it was shown that in some sense "most" other boundary conditions have a larger effect, giving a smaller entropy c_0, and making the local densities of configurations vary throughout the region. So both the 'area' and 'perimeter' terms in the theorem depend strongly on our choice of boundary conditions.

(4) Note that if two regions have the same area, perimeter and number of vertices then the log of the number of domino tilings differs by a constant in the limit, that is, the

ratio of the number of tilings is tending to a constant as $\varepsilon \to 0$. This constant depends on the shape of the regions and can in principle be computed explicitly.

(5) There has been much work done on the 'regularized' determinant of the *continuous* Laplacian, and the asymptotic distribution of its eigenvalues [Kac], [MS], [OPS]. We have not attempted here to make any connection between these two subjects, although there is a lot of evidence for a connection, see e.g. [OPS], [DD].

(6) As noted above, Temperley [T] gave a bijection between the set of spanning trees of a subgraph of $\mathbf{Z}^2$ and the set of domino tilings of a related polyomino. The corollary follows from the theorem by applying this bijection (see §2 for the definition): the graph H_ε gives rise to a polyomino P_ε of area $A_\varepsilon = \varepsilon^2(4N(H_\varepsilon) - B(H_\varepsilon) - 4)$ and perimeter $\mathrm{Perim}_\varepsilon = \varepsilon(2B(H_\varepsilon)+4)$. Plugging these values into (1) gives the formula in the corollary.

(7) The function $\mathrm{ERR}(\varepsilon)$ is unknown although it seems possible that it could be computed using Toeplitz determinants, as in [MW].

Part of the motivation for proving Theorem 1 is to validate a certain heuristic, which attempts to explain how the presence of the boundary affects the long-range structure of a random tiling. In particular, it attempts to explain how the boundary affects the densities of local configurations far from the boundary [DMB]. We call this heuristic the 'phason strain' principle.

The heuristic is as follows: the boundary causes the *average height function* of a tiling (see definition in §2.3) to deviate slightly from its entropy-maximizing value of 0. At a point in the region where the average height function has non-zero slope, the "local" entropy there is smaller than the maximal possible entropy, by an amount proportional to the square of the gradient of the average height function. The system behaves in such a way as to maximize the total entropy subject to the given boundary values of the height function, and the resulting average height function is the function which minimizes (the integral of) the square of its gradient. That is, the average height function is harmonic. This "explains" the terms $c_2(\varepsilon)\log(1/\varepsilon) + c_3(U)$ in Theorem 1.

Unfortunately the constant $\frac{1}{48}\pi$ appearing in Theorem 1 is different from the expected value of the local entropy as derived in [CKP] (where a rigorous version of the phason strain principle is proved in a different context): in [CKP] the entropy as a function of slope is shown to have the expansion

$$\mathrm{ent}(s,t) = \mathrm{ent}(0,0) - \tfrac{1}{16}\pi(s^2+t^2) + O(\text{terms of order} \geqslant 3),$$

and where (s,t) are the partial derivatives of the height function. We may conclude from this discrepancy that the computation in [CKP] can not be refined to obtain asymptotics of the same precision as Theorem 1 above (when applied to the present case, [CKP] only gives the leading term in (1)). In particular, the phason strain principle can not be considered valid in this context.

The techniques used to prove Theorem 1 can be applied to the uniform spanning tree model as well. Indeed, as mentioned above, there is a close connection between the spanning tree process on $\mathbf{Z}^2$ and the domino tiling model [T], [BP], [KPW]. Many properties of spanning trees on $\mathbf{Z}^2$ translate into computable properties of dominos. We study here one particular property of a uniform spanning tree: the distribution of the (unique) arc between two fixed points. The relevant question about tilings is to count the number of tilings of a region with a hole (single square removed). To estimate this number, we use the technique of Theorem 1: we cut the region apart up to the hole, and then sew it up again in such a way as to remove the hole. In this way we prove the well-known conjecture that the expected number of points on the tree branch within distance N from the origin grows like $N^{5/4}$. In fact we prove more:

THEOREM 3. *On the uniform spanning tree process on* $\mathbf{N}\times\mathbf{Z}$, *the expected number of vertices on the branch from* $(0,0)$ *to* ∞ *which lie within distance* N *of the origin is* $N^{5/4+o(1)}$. *For* $x>0$ *the probability of a vertex* $(x,y)=re^{i\theta}$ *to be on the branch from* $(0,0)$ *to* ∞ *is*

$$r^{-3/4(1+f(r))}\cos(\theta)^{1/4}(1+o(1))$$

where $f(r)$ *is* $o(1)$ *as* $r\to\infty$.

The branch in a uniform spanning tree has the same distribution as the loop-erased random walk (LERW), see [P]. So this proves that the growth exponent of the loop-erased random walk is $\frac{5}{4}$. This value of the exponent d has been conjectured by physicists for some time [GB], [Ma], using arguments based on conformal field theory and the assumption of conformal invariance of the "scaling limit" of the walk. Lawler [L] had previously given the bounds $1<d\leqslant\frac{4}{3}$.

Here is an outline of the paper. §2 gives the definitions and background. Most of the background comes from [Ken1] and [Ken2]: local properties of dominos can be found in [Ken1], and the conformal properties can be found in [Ken2]. In §3, we state two lemmas and use them to prove Theorem 1. Specifically, the theorem is proved by cutting up a rectilinear polygon into rectangles and using the known formula for the number of tilings of a rectangle. Lemma 6 determines how the number of tilings changes as you are making a single cut, and Lemma 7 relates this change to the change in Dirichlet energy of the average height function. The next two sections are devoted to the proofs of the lemmas. §6 recalls the formula for the number of tilings of a rectangle and proves the formula of Theorem 1 in this special case. §7 discusses the connection of domino tilings to loop-erased random walk and proves Theorem 3.

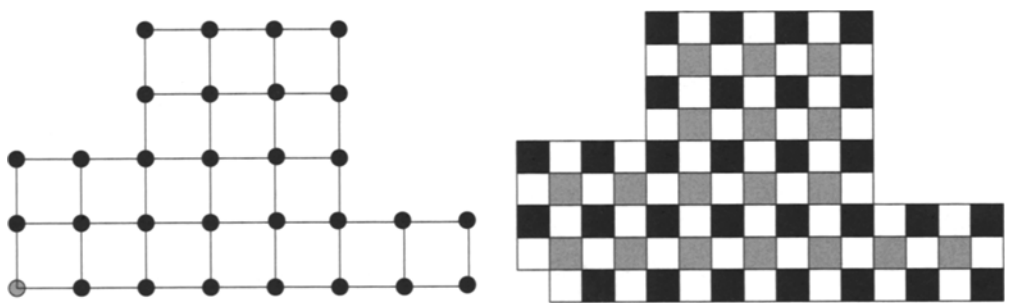


Fig. 1. The graph H and polyomino $P(H)$. The squares of B_0 are black, those of B_1 are in grey. The vertex b of H is in grey, and corresponds to the missing corner of P.

We kindly acknowledge Oded Schramm, Wendelin Werner and Bertrand Duplantier for helpful discussions, and thank Oded Schramm and Russell Lyons for proofreading.

2. Definitions and background

2.1. Temperleyan polyominos

By the *grid* $\mathcal{G}$ we mean the graph whose vertices are $2\mathbf{Z}^2$ and whose edges join all pairs of vertices at distance 2. A *lattice square* is a face of $\mathcal{G}$. A *simply-connected subgraph* of $\mathcal{G}$ is a set of vertices and edges of the grid which is the 1-skeleton of a simply-connected union of (closures of) lattice squares.

Let H be a finite simply-connected subgraph of $\mathcal{G}$, and $b\in H$ a fixed vertex adjacent to the outer face of H. We associate to H a new graph $P'=P'(H)=P'(H,b)$ as follows. There is a vertex of P' for each vertex, edge and face of H, except for the outer face and the vertex b. Two vertices u_1, u_2 of P' are connected by an edge in two cases: u_1, u_2 come from an edge e and a vertex v of H (and v is on the edge e), or u_1, u_2 come from a face f and edge e of H (and e is part of the boundary of f). In other words, $P'(H)$ is the "superposition" of H and its planar dual H', except that we discard vertex b of H and the outer vertex of H'. Let $P(H)$ denote the polyomino in $\frac{1}{2}\mathcal{G}$ whose dual (not including the vertex for the outer face) is $P'(H)$. See Figure 1 for an example.

Except for the outer face, the faces of P are squares, and come in four types, B_0, B_1, W_0, W_1. The squares in B_0 are those coming from vertices of H; the squares in B_1 are those coming from faces of H. The squares in W_0 (resp. W_1) are those coming from horizontal (resp. vertical) edges of H. The 'B' and 'W' stand for 'black' and 'white' coming from the checkerboard coloring of P. See Figure 1. We assign colors B_0, B_1, W_0, W_1 to vertices of P' corresponding to the colors of the faces of P.

Temperley [T] showed that there is a bijection between spanning trees of H and perfect matchings of $P'(H)$ (a *perfect matching*, or dimer covering, is a set of edges such that every vertex is contained in a unique edge). Perfect matchings of P' correspond to *domino tilings* (tilings with (1×2)- and (2×1)-rectangles) of $P(H)$. A polyomino (a union of lattice squares bounded by a simple closed curve) is said to be *Temperleyan* if it is of the form $P(H)$ for some simply-connected subgraph H of $\mathcal{G}$. Note that if H is a simply-connected subgraph of $\varepsilon\mathcal{G}$ for some $\varepsilon>0$, then $P(H)$ will be a polyomino in $\frac{1}{2}\varepsilon\mathcal{G}$.

The lattice square missing in P which corresponds to point $b\in H$ is called the *base square* of the polyomino. For the graph P' dual to P, b is the *base vertex* (it is not a vertex of P'). By Temperley's bijection, the number of tilings of P is independent of choice of vertex $b\in H$ as long as it is on the boundary. If P is a Temperleyan polyomino then by $H=H(P)$ we mean the unique associated simply-connected subgraph of $\mathcal{G}$ for which $P=P(H)$.

These definitions also apply to infinite graphs H, the only difference being that in this case we may if we like choose $b=\infty$, which means that we do not remove any lattice square from $P(H)$. All infinite graphs we deal with in the sequel have the property that near to infinity the boundaries are straight, that is, the boundaries have no corners outside some fixed large radius. This is to avoid certain convergence problems later. Two important examples of infinite Temperleyan polyominos are the whole plane $P(\mathbf{Z}^2)$ and the half-plane $P(\mathbf{N}\times\mathbf{Z})$. Both of these have $b=\infty$.

2.2. Conformal properties

The results of [Ken2] apply to Temperleyan polyominos. Let P be a Temperleyan polyomino with dual graph P'. We assign weights to the edges of P' so that a horizontal edge has weight 1 if its left vertex is white, -1 if its left vertex is black; a vertical edge is weighted $i=\sqrt{-1}$ if its lower vertex is white, $-i$ if its lower vertex is black. Thus the weights around a white vertex are $1, i, -1, -i$ in counterclockwise order starting from the edge leading right; around a black vertex these weights are $-1, -i, 1, i$. The adjacency matrix of the graph P' with these weights is called the *Kasteleyn matrix* K_P of P. Its determinant is the square of the number of domino tilings of P [Kas2]. The inverse of the Kasteleyn matrix is called the *coupling function* $C_P(\cdot,\cdot)$ of P. The probability of a configuration of dominos occurring in a random tiling is the absolute value of the determinant of a submatrix of the coupling function matrix [Ken1].

The coupling function has a number of important properties which we list here.

2.2.1. *Combinatorial properties of the coupling function.* First, for v_1, v_2 two vertices of P', we have $C_P(v_1,v_2)=C_P(v_2,v_1)$, and if v_1 and v_2 are both black or both white then

$C_P(v_1,v_2)=0$. Therefore we will always take the first variable of C_P to be a white vertex and the second variable to be a black vertex.

The coupling function has a concise description in terms of the Green function on H. Let G be the Green function on the graph H, that is, G is the function on $H\times H$ which satisfies $\Delta G(x,y)=\delta_x(y)-\delta_b(y)$, where Δ is the Laplacian with respect to the second variable (and recall that b is the base vertex). Here $\Delta f(v)=4f(v)-f(v+2)-f(v-2)-f(v+2i)-f(v-2i)$, except at a boundary vertex, where $\Delta f(v)$ is the degree of v times $f(v)$ minus the sum of the neighboring values. (Here $v\pm 2$, $v\pm 2i$ refer to the four neighbors of v in the graph H. Note that these vertices are at distance 2 in P'.) As stated, the function G is only well-defined up to an additive constant. We fix the constant by setting the function to be zero when $y=b$. For x,x' any two vertices of H, the function $L(y):=G(x,y)-G(x',y)$ satisfies $\Delta L(y)=\delta_x(y)-\delta_{x'}(y)$ and $L(b)=0$.

Let x and x' be adjacent vertices of H; let f,f' be the faces of H adjacent to the edge xx', with f' on the left as the edge is traversed from x to x'. Let $\hat{L}$ be the function on the faces of H which is the harmonic conjugate of $L(y)$ in the sense that for any edge $e=v_1v_2$ of H which is not the edge xx' we have $L(v_2)-L(v_1)=\hat{L}(f_1)-\hat{L}(f_2)$, where f_1 is the face to the left of the edge e (when e is traversed from v_1 to v_2) and f_2 is the face to the right. If we define $\hat{L}$ to be zero on the outer face then it is uniquely defined and harmonic except at f and f'. Moreover $\Delta\hat{L}(\,\cdot\,)=\delta_f(\,\cdot\,)-\delta_{f'}(\,\cdot\,)$, where Δ is the Laplacian on the dual H' of H. The function $\hat{L}(z)$ can also be written $\widehat{G}(f,z)-\widehat{G}(f',z)$, where $\widehat{G}$ is defined by $\Delta\widehat{G}(f,z)=\delta_f(z)-\delta_o(z)$ and $\widehat{G}(f,o)=0$, o referring to the outer face.

We now have the following description of the coupling function in terms of these Green functions. Suppose $v_1\in W_0$. Then

$$C_P(v_1,v_2)=\begin{cases} G(v_1+1,v_2)-G(v_1-1,v_2) & \text{if } v_2\in B_0,\\ i(\widehat{G}(v_1+i,v_2)-\widehat{G}(v_1-i,v_2)) & \text{if } v_2\in B_1.\end{cases} \tag{2}$$

Note that when $v_2\in B_0$, $C_P(v_1,v_2)=G(v_1+1,v_2)-G(v_1-1,v_2)$ makes sense since both $v_1\pm1$ and v_2 are in B_0 (and hence vertices of H). When $v_2\in B_1$ rather, $C_P(v_1,v_2)=i(\widehat{G}(v_1+i,v_2)-\widehat{G}(v_1-i,v_2)$ makes sense since $v_1\pm i$ and v_2 are in B_1 (faces of H). Similarly, if $v_1\in W_1$ we have

$$C_P(v_1,v_2)=\begin{cases} \widehat{G}(v_1+1,v_2)-\widehat{G}(v_1-1,v_2) & \text{if } v_2\in B_1,\\ -i(G(v_1+i,v_2)-G(v_1-i,v_2)) & \text{if } v_2\in B_0.\end{cases} \tag{3}$$

This description of the coupling function follows from the fact that K^*K is the Laplacian on H, and also acts as the Laplacian on the dual of H, see [Ken2].

An important case of the above formula for C_P is when v_1 corresponds to an edge on the boundary of H, so that one of $v_1\pm1$ or $v_1\pm i$ corresponds to the outer face of H.

Suppose for example that v_1+1 is the outer face of P'; then $\delta_{v_1+1}(v_2)=0$ by definition (for v_2 a vertex of B_1, that is, a face of H which is not the outer face), and so the term $G(v_1+1,v_2)$ can be ignored in the above formula.

2.2.2. *Asymptotic properties.* Let U be a rectilinear polygon in $\mathbf{C}$. Fix a base point $b_0\in\partial U$. Let $\{P_\varepsilon\}_{\varepsilon>0}$ be a sequence of Temperleyan polyominos $P_\varepsilon\subset\varepsilon\mathbf{Z}^2$, approximating U as $\varepsilon\to 0$ in the following sense. The P_ε are rectilinear with the same number of corners as U, one corner converging to each corner of U. Furthermore the base points $b_\varepsilon\in P_\varepsilon$ converge to b_0.

In [Ken2] we proved the following. Let $C_\varepsilon=C_{P_\varepsilon}$. Under the above convergence hypotheses, the rescaled coupling functions $C_\varepsilon/\varepsilon$ converge to a pair of complex-valued functions $F_0(v,z)$ and $F_1(v,z)$ which are meromorphic in z, in the following sense. If $\{u_\varepsilon\},\{v_\varepsilon\},\{w_\varepsilon\},\{x_\varepsilon\}$ with $u_\varepsilon,v_\varepsilon,w_\varepsilon,x_\varepsilon\in P_\varepsilon$ are four sequences of vertices of type W_0,W_1,B_0,B_1 respectively, converging to respectively $u,v,w,x\in U$, then

$$\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}C_\varepsilon(u_\varepsilon,w_\varepsilon)=\operatorname{Re}F_0(u,w),$$

$$\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}C_\varepsilon(u_\varepsilon,x_\varepsilon)=i\operatorname{Im}F_0(u,x),$$

$$\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}C_\varepsilon(v_\varepsilon,w_\varepsilon)=\operatorname{Re}F_1(v,w),$$

$$\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}C_\varepsilon(v_\varepsilon,x_\varepsilon)=i\operatorname{Im}F_1(v,x).$$

The functions F_0 and F_1 are defined by the following properties.

PROPOSITION 4 ([Ken2, Theorem 13]). *For each fixed $v\in U$ the function $F_0(v,z)$ has the properties:*

(1) *it is meromorphic as a function of $z\in U$;*

(2) *its imaginary part vanishes for $z\in\partial U$;*

(3) *it has a zero at $z=b_0$;*

(4) *it has a simple pole of residue $1/\pi$ at $z=v$, and no other poles on $\overline{U}$.*

Similarly, for each fixed $v\in U$ the function F_1 has the properties:

(1) *it is meromorphic as a function of $z\in U$;*

(2) *its real part vanishes for $z\in\partial U$;*

(3) *it has a zero at $z=b_0$;*

(4) *it has a simple pole of residue $1/\pi$ at $z=v$, and no other poles on $\overline{U}$.*

Furthermore, F_0 and F_1 are the unique *functions with these properties.*

Define the functions $F_+=F_+^U:=F_0+F_1$ and $F_-=F_-^U:=F_0-F_1$. These functions are easier to work with since they transform as nicely under conformal mappings: $F_+(v,z)\,dv$ is a meromorphic 1-form and $F_-(v,z)\,d\bar{v}$ is an antimeromorphic 1-form.

PROPOSITION 5 ([Ken2, Proposition 15]). *The function $F_+(v,z)$ is analytic in both variables. The function $F_-(v,z)$ is analytic in z and antianalytic in v. If V is another marked region and $f\colon V\to U$ a conformal isomorphism sending the base point of V to the base point of U, then we have the transformation rules*

$$F_+^V(v,z)=f'(v)F_+^U(f(v),f(z)),\tag{4}$$

$$F_-^V(v,z)=\overline{f'(v)}F_-^U(f(v),f(z)).\tag{5}$$

When $U=\mathbf{C}$ with $b_0=\infty$ we have $F_+(v,z)=2/\pi(z-v)$ and $F_-(v,z)\equiv 0$. When U is the right half-plane $\mathrm{RHP}=\{z:\mathrm{Re}(z)>0\}$ with $b_0=\infty$ we have

$$F_+(v,z)=\frac{2}{\pi(z-v)},\quad F_-(v,z)=-\frac{2}{\pi(z+\bar v)}.\tag{6}$$

This and the above transformation rules determine F_+ and F_- (and hence F_0,F_1) on any simply-connected region.

2.3. Average height function

Recall [Th] that the height function of a domino tiling is an integer-valued function on the vertices of the dominos; it is well-defined up to an additive constant. After fixing its value at some vertex, it is defined by the property that on an edge v_1v_2 which is not crossed by a domino, the height difference $h(v_2)-h(v_1)$ is $+1$ if the square to the left of the edge v_1v_2 (we mean, the square containing edge v_1v_2 and on the left when traversing v_1v_2 from v_1 to v_2) is black, and -1 if this square is white. See Figure 2. Note that the height function along the boundary is independent of the tiling.

Let U be a rectilinear polygon with base point b_0, and P_ε a sequence of Temperleyan polyominos converging to U as in the previous section. Suppose for simplicity that all the P_ε contain a fixed vertex v_0, say at a corner of P_ε. Let h_ε be the average height function of P_ε, that is, for any vertex $v\in P_\varepsilon$, $h_\varepsilon(v)$ is the average height of v over all domino tilings of P_ε, where the height at v_0 is taken to be zero. The *limiting average height function* h of a random domino tiling of U is by definition the limit as $\varepsilon\to 0$ of the functions h_ε: take $x\in U$ with $x\notin\partial U$, and let $x_\varepsilon\in P_\varepsilon$ converge to x; then

$$h(x):=\lim_{\varepsilon\to 0}h_\varepsilon(x_\varepsilon).$$

For $x\in\partial U$ but not at a corner, $h(x)$ is defined by continuity from values of h in the interior.

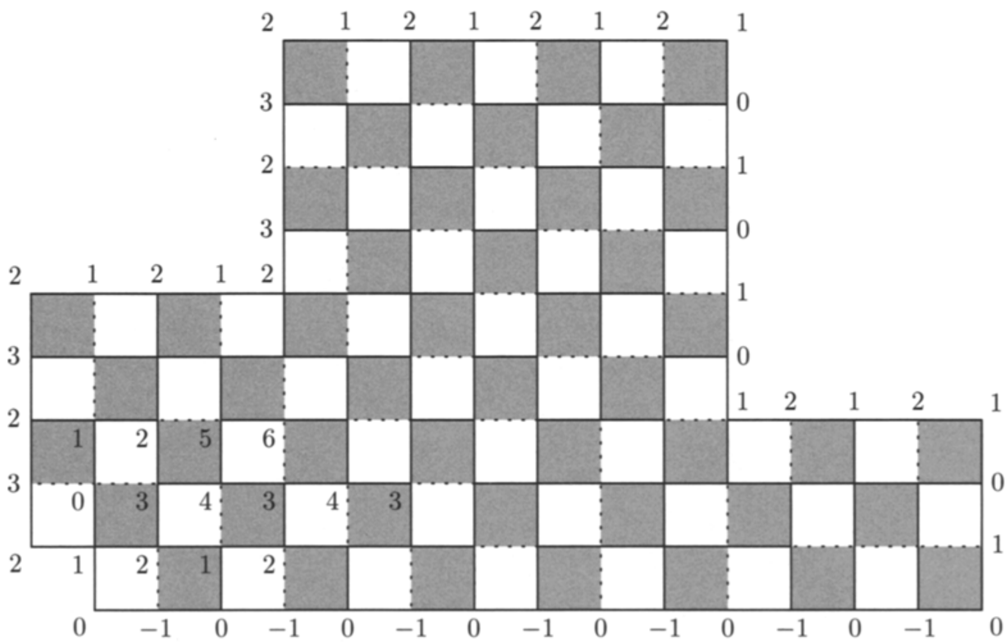


Fig. 2. Some of the heights in a domino tiling.

In [Ken2] we showed that this limit exists and has a simple expression in terms of the function F_+. Let $F_+^*(v,z)=F_+(v,z)-2/\pi(z-v)$ and let $F_+^*(v)=\lim_{z\to v} F_+^*(v,z)$. Then the limiting average height function h on U is given by the complex line integral

$$h(v)=2\,\mathrm{Im}\int_{v_0}^{v} F_+^*(u)\,du.$$

The choice of zero v_0 is immaterial since the height is only defined up to an additive constant anyway. Note that from (6), on $\mathbf{C}$ or on the right half-plane we have $F_+(v,z)=2/\pi(z-v)$, so $F_+^*\equiv 0$, and therefore $h(v)$ is constant, as expected. As another example, the map $z\mapsto(1+z)/(1-z)$ maps the unit disk to the right half-plane, mapping 1 to ∞; therefore using (4), on the unit disk with base point $z=1$ we have

$$F_+(v,z)=\frac{2(1-z)}{\pi(1-v)(z-v)} \quad \text{and} \quad F_+^*(z)=\frac{-2}{\pi(1-z)},$$

so that

$$h(v)=-\frac{4}{\pi}\,\mathrm{Im}\int_{v_0}^{v}\frac{du}{1-u}.$$

Note that in general $h(v)$ is harmonic, being the imaginary part of an analytic function.

For v on the outer boundary of U, $h(v)-h(v_0)$ is the total turning (in radians times $2/\pi$) of the boundary tangent on the path counterclockwise (cclw) around the boundary from v_0 to v. When passing the base point b_0, however, the height drops by 4. Thus a full cclw turn contributes $+4-4=0$ to h. The function h is constant on the straight edges, and there is a discontinuity at each corner of U, where h changes by ± 1 according to whether the corner is a left turn or right turn (if the base point b_0 is at a corner then the change would be -3 or -5 accordingly). These boundary values can be understood in a sense using Figure 2: the polyomino in this figure can be thought of as approximating a rectilinear octagon U, where the height on the lower boundary of U is $-\frac{1}{2}$ (the average of -1 and 0, the alternating heights on P_ε), the height on the right-most boundary is $\frac{1}{2}$ (the average of 0 and 1), and so on. Since a harmonic function is determined by its boundary values this provides a simple description of h in terms of the turning of the boundary tangent.

2.4. Dirichlet energy

Recall that the Dirichlet energy of a harmonic function h on a region U is given by

$$E(h) = \iint_U |\nabla h|^2 \, dx \, dy = \oint_{\partial U} h \, dg$$

where g is a harmonic conjugate of h, that is, g is a harmonic function so that $h+ig$ is locally an analytic function of $x+iy$.

We will be interested in the Dirichlet energy of harmonic functions which are the limiting average height functions on rectilinear polygons U. In particular, their boundary values have a finite number of jump discontinuities (at the corners); unfortunately in such a case the Dirichlet energy is infinite. To avoid this difficulty, for each sufficiently small $\delta>0$ we define the *δ-normalized Dirichlet energy* $E_\delta(h)$ as follows. Remove a δ-neighborhood of each $x\in\partial U$ for which the harmonic function h has a jump discontinuity. Let U' be the region U without these neighborhoods. The δ-normalized energy $E_\delta(h)$ is simply the integral of $|\nabla h|^2$ over U'.

If U is unbounded and if h has a jump discontinuity at ∞, we remove the neighborhood of ∞ consisting of points $|z|>1/\delta$, and compute the energy on the remaining region as before.

Here we will illustrate with an example. Let h be the bounded harmonic function on the upper half-plane which has value 0 on the x-axis to the right of the origin and 1 on the x-axis to the left of the origin. Then $h(z)=(1/\pi)\operatorname{Im}\log(z)$. The harmonic conjugate to h is $g(z)=-(1/\pi)\operatorname{Re}\log(z)$. The normalized Dirichlet energy is the integral over $\partial U'$ of $h\,dg$, which can be broken into four parts: the integral from $-\delta^{-1}$ to $-\delta$, the integral around the half-circle of radius δ, the integral from δ to δ^{-1}, and the integral around the half-circle of radius δ^{-1}. The third of these integrals is zero since h is zero on the positive

x-axis. The second and fourth are zero since dg is zero on circles about the origin. So the only contribution is from the first integral, which gives

$$E_\delta(h)=\int_{-1/\delta}^{-\delta} 1\,dg=g(-\delta)-g\left(-\frac{1}{\delta}\right)=\frac{2}{\pi}\log\frac{1}{\delta}.$$

In a more general situation, dg will not vanish along the boundaries of the δ-neighborhoods of the discontinuities, but dg will still be $O(\delta)$ there, as we now show: Let h be the average height function on a rectilinear polygon U. Take a convex corner of U, translate and rotate U so that the corner is at the origin and is bounded by the positive axes. Without loss of generality (after a linear scale) we suppose that near the origin h is 0 on the x-axis and -1 on the y-axis. Then h is the imaginary part of an analytic function $\tilde{h}(z)$ on U whose expansion at 0 is of the form

$$-\frac{2}{\pi}\log(\alpha_1 z+\alpha_2 z^2+\ldots)=-\frac{2}{\pi}\log(z)+\beta_0+\beta_1 z+\beta_2 z^2+\ldots\,.$$

In particular when $z=\delta e^{i\theta}$, we have

$$\frac{dg(z)}{d\theta}=\frac{2}{\pi}\,\mathrm{Re}(\beta_1\delta i e^{i\theta}+O(\delta^2))=O(\delta).$$

A similar argument works at a non-convex corner.

Since h has a standard form near each of its jump singularities, the δ-normalized Dirichlet energy has a very simple dependence on δ. Recall that at a convex corner the height function changes by $+1$ when moving cclw around the boundary. So the height function is of the form of that of the above example. If we change δ to a smaller δ', the change in energy at that corner is

$$\int_{i\delta}^{i\delta'} -dg=-\frac{2}{\pi}\log(\delta')+\frac{2}{\pi}\log\delta+O(\delta).$$

Thus the dependence on δ at a convex corner is $(2/\pi)\log(1/\delta)+O(\delta)$.

Similarly we can do the calculation at a concave corner. Move and rotate the corner so that it is bounded by the positive x-axis and the negative y-axis. Up to an additive constant h is 0 on the x-axis near the origin and 1 on the negative y-axis. So h is the imaginary part of an analytic function $\tilde{h}(z)=(2/3\pi)\log z+\alpha_0+\alpha_1 z+\alpha_2 z^2+\ldots$. Now if δ changes to the smaller δ', the change in energy is

$$\int_{-i\delta}^{-i\delta'} dg=-\frac{2}{3\pi}\log(\delta')+\frac{2}{3\pi}\log\delta+O(\delta).$$

Thus the dependence on δ at a concave corner is $(2/3\pi)\log(1/\delta)+O(\delta)$.

If a corner contains the base point, then if it is a convex corner the height changes by -3 rather than $+1$. The dependence on δ is therefore 9 times that for a normal convex corner, or $(18/\pi)\log(1/\delta)$. So the fact that the corner contains the base point adds $(16/\pi)\log(1/\delta)$ to its "local energy". If the corner is concave the height changes by 5 rather than 1, and so the local energy is $(50/3\pi)\log(1/\delta)$ rather than $(2/3\pi)\log(1/\delta)$, which is also an addition of $(16/\pi)\log(1/\delta)$. If the base point occurs along an edge, the height change is 4 and the energy associated is again $(16/\pi)\log(1/\delta)$.

Now from these calculations, for any rectilinear region U we can immediately compute the dependence of the δ-energy on δ, in terms of the number of vertices. A (simply-connected) region with V vertices has $\frac{1}{2}(V-4)$ concave vertices and $\frac{1}{2}(V+4)$ convex vertices, one of which we may take to be the base vertex. So the dependence on δ of its δ-energy is

$$O(\delta)+\left(\frac{2}{3\pi}\left(\frac{V-4}{2}\right)+\frac{2}{\pi}\left(\frac{V+4}{2}\right)+\frac{16}{\pi}\right)\log\frac{1}{\delta}=\left(\frac{4(V-4)}{3\pi}+\frac{24}{\pi}\right)\log\frac{1}{\delta}+O(\delta).$$

Note that $-\frac{1}{48}\pi$ times the above δ-energy gives the logarithmic terms in Theorem 1 (if we replace δ with ε).

3. Cutting a lattice region into rectangles

Here we prove Theorem 1. Let U be a rectilinear polygon with base point $b_0\in\partial U$. Let $\gamma_0=\gamma_0(t)$ be a straight (horizontal or vertical) unit speed path in U from ∂U to ∂U which avoids b_0 and does not touch ∂U except at its endpoints.

Let $P_\varepsilon\subset\varepsilon\mathbf{Z}^2$ be a Temperleyan polyomino approximating U as described in §2.2.2. Let γ_ε be a strip of width ε of lattice squares of P_ε, lying within $O(\varepsilon)$ of γ_0 and traversing P_ε from the boundary to the boundary, avoiding the base square of P_ε. Furthermore we require that γ_ε contain no square in B_0 (that is, contains only white squares and squares in B_1), and if either extremity of γ_0 is at a concave corner of U then the corresponding extremity or extremities of γ_ε are at the corresponding corners of P_ε.

Because of the boundary conditions on P_ε, γ_ε has length which is an odd multiple of ε. If we remove γ_ε from P_ε, then what remains is a union of two disjoint polyominos P_1 and P_2. Let P_1 be the polyomino which contains the base square of P_ε; then P_1 is a Temperleyan polyomino. The other polyomino P_2 will become a Temperleyan polyomino if we remove a single square s of type B_0 in P_2 adjacent to one of the endpoints of γ_ε.

Note that the union $\gamma_\varepsilon\cup\{s\}$ has a unique domino tiling. The number of tilings of P_ε equals the product of the number of tilings of P_1 and the number of tilings of P_2, divided by the probability $\Pr(\gamma_\varepsilon\cup\{s\})$ that the tiling of $\gamma_\varepsilon\cup\{s\}$ occurs in a uniform tiling of P_ε.

We can repeat this procedure on P_1 and P_2, cutting them apart into simpler and simpler pieces until we arrive at a collection of Temperleyan rectangles (a Temperleyan rectangle is an odd-by-odd rectangle with corners in B_0 and one corner square removed). Temperley [T] provides us with a formula for the exact number of tilings of a Temperleyan rectangle (see Proposition 13). Working by induction, to compute the number of tilings of P_ε it suffices to be able to compute the probability of finding, in a random tiling of P_ε, a tiling of $\gamma_\varepsilon \cup \{s\}$. We cannot compute this probability exactly but we can approximate it sufficiently closely (Lemma 6).

The region $\gamma_\varepsilon \cup \{s\}$ has a unique tiling which is a chain of dominos. Starting from the boundary of P_ε, let $a_1, a_2, \ldots, a_N$ be the set of consecutive dominos making up the tiling of $\gamma_\varepsilon \cup \{s\}$. The a_i are laid end-to-end except for a_N which is perpendicular to the others. The probability of all the a_i being present in a tiling of P_ε is a product, as j runs from 1 to N, of the probability that a_j is present, given that $a_1, \ldots, a_{j-1}$ are already present:

$$\Pr(a_1, \ldots, a_N) = \prod_{j=1}^{N} \Pr(a_j \,|\, a_1, \ldots, a_{j-1}). \tag{7}$$

Suppose that $a_1, \ldots, a_{j-1}$ are present already. These dominos form a strip running from the boundary of P_ε to a point in the interior of P_ε. The region $P_\varepsilon^{(j)} \stackrel{\text{def}}{=} P_\varepsilon \setminus \{a_1, \ldots, a_{j-1}\}$ is again a Temperleyan polyomino, by our hypothesis that the a_i contain only black vertices of type B_1. So computing $\Pr(a_j \,|\, a_1, \ldots, a_{j-1})$ is a matter of computing the probability of a_j in a random tiling of this region $P_\varepsilon^{(j)}$.

Let U_j be the region U with a slit cut out along the segment $\gamma_0([0, 2(j-1)\varepsilon])$, and translated by $-2(j-1)\varepsilon$ so that the tip of the cut is at the origin. Then $U_0 = U$ (up to translation) and U_N is a union of two rectilinear polygons.

We may suppose (after applying a rotation if necessary) that the path γ_0 is horizontal and goes from left to right. For $0 < j < N$ let f_j be the unique conformal isomorphism sending the right half-plane $\{z : \mathrm{Re}(z) > 0\}$ to U_j which sends 0 to 0 (end of the cut), ∞ to the base point b_0, and has expansion $f_j(z) = z^2 + O(z^3)$ at the origin (we do not define f_0 or f_N).

LEMMA 6. *Let $F_1^{(j)}$ be the function F_1 (coupling function limit) on the region U_j. We have*

$$\frac{\Pr(a_j \,|\, a_1, \ldots, a_{j-1})}{\sqrt{2} - 1} = 1 + \left(\frac{\pi}{\sqrt{2}} \varepsilon F_1^{(j)}(\varepsilon, 2\varepsilon) - 1 \right) + \mathrm{err}(j)$$

where $\mathrm{err}(j)$ *is* $o(\max(1/j, 1/(N-j)))$. *The term in brackets on the right-hand side is* $O(\varepsilon)$ *when j is not close to* 0 *or* N. *Let $\xi > 0$ be a small constant and $K = K(\varepsilon) := \xi/\varepsilon$.*

Then this can be written

$$\frac{\Pr(a_j\,|\,a_1,\dots,a_{j-1})}{\sqrt{2}-1}=\begin{cases}1+\dfrac{C_0}{j}+\mathrm{err}_1(j)+O(\varepsilon) & \text{if } j\leqslant K,\\ 1+\frac{1}{6}\varepsilon S\sqrt{f_j}\,(0)+\dfrac{1}{\xi}o(\varepsilon) & \text{if } K<j<N-K,\\ 1+\dfrac{C_1}{N-j}+\mathrm{err}_2(N-j)+O(\varepsilon) & \text{if } N-K\leqslant j,\end{cases}\tag{8}$$

where $S\sqrt{f_j}$ *is the Schwarzian derivative of* $\sqrt{f_j}$, *and* $\mathrm{err}_1(x)$ *and* $\mathrm{err}_2(x)$ *are* $o(1/x)$. *Here* err_1 *depends only on whether* γ_0 *starts at an edge or at a concave vertex, and* err_2 *depends only on whether* γ_0 *ends on an edge or at a vertex. The constants* C_0, C_1 *are determined as follows. If* γ_0 *starts on an edge, then* $C_0=-\frac{1}{8}$; *if* γ_0 *starts at a corner then* $C_0=-\frac{5}{72}$. *If* γ_0 *ends at an edge then* $C_1=-\frac{3}{8}$; *if* γ_0 *ends at a corner then* $C_1=-\frac{23}{72}$.

As we will see in the proof, the first and third expressions on the right-hand side of (8) are special cases of the middle expression, except for the error terms.

The probability of Lemma 6 can be related to the change in normalized Dirichlet energy of the limiting average height function of U_j:

LEMMA 7. *Let* ξ *and* K *be as in the previous lemma. For* $N>j>1$ *the difference in the* δ*-normalized Dirichlet energy of the limiting average height function between* U_j *and* U_{j-1} *is*

$$E_\delta(h_j)-E_\delta(h_{j-1})=\begin{cases}-\dfrac{48C_0}{\pi j}+O(\varepsilon) & \text{if } j\leqslant K,\\ -\dfrac{8\varepsilon}{\pi}S\sqrt{f_j}\,(z_j)+\dfrac{1}{\xi}o(\varepsilon) & \text{if } K<j<N-K,\\ -\dfrac{48C_1}{\pi(N-j)}+O(\varepsilon) & \text{if } N-K\leqslant j,\end{cases}$$

where C_0 *and* C_1 *are defined as in the previous lemma. When* $j=1$ *or* $j=N$, *the difference in energy* $E_\delta(h_j)-E_\delta(h_{j-1})$ *is a constant depending (for* $j=1$*) only on whether the cut starts at a corner or on an edge, and (for* $j=N$*) on whether the cut ends at a corner or on an edge.*

These two lemmas, and a computation of the number of tilings of a Temperleyan rectangle, give us the main result:

Proof of Theorem 1. The proof is by induction on the number of cuts required to cut P_ε apart into Temperleyan rectangles. In the case P_ε is a Temperleyan rectangle, $V=4$ and Proposition 16 below shows that (1) equals the log of the number of tilings, up to an error $O(\varepsilon)$.

For a general P_ε as in the statement, let $\gamma_\varepsilon \cup \{s\}$ be a strip as explained above, cutting P_ε apart into two Temperleyan regions P' and P''. We compute the log of the probability of $\gamma_\varepsilon \cup \{s\}$ occurring in a tiling of P_ε. This probability depends on whether γ_ε starts on an edge or at a corner, and whether it ends on an edge or at a corner.

Let $a_1, ..., a_N$ be the chain of dominos of $\gamma_\varepsilon \cup \{s\}$, and let $P_\varepsilon^{(j)} = P_\varepsilon \cup \{a_1, ..., a_j\}$. By Lemma 6 we have

$$
\begin{aligned}
\log \frac{\Pr(\gamma_\varepsilon \cup \{s\})}{(\sqrt{2}-1)^N} &= \sum_j \log \frac{\Pr(a_j \,|\, a_1, ..., a_{j-1})}{\sqrt{2}-1} \\
&= \sum_{j \leqslant K} \log\left(1 + \frac{C_0}{j} + \mathrm{err}_1(j) + O(\varepsilon)\right) \\
&\quad + \sum_{K<j<N-K} \log\left(1 + \tfrac{1}{6}\varepsilon S \sqrt{f_j}(0) + \frac{1}{\xi} o(\varepsilon)\right) \\
&\quad + \sum_{N-K \leqslant j} \log\left(1 + \frac{C_1}{N-j} + \mathrm{err}_2(N-j) + O(\varepsilon)\right) \\
&= \sum_{j \leqslant K} \frac{C_0}{j} + \mathrm{err}_3(j) + O(\varepsilon) + \sum_{K<j<N-K} \tfrac{1}{6}\varepsilon S \sqrt{f_j}(0) + \frac{1}{\xi} o(\varepsilon) \\
&\quad + \sum_{N-K \leqslant j} \frac{C_1}{N-j} + \mathrm{err}_4(N-j) + O(\varepsilon).
\end{aligned}
\tag{9}
$$

In this expression the terms $\mathrm{err}_3(j)$ and $\mathrm{err}_4(N-j)$ sum to give an error $\mathrm{ERR}_0(\varepsilon)\log(1/\varepsilon)$ where $\mathrm{ERR}_0(\varepsilon)$ is $o(1)$. The remaining terms are equal by Lemma 7 to $-\frac{1}{48}\pi$ times the change in Dirichlet energy on γ_ε, up to the error terms $O(\varepsilon)$ and $o(\varepsilon)/\xi$. The terms $O(\varepsilon)$, however, sum to $KO(\varepsilon) = O(\xi)$, which is negligible. Similarly the terms $o(\varepsilon)/\xi$ sum to $o(\varepsilon)/\varepsilon\xi$, which tends to zero if ξ tends to zero sufficiently slowly.

Since each tile (except the last) decreases the area by 2 and increases the perimeter by 4, noting that $-2c_0 + 4c_1 = \log(\sqrt{2}-1)$ which cancels the denominator of (9), the formula (1) is correct up to a justification of the term $\mathrm{ERR}_0(\varepsilon)$.

The error $\mathrm{ERR}_0(\varepsilon)$ is the sum of two error terms, one coming from the beginning of γ_ε and one from the end. Let $\mathrm{ERR}_1(\varepsilon)$ and $\mathrm{ERR}_2(\varepsilon)$ be $\sum_{j \leqslant K} \mathrm{err}_3(j)$ when γ_ε starts at an edge or a corner respectively. Let $\mathrm{ERR}_3(\varepsilon)$ and $\mathrm{ERR}_4(\varepsilon)$ be $\sum_{j \geqslant N-K} \mathrm{err}_4(N-j)$ when γ_ε ends at an edge or a corner respectively. We have that $\mathrm{ERR}_1(\varepsilon) + \mathrm{ERR}_4(\varepsilon) = \mathrm{ERR}_2(\varepsilon) + \mathrm{ERR}_3(\varepsilon)$ since when γ_ε is traversed in either direction the same error occurs (that is, the probability of $\gamma_\varepsilon \cup \{s\}$ is independent of the position of s as long as it is of type B_0 and on the correct side of γ_ε; therefore γ_0 can be traversed in either direction, giving the same probability for $\gamma_\varepsilon \cup \{s\}$). Moreover by Proposition 16, $\mathrm{ERR}_1(\varepsilon) + \mathrm{ERR}_3(\varepsilon) = 0$, since when we cut a rectangle apart into two rectangles there is no error (or rather, this error, when multiplied by $\log(1/\varepsilon)$, is still $O(\varepsilon)$).

Therefore if γ_ε begins at a corner and ends at an edge, or vice versa, the error ERR_0 is $\mathrm{ERR}_2-\mathrm{ERR}_1$. When γ_ε begins and ends at a corner the error ERR_0 is $2(\mathrm{ERR}_2-\mathrm{ERR}_1)$. When we cut P_ε apart to make rectangles, each concave corner gets cut exactly once, so the total accumulated errors will be the number of concave corners times a fixed error $\mathrm{ERR}_2(\varepsilon)-\mathrm{ERR}_1(\varepsilon)$. Setting $\mathrm{ERR}(\varepsilon)=\mathrm{ERR}_2(\varepsilon)-\mathrm{ERR}_1(\varepsilon)$ completes the proof. $\square$

As a shortcut for computing the log probability of the cut $\gamma_\varepsilon\cup\{s\}$, from Lemma 6 the log of (7) is $N\log(\sqrt{2}-1)$ plus the sum over j of $\log\bigl(1+\bigl((\pi/\sqrt{2})\,\varepsilon F_1^{(j)}(\varepsilon,2\varepsilon)-1\bigr)\bigr)$, plus the error term. When ε is small the sum over j can be replaced by an integral using the variable $t=2\varepsilon j$:

$$\begin{aligned}\Pr(\gamma_\varepsilon\cup\{s\}) = N\log\bigl(\sqrt{2}-1\bigr)+\int_{2\varepsilon}^{2\varepsilon(N-1)}\Bigl(\frac{\pi}{\sqrt{2}}\,\varepsilon F_1^{(t/2\varepsilon)}(\varepsilon,2\varepsilon)-1\Bigr)\cdot\frac{dt}{2\varepsilon}&\\ +\mathrm{ERR}_0(\varepsilon)\log\frac{1}{\varepsilon}+\mathrm{const}+o(1).&\end{aligned}\tag{10}$$

Here $\mathrm{ERR}_0(\varepsilon)$ depends on whether or not γ_0 begins or ends at a corner, but not on U. This form will be useful in §7.

4. The probability of the next domino on a cut

4.1. The slit plane

Define the *slit plane* SP as the plane minus the left half of the x-axis: $\mathrm{SP}=\mathbf{R}^2-(-\infty,0]$. For $\varepsilon>0$ define the polyomino SP_ε to be the (infinite) polyomino obtained from $P(2\varepsilon\mathbf{Z}^2+(\varepsilon,\varepsilon))$ by removing the lattice squares centered at $(-k\varepsilon,0)$ for all $k\geqslant 0$. We assume that the lattice square centered at the origin is of type B_1. The infinite polyomino SP_ε is Temperleyan, with base point b at infinity (SP_ε is gotten from the graph H_ε which is obtained from $2\varepsilon\mathbf{Z}^2+(\varepsilon,\varepsilon)$ by removing edges $(-2k\varepsilon+\varepsilon,-\varepsilon)(-2k\varepsilon+\varepsilon,\varepsilon)$ for all $k>0$).

For later use, we compute the asymptotic coupling function on the slit plane SP.

LEMMA 8. *The asymptotic coupling function on* SP *with base point at* ∞ *satisfies*

$$F_+^{\mathrm{SP}}(v,z)=\frac{1}{\pi\sqrt{v}}\,\frac{1}{\bigl(\sqrt{z}-\sqrt{v}\bigr)},$$
$$F_-^{\mathrm{SP}}(v,z)=\frac{-1}{\pi\sqrt{\bar v}}\,\frac{1}{\bigl(\sqrt{z}+\sqrt{\bar v}\bigr)}.$$

Proof. The map $z \mapsto \sqrt{z}$ maps SP to the right half-plane RHP and ∞ to ∞. On RHP we have

$$F_+^{\mathrm{RHP}}(v,z) = \frac{2}{\pi}\frac{1}{(z-v)} \quad \text{and} \quad F_-^{\mathrm{RHP}}(v,z) = \frac{-2}{\pi}\frac{1}{(z+\bar{v})}.$$

The result follows using the transformation rules (4) and (5). □

Let U, U_j be as in §3. Let $f=f_j$ be as in that section. Define $b=b_j$ and $c=c_j$ to be the coefficients in the expansion

$$f(z) = z^2 + bz^3 + cz^4 + O(z^5). \tag{11}$$

Note that $b \in i\mathbf{R}$ and $c \in \mathbf{R}$ since f maps the imaginary axis to the real axis. The inverse of f has the expansion

$$f^{-1}(z) = z^{1/2} - \tfrac{1}{2}bz + \left(\tfrac{5}{8}b^2 - \tfrac{1}{2}c\right) z^{3/2} + O(z^2). \tag{12}$$

The expansion of $\sqrt{f}$ is

$$\sqrt{f(z)} = z + \tfrac{1}{2}bz^2 + \left(\tfrac{1}{2}c - \tfrac{1}{8}b^2\right) z^3 + O(z^4), \tag{13}$$

and the Schwarzian derivative of $\sqrt{f}$ at the origin is defined as

$$S\sqrt{f}(0) := \frac{(\sqrt{f})'''}{(\sqrt{f})'} - \frac{3}{2}\left(\frac{(\sqrt{f})''}{(\sqrt{f})'}\right)^2 = 3c - \frac{9}{4}b^2. \tag{14}$$

4.2. Proof of Lemma 6

Here we prove Lemma 6. After a translation we may assume that the right-hand square of the domino a_{j-1} is the square centered at the origin in $\varepsilon\mathbf{Z}^2$. Then $\Pr(a_j \mid a_1, ..., a_{j-1}) = C_j(\varepsilon, 2\varepsilon)$ where C_j is the coupling function on $P_\varepsilon^{(j)}$ translated as above. We must then approximate $C_j(\varepsilon, 2\varepsilon)$. Recall that ε is a lattice square of type W_1 and 2ε is of type B_1.

Recall from (3) that for $v \in W_1$, the coupling function $C_j(v,z)$ on $P_\varepsilon^{(j)}$ satisfies $\Delta C_j(v,z) = \delta_{v+\varepsilon}(z) - \delta_{v-\varepsilon}(z)$, for $z \in H'$ the dual graph of $H = H(P_\varepsilon^{(j)})$, where Δ is the Laplacian on H'. Furthermore $C_j(v,z)$ is zero when z is the outer face of H. In particular, $\Delta C_j(\varepsilon, z) = \delta_{2\varepsilon}(z)$, that is, $C_j(\varepsilon, z)$ is the discrete Green function on H' centered at ε. To compute $C_j(\varepsilon, 2\varepsilon)$, we will use a theorem of Kesten comparing the values of a continuous and a discrete harmonic function near the tip of a cut:

THEOREM 9 ([Kes]). *Let D_{slit} be the slit disk $\{|z|<1\}\setminus(-1,0]$ and let $\tilde{g}$ be a continuous harmonic function on D_{slit} with piecewise continuous boundary values and boundary values 0 on the slit. Let g_ε be a discrete harmonic function on $\varepsilon\mathbf{Z}^2\cap D_{\text{slit}}$ with boundary values within $O(\varepsilon)$ of the boundary values of $\tilde{g}$, and boundary values 0 on the slit. Then for any $(k,l)\in\mathbf{Z}^2$,*

$$g_\varepsilon(k\varepsilon, l\varepsilon) = \lambda_{k,l}\tilde{g}(k\varepsilon, l\varepsilon) + o(\sqrt{\varepsilon})$$

where the constant $\lambda_{k,l}$ only depends on k and l.

We cannot apply this theorem directly since in our case $C_j(\varepsilon, z)$ is not harmonic at the point $z=2\varepsilon$. Let, however, $C_{\mathrm{SP}_\varepsilon}(v,z)$ be the coupling function on the discrete slit plane SP_ε. The function $g(z) := C_j(\varepsilon,z) - C_{\mathrm{SP}}(\varepsilon,z)$ is harmonic for $z\in B_1$ in a neighborhood of the origin in H', *including* the point $z=2\varepsilon$, since the Laplacians of C_j and $C_{\mathrm{SP}_\varepsilon}$ are equal there. Furthermore $g(z)=0$ (Dirichlet boundary conditions) when z is on the cut γ_0. In fact $g(z)$ is discrete harmonic on all of $P_\varepsilon^{(j)}$ except possibly where the interior of $P_\varepsilon^{(j)}$ meets the negative x-axis somewhere to the left of the slit (but since $C_{\mathrm{SP}_\varepsilon}$ admits an analytic continuation around the origin, we can define $g(z)$ so as to be harmonic on all of $P_\varepsilon^{(j)}$).

We can therefore use Kesten's theorem applied to $g(z)$ to compute $g(2\varepsilon) = \lambda_{1,0}\tilde{g}(2\varepsilon)$, where $\tilde{g}(z)$ is the continuous harmonic function on $P_\varepsilon^{(j)}$ with the same boundary values as $g(z)$. It remains to compute $\tilde{g}(z)$ and $C_{\mathrm{SP}_\varepsilon}$.

LEMMA 10. *On SP_ε with $z\in B_1$, and z not within $O(1)$ of the origin, we have*

$$C_{\mathrm{SP}_\varepsilon}(\varepsilon,z) = \tau\varepsilon F_1^{\mathrm{SP}}(\varepsilon,z) + o(\varepsilon) = \tau\sqrt{\frac{\varepsilon}{z}} + o(\varepsilon),$$

for some constant τ.

For the proof see the appendix. Now to compute $\tilde{g}(z)$, from the lemma it is the (continuous) harmonic function with boundary values $-\tau\varepsilon F_1^{\mathrm{SP}}(\varepsilon,z)+o(\varepsilon)$ on ∂U_j. Note, however, that the function $\varepsilon F_1^{(j)}(v,z) - \varepsilon F_1^{\mathrm{SP}}(v,z)$ as a function of z is continuous, harmonic, with boundary values $-\varepsilon F_1^{\mathrm{SP}}(v,z)$ on ∂U_j. Multiplying by τ, at $v=\varepsilon$ we must have $\tilde{g}(z) = \tau\varepsilon(F_1^{(j)}(\varepsilon,z) - F_1^{\mathrm{SP}}(\varepsilon,z)) + o(\varepsilon)$.

Let $N_L(0)$ be the neighborhood of radius L of the origin, where L is chosen so small that $N_L(0)\subset P_\varepsilon^{(j)}$. More precisely, if $j\leqslant K$ or $j\geqslant N-K$ we take $L=\min\{j\varepsilon, (N-j)\varepsilon\}$, and for $K<j<N-K$ take $L=\xi$.

On $\partial N_L(0)$, $\varepsilon F_1^{\mathrm{SP}}(\varepsilon,z)$ is $O(\sqrt{\varepsilon/L})$ (see Lemma 8), and $\varepsilon F_1(\varepsilon,z)$ is comparable to $\varepsilon F_1^{\mathrm{SP}}(\varepsilon,z)$ and therefore also $O(\sqrt{\varepsilon/L})$. Therefore on $\partial N_L(0)$, $\sqrt{L/\varepsilon}\,(\tau\varepsilon F_1(\varepsilon,z) - \tau\varepsilon F_1^{\mathrm{SP}}(\varepsilon,z))$ is bounded, and we have by Kesten's theorem (for the disk of radius L)

$$\sqrt{\frac{L}{\varepsilon}}\,g(2\varepsilon) = \lambda_{1,0}\tau\sqrt{\frac{L}{\varepsilon}}\,(\varepsilon F_1(\varepsilon,2\varepsilon) - \varepsilon F_1^{\mathrm{SP}}(\varepsilon,2\varepsilon)) + o\left(\sqrt{\frac{\varepsilon}{L}}\right),$$

or

$$g(2\varepsilon) = \lambda_{1,0}\tau(\varepsilon F_1(\varepsilon, 2\varepsilon) - \varepsilon F_1^{\mathrm{SP}}(\varepsilon, 2\varepsilon)) + o(\varepsilon/L).$$

Now from Lemma 8 (using $F_1 = \frac{1}{2}(F_+ - F_-)$) we have $\varepsilon F_1^{\mathrm{SP}}(\varepsilon, 2\varepsilon) = \sqrt{2}/\pi$, and plugging in the value of $\lambda_{1,0}\tau$ from (16) below yields the first result.

Let $f = f_j$: RHP$\to U_j$ be as in the statement. Let $s(z) = f^{-1}(z)$. Then from the transformation rules (4), (5) and (6) we have

$$F_+^U(v,z) = \frac{2s'(v)}{\pi(s(z)-s(v))} \quad \text{and} \quad F_-^U(v,z) = -\frac{2\overline{s'(v)}}{\pi(s(z)+\overline{s(v)})}.$$

Consequently

$$\varepsilon F_1(\varepsilon, 2\varepsilon) = \frac{\varepsilon}{\pi}\left(\frac{s'(\varepsilon)}{s(2\varepsilon)-s(\varepsilon)} + \frac{\overline{s'(\varepsilon)}}{s(2\varepsilon)+\overline{s(\varepsilon)}}\right).$$

Using the expansion (12) and (14), and the facts that $b \in i\mathbf{R}$, $c \in \mathbf{R}$, this reduces to

$$\frac{\sqrt{2}}{\pi} + \frac{\sqrt{2}}{6\pi} S\sqrt{f_j}(0)\varepsilon + O(\varepsilon^{3/2}).$$

So then

$$\Pr(a_j \mid a_1, \ldots, a_{j-1}) = \sqrt{2} - 1 + \lambda_{1,0}\tau \frac{\sqrt{2}}{6\pi} S\sqrt{f_j}(0)\varepsilon + o\left(\frac{\varepsilon}{\xi}\right).$$

Plugging in for $\lambda_{1,0}\tau$ gives the center result in (8).

When j is close to 0 we must replace ε in Theorem 9 by $1/j$ since j is the combinatorial distance (number of lattice points in H) to the boundary of the region. We have

$$\Pr(a_j \mid a_1, \ldots, a_{j-1}) = \sqrt{2} - 1 + \lambda_{1,0}\tau \frac{\sqrt{2}}{6\pi} S\sqrt{f_j}(0)\varepsilon + o\left(\frac{1}{j}\right). \tag{15}$$

A similar expression holds when j is close to N.

Now when j is small or close to N, the Schwarzian derivative of f_j is blowing up in a standard way which is independent of U up to an error $O(1)$:

LEMMA 11. *When j is small, the germ of f_j at the origin is independent of U (depending only on whether or not the cut starts at a corner or on an edge) up to an error $O(1)$. Similarly when j is near N, the germ of f_j at the origin is independent of U up to an error $O(1)$, only depending on whether the cut ends at an edge or a corner.*

The proof is in the appendix. Below when we write $O_z(1)$ we mean an analytic function in z each of whose coefficients is $O(1)$ (as $j \to 0$ or $j \to N$). In particular, $z^3 O_z(1)$ is an analytic function whose 2-jet vanishes.

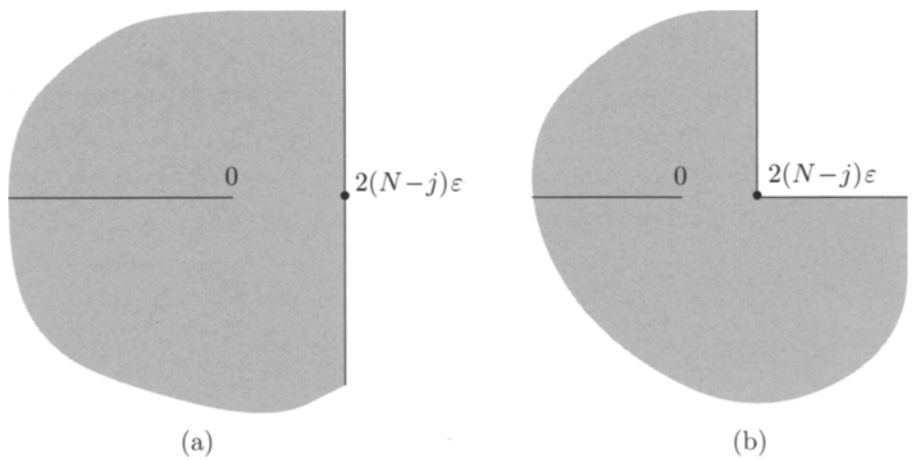


Fig. 3. When the cut ends at an edge (a) and at a corner (b).

Recall that f_j is normalized as in (11). If γ_0 starts at an edge of U, then when j is small f_j is approximated by the map $2\sqrt{\varepsilon j z^2+(\varepsilon j)^2}-2\varepsilon j$ which maps RHP to $\{x+iy:x>-2\varepsilon j\}-[-2\varepsilon j,0]$. That is, from Lemma 11 we have

$$f_j(z)=2\sqrt{\varepsilon j z^2+(\varepsilon j)^2}-2\varepsilon j+z^3O_z(1)=z^2-\frac{z^4}{4\varepsilon j}+z^3O_z(1).$$

Therefore $S\sqrt{f_j}(0)=-3/4\varepsilon j+O(1)$. Plugging this into (15) yields

$$\Pr(a_j\,|\,a_1,\dots,a_{j-1})=\sqrt{2}-1-\frac{\sqrt{2}-1}{8j}+o\left(\frac{1}{j}\right)+O(\varepsilon)$$

where the $o(1/j)$-term is independent of U: the term $o(1/j)$ only depends on the error in Kesten's theorem as applied to the coupling function on $\{x+iy:x>-2\varepsilon j\}-[-2\varepsilon j,0]$.

If γ_0 starts at a corner of U, then f_j is approximated by a map of the form

$$f(z)=C+C'(z+2iB)\sqrt{z-iB}+z^3O_z(1),$$

for constants C, C' and B (this follows from the Schwarz–Christoffel formula: the derivative of f is a constant times $z/\sqrt{z-iB}$ for some real B). The conditions that $f(iB)=-2\varepsilon j$ and $f(z)=z^2+O(z^3)$ determine C, C', B, and a short computation gives

$$f_j(z)=z^2-\frac{2i\sqrt{3}}{9\sqrt{\varepsilon j}}z^3-\frac{z^4}{4\varepsilon j}+O(z^6)+z^3O_z(1),$$

whence

$$S\sqrt{f_j}(0)=-\frac{5}{12\varepsilon j}+O(1).$$

A similar argument holds when j is near to N. If γ_0 ends at a straight edge of U, then near the origin f_j is approximated by the map

$$f_j(z)=C-\frac{C'}{\sqrt{z^2+B^2}}+z^3O_z(1),$$

which maps RHP to the region shown in Figure 3(a). Again C,C',B are determined by $f(\infty)=2\varepsilon(N-j)$ and $f(z)=z^2+O(z^3)$, and a calculation gives $S\sqrt{f_j}(0)=-9/4\varepsilon j+O(1)$.

Lastly, if γ_0 ends at a corner of U, then near the origin f_j is approximated by a map of the form

$$f_j(z)=C-\frac{C'}{\sqrt{z-iB}\,(z+2iB)},$$

which maps RHP to the region shown in Figure 3(b). Using $f(z)=z^2+O(z^3)$ and $f(\infty)=2\varepsilon(N-j)$ determines C,C',B to give

$$f(z)=z^2+\frac{2i\sqrt{3}}{9\sqrt{\varepsilon j}}z^3-\frac{3}{4\varepsilon j}z^4+O(z^5)+z^3O_z(1),$$

which gives $S\sqrt{f_j}(0)=-23/12\varepsilon j+O(1)$.

This completes the proof of Lemma 6. □

To determine $\lambda_{1,0}\tau$, we use Proposition 13 below. When we cut an $(m\times n)$-rectangle R into two rectangles R_1 and R_2 with a horizontal cut of length m, the sum of the logs of number of tilings of R_1 and R_2, minus the log of the number of tilings of R, which is the probability of the cut, is $m\log(\sqrt{2}-1)-\frac{1}{2}\log m+O(1)$. On the other hand, from the above proof this is

$$m\log(\sqrt{2}-1)+\lambda_{1,0}\tau\frac{\sqrt{2}}{6\pi(\sqrt{2}-1)}\left(-\frac{3}{4}-\frac{9}{4}\right)\log(m)+o(\log m).$$

This gives

$$\lambda_{1,0}\tau=\frac{(\sqrt{2}-1)\pi}{\sqrt{2}}. \tag{16}$$

5. The change in the Dirichlet energy

We prove here Lemma 7. The computation is in two steps. We first show that the change in the δ-normalized Dirichlet energy only depends (up to an error $o(\varepsilon)$) on the 4-jet of f at the origin, that is, only depends on b and c of (11). Then we do an explicit computation of the Dirichlet energy for a particular 2-parameter family of regions, whose corresponding maps f have 4-jets covering every possible value of (b,c).

5.1. Dependence on 4-jet

Since in §2.4 we computed the dependence on δ of E_δ, we can assume without loss of generality in this section that δ is small compared to ε.

Let h_j denote the limiting average height function for P_j, and h_{j+1} be the limiting average height function for P_{j+1}.

LEMMA 12. *Up to an additive constant we have* $h_j\circ f_j=(2/\pi)\operatorname{Im}\log f_j'(z)$. *In particular,* $h_j=h_j\circ f_j\circ f_j^{-1}=-(2/\pi)\operatorname{Im}\log(f_j^{-1})'(z)$.

The proof of the first statement follows from the definition of h_j in the last paragraph of §2.3. The second statement follows trivially from the first.

Let g_j be a harmonic conjugate on U_j of h_j, so that $\tilde{h}_j:=h_j+ig_j$ is analytic on U_j. Similarly let g_{j+1} be a harmonic conjugate for h_{j+1} on U_{j+1}.

Since the boundary of U_j is a subset of the boundary of U_{j+1}, the change in normalized Dirichlet energy can be written (recall that U_j' is the region U_j minus a δ-neighborhood of its vertices)

$$\oint_{\partial U_{j+1}'} h_{j+1}\,dg_{j+1}-\oint_{\partial U_j'} h_j\,dg_j=\oint_{\partial U_j'}(h_{j+1}\,dg_{j+1}-h_j\,dg_j)+\oint_X h_{j+1}\,dg_{j+1} \tag{17}$$

$$=\oint_{\partial U_j'}(h_{j+1}-h_j)\,dg_{j+1}-\oint_{\partial U_j'} h_j\,d(g_{j+1}-g_j)+\oint_X h_{j+1}\,dg_{j+1} \tag{18}$$

where $X=\partial U_{j+1}'-\partial U_j'$ is the path which consists of the four pieces

$$\partial N_\delta(0)\cup[-\delta,2\varepsilon+\delta]\cup\partial N_\delta(2\varepsilon)\cup[-\delta,2\varepsilon-\delta]^*$$

where the superscript * is a reminder that the second segment is traced in the reverse direction.

We will show that each of the three integrals of (18) depends (up to controlled errors) only on the 4-jets of f_j and f_{j+1}.

The first integral in (18) can be estimated as follows. On the boundary of U_j we have $h_{j+1}=h_j$, and for each corner c of U_{j+1} we have $dg_{j+1}=O(\delta)$ on $\partial N_\delta(c)$ (see §2.4). It remains to consider the boundary $\partial N_\delta(0)$, where 0 is the tip of the cut in U_j (which is not a corner of U_{j+1}). The x-axis divides $\partial N_\delta(0)$ into two parts. Since g_{j+1} is smooth on each half, when $\delta\to 0$ the integral of $(h_{j+1}-h_j)\,dg_{j+1}$ on each half tends to zero. In conclusion the first integral in (18) tends to zero with δ.

To compute the second integral in (18), we show that we can replace the path of integration by a smaller path of radius $O(\varepsilon)$; on this new path we will be able to replace dg_j and dg_{j+1} with their 4-jets.

Now $h_{j+1}-h_j$ is the harmonic function which (after choosing an appropriate base point) is zero on ∂U_j, and 1 and -1 on the upper and lower side of $[0,2\varepsilon]$ respectively. On ∂U_j we therefore have

$$g_{j+1}-g_j=\operatorname{Im}(\tilde{h}_{j+1}-\tilde{h}_j)=-i(\tilde{h}_{j+1}-\tilde{h}_j)$$

since the real part of this function vanishes.

The second integral in (18) can then be written (using $h_j=-(2/\pi)\operatorname{Im}\log(f_j^{-1})'(z)$ from Lemma 12)

$$\oint_{\partial U'} h_j\,d(g_{j+1}-g_j)=-\frac{2}{\pi}\operatorname{Im}\oint_{\partial U'}\log(f_j^{-1})'(z)(-i)(\tilde{h}_{j+1}-\tilde{h}_j)'\,dz, \tag{19}$$

where the $'$ refers to the z-derivative. This integral is the imaginary part of a contour integral. The integrand is analytic on $U'-[0,2\varepsilon]$; in particular, it is analytic on U' minus a neighborhood of the origin of radius 3ε. Therefore we can replace the path of integration by a path which winds around the boundary of the slit disk $D_{\text{slit}}=\{z\in U':|z|<3\varepsilon\}$.

Now on D_{slit} we have (see (12))

$$\log(f_j^{-1})'(z)=\log\left(\frac{1}{2\sqrt{z}}-\frac{b}{2}+\gamma\sqrt{z}+zO_z(1)\right)=\log((f_{j,\{4\}}^{-1})')(z)+z^{3/2}O_z(1)$$

where $f_{j,\{4\}}$ is the 4-jet of f_j at the origin (γ is a constant depending on b and c only). Similarly

$$\begin{aligned}\log(f_{j+1}^{-1})'(z)&=\log\left(\frac{1}{2\sqrt{z-2\varepsilon}}-\frac{b_{j+1}}{2}+\gamma_{j+1}\sqrt{z-2\varepsilon}+O(z-2\varepsilon)\right)\\&=\log((f_{j+1,\{4\}}^{-1})')(z)+(z-2\varepsilon)^{3/2}O_{z-2\varepsilon}(1),\end{aligned}$$

where we recall that since we are using coordinates in U_j,

$$f_{j+1}(z)=2\varepsilon+z^2+b_{j+1}z^3+c_{j+1}z^4+\dots\,.$$

These equations give, on ∂D_{slit},

$$dg_j=-\frac{2}{\pi}\operatorname{Re}\big(d\log((f_{j,\{4\}}^{-1})')+\sqrt{\varepsilon}\,O_z(1)\,dz\big)=dg_{j,\{4\}}+\sqrt{\varepsilon}\,O_z(1)\,dz \tag{20}$$

and similarly $dg_{j+1}=dg_{j+1,\{4\}}+\sqrt{\varepsilon}\,O_z(1)\,dz$.

Having found dg_j and dg_{j+1}, the integral (19) becomes

$$\oint h_j(dg_{j+1}-dg_j)=\oint_{\partial D_{\text{slit}}}(h_{j,\{4\}}+z^{3/2}O_z(1))\big(dg_{j+1,\{4\}}-dg_{j,\{4\}}+\sqrt{\varepsilon}\,O(1)\,dz\big). \tag{21}$$

Since h_j is bounded,

$$\sqrt{\varepsilon}\, O(1) \oint h_{j,\{4\}}\, dz = O(\varepsilon^{3/2}),$$

and similarly on the path of integration the only singularity of $dg_{j+1,\{4\}}$ or $dg_{j,\{4\}}$ occurs near the origin, where the functions are at worst $O(z^{-1})$, so

$$\oint z^{3/2} O_z(1)(dg_{j+1,\{4\}} - dg_{j,\{4\}}) = \oint z^{1/2} O_z(1)\, dz = O(\varepsilon^{3/2}).$$

Thus we can pull the errors in (21) out of the integral, giving $O(\varepsilon^{3/2})$. Therefore the integral (19) only depends on $b_j, c_j, b_{j+1}, c_{j+1}$ up to an error $O(\varepsilon^{3/2})$.

Lastly we estimate the third integral in (18). As was the case for the first integral, on the two halves of $N_\delta(0)$ the integral of dg_{j+1} tends to zero with δ. Now on $N_\delta(2\varepsilon)$ we have $dg_{j+1} = O(\delta)$, and h_{j+1} is constant (1 or -1) on the two "sides" of $[0, 2\varepsilon]$; so

$$\begin{aligned} \oint_X h_{j+1}\, dg_{j+1} &= o(1) + \int_\delta^{2\varepsilon-\delta} dg_{j+1}^+ - \int_{2\varepsilon-\delta}^{\delta} dg_{j+1}^- \\ &= o(1) + g_{j+1}(\delta)^+ - g_{j+1}(2\varepsilon-\delta)^+ - g_{j+1}(2\varepsilon-\delta)^- + g_{j+1}(\delta)^-, \end{aligned} \tag{22}$$

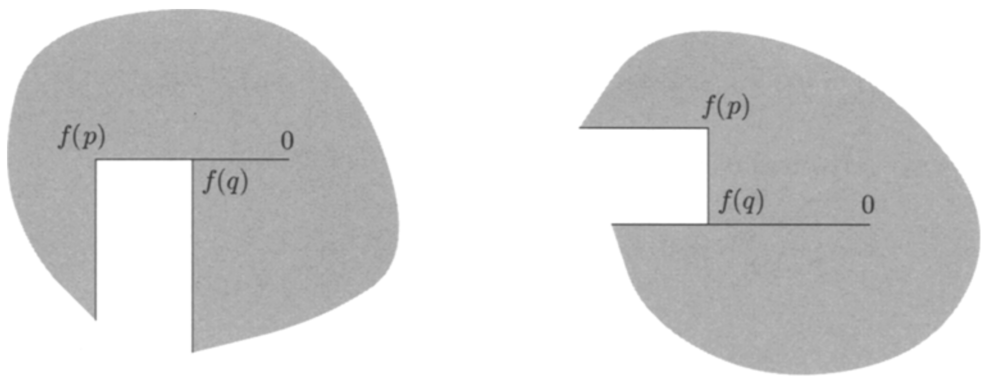
where the $+$ and $-$ refer to the two limits from above and below the axis, and the $o(1)$-term tends to zero with δ. Using $g_{j+1}(z) = (2/\pi)\,\mathrm{Re}\log(f_{j+1}^{-1})'(z)$, we see that the values at δ and $2\varepsilon - \delta$ of g_{j+1} depend only on the 4-jet up to an error $O(\varepsilon^{3/2})$.

We have shown that the change in energy depends only on the 4-jet of f_j and f_{j+1}. Now when j is not within K of 0 or N, $b_{j+1} = b_j + O(\varepsilon)$ and $c_{j+1} = c_j + O(\varepsilon)$ by Lemma 11. Changing b_{j+1} and c_{j+1} by $O(\varepsilon)$ changes the second and third integrals (21) and (22) by at most $O(\varepsilon^{3/2})$; so when $K < j < N - K$ the change in energy in fact depends only on the 4-jet of f_j.

When $j < K$ or $N - K < j$, the energy depends on the 4-jet of *both* f_j and f_{j+1}. In these cases, however, both f_j and f_{j+1} are independent of U up to $O(1)$; in the next section we will see that the energy depends only on j (through the Schwarzian derivative of $\sqrt{f_j}$) up to an error $o(1/j)$. This completes the proof that the change in Dirichlet energy only depends on the 4-jet of f at the origin, up to terms in $o(\varepsilon)$.

5.2. Computation for a specific family of functions

At this point we could simply work out the integrals (21) and (22), but it is less computationally painful to do an explicit calculation for a particular family of functions. This will also give a more accurate estimate of the energy when j is near 0 or N.



(a) (b)

Fig. 4. The image of RHP under the function $f_{p,q}$ is shown in gray.

For $p,q\in i\mathbf{R}$ and $\{0,p,q\}$ distinct, define

$$f_{p,q}(z)=2\sqrt{\frac{q}{p}}\int_0^z u\sqrt{\frac{u-p}{u-q}}\,du. \tag{23}$$

If $-ip>0>-iq$ and $|p|>|q|$ this is a map from RHP injectively onto the region $U_{p,q}$ shown in gray in Figure 4(a); this follows from the Schwarz–Christoffel formula [Ah] (f maps RHP to the polygonal region with angle 2π at the origin, $\frac{3}{2}\pi$ at $f(p)$, $\frac{1}{2}\pi$ at $f(q)$, and 2π at ∞. If $-ip>0>-iq$ and $|p|<|q|$ this is still a well-defined map which is not injective on RHP (although it is locally injective). This lack of injectivity is of no consequence to us. When p and q are on the same side of the origin the image is as shown in Figure 4(b). Again the map may or may not be injective according to whether or not p is between q and 0.

The 4-jet of $f_{p,q}$ at the origin is

$$f_{p,q}(z)=z^2+\left(\frac{1}{q}-\frac{1}{p}\right)\frac{z^3}{3}+\left(\frac{3}{4q^2}-\frac{1}{2pq}-\frac{1}{4p^2}\right)\frac{z^4}{4}+O(z^5). \tag{24}$$

For any f of the form (11) there are p,q for which $f_{p,q}$ has the same 4-jet at the origin as f, provided that $b\neq 0$: it suffices to take

$$p=\frac{12b}{16c-27b^2} \quad \text{and} \quad q=\frac{12b}{16c+9b^2}.$$

Note that when p,q take these values, two of the points $\{0,p,q\}$ are equal only when $b=0$ (the cases when one of p,q is infinite are allowed). The case $b=0$ will be dealt with later.

The integral in (23) can be explicitly evaluated, giving

$$f_{p,q}(z)=\frac{1}{4}\sqrt{\frac{q}{p}}\left(\sqrt{(z-p)(z-q)}\,(4z-2p+6q)-2(p^2+2pq-3q^2)\log\left(\sqrt{z-p}+\sqrt{z-q}\,\right)\right)-C_2$$

where C_2 is chosen so that $f_{p,q}(0)=0$:

$$C_2=\frac{1}{4}\sqrt{\frac{q}{p}}\left(\sqrt{pq}\,(-2p+6q)-(p^2+2pq-3q^2)\log\left(-(\sqrt{p}-\sqrt{q}\,)^2\right)\right).$$

We have explicitly

$$f(p)=-\frac{1}{4}q(2p-6q)+\frac{1}{4}\sqrt{\frac{q}{p}}\,(p^2+2pq-3q^2)\log\left(\frac{\sqrt{p}+\sqrt{q}}{\sqrt{q}-\sqrt{p}}\right)$$

and

$$f(q)=-\frac{1}{4}q(2p-6q)+\frac{1}{4}\sqrt{\frac{q}{p}}\,(p^2+2pq-3q^2)\log\left(\frac{\sqrt{p}+\sqrt{q}}{\sqrt{p}-\sqrt{q}}\right).$$

Thus

$$f(p)-f(q)=\frac{\pi i}{4}\sqrt{\frac{q}{p}}\,(p^2+2pq-3q^2) \tag{25}$$

where we chose the sign of the square root and branch of log to correspond to the situation of Figure 4.

Suppose $-ip>0>-iq$. Then the δ-normalized Dirichlet energy for the region $U_{p,q}=f_{p,q}(\mathrm{RHP})$ is $\oint_{U'}h\,dg$, where h is 0 on the vertical boundary from $-i\infty$ to $f(q)$, -1 on the horizontal boundary from $f(q)$ to 0, 1 on the boundary from 0 to $f(p)$, and 2 on the vertical boundary from $f(p)$ to ∞ (note that h has different values on the two "sides" of $[f(q),0]$).

To compute E_δ, we pull h back to RHP. The preimage of $N_\delta(f(p))$ is to first order a disk around p (that is, it converges to a round disk when $\delta\to 0$); let δ_p denote its radius. Similarly let δ_q and δ_0 be the radii of the preimages of the disks around $f(q)$ and $f(0)$ respectively. Let δ_∞ be defined by: $1/\delta_\infty$ is the radius of the preimage of the disk of radius $1/\delta$ around 0.

Then (recalling that g is constant up to lower-order terms on the boundary of the δ-neighborhood of the singularities)

$$\begin{aligned}E_\delta&=-\int_{q+i\delta_q}^{-i\delta_0}d(g\circ f_{p,q})+\int_{\delta_0}^{p-i\delta_p}d(g\circ f_{p,q})+2\int_{p+i\delta_p}^{1/\delta_\infty}d(g\circ f_{p,q})\\&=-2g(f_{p,q}(\delta_0))+g(f_{p,q}(q+i\delta_q))-g(f_{p,q}(p+i\delta_p))+2g(f_{p,q}(1/\delta_\infty)).\end{aligned} \tag{26}$$

From (24) we have $\delta_0^2 = \delta$ up to higher-order terms. For z close to p we have from (23)

$$f_{p,q}(z) = f_{p,q}(p) + \frac{4}{3}\frac{\sqrt{qp}}{\sqrt{p-q}}(z-p)^{3/2} + O((z-p)^2),$$

and so up to higher-order terms we have

$$\delta_p = \left(\frac{3}{4}\delta\sqrt{\frac{p-q}{pq}}\right)^{2/3}. \tag{27}$$

Similarly near q,

$$f_{p,q}(z) = f_{p,q}(q) + 4\sqrt{\frac{q}{p}}\,q\sqrt{q-p}\,(z-q)^{1/2} + O(z-q),$$

and so

$$\delta_q = \left(\frac{\delta}{4}\sqrt{\frac{p}{q}}\frac{1}{q\sqrt{q-p}}\right)^2. \tag{28}$$

For large z we have $f'_{p,q}(z) = 2\sqrt{q/p}\,z + O(1)$ which implies $f_{p,q}(z) = \sqrt{q/p}\,z^2 + O(z)$ and so

$$\delta_\infty = \left(\delta\sqrt{\frac{q}{p}}\right)^{1/2}. \tag{29}$$

Plugging these into (26) with $g \circ f_{p,q} = -(2/\pi)\operatorname{Re}\log f'_{p,q}$ gives the Dirichlet energy to be

$$\begin{aligned} E_\delta = -2\left(-\frac{2}{\pi}\operatorname{Re}\log\left(\delta^{1/2}\sqrt{\frac{p}{q}}\right)\right) - \frac{2}{\pi}\operatorname{Re}\log\left(\frac{q\sqrt{q-p}}{\frac{\delta}{4}\sqrt{\frac{p}{q}}\frac{1}{q\sqrt{q-p}}}\right) \\ + \frac{2}{\pi}\operatorname{Re}\log\left(\frac{p}{\sqrt{p-q}}\left(\frac{3}{4}\delta\sqrt{\frac{p-q}{pq}}\right)^{1/3}\right) - \frac{4}{\pi}\operatorname{Re}\log\left(\frac{1}{\delta^{1/2}}\left(\frac{q}{p}\right)^{1/4}\right) \end{aligned} \tag{30}$$

$$= \frac{20}{3\pi}\log\delta + \frac{1}{3\pi}\log\left(\frac{9p^{11}}{4^8 q^{19}|p-q|^8}\right). \tag{31}$$

By symmetry we get the same energy when $-ip < 0 < -iq$. A similar calculation shows that the same energy is obtained in the remaining cases when p and q are on the same side of the origin.

When we extend the cut by a small amount 2ε, the new region has a new uniformizing function of the same form $f_{p',q'}$, where p' and q' are defined by changing p and q so that $f(p)$ and $f(q)$ each change by -2ε. Let dp and dq denote the changes in p and q for an infinitesimal ε. We must first have $0=d(f(p)-f(q))$. This relates the change in p to the change in q. From (25) we have the equation

$$0=d(f(p)-f(q))=\left(\sqrt{\frac{q}{p}}\,(2p+2q)+(p^2+2pq-3q^2)\left(\frac{-\sqrt{q}}{2p^{3/2}}\right)\right)dp$$
$$+\left(\sqrt{\frac{q}{p}}\,(2p-6q)+\frac{1}{\sqrt{pq}}\,(p^2+2pq-3q^2)\right)dq,$$

or

$$dq=-\frac{q}{p}\left(\frac{3p^2+2pq+3q^2}{p^2+6pq-15q^2}\right)dp.$$

Since $\sqrt{q/p}\,(p^2+2pq+3q^2)$ does not change, we have

$$d(f(p))=-\frac{q}{2}\,dp+\left(3q-\frac{1}{2}p\right)dq+\frac{1}{4}\sqrt{\frac{q}{p}}\,(p^2+2pq-3q^2)$$
$$\times\left(\left(\frac{1}{\sqrt{p}+\sqrt{q}}-\frac{1}{\sqrt{p}-\sqrt{q}}\right)\frac{dp}{2\sqrt{p}}+\left(\frac{1}{\sqrt{p}+\sqrt{q}}+\frac{1}{\sqrt{p}-\sqrt{q}}\right)\frac{dq}{2\sqrt{q}}\right)$$
$$=\frac{-16pq^2}{p^2+6pq-15q^2}\,dp.$$

Now the change in E_δ when $d(f(p))=-2\varepsilon$ and $d(f(p))=d(f(q))$ is

$$-2\varepsilon\frac{dE_\delta}{dp}\frac{dp}{d(f(p))}=\frac{4(5p+7q)(p-q)}{\pi p(p^2+6pq-15q^2)}\cdot\frac{p^2+6pq-15q^2}{-16pq^2}=\frac{\varepsilon(5p+7q)(p-q)}{2\pi p^2q^2}.$$

Finally the Schwarzian derivative of $\sqrt{f_{p,q}}$ is

$$\frac{3}{4}\left(\frac{3}{4q^2}-\frac{1}{2pq}-\frac{1}{4p^2}\right)-\frac{1}{4}\left(\frac{1}{q}-\frac{1}{p}\right)^2=\frac{(5p+7q)(p-q)}{16p^2q^2}$$

which is $\frac{1}{8}\pi$ times the change in E_δ. This completes the proof when $j\in[K,N-K]$ except in the case $b=0$. For the case $b=0$ see the function f_q of §5.3.1 below. The case $j\notin[K,N-K]$ is dealt with in the next section.

5.3. Beginning and ending of a cut

As above, the change in energy near the beginning of a cut only depends on the germ of f near the tip of the cut. Near the beginning of the cut, the germ of f only depends on whether or not the cut starts at a corner or at an edge (Lemma 11). We computed the limiting functions f when j is near the beginning or ending of the cut in the proof of Lemma 6.

5.3.1. *Beginning on an edge.* Suppose first that the cut begins on a straight edge of U. We compute the change in E_δ between the time when there is no cut to the time when the cut has length $2\varepsilon j$.

Let h_q be the function on $\mathrm{RHP}-[0,q]$ whose boundary values are 0 on the y-axis, 1 on the upper boundary of the cut $[0,q]$, and -1 on the lower boundary of the cut. Up to an additive constant, this is the height function on U near the beginning of the cut.

We can compute the δ-normalized Dirichlet energy of h_q in the same way as we did for the functions $f_{p,q}$ of the previous section. Specifically, the map from RHP to $U_q=\mathrm{RHP}-[0,q]$ is

$$f_q(z)=\sqrt{2qz^2+q^2}-q=z^2-\frac{1}{2q}z^4+O_z(z^6),$$

and the Dirichlet energy is

$$2g\circ f_q(\delta_0)-g\circ f_q\big(\sqrt{-q/2}+\delta_1\big)-g\circ f_q\big(-\sqrt{-q/2}+\delta_2\big), \tag{32}$$

where (to first order) f_q maps the δ_0-neighborhood of the origin to the δ-neighborhood of the origin, and the neighborhoods of $\pm\sqrt{-q/2}$ of radius δ_1 to the δ-neighborhood of $-q$.

As before $\delta_0=\delta^{1/2}$; one computes $\delta_1=\delta_2=\delta^2/4q\sqrt{q/2}$. Plugging into (32) with $g\circ f_q=-(2/\pi)\,\mathrm{Re}\log f_q'$ gives

$$E_\delta=\frac{6}{\pi}\log\frac{1}{\delta}+\frac{6}{\pi}\log q+\frac{2}{\pi}\log 2. \tag{33}$$

When $q=2\varepsilon j$, that is, $U_q=\mathrm{RHP}-[0,2\varepsilon j]$, this gives the energy change due to the first j steps of the cut. The change in Dirichlet energy between j and $j+1$ is $6/\pi j$, which is $-(8\varepsilon/\pi)S\sqrt{f_j}(0)$ as required.

Note that if we set $\delta=\varepsilon$ in (33), the ε-normalized energy due to the beginning of the cut is

$$\frac{6}{\pi}\log j+\mathrm{const}+O(\varepsilon).$$

In particular, when $j=K=\xi/\varepsilon$ this is $(6/\pi)\log(1/\varepsilon)+O(1)$.

5.3.2. *Starting at a concave corner.* In case the cut starts at a corner of U (of angle $\frac{3}{2}\pi$), the change in energy can be computed in a similar manner. Again we compute the change in energy between the time when there is no cut to the time when the cut has length $2\varepsilon j$.

Let TQP be the three-quarters plane, $\mathrm{TQP}=\{(x,y):x>0 \text{ or } y>0\}$, and define $U_q=\mathrm{TQP}-[0,q]$. We first compute E_δ on TQP. We can take the average height function h on TQP to be 1 on the negative x-axis, and 0 on the negative y-axis, that is,

$h = \text{const} + (2/3\pi)\operatorname{Im}\log(z)$. Then $E_\delta(h)$ on TQP is $\oint h\,dg = g(\delta) - g(1/\delta)$, and up to higher-order terms $g(\delta) = (2/3\pi)\log(1/\delta)$ and $g(1/\delta) = -(2/3\pi)\log(1/\delta)$. So the energy is $(4/3\pi)\log(1/\delta)$.

The map from RHP to U_q is

$$f_q(z) = 2\sqrt{-q}\int_0^z \frac{w\,dw}{\sqrt{w-q}} = 2\sqrt{-q}\left(\frac{2}{3}(w-q)^{3/2} + 2q(w-q)^{1/2}\right) + \frac{8q^2}{3}.$$

At the origin we have

$$f_q(z) = z^2 + \frac{1}{3q}z^3 + \frac{3}{16q^2}z^4 + O(z^5).$$

Now

$$E_\delta = \oint_{U_q'} h\,dg = -g\circ f_q(1/\delta_\infty) - g\circ f_q(q + i\delta_q) + 2g\circ f_q(\delta_0).$$

Again $\delta_0 = \delta^{1/2}$, and near ∞,

$$\delta^{-1} = f(1/\delta_\infty) = \tfrac{4}{3}\sqrt{-q}\,\delta_\infty^{-3/2}.$$

Near q we have $f(z) = f(q) + 4q\sqrt{-q}\sqrt{z-q} + O(z-q)$, so

$$|4q\sqrt{-q}\,|\delta_q^{1/2} = \delta.$$

Plugging all this in we have

$$\begin{aligned} E_\delta(h_q) &= \frac{2}{\pi}\log\left(2\sqrt{q}\left(\frac{3}{4\sqrt{q}\,\delta}\right)^{1/3}\right) + \frac{2}{\pi}\log\left(\frac{8q^3}{\delta}\right) - 2\frac{2}{\pi}\log(2\sqrt{\delta}) \\ &= \frac{14}{3\pi}\log\frac{1}{\delta} + \frac{20}{3\pi}\log q + \text{const}. \end{aligned}$$

The difference in energy between U_q and TQP is then

$$\frac{10}{3\pi}\log\frac{1}{\delta} + \frac{20}{3\pi}\log q + \text{const}.$$

Now $f(q) = \frac{8}{3}q^2$, which is $-2\varepsilon j$ when $q = \sqrt{3\varepsilon j/4}$. Plugging this value of q in gives the energy due to the cut of P_j to be

$$\frac{10}{3\pi}\log\frac{1}{\delta} + \frac{10}{3\pi}\log(\varepsilon j) + \text{const}. \tag{34}$$

When j changes by 1 the energy changes by $10/3\pi j$, which is again $-(8\varepsilon/\pi)S\sqrt{f_q}(0)$.

Setting $\delta = \varepsilon$ in (34) gives the contribution to E_ε from the beginning of the cut as $(10/3\pi)\log j + \text{const} + O(\varepsilon)$.

5.4. Ending on the interior of an edge

A similar computation holds in this case, with f_q as in the proof of Lemma 6. There is, however, a shorter method, obtained as follows.

The change in E_δ near the end of a cut only depends on the local structure of the average height function near the end of the cut. Therefore the energy can be obtained from the known value of the energy for a rectangle. Indeed, cutting a rectangle into two rectangles with a single edge, the change in E_ε due to the beginning of the cut, plus the change due to the ending of the cut, plus the contribution from the "central" terms, must equal total energy change which is $(24/\pi)\log(1/\varepsilon)+\text{const}+O(\varepsilon)$ (see (36)). The contribution at the beginning of the cut is $(6/\pi)\log(1/\varepsilon)+\text{const}+O(\varepsilon)$ (see §5.3.1), so the contribution at the end of the cut is

$$\frac{18}{\pi}\log\frac{1}{\varepsilon}+\text{const}+O(\varepsilon)=-\frac{48}{\pi}\left(-\frac{3}{8}\log\frac{1}{\varepsilon}+\text{const}+O(\varepsilon)\right).$$

The constant is independent of U by Lemma 11.

5.5. Ending at a concave corner

The energy for the end of a cut ending at a corner of U can be computed from the previous case. Given an L-shaped region, cut it into two rectangles with a single cut beginning on an edge and ending at the corner. We can compute the change in Dirichlet energy by starting the cut at the concave corner and ending at the edge; the contribution from the start of the cut starting at the concave corner is (cf. (34)) $(10/3\pi)\log(1/\varepsilon)+\text{const}+O(\varepsilon)$, and the contribution from the end of this cut is $(18/\pi)\log(1/\varepsilon)+\text{const}+O(\varepsilon)$ by the previous paragraph. The sum of these must equal the contribution for starting from the edge $((6/\pi)\log(1/\varepsilon)+\text{const}+O(\varepsilon))$ and ending at the concave corner. Therefore the contribution due to a cut ending at a concave corner is

$$\frac{46}{3\pi}\log\frac{1}{\varepsilon}+\text{const}+O(\varepsilon)=-\frac{48}{\pi}\left(-\frac{23}{72}\log\frac{1}{\varepsilon}+\text{const}+O(\varepsilon)\right).$$

This completes the proof of Lemma 7.

6. The case of rectangles

Let P be a $((2m-1)\times(2n-1))$-rectangle with a unit square removed at the lower left corner. By Temperley's trick [T], domino tilings of P are in bijection with spanning trees of the $(m\times n)$-grid graph, rooted at the lower left corner.

The number of spanning trees of the $(m \times n)$-grid, which by Kirchhoff [Ki] is $1/mn$ times the product of the non-zero eigenvalues of the Laplacian, is

$$\frac{1}{mn}\prod_{(j,k)\neq 0}\left(4-2\cos\left(\frac{\pi k}{n}\right)-2\cos\left(\frac{\pi j}{m}\right)\right),$$

where j (resp. k) runs from 0 to $m-1$ (resp. 0 to $n-1$). The log of this formula was asymptotically computed in [DD]:

PROPOSITION 13 ([DD]). *The* log *of the number of spanning trees of an* $(n\times m)$-*rectangle is*

$$\frac{4Gmn}{\pi}+(m+n)\log(\sqrt{2}-1)-\frac{1}{2}\log(m)+\log(\eta(e^{-2\pi n/m}))-\frac{1}{4}\log(2)+O\left(\frac{1}{mn}\right).$$

Here G is Catalan's constant, and η is the Dedekind eta-function

$$\eta(q)=q^{1/24}\prod_{k=1}^{\infty}(1-q^k).$$

The area of P is $(2n-1)(2m-1)-1=4nm-2n-2m$, and the perimeter is $4n+4m-4$, so this expression may be rewritten

$$\frac{G}{\pi}\operatorname{Area}(P)+\left(\frac{G}{2\pi}+\frac{\log(\sqrt{2}-1)}{4}\right)\operatorname{Perim}(P)-\frac{1}{2}\log(m)+\log(\eta(e^{-2\pi n/m}))+C+O(m^{-2}), \tag{35}$$

where

$$C=\log\left(\frac{2^{5/4}}{1+\sqrt{2}}\right)+\frac{2G}{\pi}.$$

6.1. Dirichlet energy for a rectangle

Let U be the rectangle $[0,\frac{1}{2}]\times[0,\frac{1}{2}\tau]$ in $\mathbf{C}$, with base point at the lower left corner. The average height function h for U is, up to an additive constant, the harmonic function which is $0,1,2,3$ on the lower, right, upper and left boundaries respectively.

The function h has an explicit expression in terms of the the Weierstrass elliptic function $\wp(z)=\wp_{1,i\tau}(z)$ (see [Ah]: this is the doubly periodic function with periods 1 and $i\tau$ and a double pole at each point of the lattice).

LEMMA 14. *Up to an additive constant we have*

$$h(z)=-\frac{2}{\pi}\operatorname{Im}\log\wp'(z).$$

Proof. From [Ah] we have

$$\wp'(z)=\sum_{w\in\Gamma}\frac{-2}{(z-w)^3},$$

where $\Gamma=\mathbf{Z}+i\tau\mathbf{Z}$. Note that $\wp'(z)$ is real when $z\in[0,\frac{1}{2}]$ or $z\in[0,\frac{1}{2}]+\frac{1}{2}i\tau$, and pure imaginary when $z\in i[0,\frac{1}{2}\tau]$ or $z\in i[0,\frac{1}{2}\tau]+\frac{1}{2}$. Furthermore on a fundamental domain for Γ, $\wp'(z)$ is zero or infinite only at the corners of U (see [Ah]). So the argument of $\wp'(z)$ is constant on each edge of the rectangle U. At a corner z_j, $z_j\neq 0$, we have $\wp'(z)=c_j(z-z_j)+O((z-z_j)^2)$, and so $(-2/\pi)\,\mathrm{Im}\log\wp'(z)$ changes by $+1$ at each corner of U when going counterclockwise around U. □

Let U' be U minus the δ-neighborhood of the four corners of the rectangle. Then

$$E_\delta(h)=\int_{\partial U'}h\,dg=\int_{\delta}^{1/2-\delta}0\,dg+\int_{1/2+i\delta}^{(1+i\tau)/2-i\delta}1\,dg+\int_{(1+i\tau)/2-\delta}^{i\tau/2+\delta}2\,dg+\int_{i\tau/2-i\delta}^{i\delta}3\,dg+O(\delta)$$

where the contribution on the boundaries of the neighborhoods of the corners is $O(\delta)$. Since when δ is small g is essentially constant on the neighborhoods of radius δ of the corners, this energy is

$$E_\delta(h)=-g\big(\tfrac{1}{2}-\delta\big)-g\big(\tfrac{1}{2}+\tfrac{1}{2}i\tau-\delta\big)-g\big(\tfrac{1}{2}i\tau+\delta\big)+3g(\delta)+O(\delta).$$

At the origin, $\wp'(z)=-2/z^3+O(1)$, and so

$$\frac{2}{\pi}\log\wp'(\delta e^{i\theta})=\frac{2}{\pi}\log\left(\frac{-2}{\delta^3e^{3i\theta}}+\mathrm{const}+O(\delta)\right)=\frac{6}{\pi}\log\left(\frac{1}{\delta}\right)-\frac{6\theta i}{\pi}+\frac{2}{\pi}\log(-2)+O(\delta^3).$$

Near the other three corners of U,

$$\frac{2}{\pi}\log\wp'(\delta e^{i\theta}-z_j)=-\frac{2}{\pi}\log(c_j\delta e^{i\theta})+O(\delta)$$

where $c_j=\wp''(z_j)$. Using $g(z)=(2/\pi)\,\mathrm{Re}\log\wp'(z)$, the energy is

$$E_\delta(h)=-\frac{2}{\pi}(\log(c_1\delta)+\log(c_2\delta)+\log(c_3\delta))+\frac{18}{\pi}\log\left(\frac{1}{\delta}\right)=\frac{24}{\pi}\log\left(\frac{1}{\delta}\right)-\frac{2}{\pi}\log(c_1c_2c_3),$$

where c_1,c_2,c_3 are the derivatives $\wp''(z_j)$ at the three other corners of U.

LEMMA 15. *We have*

$$c_1c_2c_3=\tfrac{1}{2}(2\pi)^{12}\eta(e^{-2\pi\tau})^{24}.$$

Proof. From the differential equation

$$(\wp'(z))^2 = 4(\wp(z)-e_1)(\wp(z)-e_2)(\wp(z)-e_3)$$

(where $e_j = \wp(z_j)$) we obtain (upon differentiating the logarithms of both sides)

$$2\frac{\wp''(z)}{\wp'(z)} = \frac{\wp'(z)}{\wp(z)-e_1} + \frac{\wp'(z)}{\wp(z)-e_2} + \frac{\wp'(z)}{\wp(z)-e_3},$$

from which we get

$$\wp''(z_1) = 2(e_1-e_2)(e_1-e_3),$$

with similar expressions for c_2, c_3. Their product is

$$c_1c_2c_3 = 8(e_1-e_2)^2(e_1-e_3)^2(e_2-e_3)^2,$$

and by [Ap] this equals

$$\tfrac{1}{2}\Delta = \tfrac{1}{2}(2\pi)^{12}\eta(e^{-2\pi\tau})^{24}. \qquad \square$$

The energy is therefore

$$\frac{24}{\pi}\log\left(\frac{1}{\delta}\right) - \frac{2}{\pi}\log\left(\frac{1}{2}(2\pi)^{12}\eta(e^{-2\pi\tau})^{24}\right).$$

This is the δ-energy for a $\left(\frac{1}{2}\times\frac{1}{2}\tau\right)$-rectangle.

Let $U_{\alpha,\beta}$ be the rectangle $[0,\alpha]\times[0,\beta]$, where $\tau=\beta/\alpha$. We can scale $U_{\alpha,\beta}$ by $1/2\alpha$ to get U. As a consequence, the above expression is the $2\alpha\delta$-energy for $U_{\alpha,\beta}$. To get the δ-energy for $U_{\alpha,\beta}$, substitute $\delta/2\alpha$ for δ in the above expression. We get

$$E_\delta(U_{\alpha,\beta}) = \frac{24}{\pi}\log\frac{2\alpha}{\delta} - \frac{2}{\pi}\log\left(\frac{1}{2}(2\pi)^{12}\eta(e^{-2\pi\tau})^{24}\right). \tag{36}$$

PROPOSITION 16. *Let U be a rectangle; let P_ε be a sequence of Temperleyan rectangles of area A_ε and perimeter* $\mathrm{Perim}_\varepsilon$ *approximating U. Then the* log *of the number of domino tilings of P_ε is*

$$\frac{GA_\varepsilon}{\pi\varepsilon^2} + \left(\frac{G}{2\pi} + \frac{\log(\sqrt{2}-1)}{4}\right)\frac{\mathrm{Perim}_\varepsilon}{\varepsilon} - \frac{\pi}{48}r_2(\varepsilon,U) + C + O(\varepsilon^2),$$

where $r_2(\varepsilon,U)$ is the ε-normalized Dirichlet energy of the average height function (given by putting ε for δ in (36)*) and C is a universal constant.*

Proof. If U is an $(\alpha\times\beta)$-rectangle, this follows from (35) upon setting $2m-1=\alpha/\varepsilon$, $2n-1=\beta/\varepsilon$. $\square$

This proposition tells us why the 'natural' choice of δ in $E_\delta(h)$ is $\delta=\varepsilon$ or some constant multiple of ε. In fact there is a multiple which makes the constant C in the above formula vanish, although we will not need this accuracy.

7. Loop-erased random walk

The loop-erased random walk in a finite graph G has the following simple description. Let b_0, b_1 be two vertices of G. Take a simple random walk starting from b_0 and stopping at b_1. Erase from the path its loops, in chronological order. That is, if there is a loop, erase the first loop (the first time the path comes back to the same vertex twice). For the new path, if there is still a loop, erase the first loop, and so on. The remaining path is a simple path from b_0 to b_1.

There is a well-known connection between the LERW and spanning trees [P]: in a uniformly chosen spanning tree on a region $P \subset \mathbf{Z}^2$, the unique arc (branch) from a to b has the same distribution as the LERW from a to b on the same region. Pemantle also showed that on an infinite domain such as $\mathbf{Z}^2$ or the upper half of $\mathbf{Z}^2$, the LERW is well-defined (despite recurrence of the simple random walk, see [P]).

7.1. LERW and dominos

Let P be a $((2n-1)\times(2n-1))$-square Temperleyan polyomino with base square b_0 on the center of the right edge. Recall how Temperley's trick works: a domino tiling of P corresponds to a spanning tree of the graph $H(P)$ which is rooted at b_0 in the following way. Each domino covering a vertex v of $H(P)$ has white vertex covering the center of an adjacent edge e of $H(P)$. In the associated spanning tree the outgoing edge from v points in the direction e. The union of these edges forms a spanning tree with all edges oriented towards the root at b_0.

Let Q be the region obtained from P by removing a single black square $b \in B_0$ and a single white square w.

LEMMA 17. *When b is on the boundary of P, domino tilings of Q are in bijection with spanning trees of P for which the branch from b to b_0 contains the edge w (traversed in either direction).*

Proof. Let T be a domino tiling of Q. Let b' and b'' be the neighboring squares of w which are in H. Temperley's trick assigns to a vertex $x \in H$ an outgoing edge e if and only if xe is a domino of T. Temperley's trick, applied to T, gives a set of directed edges of the graph H such that each vertex has exactly one outgoing edge, except for b_0 and b which have no outgoing edges. Furthermore the union of these edges is a forest, that is, each component is a tree. There are exactly two components to this graph since each component is rooted at exactly one of b_0 or b.

We claim that b' and b'' are in different components of this forest. To see this, first note that the directed branch from b' cannot pass through b'', for otherwise the path from b' to b'' followed by the segment from b'' to w to b' would be a lattice path in H enclosing an odd number of squares of Q.(1) and so could not arise from a domino tiling of Q. Similarly the paths from b' and b'' cannot end both at b_0 or both at b because their union would contain a closed lattice path from b' to b''.

So the paths from b' and b'' end one at b_0 and one at b. On the path which ends at b, shift the dominos along it by one square towards b, and add an extra domino from w to its freed neighbor. This makes a domino tiling of P in which the tree branch from b passes through w (shifting dominos by 1 along a tree branch has the effect of changing the direction of the edges on that branch).

This process is reversible: from a tiling of P whose branch from b to b_0 passes through w, shift the dominos by one up to w to get a tiling of Q. □

When b is not on the boundary of P, it is possible that one of the paths from b' or b'' winds around b, or that these two paths both lead to b_0, enclosing b. So the lemma does not hold in that case.

By Lemma 17, to prove Theorem 3 it suffices to be able to count the number of tilings of Q, or rather, to compute the ratio of the number of tilings of Q to the number of tilings of P.

7.2. Region with a white hole

First we compute the coupling function on Q. Since Q has a hole (the hole w) which does not enclose the same number of black and white squares modulo 2, the weighted adjacency matrix A of Q' (the dual graph of Q) is not a Kasteleyn matrix for Q. In [Kas2], Kasteleyn describes how to redefine the weights of A to get a Kasteleyn matrix in this case (rather, he describes how to get weights for general planar graphs). One way to get a Kasteleyn matrix for Q' is the following. Start with the weighted adjacency matrix $A=A(Q')$ of §2.2. Take a path γ of vertices in Q from the outer boundary to a vertex of w (γ is a path of *faces* of Q'). Every edge in γ crosses an edge v_1v_2 of Q'. For each edge in γ, change the sign on the corresponding matrix entries $A_{v_1v_2}$ and $A_{v_2v_1}$. The matrix A' with these new signs is a Kasteleyn matrix for Q'. In particular, it has the property that its determinant is the square of the number of matchings, and its inverse is the coupling function C_Q for Q [Kas2], [Ken1].

Let $\widetilde{Q}'$ be the double cover of the graph Q', branched around the face containing w. That is, $\widetilde{Q}'$ is defined by the property that a closed path in Q' lifts to a closed path in $\widetilde{Q}'$ if and only if it winds an even number of times around the face w. Similarly we define

(1) From the Euler formula for a disk $F-E+V=1$, for any simple closed path in $\mathbf{Z}^2$ the sum of the number of faces, edges and vertices strictly enclosed is odd.

the double cover $\widetilde{H}$ of H as follows. Recall that w is an edge of H. Remove w from H and let $\widetilde{H}$ be the double cover of $H-\{w\}$ branched over the face which contained the edge w. The double cover of H', the dual of H, is similarly defined (and denoted $\widetilde{H}'$).

For each fixed v, the coupling function $C_Q(v,z)$ lifts to a discrete analytic function $\widetilde{C}_Q$ on $\widetilde{Q}'$: let z' be a lift of z, and v' a lift of v; then define $\widetilde{C}_Q(v',z')$ to be $\pm C_Q(v,z)$, with the $+$ sign when v',z' are on the same sheet of the cover (in the sense that there is a path from v' to z' which does not cross a lift of γ) and $-$ sign otherwise.

For a point $x\in Q'$ let x' and x'' denote its two lifts.

Let $v\in W_1$ and $z\in B_1$. For lifts v' of v and z' of z, the real part of $\widetilde{C}_Q(v',z')$ is harmonic on $\widetilde{H}'$ except when $z=v\pm\varepsilon$ or $z=w\pm\varepsilon$ (as in (2), with extra singularities at $w\pm\varepsilon$). So letting $\widetilde{G}$ denote the Green function on $\widetilde{H}'$, and $v_\pm=v\pm\varepsilon$, $w_\pm=w\pm\varepsilon$, it must be that $\operatorname{Re}\widetilde{C}_Q(v',z')$ is a linear combination of the Green functions $\widetilde{G}(v'_\pm,z')$, $\widetilde{G}(v''_\pm,z')$, $\widetilde{G}(w'_\pm,z')$, $\widetilde{G}(w''_\pm,z')$. By the antisymmetry under changing sheets and (2), $\operatorname{Re}\widetilde{C}_Q(v',z')$ is of the form

$$\begin{aligned}\widetilde{G}(v'_+,z')-\widetilde{G}(v'_-,z')-\widetilde{G}(v''_+,z')+\widetilde{G}(v''_-,z')\\ +\alpha_1(\widetilde{G}(w'_+,z')-\widetilde{G}(w'_-,z'))+\alpha_2(\widetilde{G}(w''_+,z')-\widetilde{G}(w''_-,z')),\end{aligned}$$

for some (real) constants α_1,α_2 which depend on v. These constants are determined by the following condition on C_Q: the harmonic conjugate of $\operatorname{Re}C_Q$ must be zero at the four points b'_0,b'_1,b''_0,b''_1. These four conditions give only 2 linear equations for α_1,α_2 since the value at b''_0 is by construction the negative of the value at b'_0, and similarly for b''_1.

We need to show that the coefficients α_1,α_2 converge as $\varepsilon\to0$, and compute their limit. To show this requires two steps. First, we will show that the Green functions $\widetilde{G}(y,z)$ converge to the corresponding continuous Green functions, and moreover that the terms $(\widetilde{G}(w'_1,z)-\widetilde{G}(w''_1,z))/\varepsilon$ and $(\widetilde{G}(w'_2,z)-\widetilde{G}(w''_2,z))/\varepsilon$ converge to the derivatives of the continuous Green functions. These derivatives have simple poles at the origin. Then we will show that there is a *unique* pair of analytic functions F_0,F_1 with the required properties having simple poles at the origin.

The convergence of the Green functions on $\widetilde{H}'$ is standard: As $\varepsilon\to0$, if y' does not converge to 0 then $\widetilde{G}(y',z')$ has the form

$$\widetilde{G}(y',z')=\frac{2}{\pi}\log|y'-z'|+\frac{2}{\pi}\log\frac{1}{\varepsilon}+\text{const}+O\left(\frac{\varepsilon}{|z'-y'|}\right)$$

(see e.g. [S]). This is enough to conclude that as $\varepsilon\to0$, $(\widetilde{G}(y-\varepsilon,z)-\widetilde{G}(y+\varepsilon,z))/\varepsilon$ tends to a harmonic function with zero boundary values and a singularity at y of the form $\operatorname{Re}(1/\pi(z-y))$. This is the derivative $\partial\tilde{g}(y,z)/\partial y$ of the continuous Green function $\tilde{g}(y,z)$.

From the symmetry $\widetilde{G}(y,z)=\widetilde{G}(z,y)$ we can conclude the same when y tends to 0 as $\varepsilon\to 0$. In conclusion, $(G(w_1',z)-G(w_1'',z))/\varepsilon$ and $(G(w_2',z)-G(w_2'',z))/\varepsilon$ converge to harmonic functions with simple poles at the origin (and single-valued harmonic conjugates).

To show that α_1 and α_2 converge, we claim that it suffices to show unicity of the limit of $\widetilde{C}_Q$. That is, since for each ε, α_1 and α_2 are solutions of a linear system whose coefficients converge (being functions of the Green function derivatives), either in the limit the linear equation becomes singular (in which case there is non-unicity of the limit) or it does not, in which case the solution is unique and the solutions for finite ε converge to this solution.

7.3. Unicity of the limit

Rather than work on a square region, it suffices to work on the unit disk with the white hole at the origin: the transformation rules (4) and (5) allow us to move the result back to the square Q.

Let U be the unit disk with marked points b_0, b_1 on the boundary. Let $\widetilde{U}$ be the double cover of U branched over the origin (the map $f(z)=z^2$ maps $\widetilde{U}$ to U). The lifts to $\widetilde{U}$ of the asymptotic coupling functions $F_0(v,z)$, $F_1(v,z)$ on U have zeros at $z=\pm\sqrt{b_0}, \pm\sqrt{b_1}$, poles at $z=\pm v$, and simple poles at $z=0$. (A priori the poles at the origin may have residue 0, in which case they are not poles at all.) These functions also are antisymmetric: for $v,z\in\widetilde{U}$ we have $\widetilde{F}_0(v,z)=-\widetilde{F}_0(v,-z)=\widetilde{F}_0(-v,-z)$ and the same for $\widetilde{F}_1$. Furthermore, since $\widetilde{F}_0$ and $\widetilde{F}_1$ are respectively pure imaginary and real on the boundary of $\widetilde{U}$, they extend by Schwarz reflection to meromorphic functions on the entire Riemann sphere, with additional simple poles at $\pm 1/\bar{v}$ and ∞ (the reflections of $\pm v$ and 0). In particular, they are rational functions. Since we know exactly the location of their (six) poles, and four of their zeros, it remains to find the other two zeros. The antisymmetry under $z\mapsto -z$ and symmetry under $z\mapsto 1/\bar{z}$ implies that the two other zeros are on the unit circle: otherwise the orbit of the zeros under these two symmetries would have four elements. Thus the functions have the form:

$$\widetilde{F}_0(v,z)=c_0\frac{\left(\frac{z}{\sqrt{b_0}}-\frac{\sqrt{b_0}}{z}\right)\left(\frac{z}{\sqrt{b_1}}-\frac{\sqrt{b_1}}{z}\right)\left(\frac{z}{b_3}-\frac{b_3}{z}\right)}{(z^2-v^2)(z^{-2}-\bar{v}^2)},$$

$$\widetilde{F}_1(v,z)=ic_1\frac{\left(\frac{z}{\sqrt{b_0}}-\frac{\sqrt{b_0}}{z}\right)\left(\frac{z}{\sqrt{b_1}}-\frac{\sqrt{b_1}}{z}\right)\left(\frac{z}{b_4}-\frac{b_4}{z}\right)}{(z^2-v^2)(z^{-2}-\bar{v}^2)},$$

for real c_0, c_1, and where $b_3=b_3(v)$ (resp. $b_4=b_4(v)$) must have absolute value 1. Now

b_3 and b_4 are chosen so that the residues of $\widetilde{F}_0$ and $\widetilde{F}_1$ are real at $z=v$. The constants $c_0=c_0(v)$, $c_1=c_1(v)$ are real and chosen so that the residue of each function at $z=v$ is $1/\pi$.

We can explicitly solve for b_3 as follows. The residue of $\widetilde{F}_0$ at $z=v$ is

$$c_0\frac{\left(\frac{v}{\sqrt{b_0}}-\frac{\sqrt{b_0}}{v}\right)\left(\frac{v}{\sqrt{b_1}}-\frac{\sqrt{b_1}}{v}\right)\left(\frac{v}{b_3}-\frac{b_3}{v}\right)}{2v(v^{-2}-\bar{v}^2)}.$$

This is supposed to be real; equating this and its complex conjugate yields

$$\frac{\left(\frac{v}{\sqrt{b_0}}-\frac{\sqrt{b_0}}{v}\right)\left(\frac{v}{\sqrt{b_1}}-\frac{\sqrt{b_1}}{v}\right)\left(\frac{v}{b_3}-\frac{b_3}{v}\right)}{2v(v^{-2}-\bar{v}^2)}$$
$$=\frac{\left(\bar{v}\sqrt{b_0}-\frac{1}{\sqrt{b_0}\,\bar{v}}\right)\left(\sqrt{b_1}\,\bar{v}-\frac{1}{\bar{v}\sqrt{b_1}}\right)\left(b_3\bar{v}-\frac{1}{\bar{v}b_3}\right)}{2\bar{v}(\bar{v}^{-2}-v^2)}.$$

Solving for b_3 gives $b_3^2=X_0/\overline{X}_0$, where

$$X_0=\left(\frac{v}{\sqrt{b_0}}-\frac{\sqrt{b_0}}{v}\right)\left(\frac{v}{\sqrt{b_1}}-\frac{\sqrt{b_1}}{v}\right)v^2+\left(\bar{v}\sqrt{b_0}-\frac{1}{\bar{v}\sqrt{b_0}}\right)\left(\bar{v}\sqrt{b_1}-\frac{1}{\bar{v}\sqrt{b_1}}\right).$$

Similarly one finds $b_4^2=-X_1/\overline{X}_1$, where

$$X_1=-v^2\left(\frac{v}{\sqrt{b_0}}-\frac{\sqrt{b_0}}{v}\right)\left(\frac{v}{\sqrt{b_1}}-\frac{\sqrt{b_1}}{v}\right)+\left(\bar{v}\sqrt{b_0}-\frac{1}{\bar{v}\sqrt{b_0}}\right)\left(\bar{v}\sqrt{b_1}-\frac{1}{\bar{v}\sqrt{b_1}}\right).$$

A lengthy calculation yields

$$\widetilde{F}_+(v,z)=\frac{4z\left(\frac{z}{\sqrt{b_0}}-\frac{\sqrt{b_0}}{z}\right)\left(\frac{z}{\sqrt{b_1}}-\frac{\sqrt{b_1}}{z}\right)}{\pi(z^2-v^2)\left(\frac{v}{\sqrt{b_0}}-\frac{\sqrt{b_0}}{v}\right)\left(\frac{v}{\sqrt{b_1}}-\frac{\sqrt{b_1}}{v}\right)}$$

(which is meromorphic in both v and z as expected) and

$$\widetilde{F}_-(v,z)=-\frac{4z\left(\frac{z}{\sqrt{b_0}}-\frac{\sqrt{b_0}}{z}\right)\left(\frac{z}{\sqrt{b_1}}-\frac{\sqrt{b_1}}{z}\right)}{\pi(1-z^2\bar{v}^2)\left(\bar{v}\sqrt{b_0}-\frac{1}{\sqrt{b_0}\,\bar{v}}\right)\left(\bar{v}\sqrt{b_1}-\frac{1}{\sqrt{b_1}\,\bar{v}}\right)}$$

(antimeromorphic in v as expected).

The map from U to $\widetilde{U}$ is $f(z)=\sqrt{z}$. Using the transformation rules (4) and (5) we have on the original region U,

$$F_+^U(v,z) = \frac{2(z-b_0)(z-b_1)}{\pi(z-v)(v-b_0)(v-b_1)}\sqrt{\frac{v}{z}},$$

$$F_-^U(v,z) = -\frac{2(z-b_0)(z-b_1)}{\pi(1-z\bar{v})(\bar{v}b_0-1)(\bar{v}b_1-1)}\sqrt{\frac{\bar{v}}{z}}.$$

The fact that the transformation rules apply in this case follows from the fact that the results $F_\pm^U$ have all the required properties of the coupling function limits (and are the unique functions with these properties).

7.4. On RHP

The map $f(z)=(z-1)/(z+1)$ maps RHP to the unit disk, sending 1 to 0, 0 to -1 and ∞ to 1. Let RHP$^\circ$ be RHP with a 'white hole' at 1. The limiting coupling functions on RHP$^\circ$ with zeros at $b_5:=f^{-1}(b_0)$ and $b_6:=f^{-1}(b_1)$ are

$$F_+^{\mathrm{RHP}^\circ}(v,z) = \frac{2(z-b_5)(z-b_6)}{\pi(z-v)(v-b_5)(v-b_6)}\sqrt{\frac{v^2-1}{z^2-1}},$$

$$F_-^{\mathrm{RHP}^\circ}(v,z) = \frac{2(z-b_5)(z-b_6)}{\pi(z+\bar{v})(\bar{v}+b_5)(\bar{v}+b_6)}\sqrt{\frac{\bar{v}^2-1}{z^2-1}}.$$

We will assume $b_6=\infty$; in this case the formulas on RHP$^\circ$ become

$$F_+^{\mathrm{RHP}^\circ}(v,z) = \frac{2(z-b_5)}{\pi(z-v)(v-b_5)}\sqrt{\frac{v^2-1}{z^2-1}}, \tag{37}$$

$$F_-^{\mathrm{RHP}^\circ}(v,z) = -\frac{2(z-b_5)}{\pi(z+\bar{v})(\bar{v}+b_5)}\sqrt{\frac{\bar{v}^2-1}{z^2-1}}. \tag{38}$$

7.5. The number of tilings of Q

Let U be the $(K\times K)$-square $U=[0,K]\times[-\frac{1}{2}K,\frac{1}{2}K]\subset$RHP. Suppose that $Q=Q_\varepsilon$ approximates U as $\varepsilon\to 0$, and suppose that the holes b, w, b_0 of Q_ε converge to points of the same name $b, w, b_0\in U$.

Since we are concerned primarily with the growth rate of the LERW, to simplify the calculations we will make the further assumption that $b, w\in U$ are within distance 1 of the origin, b_0 is on the right edge of U, and K is large. Using Lemma 11 we can then approximate to error $O(1/K)$ the coupling function limits F_0^U, F_1^U for points near the origin by the (much simpler) coupling function limits on RHP$-\{b,w\}$.

Suppose that w is on the x-axis at x-coordinate $\alpha>0$, and let $b=\beta i$ where β is real, $|\beta|<1$.

As in Lemma 17 let P be the region Q with the holes b and w filled in. Let $S\subset P$ be the chain of horizontal dominos from the origin to w.

Notationally, for a region A let $N(A)$ denote the number of tilings of A. The ratio of the number of tilings of Q to the number of tilings of P is a product of three terms

$$\frac{N(Q)}{N(P)}=\frac{N(Q)}{N(Q-S)}\cdot\frac{N(Q-S)}{N(P-S)}\cdot\frac{N(P-S)}{N(P)}, \tag{39}$$

each of which we can now approximate.

7.5.1. *The first ratio.* The inverse of the first term in (39), $N(Q-S)/N(Q)$, is the probability of S occurring in a tiling of Q. This is computed in a similar fashion to the proof of Theorem 1. Let $a_1,...,a_m$ be the set of dominos in the chain S. Then the probability of a_j occurring given $a_1,...,a_{j-1}$ is given by (see Lemma 6)

$$\frac{\Pr(a_j\,|\,a_1,...,a_j)}{\sqrt{2}-1}=1+\left(\frac{\pi}{\sqrt{2}}\,\varepsilon F_0^{U_j}(\varepsilon,2\varepsilon)-1\right)+o(\varepsilon),$$

where U_j is the translate of the region $\mathrm{RHP}-[0,2\varepsilon j]$ (with a white hole at w and black hole at b), translated by $-2\varepsilon j$ so that the tip of the cut $(2\varepsilon j)$ is at the origin; U_j corresponds to the polyomino $Q_j=Q-\{a_1,...,a_{j-1}\}$. The limiting coupling function $F_0^{U_j}$ on U_j can be computed from (37) and (38). Indeed, let $t=2\varepsilon j\in[0,\alpha]$; then the map

$$f(z)=\frac{\sqrt{(z+t)^2-t^2}}{\sqrt{\alpha^2-t^2}}$$

maps the region U_j to RHP and sends the singularities $\alpha-t$ to 1 and $i\beta-t$ to $b_5=f(i\beta-t)$.

A computation using (37) and (38) gives

$$\begin{aligned}\varepsilon F_0^{U_j}(\varepsilon,2\varepsilon)&=\tfrac{1}{2}\varepsilon f'(\varepsilon)\big(F_+^{\mathrm{RHP}^\circ}(f(\varepsilon),f(2\varepsilon))+F_-^{\mathrm{RHP}^\circ}(f(\varepsilon),f(2\varepsilon))\big)\\&=\frac{\sqrt{2}}{\pi}+\frac{\sqrt{2}\,(5t^2-\alpha^2)\varepsilon}{4\pi t(\alpha^2-t^2)}+O(\varepsilon^{3/2}).\end{aligned}$$

As a consequence

$$\frac{\Pr(a_j\,|\,a_1,...,a_j)}{\sqrt{2}-1}=1+\frac{5t^2-\alpha^2}{4t(\alpha^2-t^2)}\varepsilon+o(\varepsilon).$$

When we sum the logs of $\Pr(a_j\,|\,a_1,...,a_{j-1})$ for j running from 1 to m we get (see equation (10))

$$\begin{aligned}\log\Pr(S)&=m\log(\sqrt{2}-1)+\int_{2\varepsilon}^{\alpha-2\varepsilon}\frac{(5t^2-\alpha^2)\varepsilon}{4t(\alpha^2-t^2)}\cdot\frac{dt}{2\varepsilon}+o(\log\varepsilon)+o(1)\\&=m\log(\sqrt{2}-1)+\frac{1}{8}\log\frac{1}{\varepsilon}+\frac{1}{8}\log\alpha+o(1)+o(\log\varepsilon)\end{aligned}$$

where the $o(\log\varepsilon)$-term is independent of α and β.

Surprisingly, there is no dependence on β to first order.

7.5.2. *The second ratio.* Since b and w are on the boundary of the polyomino $P-S$, the second ratio in (39) is exactly $|C^{P-S}(w,b)|$, where C^{P-S} is the coupling function on $P-S$.

The coupling function on $P-S$ can be computed using (4), (5) and (6) and the map $f(z)=\sqrt{(z+\alpha)^2-\alpha^2}$ which maps RHP$-[0,\alpha]$ (translated by $-\alpha$ so that the tip of the cut is at the origin) to RHP, and ∞ to ∞. We must compute $\varepsilon F_0^{P-S}(\varepsilon,-\alpha+i\beta)$. We have

$$\begin{aligned}\varepsilon F_0^{P-S}(\varepsilon,-\alpha+i\beta) &= \frac{\varepsilon}{\pi}f'(\varepsilon)\left(\frac{1}{f(-\alpha+i\beta)-f(\varepsilon)}+\frac{1}{f(-\alpha+i\beta)+f(\varepsilon)}\right)\\ &= \frac{\sqrt{-2\varepsilon\alpha}}{\pi\sqrt{\alpha^2+\beta^2}}+O(\varepsilon).\end{aligned}$$

7.5.3. *The third ratio.* The computation of the third ratio in (39) is similar to the first ratio, except that now there are no singularities at b, w. We may use the coupling function on RHP, given by (6), to compute the coupling function on RHP$-[0,t]$ (again shifted so that the tip of the cut is at the origin). The appropriate function is $f(z)=\sqrt{(z+t)^2-t^2}$, and we arrive at

$$\varepsilon F_0^{U_t}(\varepsilon,2\varepsilon)=\frac{\sqrt{2}}{\pi}-\frac{\sqrt{2}}{4\pi t}\varepsilon+O(\varepsilon^2).$$

The log of the probability of S occurring is then

$$m\log(\sqrt{2}-1)+\int_{2\varepsilon}^{\alpha}-\frac{\varepsilon}{4t}\cdot\frac{dt}{2\varepsilon}=m\log(\sqrt{2}-1)-\frac{1}{8}\log\alpha-\frac{1}{8}\log\frac{1}{\varepsilon}+o(\log\varepsilon)+o(1).$$

7.5.4. *The product.* Now combining these three results, the log of the ratio (39) is

$$\begin{aligned}\log(R) &= -m\log(\sqrt{2}-1)-\frac{1}{8}\log\frac{1}{\varepsilon}-\frac{1}{8}\log\alpha+\log\frac{\sqrt{2\varepsilon\alpha}}{\pi\sqrt{\alpha^2+\beta^2}}\\ &\qquad+m\log(\sqrt{2}-1)-\frac{1}{8}\log\alpha-\frac{1}{8}\log\frac{1}{\varepsilon}+o(\log\varepsilon)+o(1)\\ &= -\frac{3}{4}\log\frac{1}{\varepsilon}+\frac{1}{4}\log\alpha-\frac{1}{2}\log(\alpha^2+\beta^2)+o(\log\varepsilon)+o(1),\end{aligned}$$

where the $o(\log\varepsilon)$-term is independent of α,β (the $o(1)$-term depends on α,β).

This is the log of the ratio of the number of tilings of Q and the number of tilings of P. For finite K there is an additional additive error term $O(1/K)$.

Letting $\alpha+i\beta=re^{i\theta}$ in polar coordinates, the probability that w is on the LERW from b to b_0 in P_ε is

$$\left(\frac{\varepsilon}{r}\right)^{3/4(1+o(1))}\cos(\theta)^{1/4}(1+o(1))$$

where the $o(1)$-term in the exponent is independent of θ.

The expected number of edges on the LERW from b to b_0, and which are at distance $\leqslant R$ from b, is then the integral of $(\varepsilon/r)^{3/4(1+o(1))}\cos(\theta)^{1/4}$, times the number of edges per unit area $1/\varepsilon^2$, times the area form $r\,dr\,d\theta$, as r runs from 0 to R. This gives $(R/\varepsilon)^{5/4(1+o(1))}$. This completes the proof of Theorem 3.

8. Open problems

There are a number of places where our arguments could stand improving, or where there are interesting avenues for further research. We list some of the outstanding ones here.

(1) There remains in Theorem 1 the error term $\mathrm{ERR}(\varepsilon)$ which is somewhat annoying. It seems reasonable to suspect that this term is in fact $O(1)$ but we do not have any method of computing this term at present.

(2) Is there an extension of Theorem 1 to the non-simply-connected case, as well as to the case of surfaces of higher genus? A more general version of Lemma 6 as well as its proof should be easy given the coupling function F_0. We do not know at present an easy extension of Lemma 7 to this case though.

(3) Can one find a more natural proof of Lemma 7? There may be a proof using invariance properties of the Schwarzian derivative, but we did not see it.

(4) Theorem 1 could be generalized to regions with polygonal boundaries which are horizontal, vertical, or have slope ± 1: this is because there is an exact formula for the determinant of the Laplacian on a triangular region $\{(x,y): x\geqslant 0,\, y\geqslant 0,\, x+y\leqslant n\}$, see [KPW].

More generally, one can ask about the relationship between the term $-\frac{1}{48}c_2(\varepsilon,U)$ of (1) and the ζ-function regularization of the determinant of the Laplacian. It appears from [OPS] that they are the same up to a multiplicative constant. Is there a simple explanation for this fact?

(5) Is $\frac{5}{4}$ the almost sure growth exponent for the LERW? The major open question about the LERW is: Is there a scaling limit of LERW, and if so is it conformally invariant?

9. Appendix

9.1. Green's function for a slit plane

We prove here Lemma 10.

Proof. As noted in the first two paragraphs of the proof of Lemma 6, we must compute the discrete Green function $G(0,z)$ on $\mathbf{Z}^2-(-\infty,-1]$, that is, the function of z harmonic on $\mathbf{Z}^2-(-\infty,0]$ with Dirichlet boundary values 0 on the boundary $(-\infty,-1]$ which satisfies $\Delta G(0,z)=\delta_0(z)$ for $z\in\mathbf{Z}^2-(-\infty,-1]$, and asymptotically $G(0,z)\to 0$ when $|z|\to\infty$.

Let $f_n(z)$ be the harmonic function on $\mathbf{Z}^2-(-\infty,-1]$ whose boundary values are 0 for $z\in[-n,-1]$ and 1 for $z\in(-\infty,-n-1]$. Let $g_n(z)=f_n(z)-f_{n+1}(z-1)$. By definition, $g_n(z)$ is harmonic on $\mathbf{Z}^2-(-\infty,0]$ with boundary value $f_n(0)$ at $z=0$ and boundary value 0 for $z\in(-\infty,-1]$. Therefore g_n is a constant times the desired Green function $G(0,z)$. This constant is 1 over the Laplacian of g_n at 0, that is, $-1/(f_{n+1}(0)+f_{n+1}(i)+f_{n+1}(-i))$ (note that $f_n(z)$ is harmonic at 0, and $f_{n+1}(z-1)$ is zero at $z=0$ and $z=-1$). So we have

$$G(0,z)=-\frac{f_n(z)-f_{n+1}(z-1)}{f_{n+1}(0)+f_{n+1}(i)+f_{n+1}(-i)}.$$

The asymptotic values of $f_n(z)$ for large n and z are

$$\frac{1}{\pi}\operatorname{Im}\log\frac{1+i\sqrt{z/n}}{1-i\sqrt{z/n}}.$$

By Theorem 9 applied to the function $f_n(z)$, we have $f_n(0)=\lambda_{1,0}\cdot 2/\pi\sqrt{n}$ and $f_n(i)=f_n(-i)=\lambda_{0,1}\cdot\sqrt{2}/\pi\sqrt{n}$. Also when $n\gg|z|\gg 1$ we have

$$f_n(z)-f_{n+1}(z-1)=\frac{1}{\pi\sqrt{nz}}+O\biggl(\frac{1}{z^{3/2}n^{1/2}}\biggr).$$

Therefore when $|z|\gg 1$, and in the limit as $n\to\infty$, we have

$$G(0,z)=-\biggl(\frac{1}{2\lambda_{1,0}+2\sqrt{2}\,\lambda_{0,1}}\biggr)\frac{1}{\sqrt{z}}+O(z^{-3/2}).$$

Scaling everything by ε (replacing z with z/ε) completes the proof. □

9.2. Local Riemann mappings

We prove Lemma 11. This is essentially a weaker version of a lemma of Hayman [H, Lemma 6.6].

First consider the case when j is small and the slit starts on an edge of U. For $q>0$ let g_q be the map $g_q(z)=\sqrt{2qz^2+q^2}-q$. Then g_q maps RHP injectively onto the translate by $-q$ of RHP$-[0,q]$. When $q=2\varepsilon j$, g_q is the standard limiting form for the f_j; we must show that the derivatives of f_j differ from those of $g_{2\varepsilon j}$ by $O(1)$. There is a constant C (depending on U but independent of j) such that for each j sufficiently small $g_{2\varepsilon j}^{-1}f_j$ maps a neighborhood of the origin in RHP to the half-disk $B_C(0)\cap$RHP. Now

$g_{2\varepsilon j}^{-1} f_j$ extends by Schwarz reflection to an injective map from a ball around the origin into $B_C(0)$; since the derivative at the origin of $g_{2\varepsilon j}^{-1} f_j$ is 1, this ball has radius comparable to C. Since $g_{2\varepsilon j}^{-1} f_j$ is bounded on this ball, the Schwarz lemma implies that its derivatives at the origin are all bounded as well. We conclude that $g_{2\varepsilon j}^{-1} f_j = z + z^2 O_z(1)$.

Now $f_j = g_{2\varepsilon j}(z + z^2 O_z(1))$ and since $g_{2\varepsilon j}$ is independent of U, the germ of f_j at the origin is independent of U up to $O(1)$.

The same argument with a different g_q works in the case where the cut starts at a corner. The argument at the end of the cut requires a simple modification. □

References

[Ah] AHLFORS, L., *Complex Analysis*, 3rd edition. McGraw-Hill, New York, 1978.

[Ap] APOSTOL, T., *Modular Functions and Dirichlet Series in Number Theory*, 2nd edition. Graduate Texts in Math., 41. Springer-Verlag, New York, 1990.

[BP] BURTON, R. & PEMANTLE, R., Local characteristics, entropy and limit theorems for spanning trees and domino tilings via transfer-impedances. *Ann. Probab.*, 21 (1993), 1329–1371.

[CKP] COHN, H., KENYON, R. & PROPP, J., A variational principle for domino tilings. To appear in *J. Amer. Math. Soc.*

[DD] DUPLANTIER, B. & DAVID, F., Exact partition functions and correlation functions of multiple Hamiltonian walks on the Manhattan lattice. *J. Statist. Phys.*, 51 (1988), 327–434.

[DMB] DESTAINVILLE, N., MOSSERI, R. & BAILLY, F., Configurational entropy of codimension-one tilings and directed membranes. *J. Statist. Phys.*, 87 (1997), 697–754.

[GB] GUTTMANN, A. & BURSILL, R., Critical exponent for the loop-erased self-avoiding walk by Monte-Carlo methods. *J. Statist. Phys.*, 59 (1990), 1–9.

[H] HAYMAN, W. K., *Multivalent Functions*. Cambridge Tracts in Math., 48. Cambridge Univ. Press, Cambridge, 1958.

[Kac] KAC, M., Can one hear the shape of a drum? *Amer. Math. Monthly*, 73:4, Part II (1966), 1–23.

[Kas1] KASTELEYN, P. W., The statistics of dimers on a lattice, I. The number of dimer arrangements on a quadratic lattice. *Physica*, 27 (1961), 1209–1225.

[Kas2] — Graph theory and crystal physics, in *Graph Theory and Theoretical Physics*, pp. 43–110. Academic Press, London, 1967.

[Ken1] KENYON, R., Local statistics of lattice dimers. *Ann. Inst. H. Poincaré Probab. Statist.*, 33 (1997), 591–618.

[Ken2] — Conformal invariance of domino tiling. To appear in *Ann. Probab.*

[Kes] KESTEN, H., Relations between solutions to a discrete and continuous Dirichlet problem, in *Random Walks, Brownian Motion, and Interacting Particle Systems*, pp. 309–321. Progr. Probab., 28. Birkhäuser Boston, Boston, MA, 1991.

[Ki] KIRCHHOFF, G., Über die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Verteilung galvanischer Ströme geführt wird. *Ann. Phys. Chem.*, 72 (1847), 497–508.

[KPW] KENYON, R., PROPP, J. & WILSON, D., Trees and matchings. *Electron. J. Combin.*, 7 (2000), Research Paper 25, 34 pp. (electronic).

[L] LAWLER, G., A lower bound on the growth exponent for loop-erased random walk in two dimensions. *ESAIM Probab. Statist.*, 3 (1999), 1–21 (electronic).

[Ma] MAJUMDAR, S. N., Exact fractal dimension of the loop-erased self-avoiding walk in two dimensions. *Phys. Rev. Lett.*, 68 (1992), 2329–2331.

[MS] MCKEAN, H. & SINGER, I., Curvature and the eigenvalues of the Laplacian. *J. Differential Geom.*, 1 (1967), 43–69.

[MW] MCCOY, B. & WU, T., *The Two-Dimensional Ising Model.* Harvard Univ. Press, Cambridge, MA, 1973.

[OPS] OSGOOD, B., PHILLIPS, R. & SARNAK, P., Extremals of determinants of Laplacians. *J. Funct. Anal.*, 80 (1988), 148–211.

[P] PEMANTLE, R., Choosing a spanning tree for the integer lattice uniformly. *Ann. Probab.*, 19 (1991), 1559–1574.

[S] SPITZER, F., *Principles of Random Walks*, 2nd edition. Graduate Texts in Math., 34. Springer-Verlag, New York–Heidelberg, 1976.

[T] TEMPERLEY, H., Enumeration of graphs on a large periodic lattice, in *Combinatorics* (Aberystwyth, 1973), pp. 202–204. London Math. Soc. Lecture Note Ser., 13. Cambridge Univ. Press, London, 1974.

[TF] TEMPLERLEY, H. & FISHER, M., Dimer problem in statistical mechanics—an exact result. *Philos. Mag.* (*8*), 6 (1961), 1061–1063.

[Th] THURSTON, W. P., Conway's tiling groups. *Amer. Math. Monthly*, 97 (1990), 757–773.

RICHARD KENYON
CNRS UMR 8628
Laboratoire de Mathématique
Bât. 425
Université de Paris-Sud
FR-91405 Orsay Cedex
France
kenyon@topo.math.u-psud.fr

Received January 11, 1999
Received in revised form October 22, 1999