ELSEVIER
Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Combinatorial Theory, Series B 92 (2004) 295–317
Journal of Combinatorial Theory Series B
www.elsevier.com/locate/jctb

Dimers, tilings and trees

Richard W. Kenyon[a], Scott Sheffield[b]

[a] *Department of Mathematics, University of British Columbia, Vancouver, BC, Canada*
[b] *Microsoft Research, One Microsoft Way, Redmond, Washington, USA*

Received 11 November 2003

Abstract

Generalizing results of Temperley (London Mathematical Society Lecture Notes Series 13 (1974) 202), Brooks et al. (Duke Math. J. 7 (1940) 312) and others (Electron. J. Combin. 7 (2000); Israel J. Math. 105 (1998) 61) we describe a natural equivalence between three planar objects: weighted bipartite planar graphs; planar Markov chains; and tilings with convex polygons. This equivalence provides a measure-preserving bijection between dimer coverings of a weighted bipartite planar graph and spanning trees of the corresponding Markov chain. The tilings correspond to harmonic functions on the Markov chain and to "discrete analytic functions" on the bipartite graph.

The equivalence is extended to infinite periodic graphs, and we classify the resulting "almost periodic" tilings and harmonic functions.

Keywords: Dimers; Polygon tilings; Spanning trees; Perfect matchings

1. Introduction

In [11], Temperley gave a bijection between the set of spanning trees of an $n \times n$ grid and the set of perfect matchings (dimer coverings) of a $(2n-1) \times (2n-1)$ grid with a corner removed. This bijection was generalized in [10] to a weight-preserving bijection (the KPW construction) from the set of in-directed spanning trees (also known as arborescences) on an arbitrary weighted, directed planar graph $\mathcal{G}_T$ to the set of perfect matchings on a related graph $\mathcal{G}_D$. The construction is useful in statistical mechanics because certain types of events in the spanning tree model can be easily computed using dimer technology,

E-mail address: sheff@microsoft.com, sheff@math.berkeley.edu (S. Sheffield).
URL: http://topo.math.u-psud.fr/∼kenyon (R.W. Kenyon).

0095-8956/$ - see front matter © 2004 Published by Elsevier Inc.
doi:10.1016/j.jctb.2004.07.001

for example winding numbers of branches and local statistics. For dimer models arising from a spanning tree model, moreover, the spanning tree formulation provides other useful information. In particular Wilson's algorithm [12] for generating spanning trees can be used to rapidly simulate dimer configurations. Moreover, the spanning tree formulation identifies natural boundary conditions ("Temperleyan" boundary conditions) for the dimer model which allows asymptotic computation of many properties, in particular conformal invariance properties of dimers [6]. However in the paper [10] it was not known if every dimer model on a bipartite planar graph corresponded to a spanning tree model on a related graph.

A seemingly unrelated construction is the construction of a "Smith diagram" from a planar resistor network [1]. This is a tiling of a plane region with squares of arbitrary sizes, which is associated in a bijective way to a critical-point-free harmonic function on the network with unit resistances (there is a square in the tiling for each edge in the graph, whose size is proportional to the current flow through the edge). This construction was generalized in [7] to planar Markov chains (graphs with transition probabilities), where a harmonic function gives a tiling with *trapezoids*. It was not known at the time what if any graphical correspondence was natural for *general* polygonal tilings.

In the current paper we extend the above equivalences and describe a correspondence between these three types of objects: weighted bipartite planar graphs, planar Markov chains, and tilings with general convex polygons.

In particular from a weighted bipartite planar graph $\mathcal{G}_D$ we can construct a tiling T of a plane region with convex polygonal tiles, and a planar Markov chain $\mathcal{G}_T$, in an essentially bijective way (that is, up to natural equivalences). There is a tile in T for each "white" vertex of $\mathcal{G}_D$, whose shape is determined by a *discrete analytic function* (see definitions below). The graph $\mathcal{G}_T$ is a graph on the 1-skeleton of the tiles, with transitions determined by their Euclidean geometry. The tilings are therefore representations of discrete analytic functions on the bipartite planar graph $\mathcal{G}_D$, which correspond to harmonic functions on the Markov chain $\mathcal{G}_T$.

An important application of this construction is that it provides a converse to the Temperley-KPW construction. That is, starting with the finite weighted bipartite planar graph $\mathcal{G}_D$, one constructs a Markov chain $\mathcal{G}_T$ and a measure-preserving bijection from the dimer model on $\mathcal{G}_D$ to the spanning tree process on $\mathcal{G}_T$. This dimer/spanning tree correspondence has a number of important consequences. Firstly, it was used in [9] in a fundamental way to classify Gibbs measures on dimer models on infinite periodic planar graphs. Secondly, since spanning trees can be sampled efficiently [12], the construction provides a way to sample efficiently from bipartite planar dimer models. Previously the only (provably efficient) way to sample general planar bipartite dimer models was to do exact computations of joint edge probabilities. A third application [8] is that it allows one to compute the asymptotics of dimer correlations and height fluctuations in terms of the Green's function on $\mathcal{G}_T$.

In Section 5, we discuss how the construction extends in the case of infinite periodic graphs. This is motivated by the study of the dimer model on periodic graphs, see [2,9]. Given any periodic planar bipartite weighted graph $\mathcal{G}_D$, we produce an essentially unique "almost periodic" planar Markov chain $\mathcal{G}_T$, which extends the dimer/spanning tree correspondence. This unicity is an important element in the classification theorem of ergodic Gibbs measures on dimer coverings of $\mathcal{G}_D$ described in [9].

2. Definitions

2.1. Dimers and measures

Let $\mathcal{G}_D = (V, E)$ be a finite bipartite planar graph. Bipartite means that the vertices V can be 2-colored, that is, colored black and white so that black vertices are only adjacent to white vertices and vice versa. Let $\nu: E \to (0, \infty)$ be a weight function on the edges. A *perfect matching*, or *dimer configuration* $M \subset E$ is a set of edges with the property that each vertex is contained in exactly one edge in M. The weight of a matching M is $\nu(M) = \prod_{e \in M} \nu(e)$. Let $\mathcal{M}(\mathcal{G}_D)$ denote the set of perfect matchings of $\mathcal{G}_D$. Let μ be the probability measure on $\mathcal{M}(\mathcal{G}_D)$ giving a matching a probability proportional to its weight: $\mu(M) = \frac{1}{Z}\nu(M)$ where $Z = \sum_{M \in \mathcal{M}(\mathcal{G}_D)} \nu(M)$.

2.2. Kasteleyn matrices

If $\mathcal{G}_D$ has n black and n white vertices, a *Kasteleyn matrix* (see [4]) for $\mathcal{G}_D$ is a real $n \times n$ matrix $K = (K_{i,j})$ whose rows index the black vertices and columns index the white vertices of $\mathcal{G}_D$, defined as follows. The entry $K_{i,j}$ is zero if there is no edge from b_i to w_j, and if there is an edge of weight $\nu(b_i w_j)$ then $K_{i,j} = \pm\nu(b_i w_j)$, where the signs are chosen so that the product of signs of edges around every interior face of K is $(-1)^{d/2+1}$, where d is the degree of the face. This property of signs is not changed if we multiply all elements in a particular column or row of K by -1 (because each vertex of $\mathcal{G}_D$ has an even number—zero or two—of edges on each face of $\mathcal{G}_D$). Moreover, such a choice of signs always exists, and by Kasteleyn's theorem, the determinant of K is (up to sign) the sum of the weights of the matchings of $\mathcal{G}_D$ [4]. By a *discrete analytic function* we mean a function f on black vertices (resp. white vertices) which satisfies $fK = 0$ (resp. $Kf = 0$). This generalizes the definition of discrete analytic function on $\mathbb{Z}^2$ defined in [3,6]. These functions play a role implicitly in Sections 4 and 5.

2.3. Gauge transformations

If we multiply the weights of all the edges in $\mathcal{G}_D$ having a fixed vertex by a constant, the measure μ does not change, since exactly one of these weights is used in every configuration. More generally, two weight functions ν_1, ν_2 are said to be *gauge equivalent* if ν_1/ν_2 is a product of such operations, that is, if there are functions f_1 on white vertices and f_2 on black vertices so that for each edge wb, $\nu_1(wb)/\nu_2(wb) = f_1(w)f_2(b)$. Gauge equivalent weights define the same measure μ.

Multiplying the ith row (resp., column) of a Kasteleyn matrix K by a positive, non-zero constant c is equivalent to multiplying by c the weights of all of the edges of $\mathcal{G}_D$ incident to b_i (resp., w_i). In other words any matrix $\tilde{K}$ obtained from K by multiplying the rows and columns of K by non-zero constants will be a Kasteleyn matrix for a graph which is gauge equivalent to $\mathcal{G}_D$.

3. T-graphs and corresponding dimer/spanning-forest models

In this section, we define a family of planar graphs called T-graphs and describe a weight-preserving correspondence between the spanning trees on a T-graph and dimer configurations on a bipartite graph derived from it. This is closely related to the result of [10], who also give a relation between spanning trees on a general planar graph and dimers on a derived graph.

In the present context, however, our derivation can be reversed: we will see in Section 4 that for every bipartite planar graph, which is *non-degenerate* in the sense that it contains no edges which fail to be used in any perfect matching of the graph (for the purposes of the dimer model, it makes sense to delete these edges), endowed with a generic choice of weights, there is a gauge-equivalent graph which can be derived from a T-graph in this way. By taking limits, the correspondence generalizes to the case when the weights are not assumed to be generic.

3.1. Complete edges that form T-graphs

The definition of T-graph on a torus—which we use in Section 5—is quite simple. A disjoint collection $L = \{L_1, L_2, \ldots, L_n\}$ of open line segments in the torus $\mathbb{R}^2/\mathbb{Z}^2$ *forms a T-graph in the torus* if $\cup_{i=1}^n L_i$ is closed. The term "T-graph" refers to the fact each endpoint of a given L_i necessarily lies on the interior of some L_j with $j \neq i$. In other words, each L_i "tees into" an L_j at each of its two endpoints.

We say a disjoint collection $L_1, L_2, \ldots, L_n$ of open line segments in $\mathbb{R}^2$ *forms a T-graph in* $\mathbb{R}^2$ if $\cup_{i=1}^n L_i$ is connected and contains all of its limit points except for some set $R = \{r_1, \ldots, r_m\}$, where each r_i lies on the boundary of the infinite component of $\mathbb{R}^2$ minus the closure $\bar{L}$ of $\cup_{i=1}^n L_i$. Elements in R are called *root vertices*. For example, a single open line segment forms a T-graph with root vertices given by the two end points. A pair of open line segments—one of which tees into the other to make the letter "T"—forms a T-graph with three root vertices. The three open edges of a triangle also form a T-graph with three root vertices. A partitioning of a convex polygon P into convex polygonal tiles using a finite number of line segments will form a T-graph with root vertices at the vertices of P if and only if it is *generic* in the sense that the endpoint of each of these line segments lies either on the interior of another line segment or on the boundary of P. (See Fig. 1.)

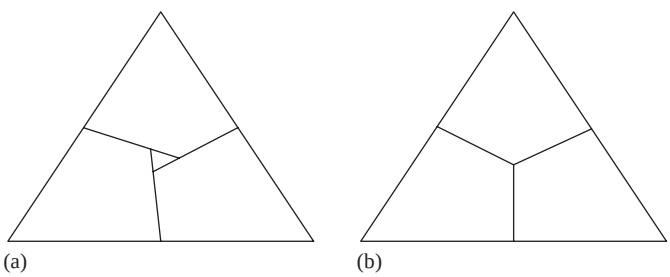

Fig. 1. (a) Line segments that form a T-graph. (b) Line segments that do not form a T-graph.

Note that each endpoint of a given L_i is *either* a root vertex or an interior point of some L_j. To distinguish the L_i from subsegments of the L_i (which we discuss later) we refer to the L_i as *complete edges*.

In subsequent subsections, we will use L to define a weighted, directed graph $\mathcal{G}_T(L)$ and a weighted, bipartite graph $\mathcal{G}_D(L)$. The ultimate goal of this section will be to derive a weight-preserving bijection between directed spanning forests on $\mathcal{G}_T(L)$ (with specified roots) and perfect matchings of $\mathcal{G}_D(L)$. When the choice of L is clear from the context, we write $\mathcal{G}_T = \mathcal{G}_T(L)$ and $\mathcal{G}_D = \mathcal{G}_D(L)$.

3.2. T-graphs and their duals

The set $V_T(L)$ of vertices of $\mathcal{G}_T$ is the set of points in $\mathbb{R}^2$ which are endpoints of at least one of the L_i. We use the term *T-graph* to refer to the graph $\mathcal{G}_T$ *together* with the corresponding set L of line segments embedded in the torus or plane. In other words, a T-graph is not merely a graph but rather a geometric construction which determines the graph $\mathcal{G}_T$ (as well as several other graphs described below). We refer to the graph $\mathcal{G}_T$ itself as the *tree-graph* of L.

A vertex v which is in the interior of a complete edge L_i (called an *interior vertex*) has exactly two edges in $\mathcal{G}_T$ directed outwards from it: these edges point towards the two immediate neighbors, v_1 and v_2, along L_i (one on each side of v). The weights on the edge from v to these two v_i are chosen in such a way that the two weights add up to one and are inversely proportional to the Euclidean distances $|v - v_i|$. These weights correspond to the transition probabilities of a Markov chain on $V_T(L)$. The root vertices are sinks of $\mathcal{G}_T$ (they have no outgoing edges in $\mathcal{G}_T$) and are fixed points of the Markov chain. Note that (by our choice of transition probabilities) the expected change in Euclidean position during a step of the Markov chain is always zero; thus, a random walk on $\mathcal{G}_T$—viewed as a Markov chain on positions in $\mathbb{R}^2$—is a martingale. In other words, the coordinate functions on the vertices of $\mathcal{G}_T$ are harmonic functions on $\mathcal{G}_T$ away from the root vertices.

See Fig. 2 for an example of a T-graph with three roots. Note that, by convention, when we have transitions both from i to j and from j to i, rather than drawing two directed edges in the graph $\mathcal{G}_T$ we draw a single edge with two transition probabilities, one from each end.

We define $\mathcal{G}'_T = \mathcal{G}'_T(L)$, an undirected dual graph of $\mathcal{G}_T$, as follows. Let C be an arbitrary simple closed curve that encircles the $\cup_{i=1}^n L_i$ and contains each of the root vertices $r_1, \ldots, r_m$ in clockwise order. The vertices of $\mathcal{G}'_T$ are the bounded faces of $\mathcal{G}_T \cup C$ (bounded connected components of $\mathbb{R}^2 \backslash (\cup_{i=1}^n L_i \cup C)$). Faces of $\mathcal{G}_T \cup C$ adjacent to C are called *outer faces* of $\mathcal{G}_T$: they correspond to *outer vertices* of $\mathcal{G}'_T$. Two vertices of $\mathcal{G}'_T$ are connected by an edge a of $\mathcal{G}'_T$ if the corresponding faces of $\mathcal{G}_T$ are adjacent across an edge of $\mathcal{G}_T$. For an edge e of $\mathcal{G}_T$ we denote by e^* its corresponding dual edge.

Lemma 3.1. *If $L_1, L_2, \ldots, L_n$ form a T-graph, then $\mathcal{G}_T$ has exactly $n+1$ faces* (*including outer faces*). *Hence $\mathcal{G}'_T$ has $n+1$ vertices.*

Proof. This follows from Euler's formula. The line segments L_i decompose the interior of C into some number n_2 of open faces (open 2-cells), $n_1 = n$ open complete edges, and

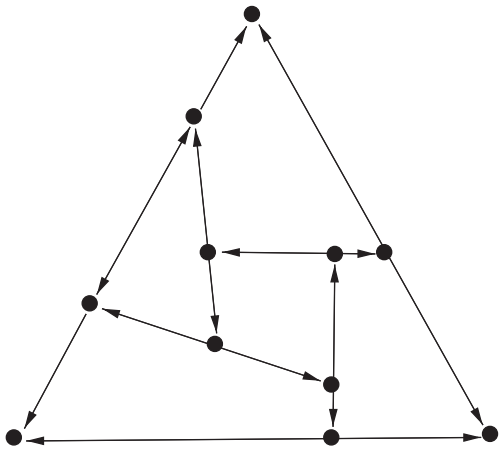

Fig. 2. The directed T-graph $\mathcal{G}_T$. The root vertices are the corners of the triangle.

$n_0 = 0$ vertices. Since $n_2 - n_1 + n_0 = 1$, the Euler characteristic of the disc, the result follows. □

3.3. Spanning trees

A *spanning tree* of a graph G is a subset of edges which is connected, contains no cycle, and passes through every vertex. If the edges of G are directed, a *directed spanning tree*, or arborescence, is a spanning tree in which every vertex but one (called the root vertex) has a unique outgoing edge. Given a subset of vertices of G called root vertices, a *directed spanning forest* is a set of edges with no cycles, passing through all vertices, each non-root vertex having a unique outgoing edge, and each component of which is connected to a unique root vertex.

We will employ the following correspondence between (non-directed) spanning trees in $\mathcal{G}'_T$ and their (non-directed) *dual spanning forests* in $\mathcal{G}_T$. Using the correspondence between edges of $\mathcal{G}'_T$ and edges of $\mathcal{G}_T$, we can think of edge subsets of both $\mathcal{G}_T$ and $\mathcal{G}'_T$ as subsets of the set of all edges of $\mathcal{G}_T$. Using this interpretation, we state the following lemma (which is illustrated in Fig. 3):

Lemma 3.2. *The complement of a spanning tree $\mathcal{T}$ of $\mathcal{G}'_T$ is a spanning forest $\mathcal{F}$ of $\mathcal{G}_T$, with roots at the root vertices. Similarly, the complement of a spanning forest $\mathcal{F}$ of $\mathcal{G}_T$, with roots at the root vertices, is a spanning tree $\mathcal{T}$ of $\mathcal{G}'_T$.*

Proof. We sketch the standard tree dualization argument. If $\mathcal{F}$ is a spanning forest of $\mathcal{G}_T$, with roots at root vertices, its complement $\mathcal{T}$ cannot contain any cycles in $\mathcal{G}'_T$ (since such a cycle would separate at least one interior vertex of $\mathcal{G}_T$ from the root vertices), and it must be connected (since otherwise, the set of edges separating two components would either

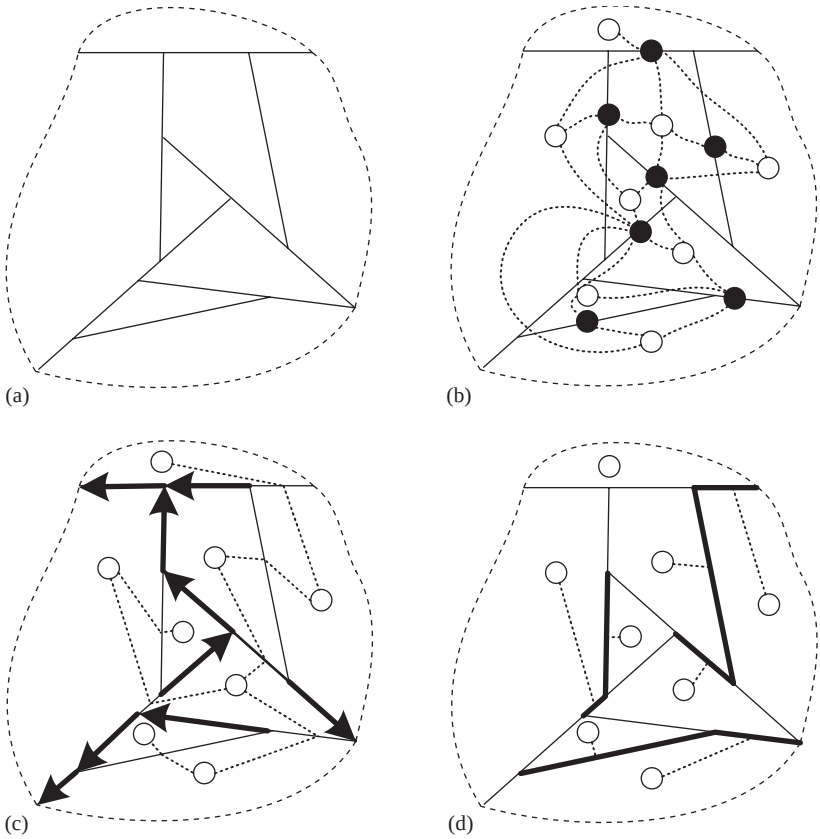

Fig. 3. (a) Edges of a T-graph L and the surrounding curve C. (b) The graph $\mathcal{G}_D^r$. This graph is the incidence graph for the set of complete edges of L (drawn as black vertices) and faces of L (drawn as white vertices). (c) A spanning forest of $\mathcal{F}$ of $\mathcal{G}_T$, drawn with thick arrows, and the dual spanning tree $\mathcal{T}$ on $\mathcal{G}_T'$, drawn with dotted lines connecting vertices of $\mathcal{G}_T'$ (which are faces of $\mathcal{G}_T$, represented as white vertices). Each such dotted line crosses a segment of a complete edge that is *not* used in $\mathcal{F}$ (there is exactly one such segment for each complete edge). (d) The marked matching corresponding to $\mathcal{F}$ when the dual root is taken to be the (unmatched) uppermost white vertex.

form a cycle in $\mathcal{G}_T$ or a path connecting two root vertices in $\mathcal{G}_T$); hence it is a spanning tree. Similarly, if $\mathcal{T}$ is a spanning tree of $\mathcal{G}_T'$, its complement $\mathcal{F}$ cannot contain cycles of $\mathcal{G}_T$ (since such a cycle would separate at least one inner face of $\mathcal{G}_T$ from the outer faces) and each connected component of $\mathcal{F}$ contains at least one root vertex (since otherwise the set of edges separating that component of $\mathcal{F}$ from its complement would form a cycle in $\mathcal{G}_T$). $\square$

3.4. Dimer graphs from T-graphs

Now we will define the weighted, bipartite (non-directed) graph $\mathcal{G}_D = \mathcal{G}_D(L)$. First, we define a slightly larger $\mathcal{G}_D^r = \mathcal{G}_D^r(L)$, whose black vertices are the n complete edges L_i and whose white vertices are the $n+1$ faces of $\mathcal{G}_T$ (including outer faces).

Table 1
Summary of graphs constructed from a collection of edges L that forms a T-graph

Graph	Vertex set
$\mathcal{G}_T = \mathcal{G}_T(L)$	Points that are endpoints of some L_i.
$\mathcal{G}'_T$	Faces and outer faces of $\mathcal{G}_T$ (i.e., bounded components of $\mathbb{R}^2 \backslash (L \cup C)$).
$\mathcal{G}^r_D$	Faces and outer faces of $\mathcal{G}_T$ (one partite class).
	Complete edges in L (other partite class).
$\mathcal{G}_D$	Same as $\mathcal{G}^r_D$ but with one outer face (the dual root) omitted.
$\mathcal{G}'_D$	Faces and some outer faces of $\mathcal{G}_D$ (which correspond to vertices of $\mathcal{G}_T$).

The graph $\mathcal{G}'_D$ (which is not exactly the same as $\mathcal{G}_T$) is defined precisely in Section 4.2.

A white vertex w of $\mathcal{G}^r_D$ is adjacent to a black vertex b of $\mathcal{G}^r_D$ if the face F corresponding to w contains a portion of the L_i corresponding to b as its boundary. The weight $v((w, b))$ is then given by the Euclidean length of the portion of the line segment. The graph thus defined is planar. To see this, note that it can be drawn on top of the tiling $\bar{L}$ as follows: put a white vertex in the interior of each face, and a black vertex in the center of each complete edge. When w and b are connected, draw a line from w inside the corresponding face towards the complete edge corresponding to b, and then along this complete edge, staying just to one side, until the center is reached. It is not hard to see that this can be done in such a way that the paths do not intersect.

The graph $\mathcal{G}_D$ is formed from $\mathcal{G}^r_D$ by (arbitrarily) picking one of the outer white vertices of $\mathcal{G}^r_D$ and removing it; we will refer to the removed vertex as the *dual root* of $\mathcal{G}_D$. Now, $\mathcal{G}_D$ is a weighted bipartite graph with n white and n black vertices. To every edge $e = (w, b)$ in a perfect matching of $\mathcal{G}_D$ (where w corresponds to a face F and b to a complete edge L_i), we denote by S_e the segment of the L_i which borders F. Because of our choice of weights, $\mu(M)$ (where μ is the probability measure on perfect matchings defined in the introduction) is proportional to $\prod_{e \in M} |S_e|$ where $|S_e|$ is the Euclidean length of S_e. Now, the edge segment S_e may have vertices of $\mathcal{G}_T$ in its interior; these vertices divide S_e into subsegments, each of which has vertices of $\mathcal{G}_T$ as its endpoints and hence corresponds to an edge of $\mathcal{G}_T$. A *marked matching* of $\mathcal{G}_D$ is a matching M of $\mathcal{G}_D$ together with a specified subsegment S'_e of S_e (which, again, we may interpret as an edge in $\mathcal{G}_T$) for each $e \in M$. We extend μ to give a measure on random marked matchings as follows: to sample a random marked matching, first choose a random matching. Then for each edge e, choose an S'_e from among the subsegments of S_e, where probability of each subsegment is proportional to its length. If M' is a marked matching, then $\mu(M')$ is proportional to $\prod_{e \in M} |S'_e|$ (Table 1).

3.5. *From dimers to trees*

Let $\mathcal{T}_{M'} = \{S'_e : e \in M\}$ be the set of edges corresponding to a marked matching M' of $\mathcal{G}_D$. Each S'_e corresponds to an edge of $\mathcal{G}'_T$, so we can think of $\mathcal{T}_{M'}$ as a subgraph of $\mathcal{G}'_T$. We direct each such edge of $T_{M'}$ (corresponding to some S'_e) of this graph from the face which corresponds to a vertex in e towards the face which does not. We use this interpretation of $\mathcal{T}_{M'}$ (as a directed subgraph of $\mathcal{G}'_T$) in the following lemma:

Lemma 3.3. *If M' is a marked matching, then $\mathcal{T}_{M'}$ is an in-directed spanning tree of $\mathcal{G}'_T$, rooted at the dual root. The dual $\mathcal{F}_{M'}$ of $\mathcal{T}_{M'}$ is thus a spanning forest of $\mathcal{G}_T$ (when $\mathcal{G}_T$ is viewed as an undirected graph).*

Proof. It is sufficient to prove that $\mathcal{T}_{M'}$ has no directed cycles, since it contains exactly one edge pointing away from each face of $\mathcal{G}_T$ (excluding the dual root). This is accomplished using Euler's formula. Suppose that $\mathcal{T}_{M'}$ had a directed cycle $F_0, F_1, \ldots, F_j = F_0$ of faces of $\mathcal{G}_T$. Let S_i be the segment S'_e separating F_i and F_{i+1}. Let C' be a simple closed curve which starts in the interior of F_0, passes through S_0 at one point, moves through the interior of F_1, passes through S_1 at a single point, etc. until it returns to F_0. Except for its intersections with the S_i's, each at a single point, C' is entirely contained in the union of the interiors of the F_i. The intersection of the L_i with the interior of C' gives a decomposition of this interior into n_2 two-cells (where n_2 is the number of faces partial or completely contained inside the loop C'), n_1 open one-cells and $n_0 = 0$ vertices. Thus, by Euler's formula $n_2 - n_1 + n_0 = n_2 - n_1 = 1$. In particular $n_2 + n_1$ is odd.

However, the sequence $w_0, b_0, w_1, b_1, \ldots, b_{j-1},\ w_j = w_0$ (where w_i is the white vertex of $\mathcal{G}_D$ corresponding to F_i and b_i is the black vertex corresponding to the complete edge containing S_i) is a cycle in $\mathcal{G}_D$, alternating edges of which are contained in M. The set of vertices in $\mathcal{G}_D$ enclosed by this cycle must be matched only with each other in a perfect matching (since the cycle disconnects these vertices from the rest of the graph). This is a contradiction to the fact that $n_2 + n_1$ is odd. $\square$

Let μ_F be the measure on directed spanning forests of $\mathcal{G}_T$, rooted at the root vertices, for which $\mu_F(\mathcal{F})$ is proportional to the product of the weights of the edges in $\mathcal{F}$. Since each of the two outgoing edges of a given interior vertex has weight (by construction) inversely proportional to its Euclidean length, $\mu_F(\mathcal{F})$ is inversely proportional to the product of the lengths of the edges of $\mathcal{F}$; hence, $\mu_F(\mathcal{F})$ is also proportional to the product of the Euclidean lengths of all edges of $\mathcal{G}_T$ which do not appear (directed one way or another) in $\mathcal{F}$.

The following is the main result of this section.

Theorem 3.4. *The map $M' \to \mathcal{F}_{M'}$ gives a one-to-one correspondence between marked matchings of $\mathcal{G}_D$ and in-directed spanning forests of $\mathcal{G}_T$, rooted at R. The correspondence is measure preserving, i.e., $\mu(M') = \mu_F(\mathcal{F}_{M'})$.*

Proof. First, we would like to interpret $\mathcal{F}_{M'}$ as a directed spanning forest of $\mathcal{G}_T$ by orienting each edge of $\mathcal{F}_{M'}$ towards its root vertex. In order to do this, we must check that if M' is a perfect matching of $\mathcal{G}_D$, then the directed path along $\mathcal{F}_{M'}$, from a vertex v to a root vertex, is a directed path of $\mathcal{G}_T$. To see this, note first $\mathcal{F}_{M'}$ contains all but one segment of each of the L_i; thus, for every interior vertex v of $\mathcal{G}_T$ (interior to some L_i), $\mathcal{F}_{M'}$ includes a path from v to exactly one of the endpoints of L_i. Call this vertex v_1; each of the directed edges in the directed path from v to v_1 is a directed edge of $\mathcal{G}_T$. If v_1 is also an interior vertex of some L_j, then there is a path of edges in $\mathcal{F}_{M'}$ from v_1 to some endpoint v_2 of L_j. Iterating this process, we must eventually produce a directed path from v to a root (since $\mathcal{F}_{M'}$ has no cycles).

It now follows immediately from Lemma 3.3 and our choice of weights, that $\mu_F(\mathcal{F}_{M'})$ is proportional to $\mu(M')$, since each is proportional to the same product of edge lengths. The proof we will be complete once we show that the map $M' \to \mathcal{F}_{M'}$ is invertible.

Let $\mathcal{F}$ be an arbitrary directed spanning forest $\mathcal{F}$, rooted in R. Since only the endpoints of a given L_i have outgoing edges pointing to vertices not on L_i, each vertex of L_i belongs to a path pointing to one of the two endpoints. It follows that $\mathcal{F}$ must include all but one of the subsegments of L_i. By Lemma 3.2, the dual of $\mathcal{F}$ is a spanning tree of $\mathcal{G}'_T$, which we may view as being directed towards the dual root. Each face F of $\mathcal{G}_T$ is (besides the dual root) is directed towards another face across an edge segment of one of the L_i. Pairing of F with the edge segment produced in this way gives a marked matching M' for which $\mathcal{F}_{M'} = \mathcal{F}$. □

3.6. *T-graphs and dimers on the torus*

If $L = \{L_1, \ldots, L_n\}$ forms a T-graph on the torus, then we can construct $\mathcal{G}_T = \mathcal{G}_T(L)$ exactly as above; in this case, $\mathcal{G}_T(L)$ has no root vertices and no outer faces. Since the faces of $\mathcal{G}_T$ and open edges L_i give a decomposition of the torus into one-cells and two-cells, Euler's formula implies that $\mathcal{G}_T$ has exactly n faces. We construct $\mathcal{G}_D$ as above (with white vertices given by faces F of $\mathcal{G}_T(L)$, black vertices by the complete edges L_i, and edges occurring between F and L_i that share a line segment, weighted according to the length of that segment). We also construct $\mathcal{G}'_T$ in a similar fashion.

A *cycle-rooted spanning forest* $\mathcal{F}$ of $\mathcal{G}_T$ is a (directed) subgraph of $\mathcal{G}_T$—with one outgoing edge from each vertex of $\mathcal{G}_T$—which has no null-homotopic (directed) cycles (i.e., no cycles which—when lifted to the universal cover of the torus—start and end at the same place). The "roots" of such an $\mathcal{F}$ are the directed cycles of $\mathcal{F}$. Clearly, every such $\mathcal{F}$ has at least one (non-null-homotopic) directed cycle.

The dual of $\mathcal{F}$ is a cycle-rooted spanning forest $\mathcal{F}'$ on $\mathcal{G}'_T$. Now, if $\mathcal{F}$ has exactly j cycles, then it is not hard to see that $\mathcal{F}'$ has j cycles as well. We can view $\mathcal{F}'$ as a directed cycle-rooted spanning forest by directing each edge not on a cycle towards its cycle root; and then orienting all of the edges in a given cycle one of the two possible directions (there are 2^j ways of doing this). The proof of the following is now similar to the proof of Theorem 3.4.

Theorem 3.5. *There is a one-to-one weight preserving correspondence between perfect matchings on $\mathcal{G}_D$ and in-directed cycle-rooted spanning forests $\mathcal{F}'$ on $\mathcal{G}'_T$ whose dual cycle-rooted spanning forests $\mathcal{F}$ are in-directed, cycle-rooted spanning forests of $\mathcal{G}_T$.*

T-graphs in a torus can be extended to give periodic T-graphs on the plane, finite subsets of which correspond to finite subgraphs of infinite lattice graphs, such as the grid graph in Fig. 4.

4. T-graphs from dimer graphs

In this section, we describe a procedure for generating $\mathcal{G}_T$ from $\mathcal{G}_D$ that applies whenever the so-called Kasteleyn matrix fails to have certain degeneracies. Before we begin the

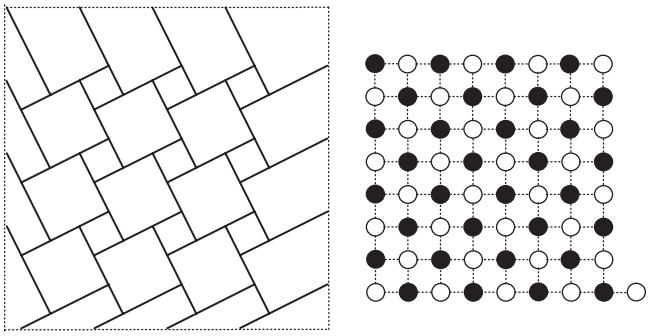

Fig. 4. A T-graph $\mathcal{G}_T$ in the plane and the corresponding graph $\mathcal{G}_D^r$.

construction, we will define Kasteleyn matrices and say a word about the kinds of graphs for which these degeneracies occur.

4.1. Cuts, and breakers

Say a square matrix K is *k-degenerate* if it has an $(n-k)\times(n-k)$ minor whose determinant is zero; otherwise it is *k-non-degenerate*. The following lemma follows from the standard correspondence between determinants of k minors of K^{-1} and $(n-k)$ minors of K:

Lemma 4.1. *K is 0-non-degenerate if and only if it is invertible. Assuming K is invertible, K is k-non-degenerate if and only if K^{-1} is $(n-k)$-non-degenerate.*

Suppose now K is a Kasteleyn matrix for a bipartite planar graph $\mathcal{G}_D$. The following is immediate:

Lemma 4.2. *If K and $\tilde{K}$ are gauge equivalent, then K is k-degenerate if and only if $\tilde{K}$ is k degenerate.*

A bipartite graph is *balanced* if it contains an equal number of black and white vertices. A *k-cut* A of a balanced bipartite graph $\mathcal{G}_D$ is subset of the vertices for which:
1. A contains at least one white vertex,
2. A contains k more black vertices than white vertices,
3. Each edge of $\mathcal{G}_D$ that connects A to its complement has a black vertex in A.

Note that if A is a k-cut, then its complement would be a k-cut if the colors black and white were reversed. In particular, the existence of k cuts does not depend on which of the two ways we choose to color the vertices. Also, if A is a k-cut of $\mathcal{G}_D$, then by adding black vertices to A and/or removing white vertices from A, we can construct m-cuts for any $k \leqslant m \leqslant n-1$. An obvious parity argument implies the following:

Lemma 4.3. *If A is a k-cut of $\mathcal{G}_D$, then any perfect matching of $\mathcal{G}_D$ contains exactly k edges which connect A to its complement; each of these edges matches a black vertex of A and a white vertex of its complement.*

A *k-breaker* is a subset S of the vertices of $\mathcal{G}_D$ with exactly k white and k black vertices for which the induced subgraph $\mathcal{G}_D \backslash S$ of $\mathcal{G}_D$ has no perfect matchings.

Lemma 4.4. *If $\mathcal{G}_D$ is a connected, balanced, bipartite graph, then $\mathcal{G}_D$*
1. *has a (-1)-cut if and only if it has no perfect matching.*
2. *has a 0-cut if and only if $\mathcal{G}_D$ contains unused edges (i.e., edges which occur in no perfect matching of $\mathcal{G}_D$).*
3. *generally has a k-cut if and only if it has a $(k+1)$-breaker.*

Proof. The first item is an immediate consequence of the Hall marriage theorem. That theorem states that $\mathcal{G}_D$ has a perfect matching if and only if there is no set B such that B has m more white vertices than black vertices and there are fewer than m edges connecting white vertices of B to its complement. A (-1)-cut is clearly such a set, with $m = 1$. Conversely, given B as described above, construct B' by removing from B all of the (at most $m-1$) white vertices of B connected to the complement of B, and if necessary, some arbitrary additional white vertices (so that $m-1$ vertices are removed in all). Then B' is a (-1)-cut.

For the second item, first, it is clear that if A is a 0-cut of B, then all of the edges connecting A to its complement will be unused. Conversely, if $\mathcal{G}_D$ has an unused edge e, then the graph G formed by removing edge e and its two vertices from $\mathcal{G}_D$ will not have any perfect matching. Therefore it will have a (-1)-cut A by part 1. The union of A and the black vertex of e is a thus a 0-cut. (Aside: if $\mathcal{G}_D$ has a *forced edge*—i.e., an edge e which occurs in *every* perfect matching of $\mathcal{G}_D$—then all the edges that share vertices with e will be unused.)

The same argument implies the third statement in the case $k = 0$. For larger k, if $\mathcal{G}_D$ has a k-cut A, then any subset of $(k+1)$ black vertices of A and $(k+1)$ white vertices of its complement is a $(k+1)$-breaker (since the remaining set of vertices in A contains more white than black vertices, but there are no edges connecting white vertices of this remaining set with its complement). Conversely, if S is a $(k+1)$-breaker, then $\mathcal{G}_D \backslash S$ has a (-1)-cut A, and the union A and the black vertices of S is a k-cut of $\mathcal{G}_D$. $\square$

Lemma 4.5. *$\mathcal{G}_D$ has no k-breaker (or, equivalently, no $(k-1)$-cut) if and only if, for a generic choice of positive weights of the edges of $\mathcal{G}_D$, the Kasteleyn matrix $K = K(\mathcal{G}_D)$ is k-non-degenerate.*

Proof. The determinant of an $(n-k) \times (n-k)$ minor of the Kasteleyn matrix is a polynomial of the edge weights. Clearly, this polynomial will be zero for a given minor precisely when the set of k white and k black vertices corresponding to rows and columns not in the minor is

a k-breaker. The result follows from the fact that any non-zero polynomial in finitely many real variables is non-zero for a generic choice of inputs. $\square$

In this paper, we will mainly be interested in whether K is k-degenerate for $k \in \{0, 1, 2\}$. But we know that whenever K is a Kasteleyn matrix (of a graph having a perfect matching) it is k-non-degenerate for $k = 0$. And assuming $\mathcal{G}_D$ has no unused edges (which we may always assume throughout, since the perfect matching model will be unchanged if we remove unused edges from $\mathcal{G}_D$) K is generically 1-non-degenerate. We will address the potential failure of K to be 2-non-degenerate in a later section.

4.2. T-graphs: construction via integration of Kasteleyn flow

Let $\mathcal{G}_D$ be a finite, weighted bipartite planar graph (with positive generic weight function v) with n black vertices $b_1, b_2, \ldots, b_n$ and n white vertices $w_1, w_2, \ldots, w_n$. Suppose $\mathcal{G}_D$ has a perfect matching and no unused edges. Suppose that $\mathcal{G}_D$ has no 1-cuts—and hence each of the entries and two-by-two minors of K^{-1} is non-zero (i.e., K is 1-non-degenerate and 2-non-degenerate).

We will now construct a T-graph corresponding to $\mathcal{G}_D$ in the case that K is 2-non-degenerate.

First, we may think of K as describing a linear map from the space $\mathbb{R}^W$ of functions on white vertices to the space $\mathbb{R}^B$ of functions on black vertices. Let b_0 be a fixed vertex on the outer boundary of $\mathcal{G}_D$. Suppose that $\mathcal{G}_D$ has m black and m white vertices on its outer boundary face. Fix a generic convex $m+1$-gon Q with edge vectors $q_0, \ldots, q_m \in \mathbb{C}$ in cyclic order (and $q_0 = -\sum_{i=1}^m q_i$). Vertices of Q will be the root vertices of $\mathcal{G}_T$. Suppose that $A_w \in \mathbb{R}^W$ assumes the values $q_1, \ldots, q_m$ in cyclic order on the white vertices on the boundary face, and that A_w vanishes on all other white vertices of $\mathcal{G}_D$. Let A_b be the function on black vertices which is equal to 1 at b_0 and 0 everywhere else. Denote by $\bar{1}$ the all-ones column vector and by $\bar{1}^t$ its transpose. View A_b as a column vector and A_w as a row vector.

We claim that there is a unique matrix $\tilde{K}$, gauge equivalent to K, for which $\tilde{K}\bar{1}$ is a non-zero multiple of A_b and $\bar{1}^t\tilde{K} = A_w$. The matrix $\tilde{K}$ can be derived explicitly from K as follows. Since K is invertible, there exists a vector f for which $Kf = A_b$. Multiplying the ith column of K by the ith component of f (non-zero, because K is 1-non-degenerate) produces a K' for which $K'\bar{1} = A_b$. Next, there exists a row vector g for which $gK' = A_w$. Multiplying the jth row of K' by the jth component of g (also non-zero, since $(K')^{-1}$ is 1-non-degenerate and nonzero entries of A_w are generic) gives the desired $\tilde{K}$.

We may think of $\tilde{K}$ as describing a vector flow (2-component flow) on $\mathcal{G}_D$: sending $\tilde{K}_{i,j}$ units of flow from b_i to w_j. The net flow into each non-boundary white vertex and each black vertex (except b_0) is zero. Now, draw a dotted line from each white vertex on the outer face of $\mathcal{G}_D$ to infinity, and from b_0 to infinity, so as to divide the outer face of $\mathcal{G}_D$ into $m+1$ outer faces; take these faces and the interior faces of $\mathcal{G}_D$ as the vertices of the dual graph $\mathcal{G}'_D$ of $\mathcal{G}_D$. Then $\tilde{K}$ also describes a dual flow on $\mathcal{G}'_D$ (obtained by rotating each edge ninety degrees counter-clockwise) whose net flow around each non-boundary face of $\mathcal{G}'_D$ is zero; viewed in this light, $\tilde{K}$ is the gradient of a function $\psi : \mathcal{G}'_D \to \mathbb{C}$.

Now, we claim that each pair of (complex) components of g is linearly independent (as a pair of vectors in $\mathbb{C} = \mathbb{R}^2$). To see this, let a and b be basis column vectors, so that $(K')^{-1}(a)$ and $(K')^{-1}b$ are columns of the matrix $(K')^{-1}$. Since the determinants of the two-by-two minors of $(K')^{-1}$ are non-zero, no complex component of the vector $z = (K')^{-1}a + i(K')^{-1}b = (K')^{-1}(a + ib)$ is a real multiple of any other component of that vector (in particular, all of the components of z are non-zero). Now A_w is a generic linear combination of vectors of the above form $a + ib$, so no component of $g = K^{-1}(A_w)$ is a real multiple of any other component of g.

Since K' is real, all the components of $\tilde{K}$ in a given row are nonzero complex numbers lying on the same line through the origin, and the directions are different in each row.

Now, extend ψ linearly to the edges of $\mathcal{G}'_D$, so that ψ maps each edge to a line segment. For each black vertex b_i of $\mathcal{G}_D$, corresponding to a black face of $\mathcal{G}'_D$, the ψ image of the union of the edges incident to b_i is a line segment, whose interior we denote by L_i; the above argument implies that no two of the L_i are parallel.

Here is the main result.

Theorem 4.6. *If K is 2-non-degenerate then the $L = \{L_1, \ldots, L_n\}$ defined above forms a T-graph, whose tree-graph we denote $\mathcal{G}_T$, with root vertices at vertices of Q, and $\mathcal{G}_D = \mathcal{G}_D(L)$ (up to gauge equivalence). Moreover, if v is a vertex of $\mathcal{G}'_D$, then $\psi(v)$ is a vertex of $\mathcal{G}_T$; if v corresponds to an outer face of $\mathcal{G}_D$, then $\psi(v)$ is a root vertex of $\mathcal{G}_T$.*

Proof. First, the change in ψ, as one moves from outer face F of $\mathcal{G}_D$ around a vertex v to another outer face, is given by the flow of $\tilde{K}$ into v, which is given by q_i, the ith component of A_w, whenever v is a white vertex w_i, and zero when v is any black vertex besides b_0. By moving around the polygon in steps, it is clear that (up to an additive constant) $\psi(F)$ assumes the values of the vertices of the convex polygon in cyclic order.

Let f be an interior vertex of $\mathcal{G}'_D$. We claim that for some black face incident to f, with vertices f_1 and f_2 incident to f, $\psi(f_1) - \psi(f)$ and $\psi(f_2) - \psi(f)$ point in opposite directions. Suppose otherwise. Then $\tilde{K}$ would have to assume opposite signs on the entry corresponding to each such pair of edges (f, f_1) and (f, f_2). By the definition of a Kasteleyn matrix, $\tilde{K}$ has positive sign for an odd (resp., even) number of the edges incident to f if the total number of edges is 0 mod 4 (resp., 2 mod 4), so this is a contradiction. It follows that $\psi(f)$ is an interior vertex of at least one L_i. In particular, this implies that the endpoint of each L_i is either an interior vertex of some L_j or a root vertex.

It also implies a *maximal principle*, i.e., that for any vector u in $\mathbb{R}^2$, the function $\psi_u(x) = (\psi(x), u)$ (an inner product computed with $\psi(x)$ treated as a vector in $\mathbb{R}^2$) has no local maxima or minima at interior faces of $\mathcal{G}_D$. That is, every interior face f (viewed as an interior vertex in $\mathcal{G}'_D$) has neighbors f_1 and f_2 satisfying $\psi_u(f_1) \leqslant \psi_u(f_1) \leqslant \psi_u(f_2)$. For generic u (i.e., any u whose slope is not parallel to one of the L_i's), the inequality can be made strict.

Now, to show that the $\{L_i\}$ form a T-graph, it remains only to show that they do not intersect one another; while proving this, we will also show that $\psi(\mathcal{G}'_D)$ partitions the convex polygon Q into convex polygons (the white faces). First, the maximal principle immediately implies that $\psi(\mathcal{G}'_D)$ lies in Q. Furthermore, we claim that as one moves x

clockwise around each a white interior face w of $\mathcal{G}'_D$, $\psi(x)$ traces out a convex polygon in some fixed orientation (clockwise or counterclockwise; we refer to this direction as the *orientation of w* and denote the polygon by $\psi(w)$). If this were not the case, then there would have to be vertices f_1, f_2, f_3, f_4, in clockwise order around w and some generic u for which $\psi_u(f_1)$ and $\psi_u(f_3)$ are less than both of $\psi_u(f_2)$ and $\psi_u(f_4)$. By the maximal principle, we can find paths in p_2 and p_4 in $\mathcal{G}'_D$ from f_2 and f_4 to root vertices along which ψ_u is strictly increasing and paths p_1 and p_3 from f_1 and f_3 to root vertices along which ψ_u is strictly decreasing. Now, let p be a path in $\mathcal{G}'_D$ formed by concatenating p_1 (reversed), a dotted line from f_1 to f_3, and p_3. This path cannot intersect p_2 or p_4 (since ψ_u at any point on these two paths is greater than ψ_u at any point on p_1 or p_3). However, the Jordan curve theorem implies that p separates its complement in $\mathcal{G}'_D$ into at least two connected components and that f_2 and f_4 (which lie on either side of p across the face w) are in separate components (this remains true even for the graph $(\mathcal{G}^Q_D)'$ formed by adding to $\mathcal{G}'_D$ the edges connecting each cyclically consecutive pair of outer vertices of $\mathcal{G}'_D$). Now, the paths p_2 and p_4 both lead to root vertices at which ψ_u assumes a larger value than it does at any point along p, and these points are in the same component of $(\mathcal{G}^Q_D)'$, a contradiction. A similar argument shows that the outer faces w, joined with this, have this orientation. Another similar argument applies to black faces and shows that as one moves x around a black interior face b of $\mathcal{G}'_D$, $\psi(x)$ traverses the corresponding L_i exactly once in each direction.

Next, we argue that all white faces have the same orientation. It is enough to prove that any white faces of $\mathcal{G}'_D$ (vertices of $\mathcal{G}_D$) w_1 and w_2 incident to a common black b have the same orientation. Now, as x traverses the boundary of the face b in $\mathcal{G}'_D$, $\psi(x)$ traces out the corresponding L_i once in each direction; divide the faces incident to b into two categories according to the orientation of the edge shared with b. Clearly, if these faces do not all have the same orientation, we can find two of them, w_1 and w_2 in opposite categories that have opposite orientations. In this case, $\psi(w_1)$ and $\psi(w_2)$ will lie on the same side of b; let u be vector orthogonal to L_i; assume without loss of generality that ψ_u assumes a larger value on points on L_i than on other points of w_1, w_2. Let f_1 and f_3 be the points in $\mathcal{G}'_D$ incident to b whose images are the endpoints of b, and let f_2 and f_4 be arbitrary points of w_1 and w_2 which do not lie on b. Let p be formed by concatenating a path p_1 from f_1 to a root on which ψ_u is strictly increasing (reversed), a dotted line from f_1 to f_3, and a path p_3 from f_3 to a root vertex along which ψ_u is strictly increasing; observing that f_2 and f_4 are on opposite sides of p, we derive a contradiction through the Jordan curve argument described above.

Finally, suppose that two of the L_i intersect. Then there must be two faces w_1 and w_2 for which $\psi(w_1)$ and $\psi(w_2)$ intersect. The outer boundary of $(\mathcal{G}^Q_D)'$ is mapped with some consistent orientation to Q. Now, let $h : Q \to \mathbb{Z}$ at x be the number of white faces $\psi(w)$ which contain x in their interiors. It is clear that h assumes the value 1 near the boundary. We claim that h is equal to one throughout $Q \backslash \psi(\mathcal{G}^Q_D)'$; otherwise, there would be an x in the interior of Q (and not at the finitely many endpoints of any L_i or intersections of pairs of L_i) on the boundary of regions at which h assumes different values. Such an x must lie on some L_i, and it is not hard to see that the two white faces incident to x and L_i must have opposite orientations. $\square$

4.3. Flat-face degeneracy

Now, suppose that K is merely 1-non-degenerate and not necessarily 2-non-degenerate; then we can formally construct ψ exactly as above; in this case, however, we cannot rule out that some of the L_i may be parallel to one another—and in fact, some of the L_i may overlap. However, the same arguments given above still imply that for each white w, $\psi(w)$ is *either* a convex face with some orientation (as described above) or a line segment traversed once in each direction (like the black faces). In the latter case, we say $\psi(w)$ is a *degenerate face*. In the presence of degenerate faces, we will consider $\psi(w)$ and $\psi(b)$ to be incident to one another along an edge if and only if w and b are adjacent vertices in $\mathcal{G}'_D$.

It is clear that if a white vertex w is degenerate, then $\psi(b)$ is parallel to $\psi(w)$ for each black b adjacent to w. A maximal component of the subgraph of $\mathcal{G}'_D$ consisting of vertices on which ψ is parallel to a given line is called a *parallel component* of $\mathcal{G}'_D$. Clearly, the neighbor set of any white vertex in a parallel component is also in the parallel component.

An *extreme point* of a degenerate face w is a vertex f incident to w for which $\psi(f)$ is an endpoint of $\psi(w)$. The union of ψ-images of a parallel component is a segment which we call an *extended complete edge*. Now observe the following.

Lemma 4.7. *Each parallel component P is a* 1*-cut.*

Proof. Observe that every f which is an interior vertex of a black edge of $\mathcal{G}'_D$ in a parallel cluster is the extreme vertex for the same number of black and white faces of $\mathcal{G}'_D$. The endpoints of the extended complete edge are extreme points of one more black vertex than white vertices. Since every face has exactly two extreme vertices, the result follows. □

Similar arguments to those given in the proof of Theorem 4.6 imply that as x traverses the outside of a parallel component, $\psi(x)$ traverses the outside of the extended complete edge exactly once in each direction. Similar arguments to those of Theorem 4.6 imply that the extended complete edges form a T-graph. We say that $L = \{L_i\}$ forms a *T-graph with overlaps* if L_i satisfies all of the T-graph conditions except that parallel pairs of L_i are allowed to intersect (overlap) one another. The above analysis implies the following:

Theorem 4.8. *Theorem* 4.6 *still holds if K is merely* 1*-non-degenerate and not necessarily* 2*-non-degenerate—except that in this case, some of the white faces may be degenerate* (*and so the T-graph may have overlaps*). *Theorem* 3.4 *still applies to T-graphs with overlaps*.

Even though some of the white faces are flat in the overlapping T-graph $\mathcal{G}_T$, we can define a dual to the overlapping T-graph, containing these faces, using the graph structure of $\mathcal{G}_D$. After doing this, all of the arguments in the proof of Theorem 3.4 apply as before, so we still have a martingale on the T-graph and have a measure-preserving correspondence between spanning forests and perfect matchings.

Recall that in any perfect matching, there is always exactly one edge connecting a given 1-cut to its complement, and that edge contains a black vertex of the 1-cut. It is perhaps not surprising that when we form the T-graph, 1-cuts, in some sense, play the same role

as single black vertices. If we had simply replaced all 1-cuts in our original graph with single black vertices, then, for a generic choice of weights, the T-graph would not have any degenerate white faces.

4.4. Extending the correspondence to degenerate weighted graphs

Recall from Lemma 4.5 that if we remove the unused edges from $\mathcal{G}_D$, then the Kasteleyn matrix for $\mathcal{G}_D$ is 1-non-degenerate (and hence Theorems 3.4 and 4.8 apply) for a generic choice of weight functions v. Suppose, however, that the Kasteleyn matrix for $\mathcal{G}_D$ is not 1-non-degenerate for a particular choice of weight function v. Then we would like to take a generic sequence of weights v_i converging to v, look at the limit (or some subsequential limit) of the corresponding T-graphs, and show that the measure-preserving correspondence described in Theorem 3.4 still holds for the limiting object. The problem is that, as Fig. 1 makes clear, the limit of a sequence of T-graphs need not be a T-graph at all; in fact, some of the edge segments and faces may shrink to single points.

For practical computational applications, it may be sufficient to have the correspondence between dimers and spanning forests for a generic choice of weights. But a word of caution is in order. Consider the dimer model whose T-graph is given by the right diagram in Fig. 1; if weights v_i tend to a limit v in such a way that the T-graphs have the graph on the left as a limit, then the shrinking small triangle in the center of the diagram will become a "trap" for the random walk on the T-graph, in that the expected amount of time that a walk spends on these three vertices before existing towards a root vertex tends to infinity; sampling algorithms that rely on random walks will perform poorly for weights approximating v. In this case, however, one can simplify the limiting problem by reducing the three vertices in the small triangle at the center to single vertex. The probability tends to one that only one of the "long" directed edges (i.e., edges whose lengths are not tending to zero) extending outward from these three vertices will appear in a random tree; given a spanning tree of the "reduced" graph, it is possible to work out which "short" edges appear in the graph. The details of this and more general versions of this reduction are left to the reader.

5. Periodic and almost periodic T-graphs

5.1. Definitions for almost periodic T-graphs

In this section, we prove some results about T-graphs which are motivated by the study of ergodic Gibbs measures on tilings of infinite periodic planar graphs. More on this subject can be found in [9], who cite the results of this section. Our first aim here is to construct from periodic bipartite planar graphs (and under certain conditions on the weights) infinite T-graphs with a property called "almost periodicity."

Let $\mathcal{G}_D$ be embedded in the torus $\mathbb{R}^2/\mathbb{Z}^2$ and let $\mathcal{G}_D^\infty$ be the doubly periodic lift to $\mathbb{R}^2$ (we assume that $\mathcal{G}_D^\infty$ is connected). As before, assume that $\mathcal{G}_D$ has n white and n black vertices. Denote by $v_{j,k}$ the vertex of $\mathcal{G}_D^\infty$ which lies in the square $[j, j+1) \times [k, k+1)$ and whose projection to the torus is the vertex $v \in \mathcal{G}_D$. For the sake of simplicity we will assume throughout this section that $\mathcal{G}_D$ has no unused edges and that it has generic weights. The

non-generic weight case requires a slightly finer analysis which we choose not to go into here.

A function f on the vertices of $\mathcal{G}_D$ is (α, β)-*periodic* if $f(v_{j+x,k+y}) = \alpha^x \beta^y f(v_{j,k})$ for all $(v_{j,k}) \in \mathcal{G}_D^\infty$. Say f is *almost periodic* if it is (α, β)-periodic and α and β have modulus one (but are not necessarily roots of unity). In this case, we write $\alpha = e^{2\pi i a}$ and $\beta = e^{2\pi i b}$.

If a and b are rational, then f is doubly periodic with some period.

For a fixed (α, β) the linear space of (α, β)-periodic functions is $2n$-dimensional and is parametrized by the space of functions on one period of $\mathcal{G}_D^\infty$—which we can represent as a single copy of $\mathcal{G}_D$. It has a natural basis consisting of functions δ_v whose value is 1 at $v \in [0, 1)^2$ and zero at other vertices in the fundamental domain. Let K be a Kasteleyn matrix for $\mathcal{G}_D$ and K^∞ an infinite-dimensional Kasteleyn matrix for $\mathcal{G}_D^\infty$ which is a lift of K. We can think of K^∞ as a linear function from the set of functions on the black vertices of $\mathcal{G}_D^\infty$ to functions on the white vertices of $\mathcal{G}_D^\infty$. Since this function maps (α, β)-periodic functions to (α, β)-periodic functions, it induces a linear map from the n-dimensional space of functions on the black vertices of $\mathcal{G}_D$ to the n-dimensional space of functions on white vertices of $\mathcal{G}_D$. Denote by $K_{\alpha,\beta}$ the matrix of this linear map in the basis $\{\delta_v\}$.

The determinant $\det K_{\alpha,\beta}$ is a polynomial function of α and β; in particular for certain (α, β) (corresponding to zeros of this polynomial function) the matrix $K_{\alpha,\beta}$ has a non-trivial null space, and hence we can find (α, β)-periodic functions f and g satisfying $K^\infty f = 0$ and $g K^\infty = 0$. If the polynomial $\det K_{\alpha,\beta}$ *happens* to have a zero (α, β) that lies on the unit torus of complex variable pairs that both have modulus one, then f and g are almost periodic. If, furthermore, f and g *happen* to be nowhere zero, then we can define an infinite T-graph as follows. First, observe that the function $\tilde{K}_1^\infty(vw) = f(v)g(w)K^\infty(v, w)$ on edges vw of $\mathcal{G}_D^\infty$ is a nowhere zero flow. The dual of this flow is the gradient of a function ψ_1 on $\mathcal{G}_D'$. Similarly the dual of $\tilde{K}_2^\infty(vw) = f(v)\overline{g(w)}K^\infty(v, w)$ is the gradient of a function ψ_2 on $\mathcal{G}_D'$ (where $\bar{g}$ denotes the complex conjugate of g). We may assume (multiplying $g(w)$ by a generic modulus one complex number if necessary) that $g(w) + \overline{g(w)} = 2\mathrm{Re}g(w)$ is also nowhere zero. Then we can think of $\tilde{K} = \tilde{K}_1 + \tilde{K}_2$ as an infinite Kasteleyn matrix and $\psi = \psi_1 + \psi_2$ as the corresponding T-graph. We will call a mapping ψ from $(\mathcal{G}_D^\infty)'$to $\mathbb{R}^2$, constructed in this way, an *almost periodic T-graph mapping*. See Fig. 5.

We remark that, given a fixed v, the number of α, β on the unit torus for which $\det K_{\alpha,\beta} = 0$ also plays a fundamental role in [9], where it is shown that the minimal specific free energy ergodic Gibbs measure on perfect matchings of the infinite weighted graph $\mathcal{G}_D^\infty$ is *smooth* if the corresponding polynomial $K_{\alpha,\beta}$ has 0 roots on the unit torus and *rough* if it has 2 roots (necessarily complex conjugates) on the unit torus (in the non-generic case of a single root, it is rough only when $\frac{d}{d\alpha} \det K_{\alpha,\beta} = \frac{d}{d\beta} \det K_{\alpha,\beta} = 0$). The terms "smooth" and "rough" come from the statistical physics literature and are defined in [9]. The main goal of this section is to prove that when the choice of weights is generic, the number of modulus-one values of (α, β) that are roots of $\det K_{\alpha,\beta}$ always belongs to the set $\{0, 2\}$.

5.2. *Generic points on the variety of almost periodic T-graphs*

Write $\mathbb{R}_+$ for the set of strictly positive real numbers, $\mathbb{C}_+$ for the set of non-zero complex numbers, and write $\mathbb{P}^k$ for k-dimensional complex projective space. Suppose that $|V| = 2n$,

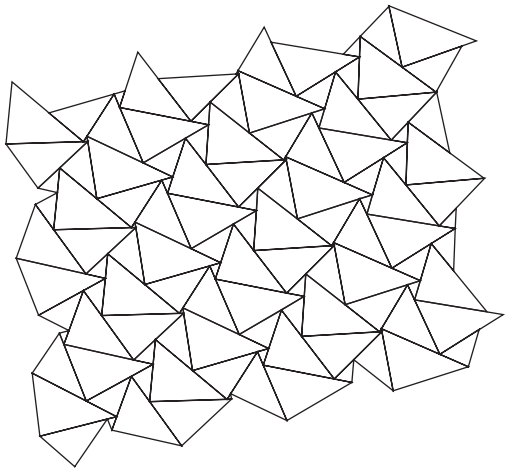

Fig. 5. Almost periodic T-graph mapping of the honeycomb graph, with periodic edge weights 4, 5, and 6 according to direction. The edges of the graph shown correspond to black vertices of the honeycomb lattice; the triangular faces of the graph shown correspond to white vertices of the honeycomb lattice.

and define a variety $X \subset \mathbb{R}_+^{|E|} \times \mathbb{C}_+^2 \times \mathbb{P}^{n-1}$ by

$$X = \{(v, \alpha, \beta, f) : K_{\alpha,\beta} f = 0\}.$$

Here f is an element in $\mathbb{P}^{n-1}$, which is a one-dimensional subspace of $\mathbb{C}^n$, and by $K_{\alpha,\beta} f = 0$ we mean that this subspace lies in the null space of $K_{\alpha,\beta}$. By abuse of notation, if f is a non-zero function on the black vertices of $\mathcal{G}_D$, we will also use f to denote the element of $\mathbb{P}^{n-1}$ given by the linear span of f. Denote by $\tilde{X}$ the subset of X consisting of points for which $|\alpha| = |\beta| = 1$. Denote by $\mathrm{Adj}(K_{\alpha,\beta})$ the *adjugate* matrix of $K_{\alpha,\beta}$, whose entries are the $(n-1) \times (n-1)$ minors of $K_{\alpha,\beta}$ (so that $K_{\alpha,\beta}\mathrm{Adj}(K_{\alpha,\beta}) = \det K_{\alpha,\beta}$). It is easily seen that $\mathrm{Adj}(K_{\alpha,\beta})$ is identically equal to zero if and only if the rank of $K_{\alpha,\beta}$ is less than $n-1$; and if the rank of $K_{\alpha,\beta}$ is exactly $n-1$, then at least one column of $\mathrm{Adj}(K_{\alpha,\beta})$ is a non-zero vector whose span is the null space of $K_{\alpha,\beta}$. The following is the main result of this section:

Theorem 5.1. *The variety $\tilde{X}$ is irreducible. For a generic choice of v, there are either zero or two quadruples (v, α, β, f) in X. When the latter is the case and v is generic, then the corresponding α, β, f are such that $\mathrm{Adj}(K_{\alpha,\beta})$ has rank $n-1$, all of its coordinates are non-zero, and f is given by any column of the (rank one) matrix $\mathrm{Adj}(K_{\alpha,\beta})$.*

Let us say a word about the significance of this theorem to T-graph classification before we prove it. By obvious symmetry, Theorem 5.1 implies that for generic v, there are either zero or exactly two quintuples (v, α, β, f, g) with $gK_{\alpha,\beta} = 0$ and $fK_{\alpha,\beta}$. Recall that our almost periodic T-graphs were defined to have gradient given by $\tilde{K}^{\infty}(vw) = 2f(v)\mathrm{Re}(g(w))K^{\infty}(v, w)$. Since f and g are uniquely determined up to complex conjugacy and multiplication by a constant factor, this implies that the almost periodic T-graph is

completely determined up to rotations (which arise from multiplying f by a modulus one constant), constant rescalings (which arise from multiplying either g or f by a real constant), reflection (which comes from complex conjugacy), translations of the image space (which arise from the fact that $\tilde{K}^{\infty}$ only determines the T-graph mapping up to an additive constant) and "translation of the domain, or a limit of such translations." To explain the last symmetry, note that multiplying both f and g by $\alpha^m\beta^n$ is equivalent to composing the T-graph mapping with translation of the domain by (m, n). If one of α, β is irrational, then we can achieve any modulus one number as a limit of numbers of the form $\alpha^m\beta^n$. We summarize these observations informally by saying that "the almost periodic T-graph mapping corresponding to v is unique up to affine orthogonal transformations of the image and translations of the domain." We say two T-graphs are equivalent if one can be obtained from the other via a symmetry of this sort. Note, of course, that if α and β are both rational, then multiplying f and g by a modulus one number is *not* necessarily the same as a domain translation, or even a limit of such translations. In this case, there is a one parameter family of T-graph equivalency classes.

We will now prove Theorem 5.1 in stages, beginning with the following lemma. First, denote by X' the projection of X onto its first three coordinates (v, α, β); i.e., X' is the zero set of the polynomial $P(v, \alpha, \beta) = \det K_{\alpha,\beta}$.

Lemma 5.2. *The variety X' is irreducible. Moreover, for a generic point(v, α, β) on X', the matrix* $\mathrm{Adj}(K_{\alpha,\beta})$ *has no zero entries, and the f for which $(v, \alpha, \beta, f) \in X$ is unique.*

Proof. Clearly, P is affine linear as a function of $v(e)$, that is $P = v(e)P_e + P'_e$, where P_e and P'_e do not involve $v(e)$. If we could write $P = P_1P_2$, then each $v(e)$ must occur in either P_1 or P_2, but not both. Since the multiplicity of the $v(e)$terms determine the multiplicity of α and β in each monomial, this implies that there is no cancellation when multiplying out P_1 times P_2 (i.e., there are no monomials that can represented as a product of a monomial in P_1 and a monomial in P_2 in two different ways). Thus, each monomial in P_1 times a monomial of P_2 corresponds to a matching. Let E_1, E_2 be the set of edges represented in P_1, P_2, respectively, and V_1, V_2 their vertices. If an edge e connected a vertex v_1 of V_1 to a vertex v_2 of V_2, then its weight could *not* occur in either P_1 or P_2, since if it occurred in a monomial of, say, P_1, then the product of that monomial with a monomial of P_2 that included a factor of $v(e')$ with e' incident to v_2 (such a monomial exists by definition) would *not* correspond to a matching, since it would involve two edges incident to v_2. Thus e must be unused, a contradiction. Thus, if $P = P_1P_2$, then one of the P_i—say, P_2—must be a function of α and β alone. Since each combination of edge weights corresponding to a matching occurs in exactly one monomial of P, we conclude that P_2 is a monomial in α and β.

Furthermore, P is irreducible when considered as a polynomial in both the edge weights and α, β, except for a monomial factor in α and β. That is, if $P = P_1(v, \alpha, \beta)P_2(v, \alpha, \beta)$ then one of the P_i consists of a single monomial in α and β. To see this, note that by the previous result, we may assume without loss of generality that P_2 is a polynomial in α, β alone; and since we are assuming $\alpha \neq 0$, $\beta \neq 0$, the variety is not changed if we divide out by this term so that P is an irreducible polynomial.

Fix an edge e and consider the polynomial P_e as defined above. Since P is irreducible and e occurs in a proper subset of the set of all matchings, the zero set of $v(e)P(e)$, intersected with X', forms a proper subvariety of X'. In other words, on a generic subset of X', none of the entries of $\mathrm{Adj}(K_{\alpha,\beta})$ corresponding to an edge in $\mathcal{G}_D$ are zero. Since $\mathrm{Adj}(K_{\alpha,\beta})$ has rank at most one, and every row and column has a non-zero entry, we conclude that every entry of $\mathrm{Adj}(K_{\alpha,\beta})$ is non-zero and f is the span of any column of $\mathrm{Adj}(K_{\alpha,\beta})$. □

Lemma 5.3. *For a generic choice of weights v, every pair α, β for which $(v, \alpha, \beta) \in X$ is such that $\mathrm{Adj}(K_{\alpha,\beta})$ has no zero entries, and the f for which $(v, \alpha, \beta, f) \in X$ is unique.*

Proof. Lemma 5.2 implies that for *generic* edge weights v, P and P_e have no common factor as functions of α and β except for monomial factors. To see this, by irreducibility note that there exist polynomials $Q_1 = Q_1(\alpha, \beta, w)$ and $Q_2 = Q_2(\alpha, \beta, w)$ such that $PQ_1 + P_eQ_2 = Q(\alpha, w)$ where Q is a nonzero polynomial depending only on α and the weights w, not on β. Similarly there exist Q_3, Q_4 such that $PQ_3 + P_eQ_4 = Q'(\beta, w)$ where Q' is a non-zero polynomial independent of α. Plugging in generic values for w, Q and Q' will still be nonzero, but any common factor of P and P_e is a common factor of Q and Q' which is impossible. So P and P_e have no common factor for generic w.

Therefore, when v is fixed generically, by Bezout's theorem P and P_e—viewed as polynomials in α and β—have a finite number of common zeros. By genericity none of these zeros lies on the unit torus (since for any positive real x, we can choose v_x so that $P(v, \alpha, \beta) = P_{v_x, x\alpha, x\beta}$; and replacing v with such a v_x, for a generic choice of x, preserves the genericity of the weights). □

Lemma 5.4. *Any almost periodic T-graph mapping ψ is unbounded as a function of $(\mathcal{G}_D^\infty)'$. Moreover if u is any vector in $\mathbb{R}^2\backslash\{0\}$, then (ψ, u) is unbounded if it is not identically equal to a constant* (*in which case ψ is degenerate—i.e., its image is contained in a line*).

Proof. Suppose that f is (α, β)-periodic and g is (γ, δ)-periodic with $\gamma = e^{2\pi i c}$ and $\delta = e^{2\pi i d}$. Then $\tilde{K}_1^\infty(v_{j,k}, w_{j+\ell,k+m})$ is a function of ℓ, m whose real and imaginary parts can both be written in the form $\cos(a\ell + bm + x)\ \cos(c\ell + dm + y)$ times a constant, for some x and y.

If ψ were bounded on $\mathcal{G}_D'$, then the corresponding martingale on the T-graph would almost surely converge (by the martingale convergence theorem), and there would thus have to be a path of vertices $v_1, v_2, \ldots$ for which $\psi(v_i)$ converges to a constant. We claim that this is impossible. It is enough to show that for some ε, the set of edges $(vw)^*$ for which $0 < \tilde{K}^\infty(v, w) < \varepsilon$ has no infinite cluster. For some $N > |\mathcal{G}_D|$, we can always find ε small enough so that the distance between any two clusters of $(\ell, m) \in \mathbb{Z}^2$ (viewed as points in $\mathbb{Z}^2$) on which $0 < \cos(a\ell + bm + x) < \varepsilon^{1/2}$ is at least $2N$ times the diameter of the largest such cluster, and similarly for clusters on which $0 < \cos(c\ell + dm + y) < \varepsilon^{1/2}$. (This is trivial if a and b are rational, since the function is periodic in that case; if they are

irrational, then we can find ε_0 for which there is no integer pair (n_1, n_2) for which $n_1 a + n_2 b$ is less than ε_0 (modulo 2π) and $|n_1 + n_2| \leqslant 2N$. Choose ε_0 small enough that there can't be two values differing by ε_0 (modulo 2π) with cosines ε apart.) Now, it is clear that the largest cluster of ℓ, m on which even one of these statements holds is at most $2N$; since the gradient of ψ has norm at least ε when neither statement holds, we conclude that ψ cannot be bounded.

The same argument shows that there cannot exist a non-zero vector $u \in \mathbb{R}^2$ for which the inner product $(\psi(v), u)$ is bounded as a function of v, unless $(\psi(v), u)$ is constant. □

Lemma 5.5. *If v is generic, then the maximum number of linearly independent, almost periodic solutions to $K^\infty f = 0$ (or similarly, solutions to $gK^\infty = 0$) is two. If there are two solutions, which are (α, β)- and (γ, δ)-periodic, then $\alpha = \bar{\gamma}$ and $\beta = \bar{\delta}$.*

Proof. For each α and β, the left null space of $K_{\alpha,\beta}$ has the same dimension as the right null space. Now, suppose that f is α, β-periodic and g is γ, δ-periodic with $\gamma = e^{2\pi i c}$ and $\delta = e^{2\pi i d}$. Then as in the proof of Lemma 5.4, $\tilde{K}^\infty(v_{j,k}, w_{j+\ell,k+m})$ is a function of ℓ, m whose real and imaginary parts can both be written in the form $\cos(a\ell + bm + x)\ \cos(c\ell + dm + y)$ times a constant, for some x and y. Let S be a cycle in $\mathcal{G}'_D$; if we lift it to $(\mathcal{G}^\infty_D)'$, then its endpoints are its starting points plus an integer pair, (n_1, n_2). Now, we would like to determine the asymptotics of ψ_1 and ψ_2 (whose derivative is the dual of $\tilde{K}^\infty$) along S^∞ (a periodic lifting of S to $\mathcal{G}^\infty_D$). Expanding the cosines in exponentials, this involves adding up $|S|$ separate sequences (functions of ℓ) of the form:

$$\sum_{\ell=1}^{n_1} e^{2\pi i[(x+\ell a)\pm(y+\ell c)]}$$

and $|S|$ sequences of the corresponding form for m.

Clearly, ψ will remain bounded independently of x and y, provided $a \neq \pm c \bmod 2\pi$ and $b \neq \pm d \bmod 2\pi$. In fact we must take the same sign for both equalities: unless $(a, b) = \pm(c, d) \bmod 2\pi$ it is possible to find an independent pair of integer vectors (m_1, n_1) and (m_2, n_2) for which $am_1 + bn_1 \neq \pm(cm_1 + dn_1) \bmod 2\pi$ and similarly $am_2 + bn_2 \neq \pm(cm_2 + dm_2) \bmod 2\pi$. Taking S_1 and S_2 to be corresponding paths, we may deduce that ψ is bounded unless (α, β) and (γ, δ) are either equal to one another or conjugates; by Lemma 5.4 (α, β) and (γ, δ) are either equal to one another or conjugates.

Now suppose we have $(a, b) = \pm(c, d)$. Then for the sums corresponding to steps in S,

$$\sum_{\ell=1}^{n_1} \cos(x + \ell a)\cos(y \pm \ell a)$$

is *approximately linear* as a function of n_1, that is, equal to a linear function plus a bounded function. If there were three linearly independent solutions f_1, f_2, f_3 to $Kf = 0$, and ψ_1, ψ_2, ψ_3 are formed using g and f_1, f_2, f_3, then a linear combination of the ψ_1, ψ_2, ψ_3 would be approximately the linear function zero (i.e., bounded), a contradiction, by Lemma 5.4.

Finally, since it is clear that (α, β) is not real (i.e., not equal to ± 1) for a generic choice of v, so any almost periodic f or g will be a strictly non-real function, that is, linearly independent from its complex conjugate, which is also a zero of K^{∞}. $\square$

Now, Theorem 5.1 now follows immediately from Lemmas 5.3 and 5.5.

Acknowledgements

We thank Andrei Okounkov for several ideas and discussions.

References

[1] R.L. Brooks, C.A.B. Smith, A.H. Stone, W.T. Tutte, Dissecting a rectangle into squares, Duke Math. J. 7 (1940) 312–340.
[2] R. Burton, R. Pemantle, Local characteristics, entropy and limit theorems for uniform spanning trees and domino tilings via transfer-impedances, Ann. Probab. 21 (1993) 1329–1371.
[3] R.J. Duffin, Basic properties of discrete analytic functions, Duke Math J. 23 (1956) 335–363.
[4] P. Kasteleyn, Graph theory and crystal physics, in: Graph Theory and Theoretical Physics, Academic Press, London, 1967, pp. 43–110.
[5] R. Kenyon, The Laplacian and Dirac operators on critical planar graphs, Invent. Math. 150 (2002) 409–439.
[6] R. Kenyon, Conformal invariance of domino tiling, Ann. Probab. 28 (2000) 759–795.
[7] R. Kenyon, Tilings and discrete Dirichlet problems, Israel. J. Math 105 (1998) 61–84.
[8] R. Kenyon, Height fluctuations in the lozenge tiling model. arXiv:math-ph/0405052.
[9] R. Kenyon, A. Okounkov, S. Sheffield, Dimers and Amoebae, Ann. Math., to appear, arXiv:math-ph/0311005.
[10] R. Kenyon, J. Propp, D. Wilson, Trees and Matchings, Electron. J. Combin. 7 (2000), Research paper 25, 34pp.
[11] H. Temperley, Combinatorics: Proceedings of the British Combinatorial Conference 1973, London Mathematical Society Lecture Notes Series, vol. 13, 1974, pp. 202–204.
[12] D. Wilson, Generating random spanning trees more quickly than the cover time. Proceedings of the 28th Annual ACM Symposium on the Theory of Computing, Philadelphia, PA, 1996, ACM, New York, 1996, pp. 296–303.