LECTURE NOTES ON THE DIMER MODEL

FABIO LUCIO TONINELLI

Abstract. Unrevised draft: can contain mistakes! If you spot any, let me know.

Contents

These lecture notes are about the so-called dimer model. This is a two-dimensional statistical mechanics model that, as discovered by Kasteleyn

and Temperley-Fisher in the '60s [14, 15, 16, 10, 29], is "exactly solvable" in the sense that in finite volume, its partion functions and correlations can be written as determinants or Pfaffians. Recently, interest in the dimer model has been revived by the discovery of its properties of conformal invariance and its link to the massless Gaussian field (or Gaussian Free Field) on one side, and because of its tight relation with other two-dimensional statistical mechanics model like the Ising model and the 6-vertex model.

See also e.g. [16, 20, 3, 7] for other reviews on dimer models.

1. Perfect matchings on bipartite planar graphs

Let $G = (V, E)$ be a planar, simple bipartite graph with edge set E and vertex set V. "Bipartite" means that we can decompose the set of vertices V as $V = V_W \cup V_B$ (W/B for white/black) with $V_W \cap V_B = \emptyset$ and such that each edge $e \in E$ has one endpoint in V_B and the other in V_W. "Simple" means that "self-loops" (edges having two coinciding endpoints) and multiple edges between two vertices are not allowed[1]. "Self-loops" are forbidden also by bipartiteness. The graph can be either finite or infinite. The dual graph is denoted as usual G^*.

A *perfect matching* of G (also referred to as or *close-packed dimer configuration* or, for our purposes, simply *dimer configuration*) is a subset $M \subset E$ of edges such that every $v \in V$ is contained in one and only one edge $e \in M$. Of course, the number of edges in M is exactly $|V|/2$. Edges in M are called *dimers*.

Assumption 1.1. *We always assume, to avoid trivialities, that:*

(i) *the set Ω_G of perfect matchings of G is not empty. In particular, $|V_W| = |V_B| = |V|/2$.*
(ii) *G is connected;*
(iii) *for every edge e, there is at least a perfect matching that contains e and one that does not contain it.*

Note that assumptions (ii) and (iii) are not really restrictive. First of all, if the graph is disconnected, then the partition function of the dimer model (defined in Section 2.1) factorizes as the product of partition functions on each connected component of G. Secondly, if an edge is contained in all matchings, we can as well remove its two endpoints, together with all edges incident to it. Then, the perfect matchings of the new graph thus obtained are in obvious bijection with those of G. In particular, note that (ii) implies that we are assuming that there are no vertices of degree 1, otherwise the unique incident edge would be always occupied by a dimer. Finally, if e is contained in no perfect matching then we may as well remove it from the edge set: the perfect matchings of the new graph coincide with those of the original one.

Later we will also consider one case where G is not planar, namely the case where it is a discrete torus, and also the case where G is planar but not biopartite (e.g. the triangular lattice). We start with the dimer model in the planar bipartite case which is the simplest.

[1] It is simple to adapt the theory when multiple edges are present. See footnote 2

1.1. **Height function.** In the bipartite case, a perfect matchings can be equivalently seen as a discrete interface and the correspondence is actually a bijection (once a *reference face* and a *reference perfect matching* are fixed). We remark right away that the interface-matching mapping is lost if G is either non-planar or non-bipartite.

The height h_M associated to a perfect matching $M \in \Omega_G$ is an integer-valued function (depending on the matching M) defined on G^*. The definition is as follows. First, fix a reference perfect matching $M_0 \in \Omega_G$ and a reference face $f_0 \in G^*$ (for instance, the outer face if the graph is finite). Then, set $h_M(f_0) = 0$. Finally, for $f, f' \in G^*$ and $C_{f\to f'}$ a nearest-neighbor path on G^* starting at f and ending at f', define

$$h_M(f') - h_M(f) = \sum_{e \in C_{f\to f'}} \sigma_e (1_{e\in M} - 1_{e\in M_0}) \tag{1.1}$$

where the sum runs over edges crossed by the path and σ_e is $+1$ if e is crossed with the white vertex on the right and -1 otherwise.

It is a crucial fact that the definition (3.4) is *independent* of the choice of $C_{f\to f'}$. The proof is very simple and is based on the fact that, on planar graphs, any two paths $C_{f\to f'}$ can be "continuously" deformed one into the another. Along the transformation, the value of the r.h.s. of (3.4) does not change. See Fig. 1. Note that the proof uses both planarity (to deform any path $C_{f\to f'}$ "continuously" into any other one) and the bipartite structure (to define the signs σ_e).

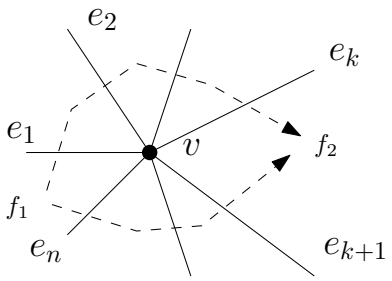

FIGURE 1. Label $e_1, \dots e_n$, say anti-clockwise, the edges incident to vertex v. If the path $C_{f\to f'}$ goes from f_1 to f_2 leaving v to its left, the contribution to $h(f') - h(f)$ is $\sum_{k+1}^{n}(1_{e_i\in M} - 1_{e_i\in M_0})$. If instead it leaves v to its right, the contribution is $\sum_1^n(1_{e_i\in M_0} - 1_{e_i\in M})$. The difference is then $\sum_1^n 1_{e_i\in M} - \sum_1^n 1_{e_i\in M_0}$ which is zero since both M and M_0 have exactly one dimer incident to v. The argument is similar if v is a white vertex. Any two paths from f to f' can be turned one into the other via a finite chain of transformations of this type at black or white vertices.

We will often omit the index M in the height function. Also, we do not indicate exiplicitly the dependence on M_0.

Remark 1.2. *From the proof of path-independence of the right-hand side of* (3.4)*, we see that one can also replace* $1_{e\in M_0}$ *by any function* $e \mapsto c(e) \in \mathbb{R}$

such that, for any vertex v,

$$\sum_{e\sim v} c(e) = 1$$

where the sum runs over all edges incident to v. In other words,

$$h_M(f') - h_M(f) = \sum_{e\in C_{f\to f'}} \sigma_e(1_{e\in M} - c(e)) \tag{1.2}$$

Of course, h is then not integer-valued any more, but this is not important.

Note also that the correspondence between M and height function is a bijection (once M_0 and the overall constant $h_M(f_0)$ are fixed). Indeed, the gradient $h_M(f) - h_M(f')$ for neighboring faces f, f', determines whether the edge separating f, f' is occupied or not.

When G is an infinite periodic lattice, e.g. $\mathbb{Z}^2$ or the honeycomb lattice $\mathcal{H}$ (the infinite lattice where faces are regular hexagons and vertices have degree 3, as in Figure 3), one usually takes as M_0 (or the function $c(\cdot)$) to be periodic. For instance, on $\mathcal{H}$ one can choose $c(e) = 1/3$ for every e and $c(e) = 1/4$ on the square lattice.

Remark 1.3. *[Dimers and tilings] There is a correspondence between perfect matchings of a planar graph and tilings of the plane. Let us see this in the case of infinite periodic graphs like $\mathbb{Z}^2$ and $\mathcal{H}$. The faces of G^* are in one-to-one correspondence with vertices of G and cover the plane. Given a matching M, join any two faces of G^* whose corresponding vertices of G are matched. This defines a tiling of the plane, each tile being the union of two faces of G^*. The correspondence is bijective. See Fig. 2 for the case of $\mathbb{Z}^2$ and Fig. 3 for the case of $\mathcal{H}$.*

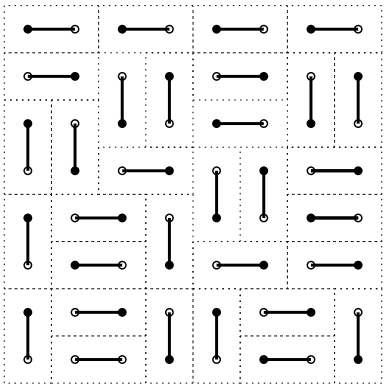

FIGURE 2. A perfect matching of a finite domain of $\mathbb{Z}^2$ and the corresponding tiling. Tiles are dominos (2×1 rectangles) obtained as union of two square faces of $(\mathbb{Z}^2)^*$.

Remark 1.4 (Double dimer configurations and loops)**.** *Given two perfect matchings M and M' of a graph G, their union (i.e. the set of edges of G that are either in M or in M' or in both) is a collection of closed, simple, non-intersecting loops (of even length) and of double edges.*

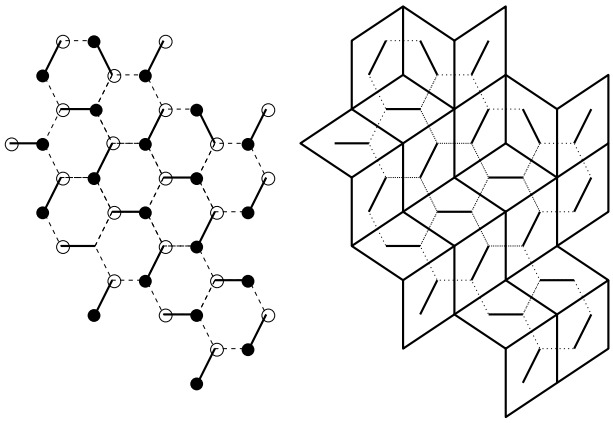

FIGURE 3. A perfect matching of a finite domain of $\mathcal{H}$ and the corresponding rhombus tiling. Tiles are rhombi obtained as union of two triangular faces of $\mathcal{H}^*$.

Definition 1.5 (Rotations along loops)**.** *Let Γ be a simple closed cycle on G, i.e. a sequence of distinct vertices $v_1, \dots, v_{2n}$ with v_i neighbor of v_{i+1} and v_{2n} neighbor of v_1 (note that a cycle is necessarily of even length on a bipartite graph). Let $e_1, \dots, e_{2n}$ be the edges along Γ, labeled consecutively, say anticlockwise. If every second edge is occupied by a dimer in a configuration M, we will say then that a rotation is possible along Γ: the rotation consists in exchanging empty and occupied edges and the new set of occupied edges thus obtained is still a perfect matching M'.*

In the particular case where Γ consists in the $2n$ edges along a face f, we will call the rotation "an elementary rotation at f".

Note that when an elementary rotation is performed at $f \neq f_0$ (with f_0 the reference face) then the height function changes by ± 1 at f and is unchanged at all other faces. If instead $f = f_0$, then the height changes by the same value ± 1 at all faces except f_0, where it is fixed to zero by definition.

2. Kasteleyn theory

In this section we will learn how to count the number of perfect matchings of G or, more generally, how to compute the "generating function" from which all correlations of the probability measures of the type (2.1) can be deduced. From that, we will deduce a formula for correlation functions for the uniform measure or, more generally, for a Boltzmann measure where every dimer has a weight that depends on the edge it occupes.

2.1. Counting perfect matchings: The bipartite case.

To each edge $e \in E$ we assign a weight $t_e > 0$ (we exclude zero weights, as that would boil down to removing the corresponding edges). In Section 2.6, we will be interested in the probability measure $\pi_{G,\underline{t}}$ on Ω_G, with

$$\pi_{G,\underline{t}}(M) \propto w_{\underline{t}}(M) := \prod_{e \in M} t_e. \tag{2.1}$$

For the moment, we want to see how to compute its partition function (or generating function)

$$Z = Z_{G,\underline{t}} = \sum_{M \in \Omega_G} \prod_{e \in M} t_e. \tag{2.2}$$

This result will be essential in Section 2.6 to obtain correlation functions of the (random) dimer model with probability law $\pi_{G,\underline{t}}$. For instance, one can obtain correlation functions by multiple derivatives of the logarithm of Z with respect to the variables t_e.

The special feature of the dimer model on a planar bipartite graph, that makes it "exactly solvable", is that its partition function is simply given by the determinant of a square $|V_B| \times |V_W|$ matrix, the so-called Kasteleyn matrix. In the non-planar case the situation is more complicated, for instance in the toroidal case we will see that we have to combine linearly four determinants (Section 2.4). In the planar but non-bipartite case, there is a similar structure where determinants are replaced by so-called Pfaffians (Section 2.5).

The Kasteleyn matrix K is a complex-valued square matrix, with rows and columns indexed by vertices in the black and white subgraphs V_B/V_W, respectively. The matrix element $K(\mathrm{b}, \mathrm{w})$ is zero if w, b are not neighbors. If instead w, b are neighbors and are the endpoints of edge e,

$$K(\mathrm{b}, \mathrm{w}) = t_e \theta_e \tag{2.3}$$

with θ_e a complex number of modulus 1, satisfying the condition that for any face f of G, calling $e_1, \dots e_{2n}$ its edges in cyclic order (clockwise or anti-clockwise), then

$$\frac{\theta_{e_1} \theta_{e_3} \dots \theta_{e_{2n-1}}}{\theta_{e_2} \theta_{e_4} \dots \theta_{e_{2n}}} = (-1)^{n+1}. \tag{2.4}$$

Note that the choice of e_1 and the choice between clockwise and anti-clockwise order is irrelevant. Recall also that we are assuming that G has no multiple edges, so that there is a unique edge e having b, w as endpoints[2].

We will prove in Section 2.3 that such a choice of $\{\theta_e\}_e$ (Kasteleyn weighting) is always possible and in general not unique (for non-uniqueness, see Section 2.2):

Theorem 2.1. *For every planar bipartite graph G, there exists at least a function $\theta : E \ni e \mapsto \{\theta_e\} \in \mathbb{C}$ with $|\theta_e| = 1$ such that (2.4) holds at each face. Actually, it is possible to choose $\theta_e \in \{-1, +1\}$ for every $e \in E$.*

Let us first look at some examples.

Example 2.2 (Kasteleyn weigthing, hexagonal lattice)**.** *Whenever each face is surrounded by $2n$ edges with n an odd integer, then we can just choose $\theta_e = +1$ for every e. This is the case for instance of the hexagonal lattice, where $n = 3$.*

[2] It is easy to adapt the theory when multiple edges are allowed. In fact, if there are n edges $e_1, \dots, e_n$ between w and b, with weights $t_{e_1}, \dots, t_{e_n}$, the partition function equals that of a graph where there is a single edge $\bar{e}$, with weight $t_{e_1} + \dots + t_{e_n}$.

Example 2.3 (Kasteleyn weigthing, square lattice)**.** *Let* $G \equiv \mathbb{Z}^2$*. One possible Kasteleyn weighting is to choose* $\theta_e = 1$ *on horizontal edges and* $\theta_e = i$ *on vertical edges. Another natural choice is the following (as in Fig. 4): if* e *has black/white endponts* b, w*, take* $\theta_e = \mathrm{w} - \mathrm{b}$ *with* b, w *interpreted as elements of the complex plane* $\mathbb{C}$ *(e.g. if* w *is just below* b*, then* $\theta_e = -i$*). See Fig. 5 for yet another choice.*

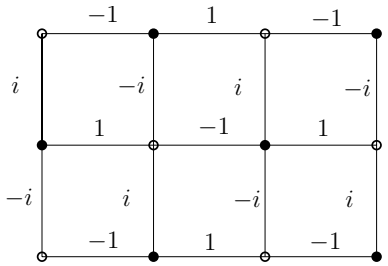

FIGURE 4. A possible Kasteleyn weighting of $\mathbb{Z}^2$.

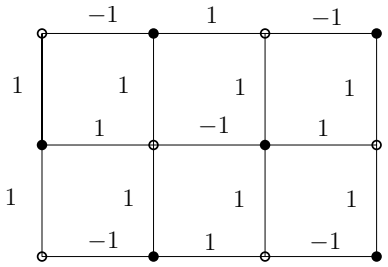

FIGURE 5. Another Kasteleyn weighting of $\mathbb{Z}^2$.

The crucial point is the following:

Theorem 2.4. [14, 16] *Let* G *be a finite, planar, bipartite, weighted graph and* K *a Kasteleyn matrix associated to it. Then,*

$$Z_{G,\underline{t}} = |\det(K)|. \tag{2.5}$$

In the non-bipartite case there is a similar statement, with a Pfaffian instead of the determinant, and actually Kasteleyn solved the non-bipartite case directly. This is discussed in Section 2.5.

One of the advantages of having a determinantal formula for the partition function is that, in interesting situations, the Kasteleyn matrix can be explicitly diagonalized or reduced to a simple block-diagonal form via a natural change of basis, and one obtains explicit formulas, even as the graph grows to infinity. Note that if we can diagonalize explicitly K, the partition function is just the product of the absolute values of its eigenvalues.

Another (computational) advantage of Eq. (3.6) is that there are algorithms that compute the determinant of a $n \times n$ matrix in polynomial time; e.g. an algorithm based on Gauss elimination runs in time $O(n^3)$. This is much faster than computing the partition function summing over all the exponentially many (in the number of vertices) configurations. By the way,

from formula (3.6) one might (wrongly!) guess that the number of configurations is of the order of the factorial of the number of black vertices (i.e. the number of permutations defining the determinant). We will see that the number actually grows only exponentially, because the matrix K is very sparse and most permutations give zero.

The determinantal formula for the partition function also allows to have a determinantal formula for all correlations function, as we will see in Section 2.6.

Remark 2.5. *Whether the partition function is* $+\det(K)$ *or* $-\det(K)$ *depends on the choice of how we label sites: changing the labelling corresponds to permuting some rows or columns, so the determinant can change sign, while clearly the partition function does not.*

Proof of Theorem 2.4. Expand the determinant as sum over the set $S_{|V_B|}$ of permutations π of $\{1, \dots, |V_B|\}$:

$$\det(K) = \sum_{\pi \in S_{|V_B|}} \sigma_\pi \prod_{i=1}^{|V_B|} K(b_i, w_{\pi(i)}), \tag{2.6}$$

with $\sigma_\pi = \pm 1$ denoting the signature of π. Each term corresponds (bijectively) to a matching M with the correct weight (in absolute value) $\prod_{e \in M} t_e$. It remains to check that the complex number (of absolute value 1)

$$\sigma_\pi \prod_i \theta(b_i, w_{\pi(i)}) \tag{2.7}$$

is the same for every π. To see that (2.7) is independent of π, observe that:

Lemma 2.6. *Given* $M, M' \in \Omega_G$ *one can connect* M *to* M' *via a finite chain of elementary rotations (cf. Definition 1.5).*

When an elementary rotation is performed at face f with boundary edges $e_1, \dots, e_{2n}$, the signature of the permutation σ changes by $(-1)^{n+1}$ (which is the signature of the cyclic permutation $\{2, 3, \dots, n, 1\}$). On the other hand, $\theta_{e_1}\theta_{e_3} \dots \theta_{e_{2n-1}}$ changes to $\theta_{e_2}\theta_{e_4} \dots \theta_{e_{2n}}$ and their ratio is also $(-1)^{n+1}$ by (2.4). Altogether, (2.7) is unchanged by an elementary rotation and by Lemma 2.6 it is constant on Ω_G. ∎

Proof of Lemma 2.6. Define the height function as in Section 1.1, taking as reference face the external face f_0 and choosing the function $c(e)$ (as in Remark 1.2) to be the probability of occupation of edge e under the uniform measure, i.e. the number of perfect matchings in Ω_G that contain e divided by the total number of perfect matchings. Recall that, according to Assumption 1.1, $c(e) > 0$ for every edge. Given a face f where $h_M(f) > h_{M'}(f)$, follow a maximal path for M' starting at f, i.e. a nearest-neighbor path on faces which crosses only edges which are *unoccupied* in M' and that have the white vertex on the right. Along any such path, $h_M - h_{M'}$ is increasing and moreover $h_{M'}$ is strictly decreasing (because $c(e)$ is strictly positive). Therefore, the path cannot form a loop, so it has to stop after a finite number of steps, at some face f'. Note that f' cannot be the external face f_0, because h_M and $h_{M'}$ coincide there, while $h_M(f') > h_{M'}(f')$. If the path cannot be continued beyond f', it means that every second edge

surrounding this face (and more precisely, those that have the white vertex on the right when crossed exiting f') is occupied in M'. Therefore, an elementary rotation is possible at f' for M'. Performing the elementary rotation at f', the height of M' increases there. Repeat the procedure until there are no faces left with $h_M > h_{M'}$ or $h_M < h_{M'}$. ∎

Remark 2.7. *We could have proved Theorem 2.4 without using explicity the existence of the height function (as will be done for the analogous result for non-bipartite graphs).*

2.2. **Gauge invariance.** Given a Kasteleyn weighting, one can obtain others via a so-called "gauge transformation". I.e. condition (2.4) is unchanged if we take a function α defined on the set of vertices and taking complex values of modulus 1 and we replace, for each edge e,

$$\theta_e \mapsto \theta_e \alpha_w \alpha_b,$$

if b, w are the endpoints of e. Note that this operation is equivalent to multiplying rows and columns of K by modulus-one numbers, an operation that does not change the absolute value of the determinant. Conversely, one can see that any two Kasteleyn weightings can be obtained one from the other via a gauge transformation, see [20, Sec. 3.2].

2.3. **Existence of a Kasteleyn weighting.** Theorem 2.4 is a consequence of the following more general result, due to Kasteleyn [16].

Theorem 2.8. *Let G be a finite planar graph (not necessarily bipartite). The graph admits a clockwise-odd orientation, i.e. there exists a choice of orientation of the edges such that for each internal face the following holds: running clockwise around the face, the number of edges that are co-oriented is odd. Moreover, for any simple cycle Γ of even length such that $G\backslash\Gamma$ admits a perfect matching, the number of edges of Γ that are oriented clockwise is odd.*

Here, more explicitly, $G \setminus \Gamma$ denotes the graph where the vertices in Γ together with any edge incident to them are removed.

Proof of Theorem 2.8. Given G, let T be a spanning tree of the dual graph G^*, rooted at the outer face. Orient the edges of G not crossing T in an arbitrary way. See Fig. 6. Then, there is a unique way of orienting the remaining edges so that globally the orientation is clockwise odd; to do so, orient the remaining edges one by one proceeding from the leaves of the tree towards the root. At each step there is a unique choice (the choice depends on how the edges not crossing T were oriented initially).

Now let Γ be a simple cycle of even length. It encloses an even number V_{int} of vertices (recall that the graph is planar and that $G\backslash\Gamma$ is assumed to admit a perfect matching, so that inner vertices are matched among themselves) and an (even or odd) number E_{int} of inner edges (we count also the edges that join vertices in V_{int} to vertices in Γ). See Fig. 7. The cycle Γ consists of an even number $E_{\partial\Gamma}$ and even number $V_{\partial\Gamma}$ of vertices and observe that $E_{\partial\Gamma} = V_{\partial\Gamma}$. We let $E = E_{int} + E_{\partial\Gamma}$ and similarly $V = V_{int} + V_{\partial\Gamma}$ be the total number of edges and vertices in the graph determined by Γ and its interior. Call $\mathcal{F}$ the collection of faces inside Γ and F its cardinality. Also,

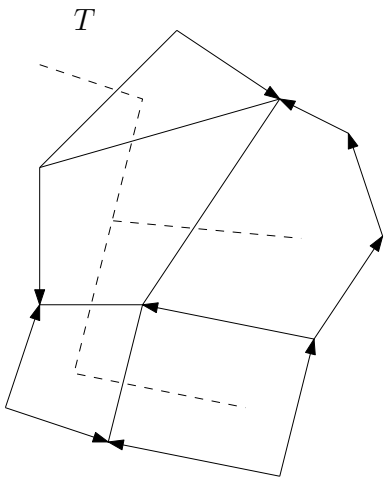

FIGURE 6. A graph G and a spanning tree T of G^*, together with an arbitrary orientation of the edges not crossing T. The remaining edges can be uniquely oriented to respect the definition of clockwise odd orientation

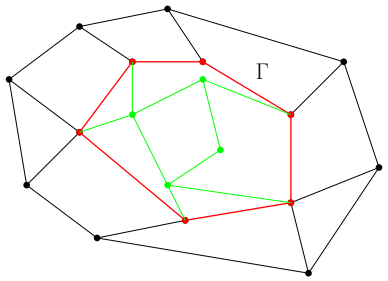

FIGURE 7. A graph G with a cycle Γ of even length (marked in red) with $E_{\partial\Gamma} = V_{\partial\Gamma} = 6$. Check that $G \setminus \Gamma$ does admit a perfect matching! Inner vertices and edges are drawn in green.

for every $F \in \mathcal{F}$ let $n(f)$ be the number of clockwise oriented edges along the boundary of f. By construction, $n(f)$ is odd for every f. We have

$$0 \bmod 2 = \sum_{f\in\mathcal{F}} (n(f)+1) = F + \sum_{f\in\mathcal{F}} n(f) = F + n(\Gamma) + E_{int} \tag{2.8}$$

where $n(\Gamma)$ is the number of clockwise oriented edges of Γ. In the third equality, we used the fact that all inner edges are counted twice, once with the clockwise orientation and once with the anti-clockwise orientation, so each of them contributes $+1$ to the sum. From (2.8) we see (because V_{int} is even)

$$\begin{aligned} 0 \bmod 2 &= n(\Gamma) + F + (E_{\partial\Gamma} + E_{int}) + (V_{\partial\Gamma} + V_{int}) && (2.9)\\ &= n(\Gamma) + F + E + V = n(\Gamma) + F - E + V. && (2.10) \end{aligned}$$

Euler's formula gives $F - E + V = 1$, so that $n(\Gamma)$ is odd as wished. ∎

Remark 2.9. *Note that we did not really use that $G \setminus \Gamma$ admits a perfect matching, but just that the number V_{int} of inner vertices is even.*

Proof of Theorem 2.4. Orient the edges of G as in Theorem 2.8. Given an edge e, set $\theta(e)$ to be $+1$ if it is oriented from black to white, and -1 in the opposite case. We want to verify that (2.4) is satisfied. Recall that G is bipartite. If a face f is surrounded by $2n$ edges, then the l.h.s. of (2.4) is the product of $\theta(e_i)/\theta(e'_i)$ for i ranging from 1 to n, where e_i, e'_i are the two boundary edges incident to the i^{th} black vertex. Note that $\theta(e_i)/\theta(e'_i)$ is -1 if both edges have the same orientation, and $+1$ if they have opposite orientation. Therefore, the product gives

$$(-1)^{\#\{i: e_i \text{ and } e'_i \text{ are equally oriented}\}} \tag{2.11}$$

$$= (-1)^{n+\#\{i: e_i \text{ and } e'_i \text{ are differently oriented}\}} \tag{2.12}$$

$$= (-1)^{n+1} \tag{2.13}$$

because the chosen orientation is clockwise-odd. ∎

Remark 2.10. *Note that* (2.4) *holds also for the external face f_0. Indeed, the cycle Γ determined by the boundary edges has even length and the number of inner vertices V_{int} enclosed by it is also even. By Theorem 2.8 and Remark 2.9, the number of boundary edges oriented clockwise (or anti-clockwise) is odd and one proceeds like in* (2.11).

2.4. **Counting perfect matchings in the bipartite, non-planar case: periodic boundary conditions.** In the proof of Theorem 2.4 we have used planarity of the graph when we claimed that given any two perfect matchings M and M', they can be turned one into the other by a chain of elementary rotations. For non-planar graphs, this is in general not true.

The matching M' can be obtained from M by a rotation around evey loop ("rotation" here means, as in Definition 1.5, that if in M the edges $e_1, e_3, \dots, e_{2n-1}$ are occupied in the loop $e_1, \dots, e_{2n}$, after the rotation the edges $e_2, e_4, \dots, e_{2n}$ are occupied instead). If the graph is planar, the rotation around a loop can be decomposed, as we have seen, as a chain of elementary rotations. On non-planar graphs this is not always possible.

In general, if the graph G is embedded on a surface of genus g, the partition function of the dimer model can be expressed in terms of 4^g determinants [30]. In practice, this is not very useful if g grows with the size of the graph. The interesting case we consider here is that of a toroidal two-dimensional graph, that can be seen as embedded on the two-dimensional torus, so that $g = 1$ and the partition function can be given as the combination of just 4 determinants [14]. The toroidal case is very useful because in this case, due to translation invariance, it is easy to reduce the Kasteleyn matrix to a simple block-diagonal form (with block sizes that are independent of the size of the graph) and to compute the large-volume limit of the partition function and of correlations. This will be done in Section 3.

More precisely, we start with an infinite, periodic, planar bipartite planar graph G. This means that G can be drawn on the plane in such a way that $\mathbb{Z}^2$ acts as a group of isomorphisms that preserves the color of vertices (and the weights of edges). The graph is then a periodic repetition of a

fundamental domain G_1 (containing n white and n black vertices) in two lattice directions $\vec{e}_1, \vec{e}_2$. The graph G_1 is just G modulo translations.

See Fig. 8 for the square and hexagonal lattice examples; see Fig. 13 for a graph where the fundamental domain is more complicated (it comprises 4 vertices of both colors). We let G_L be the corresponding toroidal graph (of

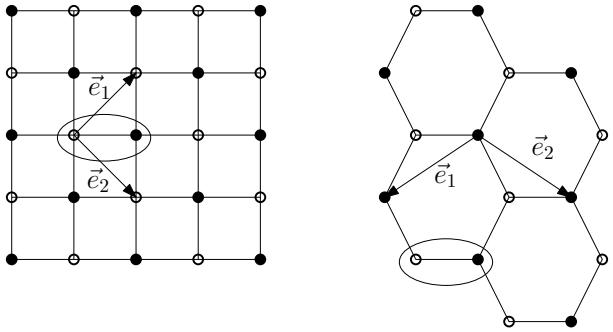

FIGURE 8. The graphs $\mathbb{Z}^2$ and $\mathcal{H}$ with the fundamental domain (encircled) and the vectors $\vec{e}_1, \vec{e}_2$.

period L) with periodic boundary conditions in both $\vec{e}_1$ and $\vec{e}_2$ directions. See Fig. 9 for the hexagonal case with $L = 4$ and Fig. 10 for the square case with $L = 3$.

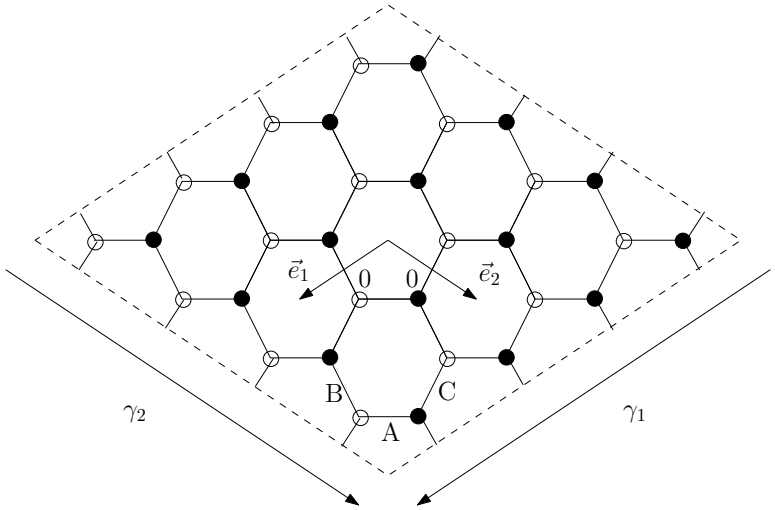

FIGURE 9. Hexagonal graph. The graph G_L with $L = 4$, together with the oriented loops γ_1, γ_2 winding in the two directions. We marked the black/white vertices with coordinates $0 = (0, 0)$.

Remark 2.11 (Height function on the torus)**.** *Clearly, G_L is not planar any more. Note that the height function is not well-defined any more, or rather, it is multi-valued: if we take definition* (3.4) *and choose $f = f'$, the height increase along a path $C_{f\to f}$ will be given by $n_1\Delta_1 + n_2\Delta_2$, with $\Delta_1 = \Delta_1(M), \Delta_2 = \Delta_2(M)$ two integers depending on the dimer configuration M*

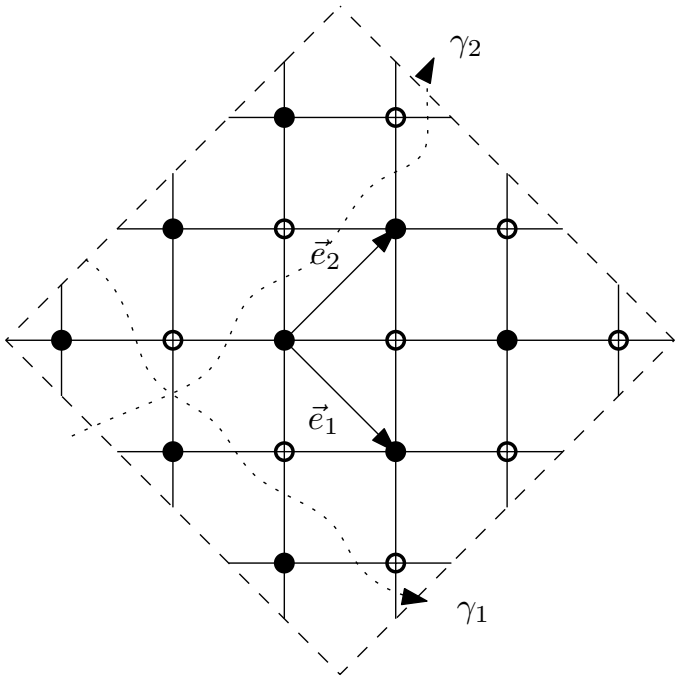

FIGURE 10. Square graph. The graph G_L with $L = 3$, together with the oriented loops γ_1, γ_2 winding in the two directions.

and n_i the winding number of $C_{f\to f}$ in the $\vec{e}_i$ direction. The reason is that we cannot always deform $C_{f\to f}$ "continuously" into the trivial path $\{f\}$. If we take a uniform perfect matching of G_L, then Δ_1, Δ_2 will be random integers.

The law of (Δ_1, Δ_2) has an interesting behavior as $L \to \infty$. In [2], for the dimer model on the honeycomb lattice with uniform weights, it has been proven that it converges to the law of two independent, discrete Gaussian random variables, i.e. the probability that $(\Delta_1, \Delta_2) = (n, m) \in \mathbb{Z}^2$ converges as $L \to \infty$ to

$$\frac{1}{Z} e^{-c(n^2+m^2)},$$

for an explicit positive constant c. Note that the values Δ_i are not rescaled with L: their typical values are $O(1)$ in the $L \to \infty$ limit. See also [8].

Exercise 2.12. *Check in the hexagonal case that Δ_1, Δ_2 do indeed depend on the configuration, putting in Fig. 9 either all dimers horizontal or all dimers oriented north-east.*

Remark 2.13. *Let C be a closed self-avoiding loop (concatenation of edges) on G_L, whose edges are alternately occupied/empty in some M. Assume that C has zero winding number. A rotation around C does not change Δ_1, Δ_2.*

Let K be the Kasteleyn matrix of the graph G_L, obtained simply by periodizing that of the infinite lattice G[3]. We assume that the modulus-one

[3] To be precise, it may happen that G_L is not a simple graph, i.e. it has multiple edges. For instance, this is the case for the honeycomb graph and for the square grid

complex numbers θ_e in (2.3), defining the Kasteleyn matrix, take values in $\{-1, 1\}$ (recall from Theorem that this is always possible). The reason why Theorem 2.4 has to be modified is that when we expand $\det K$, not all terms come with the same sign. In fact, as we see in a moment, the sign depends on

$$(\Delta_1, \Delta_2) \bmod 2.$$

To give the generalization of Theorem 2.4 in the case of the torus, let γ_1, γ_2 be two self-avoiding, oriented, simple loops on G_L^*, which have winding numbers $(1, 0)$ and $(0, 1)$ respectively (i.e. γ_1 winds horizontally and γ_2 winds vertically; see Fig. 9 and 10 in the honeycomb and square graph cases). For technical convenience, we will require also that there exists a dimer configuration M_0 such that γ_1, γ_2 cross none of the edges in M_0. For instance, in the case of the honeycomb graph with γ_i chosen as in Fig. 9, M_0 is the collection of all horizontal edges; for the square grid and γ_i chosen as in Fig. 10, M_0 is the collection of vertical edges with black vertex on top. We introduce four Kasteleyn matrices $K_{\theta\tau}$, with $\theta, \tau \in \{0, 1\}$. To obtain $K_{\theta\tau}$ from K, just multiply the matrix element $K(b, w)$ by $(-1)^\tau$ for all edges (b, w) crossed by γ_1 and by $(-1)^\theta$ if for all edges (b, w) crossed by γ_2 (it may happen that an edge crosses both, then we multiply by both factors). Note that $K_{00} = K$ and it is referred to as the Kasteleyn matrix with periodic-periodic boundary conditions. Similarly, K_{10} is said to have antiperiodic-periodic boundary conditions, and so on. Then we have

Theorem 2.14. *The partition function on the torus is*

$$Z_{G_L,\underline{t}} = \frac{1}{2} \sum_{\theta,\tau\in\{0,1\}} c_{\theta\tau} \det K_{\theta\tau} \tag{2.14}$$

where $c_{\theta\tau} = \pm 1$; three of the signs $c_{\theta\tau}$ are equal and the fourth is opposite.

The actual values of $c_{\theta\tau}$ depend on the graph G, on the choice of Kasteleyn matrix, of the labeling of vertices and possibly on the parity of L.

Proof of Theorem 2.14. For the proof in the case of the lattice $\mathbb{Z}^2$, see [14]. The general case is treated in [30, 5] as a particular case of graphs embedded on surfaces of genus g. Here we give the proof for the hexagonal lattice. In this case (and given certain specific choices of Kasteleyn matrix and of labeling of vertices), we will see that

$$\{c_{00}, c_{10}, c_{01}, c_{11}\} = \{-1, 1, 1, 1\} \tag{2.15}$$

for L even and

$$\{c_{00}, c_{10}, c_{01}, c_{11}\} = \{1, 1, 1, -1\} \tag{2.16}$$

for L odd.

When we expand the determinant of K_{00}, non-zero terms of the expansion are in bijection with perfect matchings of G_L, and their absolute value equals the weight of the corresponding dimer configuration. Recall that we chose a Kasteleyn matrix with non-zero elements taking values ± 1. For the

if $L = 1$. However, this does not happen if L is large enough and since anyway we are interested mostly in the limit $L \to \infty$, we assume that L is large enough so that G_L is simple.

hexagonal lattice the complex numbers θ_e in the Kasteleyn matrix can be taken to be all equal $+1$, and we assume that this choice has been made. Therefore, each term in the expansion of the determinant is real, but the signs are not all equal, because of the signature of the permutation. In fact, let M_0 be the reference configuration where all dimers are horizontal, so that none of them crosses the paths γ_1, γ_2.

Assume to fix ideas that the term in the expansion of the determinant corresponding to M_0 is positive (as noted before, this can be guaranteed by a suitable labellig of vertices). As we noted above, given any other configuration M, the union of M and M_0 consists in a collection of double edges and of closed, simple, non-intersecting loops of even length. Since G_L is bipartite there is a natural way to orient the loops: loops are oriented in such a way that the edges occupied by dimers of M are oriented from black to white. Call $n(M) \in \mathbb{Z}^2$ the total winding number of the loops. I.e. $n_1(M)$ is the sum over the loops of the horizontal winding number of the loop, and $n_2(M)$ is the sum of the vertical winding numbers. Say that a configuration M is of the class $(0,0)$ if n_1, n_2 are both even, of the class classes $(0,1)$ if n_1 is even and n_2 is odd, and similarly for classes $(1,0)$ and $(1,1)$.

Exercise 2.15. *Define the multi-valued height function on the torus with M_0 as reference configuration. Prove that the "height changes" $(\Delta_1(M), \Delta_2(M))$ introduced in Remark 2.11 equal $(-n_2(M), -n_1(M))$, where $n_i(M)$ is the sum over loops Γ of $M \cup M_0$ of the (positive or negative) winding number of Γ in direction i.*

To turn M into M_0, we have to perform a rotation around each one of the loops. We claim:

Claim 2.16. *In the expansion of $\det(K_{00})$, configurations M in the class (ϵ_1, ϵ_2) come with a sign $\sigma_{\epsilon_1\epsilon_2}$ with respect to M_0, where three of the $\sigma_{\epsilon_1\epsilon_2}$ are equal and the fourth is opposite. More explicitly, for the hexagonal graph, one finds*

$$\{\sigma_{00}, \sigma_{10}, \sigma_{01}, \sigma_{11}\} = \{1, -1, -1, -1\} \tag{2.17}$$

if L is even and

$$\{\sigma_{00}, \sigma_{10}, \sigma_{01}, \sigma_{11}\} = \{1, 1, 1, -1\} \tag{2.18}$$

if L is odd.

Let us accept this fact for a moment and let us conclude the proof of Theorem 2.14. When we look at the expansion of $\det(K_{\theta\tau})$, we have to keep into account the fact that the signs of the matrix elements of $K_{\theta\tau}$ are not all equal to those of K_{00}. Note however that the configuration M_0 appears with the same sign (plus) for every θ, τ, because (by assumption) none of its dimers crosses the paths γ_1, γ_2 where the Kasteleyn matrix has a sign depending on θ, τ. Also, the relative sign of the term M will be as in the expansion of $\det(K_{00})$, times $(-1)^{\theta n_1(M)}(-1)^{\tau n_2(M)}$. In other words, in the expansion of $\det(K_{\theta\tau})$, terms in the class (ϵ_1, ϵ_2) come with sign $\sigma_{\epsilon_1\epsilon_2}(-1)^{\theta\epsilon_1+\tau\epsilon_2}$ with respect to M_0. If we choose $c_{\theta\tau}$ in (2.14) so that

$$\frac{1}{2}\sum_{\theta,\tau\in\{0,1\}} c_{\theta\tau}\sigma_{\epsilon_1\epsilon_2}(-1)^{\theta\epsilon_1+\tau\epsilon_2} = +1 \tag{2.19}$$

for every choice of ϵ_1, ϵ_2 then we are done, since each term in the r.h.s. of (2.14) appears with the same sign as M_0. Note that Eq. (2.19) can be written as

$$A \cdot c = \sigma, \tag{2.20}$$

where $c = (c_{00}, c_{10}, c_{01}, c_{11})$, $\sigma = (\sigma_{00}, \sigma_{10}, \sigma_{01}, \sigma_{11})$ while A is the 4×4 matrix

$$A = \frac{1}{2}\begin{bmatrix} +1 & +1 & +1 & +1 \\ +1 & -1 & +1 & -1 \\ +1 & +1 & -1 & -1 \\ +1 & -1 & -1 & +1, \end{bmatrix}, \quad A^{-1} = A. \tag{2.21}$$

Then, $c = A \cdot \sigma$, i.e.

$$c_{\theta\tau} = \frac{1}{2} \sum_{\epsilon_1,\epsilon_2=0,1} (-1)^{\theta\epsilon_1 + \tau\epsilon_2} \sigma_{\epsilon_1\epsilon_2}.$$

It is immediately checked from the form of A that $c_{\theta\tau} \in \{-1, +1\}$ and that one of the values is opposite to the three others. More explicitly, given the values of $\sigma_{\epsilon_1\epsilon_2}$ provided by Claim 2.16, it is immediate to check that (2.15)-(2.16) hold.

■

Proof of Claim 2.16. We already know, from the proof of Theorem 2.4, that if the winding numbers $(m_1(\Gamma), m_2(\Gamma))$ of a loop Γ are zero, then two configurations M, M' related by a rotation along Γ appear with the same sign in the expansion of $\det(K_{00})$; therefore, from now on we consider only loops with non-trivial winding. If there are no such loops then M is in the class $(0, 0)$ and it comes with the same sign as M_0. Since the loops do not intersect, we have $(m_1(\Gamma), m_2(\Gamma)) = (m_1(\Gamma'), m_2(\Gamma'))$ for every pair of non-trivial loops Γ, Γ'.

In conclusion, we can assume that there are $R > 0$ non-trivial loops, all with the same winding number $(m_1, m_2) \neq (0, 0)$. Finally, since loops have no self-intersections, then [9, Prop. 1.5] $(m_1, m_2) \in \mathcal{W}$, where

$$\mathcal{W} := \{(0, \pm 1)\} \cup \{(\pm 1, 0)\} \cup \{(p, q) \in (\mathbb{Z} \setminus \{0\})^2 : \gcd(p, q) = 1\}. \tag{2.22}$$

In particular, at least one among m_1, m_2 is odd. Then, we see that R must be even if M is in the class $(0, 0)$ and odd in all other cases:

$$R \bmod 2 = 1 - (1 - \epsilon_1)(1 - \epsilon_2). \tag{2.23}$$

We have (see Fig. 11):

Claim 2.17. *A rotation around a loop of winding (m_1, m_2) corresponds to a cyclic permutation of length $L(m_1 + m_2)$, whose signature is*

$$(-1)^{L(m_1+m_2)+1}.$$

The effect of performing the rotation at all R loops is then

$$\begin{aligned} \sigma_{\epsilon_1\epsilon_2} = (-1)^{[(m_1+m_2)L+1]R} &= (-1)^{(\epsilon_1+\epsilon_2)L+R} \\ &= (-1)^{(\epsilon_1+\epsilon_2)L+1-(1-\epsilon_1)(1-\epsilon_2)} \end{aligned} \tag{2.24}$$

where we used that $(\epsilon_1, \epsilon_2) := (m_1 R, m_2 R) \bmod 2$ and (2.23). It is immediately checked that this expression gives (2.17)-(2.18). ∎

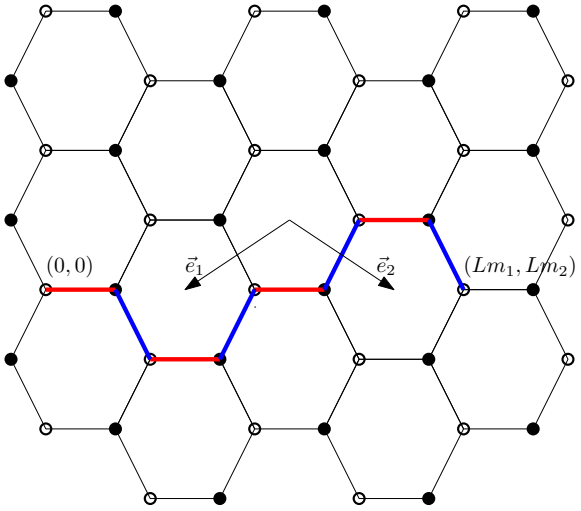

FIGURE 11. A loop of $M \cup M_0$ on the toroidal graph $G_L, L = 2$, drawn in the infinite graph G, joins white vertex $(0, 0)$ with white vertex (Lm_1, Lm_2). Blue edges belong to M, red edges to M_0. Here, the winding is $(m_1, m_2) = (-1, 1)$. The length of the loop (i.e. the number of its edges) is fixed to be $|m_1|L + |m_2|L$, because all dimers in M_0 are horizontal so that each red-blue step corresponds to a translation by $+\vec{e}_2$ or $-\vec{e}_1$.

Exercise 2.18. *Let G_L be the toroidal square graph as in Fig. 10 and choose the Kasteleyn weighting as in Fig. 5. Compute $c_{\theta\tau}$ and $\sigma_{\epsilon_1\epsilon_2}$.*

2.5. **Counting perfect matchings: The planar, non-bipartite case.** For completeness, in this section we give Kasteleyn's theorem for the partition function of the dimer model on a planar, non-bipartite graph. However, in the rest of the notes we will restrict to the bipartite case.

Let G be a planar graph, not necessarily bipartite, and we let $t_e > 0$ be the weight associated to edge e. We want to compute

$$Z_{G,\underline{t}} := \sum_{M \in \Omega_G} \prod_{e \in M} t_e. \tag{2.25}$$

Kasteleyn's theory provides a formula similar to (3.6). However, we have to find a proper generalization of the Kasteleyn matrix. Recall that in the bipartite case, K was a square matrix with lines indexed by black sites and column by white sites. For non-planar graphs, the Kasteleyn matrix K is rather a square matrix of size $|V|$, the cardinality of the vertex set of G, and not $|V|/2$ as in the bipartite case. More precisely, we first orient edges according to the clockwise-odd orientation provided by Theorem 2.8. Then,

given two vertices v, v' of G, we let

$$K(v,v') = \begin{cases} 0 & \text{if } v \not\sim v' \\ t_e & \text{if } v \sim v' \text{ and the edge } v,v' \text{ is oriented from } v \text{ to } v' \\ -t_e & \text{if } v \sim v' \text{ and the edge } v,v' \text{ is oriented from } v' \text{ to } v \end{cases} \quad (2.26)$$

Note that K is an antisymmetric matrix.

Given an anti-symmetric square matrix M of size $2n$, its Pfaffian $\mathrm{Pf}(M)$ is defined as [16]

$$\mathrm{Pf}(M) = \frac{1}{2^n n!} \sum_{\pi \in S_{2n}} \sigma_\pi M_{\pi(1)\pi(2)} \dots M_{\pi(2n-1)\pi(2n)}. \quad (2.27)$$

Note that the definition resembles a lot that of the determinant. In fact, the following relation holds [16]:

$$\det M = (\mathrm{Pf}(M))^2. \quad (2.28)$$

Note also that if we change the order of $(\pi(2i-1), \pi(2i))$ into $(\pi(2i), \pi(2i-1))$. the product of matrix elements of M changes sign (because M is antisymmetric) but the signature of the permutation also changes sign. Also, if we exchange the positions of $(\pi(2i-1), \pi(2i))$ and $(\pi(2j-1), \pi(2j))$ the product of matrix elements is unchanged, and so is also the signature. Altogether, one can rewrite the Pfaffian as

$$\mathrm{Pf(M)} = \sum_{\pi \in M[2n]} \sigma_\pi M_{\pi(1)\pi(2)} \dots M_{\pi(2n-1)\pi(2n)} \quad (2.29)$$

where now the sum is over un-ordered matchings of $1, \dots, 2n$. Kasteleyn's theorem for the partition function of the dimer model is then the following [14, 16]:

Theorem 2.19. *Let G be planar. Then,*

$$Z_{G,\underline{t}} = |\mathrm{Pf}(K)|. \quad (2.30)$$

As the reader can imagine, when G is non-bipartite and is embedded on the torus instead of the plane, Eq. (2.30) is replaced by a formula similar to (2.14), with a sum over 4 Pfaffians. See [14, 16].

Note that, if we are able to diagonalize K, then we can obtain $Z_{G,\underline{t}}$ via (2.28) as the square root of the product of its eigenvalues. As an application, Kasteleyn [14] computed (via explicit diagonalization of the matrix K) the asymptotics of the number $Z_{\Lambda_L,\underline{1}}$ of perfect matchings of a $2L \times 2L$ square grid Λ_L of $\mathbb{Z}^2$: the result is that

$$\lim_{L\to\infty} \frac{1}{L^2} \log Z_{\Lambda_L,\underline{1}} = \frac{4\kappa}{\pi}, \quad (2.31)$$

where $\kappa \approx 0.9159...$ is the so-called Catalan's constant

$$\kappa = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)^2}.$$

The subscript $\underline{1}$ in the partition function means that we are giving weight 1 to all edges, so that we are just counting the number of configurations. One can also give different weights t_h, t_v to horizontal and vertical edges, and the "free energy" in the r.h.s. of (2.31) as a function of t_h, t_v. We will work out

similar computations later, working on the torus rather than on the planar graph Λ_L.

Proof of Theorem 2.19. The proof is very similar to that of Theorem 2.4. Using the expression (2.29) for the Pfaffian, we see that each term corresponds bijectively to one perfect matching of G, and that the absolute value of the term is the weight of the configuration. To see that all terms come with the same sign, let M, M' be two configurations and remember that the loops formed by $M \cup M'$ have even length. Assume that $M \cup M'$ contains a single loop Γ, the argument in the general case being a simple generalization. We want to show that when we perform a rotation around any such loop, the relative sign between M and M' is unchanged and then one concludes immediately.

Call $v_1, \dots, v_{2n}$ the vertices along the loop Γ, taken in clockwise order, and e_i the edge from v_i to v_{i+1}, with e_{2n} the edge from v_{2n} to v_1 (by convention we will let $v_{2n+1} := v_1$). Say that in configuration M, v_{2i} is matched with v_{2i-1} so that in M' it is matched with v_{2i+1}. Assume without loss of generality that the vertices of G have been labeled so that $v_1, \dots, v_{2n}$ are the first $2n$ of them. Then, we can choose the permutation representing the term M in the expansion of the Pfaffian to be a permutation of the type $\{1, 2, \dots, 2n, \pi(2n+1), \dots, \pi(|V_G|)\}$ (with $|V_G|$ the number of vertices of G) and the permutation representing the term M' to be the permutation $\{2, 3, \dots, 2n, 1, \pi(2n+1), \dots, \pi(|V_G|)\}$. The two permutations have opposite signature, since they differ by a cyclic permutation of even length $2n$ (recall that the signature of a cyclic permutation of length k is $(-1)^{k+1}$). Let us see how the sign of the product of matrix elements of K changes. For M, the sign of the product of Kasteleyn matrix elements along Γ equals

$$(-1)^{\#\{1 \le i \le n: \text{ the edge } e_{2i-1} \text{ is oriented from } v_{2i} \text{ to } v_{2i-1}\}},$$

as is clear from definition (2.26). Similarly, for M' the sign is

$$(-1)^{\#\{1 \le i \le n: \text{ the edge } e_{2i} \text{ is oriented from } v_{2i+1} \text{ to } v_{2i}\}}.$$

The product of the two signs is (-1) to the power of the number of edges $e_1, \dots, e_{2n}$ of Γ that are oriented anti-clockwise. Given that the loop is of even length, this equals also

$$(-1)^{\#\{1 \le i \le 2n: e_i \text{ is oriented clockwise}\}}.$$

Recall now that we have chosen a clockwise-odd orientation, so that this sign is just -1. Combining with the -1 coming from the ratio of the signatures of the two permutations, we see that M and M' come with the same sign. ∎

2.6. **Random perfect matchings: correlation functions.** From now on, we restrict ourselves to the bipartite dimer model.

Denote by 1_e the indicator function that there is a dimer at e, i.e. that $e \in M$.

We are interested in computing correlation functions of the type

$$\pi_{G,\underline{t}}\left[\prod_{i=1}^{n} 1_{e_i}\right] := \frac{\sum_M w_{\underline{t}}(M) \prod_{i=1}^{n} 1_{e_i}}{\sum_M w_{\underline{t}}(M)}. \tag{2.32}$$

Let $e_i = (w_i, b_i)$. Then,

Theorem 2.20. [17] *Let the graph G be planar and bipartite.*

$$\pi_{G,\underline{t}}\left[\prod_{i=1}^{n} 1_{e_i}\right] = \left[\prod_{i=1}^{n} K(b_i, w_i)\right] \det\left(\left\{K^{-1}(w_i, b_j)_{i,j\le n}\right\}\right). \tag{2.33}$$

The formula should be read as follows. One has to compute the inverse K^{-1} of the $|V_W| \times |V_B|$ matrix K and then compute the determinant of the $n \times n$ sub-matrix obtained from K^{-1} by keeping only lines corresponding to $w_1, \dots, w_n$ and columns corresponding to $b_1, \dots, b_n$.

Proof. Note that the l.h.s. of (2.32) can be obtained as $1/Z_{G,\underline{t}}$ times the derivative w.r.t. $\log t_{e_1}, \dots, \log t_{e_n}$ of $Z_{G,\underline{t}}$. Instead of taking derivatives of the partition function, i.e. of the determinant of the Kasteleyn matrix, we will proceed differently.

We know that $Z_{G,\underline{t}} = \alpha \det K$ for some $|\alpha| = 1$. In the expansion of $\det K$, the terms we want are those proportional to $\prod_{i=1}^{n} t_{e_i}$. Expand the determinant w.r.t. row b_1:

$$\det K = (-1)^{b_1+1}\left(K(b_1, 1) \det K[\{b_1\}; \{1\}] - K(b_1, 2) \det K[\{b_1\}; \{2\}] + \dots\right) \tag{2.34}$$

where, given sets of integers A, B, we denote by $K[A; B]$ the matrix obtained by removing the rows with index in A and the columbs with index in B. With some abuse of notation, we identified b_1 with its label in $\{1, \dots, |V_B|\}$. The only term that contains $K(b_1, w_1)$ is

$$(-1)^{b_1+w_1} K(b_1, w_1) \det K[\{b_1\}; \{w_1\}]. \tag{2.35}$$

Now expand $\det K[\{b_1\}; \{w_1\}]$ w.r.t. row b_2 and so on. Altogether, the term we are interested in (the one proportional to $\prod_1^n t_{e_i}$) is

$$\prod_{1}^{n} K(b_i, w_i) \times (-1)^{\sum_1^n (b_i+w_i)} \det K[\{b_1, \dots, b_n\}; \{w_1, \dots, w_n\}]. \tag{2.36}$$

The matrix appearing in the last expression is of dimension $(|V_W| - n) \times (|V_B| - n)$.

Now we use the following fact from linear algebra, whose proof is based on the Shur complement formula:

$$\det M[A; B] = (-1)^{\sum_i (a_i+b_i)} \det M^{-1}[B^c; A^c] \times \det M \tag{2.37}$$

(this generalizes the well-known $A_{ij}^{-1} \det A = (-1)^{i+j} \det A[\{j\}; \{i\}]$). Applying (2.37) and finally dividing by the partition function, we get the claim. ∎

When the graph is embedded on the two-dimensional torus, with periodic boundary condition and period L as in Section 2.4, a similar formula holds (with analogous proof):

Theorem 2.21. *Assume that* $\det K_{\theta\tau} \neq 0$ *for every* $\theta, \tau \in \{0, 1\}$. *Then,*

$$\pi_{G_L,\underline{t}}\left[\prod_{j=1}^{n} 1_{e_j}\right] = \frac{\frac{1}{2}\sum_{\theta,\tau=\pm} c_{\theta\tau}\left[\det K_{\theta\tau}\right] \prod_{j=1}^{n} K_{\theta\tau}(e_j) \det\{K_{\theta\tau}^{-1}(\mathrm{w}_j, \mathrm{b}_k)_{j,k\le n}\}}{Z_{G_L,\underline{t}}} \tag{2.38}$$

where $c_{\theta\tau} = \pm 1$ *is the same sign that appears in* (2.14).

In fact, it may happen that one of the determinants $\det K_{\theta_0\tau_0}$ vanishes for some L or choice of weights $\underline{t}$ (this will play a role in Section 3.2). In this case, the corresponding inverse matrix $K^{-1}_{\theta_0\tau_0}$ is not well-defined. We will see how to deal with such situation.

3. Infinite-volume limit: free energy and Gibbs measures

3.1. **Free energy.** Suppose that our graph G is periodic in space. I.e. it is an infinite, planar, bipartite periodic graph as in Section 2.4. We assume that edge weights are unchanged under translations by $n\vec{e}_1 + m\vec{e}_2$, with $n, m \in \mathbb{Z}$ and we periodize the graph, with period L, in both directions. See Fig. 12 for an example on the square grid, where the fundamental domain G_1 contains four vertices, and not just two as was the case in Fig. 8. The model then depends on a finite, L-independent number k of weights $\underline{t} \equiv \{t_1, \dots, t_k\}$, e.g. $k = 2$ in the example of Fig. 12.

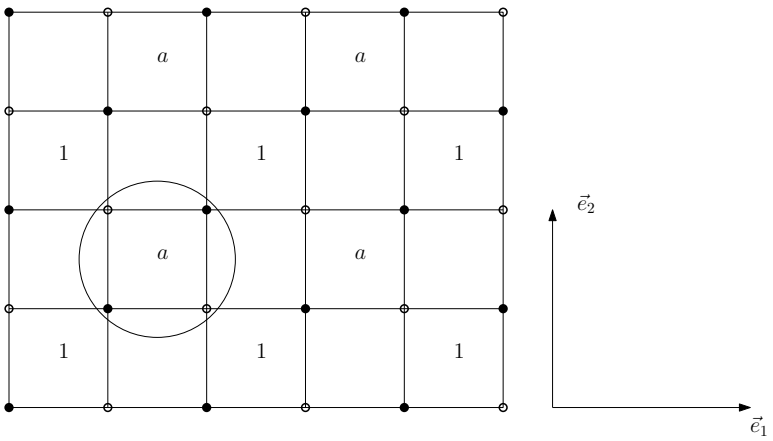

Figure 12. The graph $\mathbb{Z}^2$ with a fundamental domain containing 4 sites. We give a label "a" or "1" to a sub-lattice of faces. The weight of an edge equals, by definition, the label of the unique labelled face it belongs to. Weights are then invariant under translations by $n\vec{e}_1 + m\vec{e}_2$.

With the usual nomenclature of statistical mechanics (up maybe to a global sign) the free energy of the dimer model is defined as

$$F(\underline{t}) = \lim_{L\to\infty} \frac{1}{L^2} \log Z_{G_L,\underline{t}}. \tag{3.1}$$

We will show that not only the limit exists, but that we can actually compute it. For the moment, let us make a couple of comments. First of all, from the free energy we can exctract interesting information. For instance, when edge weights are identically 1, it gives simply the exponential asymptotics of the number of configurations. Also, it is clear that

$$\partial_{\log t_i} \frac{1}{L^2} \log Z_{G_L,\underline{t}}, \tag{3.2}$$

$i \leq k$, gives the average number of dimers on edges "of type i" (i.e. edges of weight t_i). Since the logarithm of the partition function is convex in the variables $\log t_i$, its limit is also convex. Standard results on convex functions guarantee that the l.h.s. of (3.2) converges to $\partial_{\log t_i} F(\underline{t})$, under the assumption that the free energy is differentiable at $\underline{t}$. On the other hand, convex functions are almost everywhere differentiable. Actually, in the examples we work out in detail below, we will be able to see explicitly whether F is differentiable or not.

Remark 3.1. *Usual statistical mecahnics models on periodic lattices, like the Ising model on $\mathbb{Z}^d$, have the property that the infinite-volume free energy is independent of the way the domain Λ_L tends to the whole lattice $\mathbb{Z}^d$, provided some reasonable assumptions are satisfied, notably that the cardinality of the boundary of Λ_L is negligible with respect to the cardinality of the full Λ_L as $L \to \infty$. This property* completely fails *for the dimer model. I.e. it is possible (and quite easy) to find two sequences of domains $\Lambda_L^{(j)}, j = 1, 2$ of the infinite graph G, with $|\partial\Lambda_L^{(j)}|/|\Lambda_L^{(j)}| \stackrel{L\to\infty}{\to} 0$, such that the limit free energy is different for different j. This may look like a pathology of the model, but in Section 4 we will see that it corresponds to a very intuitive phenomenon and height function interpretation of a dimer configuration provides a simple explanation of it. Technically, the reason is that dimers interact via a hard-core potential (i.e. conditionally on the event that a dimer occupies edge e, then the probabity of occupation of any edge incident to e is zero), while Ising spins do not (conditionally on the event a spin σ_i is $+$, neighboring spins have a non-zero, L-independent probability of taking any value, even at very low temperature).*

In order to explain how a variational formula gives the dependence of the free energy on the shape of the region Λ_L, we need to compute it first in the case of the toroidal graph G_L.

3.2. **Computing the free energy on the torus.** In this section we focus on the case of the dimer model on the toroidal graph G_L, and in Section 4 we explain how to use this result to compute (via a Large Deviation Principle) the free energy in more general domains. For definiteness, we restrict here to the hexagonal graph with minimal fundamental domain containing one black and one white vertex. The reader who is interested in the case of more general periodic graphs is referred to [22], where the general theory is developed (see also Section 3.5 below).

In this section, G_L denotes the graph periodized (with period L) in both directions $\vec{e}_1, \vec{e}_2$, as in Figure 9. With abuse of notation, we still denote V_W, V_B the set of white/black sites of G_L, without keeping the L dependence. Black/white sites are therefore indexed by coordinates $x = (x_1, x_2) \in \Lambda_L$, where

$$\Lambda_L := \{x : -L/2 \leq x_i \leq L/2 - 1, i = 1, 2\} \tag{3.3}$$

(we assume for simplicity that L is even). We adopt the convention that a black vertex has the same coordinates as the white vertex just to its left.

We assign weights A, B, C to edges that are horizontal, north-west oriented and north-east oriented, respectively. Of course we could set one of

them, say A, to 1 without loss of generality. We will see later that the model becomes trivial (in the limit $L \to \infty$) unless A, B, C satisfy the triangular inequality (i.e. none of the three values exceeds the sum of the two others). According to the general procedure, we define the Kasteleyn matrix, with rows/columns indexed by black/white vertices b/w of G_L, as follows. If b, w are not neighbors, then $K(b, w) = 0$. Otherwise,

$$K(b, w) = A\, 1_{x(b)=x(w)} + B\, 1_{x(w)=x(b)+(0,1)} + C\, 1_{x(w)=x(b)-(1,0)}. \tag{3.4}$$

Here, $x(v)$ denotes the coordinates $x = (x_1, x_2)$ of a vertex v. We still call $Z_{G_L,\underline{t}}$ and $\pi_{G_L,\underline{t}}$ the partition function and probability measure.

3.2.1. *Diagonalization of the Kasteleyn matrix.* In this section we diagonalize the four matrices $K_{\theta\tau}$ that appear in formula (2.14). The point is that, by translation invariance, eigenfunctions are provided simply by the Fourier basis (complex exponentials).

For $\theta, \tau \in \{0, 1\}$, let

$$\mathcal{D}_{\theta\tau} = \left\{k = (k_1, k_2), k_1 = \tfrac{2\pi}{L}(n_1 + \tfrac{\theta}{2}), k_2 = \tfrac{2\pi}{L}(n_2 + \tfrac{\tau}{2}), -\tfrac{L}{2} \le n_i < \tfrac{L}{2}\right\} \subset [-\pi, \pi]^2 \tag{3.5}$$

(recall we are assuming for simplicity that L is even) and

$$f_k(x) = \frac{1}{L} e^{-ikx} \tag{3.6}$$

to be seen as a function on V_W. The matrix $K_{\theta\tau}$ maps functions on V_W into functions on V_B.

Going back to the definition of the matrix $K_{\theta\tau}$, let us assume that the path γ_2 crosses the edges from black vertices of coordinate $(-L/2, x_2), -L/2 \le x_2 \le L/2 - 1$ to white vertices $(L/2 - 1, x_2)$, while γ_1 crosses the edges from black vertices $(x_1, L/2 - 1), -L/2 \le x_1 \le L/2 - 1$ to white vertices $(x_1, -L/2)$. One has for a black site of coordinate $x = (x_1, x_2)$

$$\begin{aligned} &[K_{\theta\tau} f_k](x) \\ &= A f_k(x) + B(-1)^{\tau \mathbf{1}_{x_2 = L/2-1}} f_k(x + (0, 1)) + C(-1)^{\theta \mathbf{1}_{x_1 = -L/2}} f_k(x - (1, 0)) \end{aligned} \tag{3.7}$$

where, by convention, $x - (1, 0) \equiv (L/2 - 1, x_2)$ if $x_1 = -L/2$ and similarly $x + (0, 1) \equiv (x_1, -L/2)$ if $x_2 = L/2 - 1$. Next, note that for $k \in \mathcal{D}_{\theta\tau}$, we have

$$(-1)^\tau e^{-ik(x_1, -L/2)} = e^{-ik(x_1, L/2-1)} e^{-ik_2}, \tag{3.8}$$

$$(-1)^\theta e^{-ik(L/2-1, x_2)} = e^{-ik(-L/2, x_2)} e^{ik_1}. \tag{3.9}$$

Therefore,

$$[K_{\theta\tau} f_k](x) = \mu(k) f_k(x) := \left(A + B e^{-ik_2} + C e^{+ik_1}\right) f_k(x). \tag{3.10}$$

In other words, f_k is an eigenfunction of $K_{\theta\tau}$ with eigenvalue $\mu(k)$. The collection $f_k(\cdot), k \in \mathcal{D}_{\theta\tau}$ is an orthonormal basis of eigenfunctions, since their number, L^2, equals the dimension of the Kasteleyn matrix $K_{\theta\tau}$.

We have then

$$\det K_{\theta\tau} = \prod_{k \in \mathcal{D}_{\theta\tau}} \mu(k) = \prod_{z, w \in \mathbb{C} : z^L = (-1)^\theta, w^L = (-1)^\tau} (A + B/w + Cz) \tag{3.11}$$

where we used that $k \in \mathcal{D}_{\theta\tau}$ if and only if $(z,w) = (e^{ik_1}, e^{ik_2})$ satisfies $(z^L, w^L) = ((-1)^\theta, (-1)^\tau)$.

Remark 3.2. *Note that, for $L \in 2\mathbb{N}$ as we are assuming, $\det K_{\theta\tau}$ (that is real, being the determinant of a real matrix) satisfies also*

$$\det K_{\theta\tau} \geq 0, (\theta,\tau) \neq (0,0) \tag{3.12}$$

while

$$\text{sign}(\det K_{00}) = \begin{cases} -1 & \text{if} \quad A,B,C \text{ satisfy the triangular inequality} \\ +1 & \text{if} \quad A,B,C \text{ do not satisfy the triangular inequality.} \end{cases} \tag{3.13}$$

In fact, note that for $(\theta,\tau) \neq (0,0)$, if $k \in \mathcal{D}_{\theta\tau}$ then also $-k \in \mathcal{D}_{\theta\tau}$ and $k \neq -k$ (here, $-k$ has to be considered modulo $(2\pi, 2\pi)$). Since $\mu(k) = \overline{\mu(-k)}$, one easily gets that $\det K_{\theta\tau} \geq 0$.

As far as K_{00} is concerned, the situation is different. All values $k \in \mathcal{D}_{00}$ except for $k = (0,0), (0,\pi), (\pi,0)$ and (π,π) come in distinct pairs $\{k,-k\}$ and give a positive contribution. On the other hand, one has

$$\mu(0,0) = A+B+C, \quad \mu(\pi,\pi) = A-B-C, \tag{3.14}$$

$$\mu(0,\pi) = A-B+C, \quad \mu(\pi,0) = A+B-C, \tag{3.15}$$

hence the claim.

With a similar reasoning one sees that, for L odd, all determinants are positive except $\det(K_{11})$, whose sign is as in the r.h.s. of (3.13).

The reader can already guess the following:

Theorem 3.3.

$$F(A,B,C) := \lim_{L\to\infty} \frac{1}{L^2} \log Z_{G_L,\underline{t}} = \frac{1}{(2\pi)^2} \int_{[-\pi,\pi]^2} \log|\mu(k)| dk_1 dk_2 \tag{3.16}$$

$$= \frac{1}{(2\pi i)^2} \int \log|P(z,w)| \frac{dz}{z}\frac{dw}{w} := \frac{1}{(2\pi i)^2} \int \log|A+B/w+Cz| \frac{dz}{z}\frac{dw}{w} \tag{3.17}$$

with the second integral performed over the torus $\mathbb{T} := \{(z,w) \in \mathbb{C}^2 : |z| = |w| = 1\}$.

Remark 3.4. *For $z \in \mathbb{C}$, let $\text{Log}\,(z)$ denote the principal branch of the logarithm on the complex plane, with cut on the negative real axis, so that $2\log|z| = \text{Log}\,(z) + \text{Log}\,(\bar{z})$. Then, the integral (3.16) can be written as*

$$\frac{1}{(2\pi i)^2} \int \text{Log}\, P(z,w) \frac{dz}{z}\frac{dw}{w} \tag{3.18}$$

Before deducing in detail Theorem 3.3 from (3.11), let us point out some potential difficulties. First of all, if A,B,C satisfy the triangular inequality, then the "characteristic polynomial" $P(z,w) := A + B/w + Cz$ has two (and only two) zeros on the torus $\mathbb{T}$, and the second zero is the complex conjugate of the first. In other words, $\mu(k)$ has two zeros, $p^+ := (p_1^+, p_2^+)$ and $p^- = -p^+$.

Exercise 3.5. *Consider the triangle of sides A,B,C and let $\theta_A, \theta_B, \theta_C$ be the corresponding angles. Then, prove by elementary geometry that the zeros of μ are at*

$$p^\pm := \pm(\pi - \theta_B, \pi - \theta_C). \tag{3.19}$$

For instance, if $A = B = C$ then $p^{\pm} = \pm(2\pi/3, 2\pi/3)$.

Then, it may happen that, for some choice of A, B, C, of L and of θ, τ, one eigenvalue of $K_{\theta\tau}$ is extremely small or even exactly zero, so that its determinant cannot behave like $\exp(L^2 F)$. By the way, this is exactly what happens for the dimer model on the periodized square lattice $\mathbb{Z}^2$, when edge weights are uniform ($t_e \equiv 1$). In this case, the matrix K_{00} has one vanishing eigenvalue for every L, while the other matrices $K_{\theta\tau}$ do not. As another example, take the dimer model on the honeycomb lattice with weights $A = B = C$. Then, one easily sees that K_{00} has one vanishing eigenvalue whenever L is multiple of 3.

Another delicate point is that, even if one can deduce from (3.11) the exponential asymptotics of each $\det(K_{\theta\tau})$, one needs care when the determinants are combined as in (2.14): since the coefficients $c_{\theta\tau}$ do not all have the same sign, there might be dangerous cancellations and $Z_{G_L,\underline{t}}$ might have a rate of exponential growth that is strictly smaller than that of any of the $\det(K_{\theta\tau})$.

As a last remark, when A, B, C do not satisfy the triangular inequality, say $A > B + C$, then the problem looks much easier because the function μ has no zeros. Unfortunately, this is the least interesting case. In fact, we will see that in this situation the density of dimers of type A (horizontal dimers) tends to 1 in the thermodynamic limit. In other words, the system is in a so-called "frozen phase" with no fluctuations at all.

Proof of Theorem 3.3. We start from (2.14). Each $\det K_{\theta\tau}$ is a sum over dimer configurations M of a weight that equals, in absolute value, the weight of M. Therefore, $|\det K_{\theta\tau}| \leq Z_{G_L,\underline{t}}$. Then, since $|c_{\theta\tau}| = 1$, we see that

$$\max_{\theta,\tau} |\det K_{\theta\tau}| \leq Z_{G_L,\underline{t}} \leq 2 \max_{\theta,\tau} |\det K_{\theta\tau}|. \tag{3.20}$$

Therefore, the free energy is the Laplace asymptotics of $\max_{\theta,\tau} |\det K_{\theta\tau}|$.

From (3.11) we see that

$$\frac{1}{L^2} \log |\det K_{\theta\tau}| = \frac{1}{L^2} \sum_{k \in \mathcal{D}_{\theta\tau}} \log |\mu(k_1, k_2)|. \tag{3.21}$$

If the function μ has no zeros on $[-\pi, \pi]^2$, which happens if A, B, C do not satisfy the triangular inquality (i.e. $A > (B + C)$ or one of the other two analogous possibilities) then it is clear that the r.h.s. of (3.21) converges to

$$\frac{1}{(2\pi)^2} \int_{[-\pi,\pi]^2} dk_1 dk_2 \log|\mu(k_1, k_2)| \tag{3.22}$$

that is also the r.h.s. of (3.16).

Let us consider instead the more interesting case where A, B, C satisfy the triangular inequality, so that the function $P(z, w)$ has two conjugate zeros. Note first of all that the integral in the r.h.s. of (3.16) is convergent, since the logarithmic singularity is integrable. Second, for every (θ, τ) call $k^{\pm}_{\theta,\tau}$, the value $k \in \mathcal{D}_{\theta\tau}$ that is closest to $p^{\pm}$; if there is more than one at minimal distance (there are at most four), choose one arbitrarily.

Remark 3.6. *Recall how the sets $\mathcal{D}_{\theta\tau}$ were defined in* (3.5)*, and the fact that $p^- = -p^+$. The following statements are easily checked:*

(i) $k^-_{\theta,\tau} = -k^+_{\theta,\tau} \mod (2\pi, 2\pi)$.
(ii) *for every* (θ,τ),

$$|k^\pm_{\theta,\tau} - p^\pm| \le \frac{\sqrt{2}\pi}{L};$$

(iii) *There exists at most one choice of* (θ,τ), *call it* $(\bar\theta,\bar\tau)$, *such that*

$$|p^\pm - k^\pm_{\bar\theta,\bar\tau}| \le \frac{\pi}{2L}.$$

Write

$$|\det K_{\theta\tau}| = |\mu(k^+_{\theta,\tau})|^2 \prod_{k\in\mathcal{D}_{\theta\tau}\setminus\{k^\pm_{\theta,\tau}\}} |\mu(k)| \tag{3.23}$$

$$= |\mu(k^+_{\theta,\tau})|^2 \times \exp\left[L^2 \times \frac{1}{L^2}\sum_{k\neq k^\pm_{\theta,\tau}} \log|\mu(k)|\right]. \tag{3.24}$$

By Riemann approximation, the sum in the exponent times $1/L^2$ can be written as the integral in the r.h.s. of (3.16), plus an error term $o(1)$ as $L\to\infty$:

Lemma 3.7. *For every choice of* θ,τ *one has*

$$\frac{1}{L^2}\sum_{k\neq k^\pm_{\theta,\tau}} \log|\mu(k)| \overset{L\to\infty}{=} \frac{1}{(2\pi)^2}\int_{[-\pi,\pi]^2} \log|\mu(k)|dk_1dk_2 + o(1). \tag{3.25}$$

See Appendix A for the proof (a little care is needed since the integrand is singular at $p^\pm$).

The term

$$|\mu(k^+_{\theta,\tau})|^2 \tag{3.26}$$

is obviously *upper* bounded by a constant but is anyway potentially dangerous: if $k^\pm_{\theta,\tau}$ are very close to $p^\pm$ (possibly coinciding with $p^\pm$), it can be very small. On the other hand, recall that we have to take first the maximum of $|\det K_{\theta\tau}|$ over (θ,τ) and then the limit $L\to\infty$. I.e., essentially, we have to take the maximum of (3.26) over θ,τ. Then, from Remark (3.6) we see that the maximum over (θ,τ) of (3.26) is bounded *below* by $const./L^2$, because $\mu(\cdot)$ vanishes linearly at $p^\pm$. Altogether,

$$\frac{1}{L^2}\log\max_{\theta,\tau}|\det K_{\theta\tau}| = F(A,B,C) + o(1) \tag{3.27}$$

and the theorem follows. ∎

3.3. **Infinite-volume correlation functions.** As long as L is finite, correlation functions on the torus are given by formula (2.38). We want to take the limit $L\to\infty$ and get rid of the combination of determinants.

We assume right away that A,B,C satisfy (strictly) the triangle condition, so that all three dimer densities are non-trivial and the angles $\theta_A,\theta_B,\theta_C$ are in $(0,\pi)$. In the converse case, we know that dimers are asymptotically all of one type, so that correlation functions are trivial (i.e. they vanish in the $L\to\infty$ limit).

We note first of all that, if w has coordinates x and b has coordinates y, then

$$K_{\theta\tau}^{-1}(w,b) = \frac{1}{L^2}\sum_{k\in\mathcal{D}_{\theta\tau}} \frac{e^{ik(y-x)}}{\mu(k)}. \tag{3.28}$$

From now on, we assume that the integer n in (2.38) is fixed independent of L. Likewise, calling x_i, y_i the coordinates of the white/black vertex of e_i, we assume that x_i, y_i are independent of L (in particular, none of the edges e_i crosses the paths γ_1, γ_2 of Figure 9). In view of (2.38), it is not surprising if we state the following (see the end of this section for the proof):

Theorem 3.8. *One has*

$$\lim_{L\to\infty} \pi_{G_L,\underline{t}}\left[\prod_{j=1}^{n} 1_{e_j}\right] = \prod_{j=1}^{n} K_{00}(e_j) \times \det\{K^{-1}(x_j - y_k)_{j,k\le n}\} \tag{3.29}$$

where

$$K^{-1}(x) := \frac{1}{(2\pi)^2}\int_{[-\pi,\pi]^2} \frac{e^{-ikx}}{\mu(k)} = \frac{1}{(2\pi i)^2}\int \frac{z^{-x_1}w^{-x_2}}{P(z,w)}\frac{dz}{z}\frac{dw}{w}. \tag{3.30}$$

We call $\pi_{G,\underline{t}}$ the infinite-volume measure.

While the statement looks almost obvious, there is a rather delicate point related to the values of k close to $k_{\theta,\tau}^{\pm}$.

Proof of Theorem 3.8. Let us start from (2.38) and let us assume for the moment being that all four matrices $K_{\theta\tau}$ are invertible. We recall also that $K_{\theta\tau}^{-1}(\mathrm{w},\mathrm{b})$ is given by (3.28).

Write $M^{(\theta,\tau)}, \tilde{M}^{(\theta,\tau)}$ for the $n\times n$ matrices

$$M^{(\theta,\tau)} = \{K_{\theta\tau}^{-1}(\mathrm{w}_j,\mathrm{b}_k)\}_{1\le j,k\le n} \tag{3.31}$$

and

$$\tilde{M}^{(\theta,\tau)} = \{\tilde{K}_{\theta\tau}^{-1}(\mathrm{w}_j,\mathrm{b}_k)\}_{1\le j,k\le n} \tag{3.32}$$

where $\tilde{K}_{\theta\tau}^{-1}(\mathrm{w},\mathrm{b})$ is defined as in (3.28), except that the sum is restricted to

$$k\in\mathcal{D}_{\theta\tau}\setminus\{k_{\theta,\tau}^{\pm}\}.$$

It is easy to prove (see Appendix A):

$$\lim_{L\to\infty}\tilde{K}_{\theta\tau}^{-1}(\mathrm{w}_j,\mathrm{b}_k) = K^{-1}(x_j-y_k), \tag{3.33}$$

where the r.h.s. was defined in (3.30). One has then

$$M^{(\theta,\tau)} = \tilde{M}^{(\theta,\tau)} + \sum_{\omega=\pm}\frac{1}{L^2\mu(k_{\theta,\tau}^{\omega})}u_{\omega}^{(\theta,\tau)}[v_{\omega}^{(\theta,\tau)}]^{\mathrm{T}} \tag{3.34}$$

where $u_{\omega}^{(\theta,\tau)}, v_{\omega}^{(\theta,\tau)}$ are the $n-$dimensional column vectors whose j^{th} components are

$$[u_{\omega}^{(\theta,\tau)}]_j = e^{-ik_{\theta,\tau}^{\omega}x_j} \tag{3.35}$$

$$[v_{\omega}^{(\theta,\tau)}]_j = e^{ik_{\theta,\tau}^{\omega}y_j} \tag{3.36}$$

while v^{T} denotes the tranpose of v. In other words, we write $M^{(\theta,\tau)}$ as $\tilde{M}^{(\theta,\tau)}$ plus two rank-one perturbations. We recall the so-called "Matrix

Determinant Lemma" that states that, if M is any $n \times n$ matrix and u, v are two $n-$dimensional column vectors,

$$\det(M + u\, v^{\mathrm{T}}) = \det M + v^{\mathrm{T}} \operatorname{adj}(M)\, u, \tag{3.37}$$

with $\operatorname{adj}(M)$ the adjugate matrix of M, i.e. the $n \times n$ matrix whose (i,j) element equals $(-1)^{i+j}$ times the (i,j)-minor of M.

Applying this identity twice, we see that (we write for simplicity u_ω, v_ω instead of $u_\omega^{(\theta,\tau)}, v_\omega^{(\theta,\tau)}$)

$$\begin{aligned}\det(K_{\theta\tau}) \det(M^{(\theta,\tau)}) &= \det(K_{\theta\tau}) \det(\tilde{M}^{(\theta,\tau)}) \\ &+ \frac{\det(K_{\theta\tau})}{L^2 \mu(k^+_{\theta,\tau})} v_+^{\mathrm{T}} \operatorname{adj}(\tilde{M}^{(\theta,\tau)})\, u_+ \\ &+ \frac{\det(K_{\theta\tau})}{L^2 \mu(k^-_{\theta,\tau})} v_-^{\mathrm{T}} \operatorname{adj}(\hat{M}^{(\theta,\tau)})\, u_-, \\ \hat{M}^{(\theta,\tau)} &= \tilde{M}^{(\theta,\tau)} + \frac{1}{L^2 \mu(k^+_{\theta,\tau})} u_+ v_+^{\mathrm{T}} \end{aligned} \tag{3.38}$$

Let us first look at the first term in the r.h.s. of (3.38) (the one not involving u_ω, v_ω) and *let us disregard the other terms for the moment.* Plugging the first term into (2.38), we see that the multi-edge probability in the l.h.s. of (3.29) equals

$$\frac{\frac{1}{2} \sum_{\theta,\tau=\pm} c_{\theta\tau} \left[\det K_{\theta\tau}\right] \prod_{j=1}^n K_{\theta\tau}(e_j) \det\{\tilde{M}^{(\theta,\tau)}\}}{Z_{G_L,\underline{t}}}. \tag{3.39}$$

Note that $K_{\theta\tau}(e_j) = K_{00}(e_j)$ for L large enough, because of how we chose the paths γ_1, γ_2 just before (3.7): in fact, the edges $e_j, j \le n$ are fixed independently of L, while the paths γ_1, γ_2 are moved to infinity as $L \to \infty$, so that they do not cross $e_j, j \le n$. Recall also that $|\det(K_{\theta\tau})|/Z_{G_L,\underline{t}}$ is bounded above by 1 (cf. (3.20)). Finally, recall (3.33). Altogether, this implies that (3.39) tends as $L \to \infty$ to the r.h.s. of (3.29).

It remains to prove that the last two terms in the r.h.s. of (3.38) give a negligible contribution to $\pi_{G_L,\underline{t}}[\prod_{j=1}^n 1_{e_j}]$, as $L \to \infty$. We look for instance at the former term; the one containing $\hat{M}$ can be treated the same way. Note that

$$v_+^{\mathrm{T}} \operatorname{adj}(\tilde{M}^{(\theta,\tau)})\, u_+ \overset{L\to\infty}{=} O(1),$$

because every element of the matrix $\tilde{M}^{(\theta,\tau)}$ and of the vectors $u_+^{(\theta,\tau)}, v_+^{(\theta,\tau)}$ is uniformly bounded, and their dimension n does not grow with L. Therefore, to conclude the proof of the theorem it is sufficient to show that

$$\frac{1}{L^2 Z_{G_L,\underline{t}}} \left| \frac{\det(K_{\theta\tau})}{\mu(k^-_{\theta,\tau})} \right| \overset{L\to\infty}{=} o(1). \tag{3.40}$$

Since $\mu(\cdot)$ vanishes linearly at $p^\pm$ and, one has from item (ii) in Remark 3.6

$$\mu(k^\pm_{\theta\tau}) = O(1/L) \quad \text{for every} \quad (\theta,\tau). \tag{3.41}$$

On the other hand, call $(\hat\theta,\hat\tau)$ the value of (θ,τ) such that $|p^\pm - k^\pm_{\theta\tau}|$ is maximal. From item (iii), we see that and

$$|\mu(k^\pm_{\hat\theta\hat\tau})| \geq \frac{c}{L}. \tag{3.42}$$

We also need the following technical fact (see Appendix A for the proof):

Lemma 3.9. *There exist two L-independent constants $0 < c_- < c_+ < \infty$ such that, for every choices of (θ,τ) and (θ',τ')*

$$c_- \left|\frac{\mu(k^+_{\theta\tau})\mu(k^-_{\theta\tau})}{\mu(k^+_{\theta'\tau'})\mu(k^-_{\theta'\tau'})}\right| \leq \left|\frac{\det(K_{\theta\tau})}{\det(K_{\theta'\tau'})}\right| \leq c_+ \left|\frac{\mu(k^+_{\theta\tau})\mu(k^-_{\theta\tau})}{\mu(k^+_{\theta'\tau'})\mu(k^-_{\theta'\tau'})}\right|. \tag{3.43}$$

Then, in view of $Z_{G_L,\underline{t}} \geq |\det(K_{\hat\theta\hat\tau})|$ (cf. (3.20)), the l.h.s. of (3.40) is upper bounded by

$$\frac{1}{L^2|\mu(k^-_{\theta\tau})|}\left|\frac{\det(K_{\theta\tau})}{\det(K_{\hat\theta\hat\tau})}\right| \leq c_+ \left|\frac{\mu(k^+_{\theta\tau})}{L^2\mu(k^+_{\hat\theta\hat\tau})\mu(k^-_{\hat\theta\hat\tau})}\right|. \tag{3.44}$$

Using (3.42) and (3.41), we see that the r.h.s. is $o(1)$ as $L\to\infty$ and (3.40) follows.

In the proof of the Theorem, we have assumed that none of the determinants $\det(K_{\theta\tau})$ is exactly zero, since we started from (2.38). However, this restriction is easily removed. Recall from Remark 3.6 that at most one of the four determinants is zero for a given value of L, and let (θ_0,τ_0) be its label. Note first of all that, while (3.38) is ill-defined when $\mu(k^\pm_{\theta_0,\tau_0}) = 0$ (indeed, the l.h.s. is $0\times\infty$), this expression has a well-defined limit if $\mu(k^\pm_{\theta_0,\tau_0}) \to 0$. In fact, it is immediate to see that this limit is just

$$\frac{(v_-^{\mathrm{T}}u_+)(v_+^{\mathrm{T}}u_-)}{L^4} \prod_{k\in\mathcal{D}_{\theta_0\tau_0}\setminus\{k^\pm_{\theta_0,\tau_0}\}} \mu(k). \tag{3.45}$$

We claim in fact that, when $\det(K_{\theta_0\tau_0}) = 0$, (2.38) still holds, with

$$\det(K_{\theta_0\tau_0})\det\{K^{-1}_{\theta_0\tau_0}(\mathrm{w}_j,\mathrm{b}_k)_{j,k\leq n}\}$$

replaced by (3.45). Given this, the proof of Theorem 3.8 works exactly the same as in the case where all determinants are non-zero. The reason why the claim is true is that the probability in the l.h.s. of (3.45) is continuous w.r.t. the edge weights, *at fixed L*. If it so happens that $\det(K_{\theta_0\tau_0}) = 0$ for a given L, one can slightly change the weights $\underline{t}$ (i.e., in the case of the hexagonal graph we are looking at, the numbers A,B,C are replaced by $A+\epsilon A', B+\epsilon B', C+\epsilon C'$, with ϵ small enugh) in such a way that $\det(K_{\theta_0\tau_0}) \neq 0$ while the other three determinants are still non-zero. Letting $\epsilon\to 0$, one obtains the claim. ∎

Remark 3.10. *The proof of Theorem 3.8 we have given here differs from that in* [22]. *In fact, to the best of our understanding, the delicate issue of the values of k close to $k^\pm_{\theta,\tau}$ was not really solved there. For instance, it is stated in* [22, Sec. 4.3] *that, if $k^\pm_{\theta,\tau}$ is (dangerously) at distance $\ll L^{-2}$ from $p^\pm$ for a given L and (θ,τ), then by changing L by $O(\sqrt{L})$ this is no more the case. This is false in general. For instance, for the dimer model*

on $\mathbb{Z}^2$ with constant edge weights, the characteristic polynomial $\mu(\cdot)$ is (cf. Example (3.22) *below)*

$$\mu(k) = 1 + ie^{ik_1} - e^{i(k_1+k_2)} - ie^{ik_2}. \tag{3.46}$$

It is immediately verified that $p^+ = (0,0), p^- = (\pi,\pi)$ so that for $(\theta,\tau) = (0,0)$, $k^{\pm}_{0,0}$ is at distance zero from $p^{\pm}$, for every value of L.

As another example, for the dimer model on the honeycomb graph with $A = B = C = 1$, one sees (Exercise!) that $\det K_{00} = 0$ whenever L is a multiple of 3.

3.3.1. *Asymptotic dimer densities.* We give here a first application of Theorem 3.8, by computing the asymptotic (for $L \to \infty$) densities of dimers of the three types, i.e. the $L \to \infty$ limit of the probability that an edge of type A, B, or C is occupied by a dimer.

To this purpose, take $n = 1$ in Theorem 3.8 and let e_A, e_B, e_C be three edges of type A, B, C respectively. We have

$$\rho_A := \pi_{G,\underline{t}}(e_A \in M) = \frac{A}{(2\pi i)^2} \int \frac{1}{(A + B/w + Cz)} \frac{dz}{z} \frac{dw}{w} \tag{3.47}$$

$$\rho_B := \pi_{G,\underline{t}}(e_B \in M) = \frac{B}{(2\pi i)^2} \int \frac{1}{(A + B/w + Cz)} \frac{dz}{z} \frac{dw}{w^2} \tag{3.48}$$

$$\rho_C := \pi_{G,\underline{t}}(e_C \in M) = \frac{C}{(2\pi i)^2} \int \frac{1}{(A + B/w + Cz)} dz \frac{dw}{w}. \tag{3.49}$$

The easiest, and least interesting case, is once more that where the triangular inequality is not satisfied for A, B, C. Say for instance that $A > B + C$. Then, the contour of integration $|z| = 1$ contains only the pole $z = 0$. Applying the residue theorem,

$$\rho_A = \frac{A}{2\pi i} \int \frac{1}{(A + B/w)} \frac{dw}{w}. \tag{3.50}$$

The integration over w is done similarly and the final result is that

$$\rho_A = 1.$$

In other words, the asymptotic density of horizontal dimers is 1 and the system is said to be in a "frozen phase"

Exercise 3.11. *Check that the two other integrals give zero.*

The asymptotic dimer densities are more interesting when the triangle condition is satisfied:

Exercise 3.12. *Check (again by an application of the residue theorem) the following: if A, B, C satisfy the triangle condition, then*

$$\rho_A = \theta_A/\pi, \rho_B = \theta_B/\pi, \rho_C = \theta_C/\pi$$

where, given a triangle of side lengths A, B, C, we denote $\theta_A, \theta_B, \theta_C$ the angles opposite to sides of lenght A, B, C respectively.

Hints (for ρ_A; the other computations are analogous):

- *the poles in z are $z = 0$ and $z = z_0(w) = -(A + B/w)/C$;*
- *check that if both are inside the contour $|z| = 1$, then the integral in z gives zero (deform the contour to a circle of radius R and send R to ∞);*

- *what remains is then*
$$\frac{1}{2\pi i}\int_{|w|=1:|z_0(w)|>1}\frac{1}{w+B/A}dw; \tag{3.51}$$
- *with $w=e^{i\theta}$, note that the domain of integration becomes the portion of contour from $w_0:=e^{i(\pi-\theta_C)}$ to $\overline{w_0}$;*
- *do the integral and use simple trigonometry to obtain the result $(\pi-\theta_B-\theta_C)/\pi=\theta_A/\pi$.*

Exercise 3.13. *Compute the asymptotic dimer densities ρ_A,ρ_B,ρ_C by taking the derivatives of $F(A,B,C)$ with respect to $\log A,\log B,\log C$ respectively, and check that one obtains the same result as above. This is a bit tricky because of the absolute value in* (3.16)*, or because of the branch cut in* (3.18)*.*

3.4. **Large distance behavior of K^{-1}.** From the definition, we see that $K^{-1}(x)$ is the Fourier coefficient of frequencies $x\in\mathbb{Z}^2$ of the function $1/\mu(\cdot)$, that is a periodic function on the torus $[-\pi,\pi]^2$. If $1/\mu$ were C^∞, it would be a standard fact to see that $K^{-1}(x)$ decays faster than any power of $|x|$, for $|x|\to\infty$.

Exercise 3.14. *Prove this, using integration by parts.*

Instead, we know that $1/\mu$ has two simple poles at p^+ and $p^-=-p^+$. As a consequence, $K^{-1}(\cdot)$ decays much more slowly at large distance. Let us see precisely how.

Remark first of all that
$$\mu(k)=\alpha^\pm(k_1-p_1^\pm)+\beta^\pm(k_2-p_2^\pm)+O(|k-p^\pm|^2) \tag{3.52}$$
$$\alpha^\pm:=iCe^{ip_1^\pm},\qquad \beta^\pm:=-iBe^{-ip_2^\pm} \tag{3.53}$$
around $k=p^\pm$. Note that
$$\overline{\alpha^+}=-\alpha^-,\quad \overline{\beta^+}=-\beta^-, \tag{3.54}$$
so that we may write
$$\alpha^\omega=\omega\alpha_1+i\alpha_2,\quad \beta^\omega=\omega\beta_1+i\beta_2,\quad \omega=\pm, \tag{3.55}$$
with α_j,β_j real. If one argues that the leading contribution to $K^{-1}(x)$ is given by the integral around the poles and replaces μ by its linear approximation around them, i.e. replaces the integral in (3.30) with
$$\frac{1}{(2\pi)^2}\sum_{\epsilon=\pm}\int_{\mathbb{R}^2}e^{-ikx}\frac{1}{\alpha^\epsilon(k_1-p_1^\epsilon)+\beta^\epsilon(k_2-p_2^\epsilon)}, \tag{3.56}$$
then (ignoring the fact that the integral is not absolutely convergent) the integral in (3.56) can be computed via the residue theorem and the result correctly gives the leading term in Eq. (3.57) below.

It is actually not hard to justify these approximations:

Proposition 3.15. *One has*
$$K^{-1}(x)=\frac{1}{\pi}\Re\left[\frac{e^{-ip^+x}}{\phi(x)}\right]+O(|x|^{-2}) \tag{3.57}$$

as $|x| \to \infty$*, where*

$$\phi(x) := -\beta^+ x_1 + \alpha^+ x_2. \tag{3.58}$$

Remark 3.16. *Note that, using* (3.19)*, one has*

$$\frac{\alpha^+}{\beta^+} = \frac{C}{B} e^{i\theta_A}.$$

Under our assumption that the triangular inequality holds (strictly) for A, B, C*, the angle* θ_A *is in* $(0, \pi)$*, and therefore the ratio* α^+/β^+ *is not real. In other words,* α^+, β^+ *are not colinear and the map* $\mathbb{R}^2 \ni z \mapsto \phi(z) \in \mathbb{C}$ *spans the whole complex plane. For the same reason, whenever* $|x|$ *tends to infinity, we can bound*

$$c_1|x| \le |\phi(x)| \le c_2|x| \tag{3.59}$$

for some finite and non-zero constants c_1, c_2*. This would not be the case if* α^+, β^+ *were colinear. Indeed, if* $\alpha^+ = \gamma\beta^+$ *with* $\gamma \in \mathbb{R}$*, then* $\phi(\lfloor \gamma x_1 \rfloor, x_1) = O(1)$ *even for large* x_1*.*

Remark 3.17 (Dimer-dimer correlations)**.** *Let* e_1, e_2 *be two (L-independent) edges of* G_L*, with coordinates* x_i, y_i *for their white/black endpoints. From Theorem 3.8 we see that*

$$\lim_{L\to\infty} \pi_{G_L,\underline{t}}[e_i \in M] = K_{00}(e_i)K^{-1}(x_i - y_i), \quad i = 1, 2 \tag{3.60}$$

and

$$\lim_{L\to\infty} \left\{\pi_{G_L,\underline{t}}[e_1 \in M, e_2 \in M] - \pi_{G_L,\underline{t}}(e_1 \in M)\pi_{G_L,\underline{t}}(e_2 \in M)\right\} \tag{3.61}$$

$$= -K_{00}(e_1)K_{00}(e_2)K^{-1}(x_1 - y_2)K^{-1}(x_2 - y_1). \tag{3.62}$$

From (3.59) *we see that the dimer-dimer correlation (that we denote also* $\pi_{G_L,\underline{t}}(e_1 \in M; e_2 \in M)$*) decays in absolute value like the inverse of the square of the distance beween the two edges, with certain oscillatory pre-factors (due to the complex exponential* $\exp(-ip^+x)$*) that will play an important role later (see Section 5).*

Proof of Proposition 3.15. For $|x| \to \infty$, either x_1 or x_2 is large. Assume the former is the case, and assume also that x_1 is negative (the opposite case can be treated analogously). We start from the integral (3.30). Let us integrate on z first. The only pole is at $z_0(w) = -(A + B/w)/C$ (because $-x_1 - 1 \ge 0$). One gets then

$$\frac{1}{2\pi i C} \int_{w:|z_0(w)|<1} z_0(w)^{-x_1-1} w^{-x_2} \frac{dw}{w} \tag{3.63}$$

$$= \frac{1}{2\pi C} \int_{\pi-\theta_C}^{\pi+\theta_C} d\theta e^{-i\theta x_2} \left[-\frac{A + Be^{-i\theta}}{C}\right]^{-x_1-1} \tag{3.64}$$

$$= \frac{1}{\pi C} \Re \left\{ \int_{\pi-\theta_C}^{\pi} d\theta e^{-i\theta x_2} \left[-\frac{A + Be^{-i\theta}}{C}\right]^{-x_1-1} \right\} \tag{3.65}$$

where in the second expression we wrote $w = e^{i\theta}$ and noticed that $|w_0(z)| < 1$ is equivalent to $\theta \in (\pi - \theta_C, \pi + \theta_C)$. Set $\theta = \pi - \theta_C + u$ and observe that

$[-(A+Be^{-i(u+\pi-\theta_C)})/C]$ equals $\exp(i(-\theta_B+\pi))$ for $u=0$ and has absolute value smaller than 1 for $u\in(0,\pi]$. Also,

$$-\frac{A+Be^{-i(u+\pi-\theta_C)}}{C}=\exp(i(-\theta_B+\pi))e^{-iu\beta^+/\alpha^++O(u^2)} \tag{3.66}$$

and, by the previous remark, the second complex exponential has absolute value smaller than 1. Equation (3.63) then reduces to

$$\frac{1}{\pi C}\Re\left\{e^{-i(\pi-\theta_C)x_2}e^{i(\theta_B-\pi)(x_1+1)}\int_0^{\theta_C}du\,e^{-iux_2+i(u+O(u^2))\beta^+/\alpha^+(x_1+1)}\right\} \tag{3.67}$$

Since x_1 is large and negative, the integral is dominated by the contribution for $u\lesssim 1/|x_1|$. We can therefore neglect the $O(u^2)$ error and extend the domain of integration up to infinity, getting

$$\frac{1}{\pi C}\Re\left\{e^{-i(\pi-\theta_C)x_2}e^{i(\theta_B-\pi)(x_1+1)}\int_0^{+\infty}du\,e^{-iux_2+iu\beta^+/\alpha^+(x_1+1)}\right\} \tag{3.68}$$

$$=\frac{1}{\pi C}\Re\left[e^{-i(\pi-\theta_C)x_2}e^{i(\theta_B-\pi)(x_1+1)}\frac{\alpha^+}{ix_2\alpha^+-i\beta^+x_1}\right]+O(|x|^{-2}). \tag{3.69}$$

Recalling the definitions of α^+ and of p^+, we see that the r.h.s. is the same as that of (3.57).

■

Exercise 3.18. *Prove that $x=(n,n)$ with $n\in\mathbb{Z}\setminus\{0\}$, then*

$$K^{-1}(x)=-\frac{\sin(n(\theta_B+\theta_C))}{\pi nA}. \tag{3.70}$$

Hint: Say that $n<0$. Start from the first line of (3.63) with $x_1=x_2=n$ and compute the integral in the complex plane, using that the primitive of

$$w^{-n-1}z_0(w)^{-n-1}$$

is

$$\frac{C}{An}w^{-n}z_0(w)^{-n}.$$

Exercise 3.19 (Dimers and determinantal proint process)**.** *Deduce from the previous exercise that the set of occupied horizontal dimers in a given horizontal column is a determinantal point process with kernel given by the so-called "sine kernel"*

$$S(n):=-\frac{\sin(n(\theta_B+\theta_C))}{\pi n},n\neq 0,\qquad S(0)=\rho_A.$$

I.e. let $\bar e_n$ be the horizontal edge with black and white endpoints both of coordinate (n,n). Then

$$\pi(\bar e_{n_1},\dots,\bar e_{n_k}\in M)=\det[S(n_i-n_j)]_{i,j\le k}. \tag{3.71}$$

The kernel $S(\cdot)$ is called "sine kernel"; note that $S(n)=S(-n)$, so that the kernel is symmetric. Determinantal point processes (with symmetric kernel) have a lot of interesting structure, see [28]*. Let us give two properties here:*

- *Let $\mathcal{N}_N$ the number of particles in $[1,\dots,N]$ (in our case, of horizontal dimers on edges $\bar e_n, 1\le n\le N$). If the variance of $\mathcal{N}_N$ tends to infinity as $N\to\infty$, then the rescaled number $\overline{\mathcal{N}_N}:=$*

$(\mathcal{N}_N - \mathbb{E}\mathcal{N}_N)/\sqrt{\mathrm{Var}\mathcal{N}_N}$ tends, as $N \to \infty$, to a standard Gaussian Normal random variable. See [28, Th. 8]*. In the case for the sine kernel it can be seen by hand* [20, Sec. 6.3]*, and also follows from Section 5 below, that the variance grows logarithmically with N. Note that the particle occupation variables are not at all i.i.d. (in which case the variance would grow like N instead); this makes this CLT all the more non-trivial.*

- *Let $\lambda_1, \dots, \lambda_N$ the eigenvalues of the matrix $S_N = \{S(i-j)\}_{1 \le i,j \le N}$. The λ_i are real, since the matrix is symmetric. Also, they are non-negative, since its minors (being probabilities of events) are all non-negative. Finally, they do not exceed the value 1. This is proven in the general setting e.g. in* [1, Th. 22]*; in our case it can also be presumably proven by hand. In conclusion, $\lambda_i \in [0,1]$. Then, one can see that $\mathcal{N}_N$ has the same law of the sum $Y = \sum_{i=1}^N X_i$, where the X_i are independent Bernoulli variables, with parameters λ_i. (Check as an exercise that Y and $\mathcal{N}_N$ have the same mean and variance).*

3.5. **General bipartite, periodic toroidal graphs.** We worked out in detail the diagonalization of the Kasteleyn matrix, the computation of the free energy and of infinite-volume correlations (and their long-distance asymptotics) in the case of the hexagonal lattice $\mathcal{H}$ with "minimal fundamental domain" consisting of two vertices. Let us explain here how things work in the more general case of bipartite, periodic planar graphs, where in general the fundamental domain contains n white and n black vertices.

Look for instance at the square-octagon graph of Fig. 13. The minimal fundamental domain, encircled, contains 4 white and 4 black vertices.

As usual, we periodize the graph with period L in both directions. Note that the graph G_L thus obtained contains nL^2 vertices of each color. Call as usual $K_{\theta\tau}, \theta, \tau \in \{0,1\}$ the four $nL^2 \times nL^2$-dimensional Kasteleyn matrices on the torus, constructed as in Section 2.4. In the case of the hexagonal graph, where $n = 1$, we have diagonalized K as $\hat{K} = P^{-1}KP$ where P is the matrix whose columns are all the distinct Fourier eigenfunctions $f_k(\cdot)$, $k \in \mathcal{D}_{\theta,\tau}$ and consequently the lines of P^{-1} are given by $\bar{f}_k(\cdot)$. In other words, in the Fourier basis the matrix $\hat{K}$ is diagonal and the diagonal elements are $\mu(k)$. Then, the determinant of $K_{\theta\tau}$ is just the product over $k \in \mathcal{D}_{\theta,\tau}$ of $\mu(k)$. The situation is slightly more complicated for $n > 1$. I.e., for $k \in \mathcal{D}_{\theta,\tau}$ and $1 \le a \le n$ we define

$$f_{a,k} : (x \in \Lambda_L, a' \in \{1, \dots, n\}) \mapsto f_{a,k}(x, a') := \frac{1}{L} e^{-ikx} \mathbf{1}_{a'=a}, \qquad (3.72)$$

to be seen as a function of white sites (indexed by coordinates x and a'). Note that for every choice of θ, τ there are nL^2 such functions, and they are orthonormal. However, they are not necessarily eigenfunctions of $K_{\theta\tau}$, but almost: the function $[K_{\theta\tau} f_{k,a}](\cdot)$ is a linear combination of $f_{k,a'}(\cdot)$ for different a', *but for the same k*. More precisely, label in some arbitrary way as $k_1, \dots k_{L^2}$ the elements of $\mathcal{D}_{\theta,\tau}$ and let P be the matrix whose first n columns are the functions $f_{k_1,1}(\cdot), \dots, f_{k_1,n}(\cdot)$ and so on until the last one that is $f_{k_{L^2},n}$. Then:

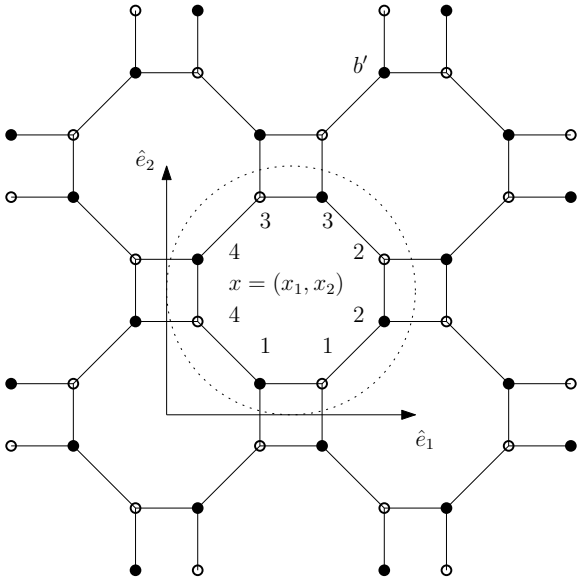

FIGURE 13. The square-octagon graph with its fundamental domain. All edge weights equal 1, for simplicity. The graph is invariant under translations by $n\hat{e}_1+m\hat{e}_2$, $n, m \in \mathbb{Z}$. A vertex is identified by: (1) its color, (2) the coordinates (x_1, x_2) of the fundamental domain it belongs to and (3) by the index $a \in \{1, \dots, 4\}$ specifying its position inside the fundamental domain. For instance, the black vertex b has coordinates $(x_1, x_2 + 1)$ and $a = 2$.

Theorem 3.20. *The matrix $P^{-1}K_{\theta\tau}P$ is block-diagonal, with block of size n. Each block corresponds to a value $k_i, i \leq L^2$. The block corresponding to the value k is a $n \times n$ matrix $M(k) = \{M(a, a'; k)\}_{a,a' \leq n}$ where*

$$M(a, a'; k) = \sum_{e : a \overset{e}{\sim} a'} K_e e^{-ikx_e}. \tag{3.73}$$

In this formula, the sum runs over all edges joining the black vertex b of type a in the fundamental domain of coordinates $x = (0, 0)$ to a white vertex w of type a' (w can be either in the same fundamental domain or in another one); K_e is the Kasteleyn matrix element of $K_{0,0}$ corresponding to edge (w, b); $x_e \in \mathbb{Z}^2$ is the coordinate of the fundamental domain to which w belongs.

Moreover,

$$\det(K_{\theta\tau}) = \prod_{k \in \mathcal{D}_{\theta\tau}} \mu(k), \qquad \mu(k) := \det M(k). \tag{3.74}$$

Proof. We will be sketchy, as the proof is similar to what we did for the honeycomb lattice already. The matrix element of $P^{-1}K_{\theta\tau}P$ on row labelled k, a and column k', a' is

$$\frac{1}{L^2} \sum_b \sum_w K_{\theta\tau}(b, w) e^{ikx(b) - ik'x(w)} 1_{a(b)=a} 1_{a(w)=a'} \tag{3.75}$$

where the sum runs over black vertices b, whose coordinates are denoted $x(b)$ and whose type is denoted $a(b)$ and white vertices w of coordinates $x(w)$ and type $a(w)$. Given the black vertex b of coordinate $x(b)$ and type a, we have that

$$\sum_w K_{\theta\tau}(b,w)e^{-ik'x(w)}1_{a(w)=a'} = e^{-ik'x} \times M(a,a';k'), \tag{3.76}$$

with $M(\cdot)$ defined in (3.73). Here we have used translation invariance of the graph and the fact that $k \in \mathcal{D}_{\theta\tau}$. At this point, the sum over b (constrained to be of type a) can be performed in (3.75) and one gets the indicator function that $k = k'$ times $M(a,a';k)$ as desired. ∎

Example 3.21 (Hexagonal graph)**.** *Just to check things, let us look at the usual hexagonal graph, see Fig. 9. Here, $n = 1$ so that $M(k)$ is just a number. There are three edges connecting the black vertex of coordinates* 0 *to a white vertex: the horizontal edge (and this contributes A to* (3.73)*), the north-west oriented edge that connects to the white vertex of coordinates* $(0,1)$ *and contributes Be^{-ik_2}, and the north-east oriented edge that connects to the white vertex* $(-1,0)$ *and contributes Ce^{ik_1}. Altogether $M(k)$ equals what we called $\mu(k)$ earlier and* (3.74) *reduces to the formula we already knew.*

Example 3.22. *[Square graph] Another well-studied example is that of the square lattice $\mathbb{Z}^2$ with minimal fundamental domain consisting of two vertices. See Figure 14. Note that we may set $t_4 := 1$ by multiplying all weights by a common factor. Also, we will see in Section 4.1 that the case $t_1t_3 = t_2$ is particularly natural. With the choice of Kasteleyn matrix elements as in the figure, check as an exercise that*

$$\mu(k) = t_1 + it_2e^{ik_1} - t_3e^{i(k_1+k_2)} - it_4e^{ik_2}. \tag{3.77}$$

Example 3.23 (Square-octagon graph)**.** *Consider the graph of Fig. 13 and assume for simplicity that all edge weights equal* 1*. A possible Kasteleyn matrix for the infinite graph consists in taking all non-zero matrix elements equal* 1*, except for horizontal edges with the black vertex on the left, for which the matrix element is* -1*. Then, check that*

$$M(k) = \begin{bmatrix} -1 & 0 & e^{ik_2} & 1 \\ 1 & 1 & 0 & -e^{-ik_1} \\ e^{-ik_2} & 1 & 1 & 0 \\ 0 & e^{ik_1} & 1 & 1 \end{bmatrix} \tag{3.78}$$

$$\mu(k) = -5 + e^{ik_1} + e^{-ik_1} + e^{ik_2} + e^{-ik_2} \tag{3.79}$$

Assuming that $\det K_{\theta\tau} \neq 0$, so that in particular each $M(k)$ is invertible, the inverse matrix $K_{\theta\tau}^{-1}$ is simply given by

$$K_{\theta,\tau}^{-1}(\mathrm{w},\mathrm{b}) = \frac{1}{L^2}\sum_{k\in\mathcal{D}_{\theta\tau}} [M(k)]^{-1}_{a,a'}e^{-ik(x-x')}. \tag{3.80}$$

where w (resp. b) is a white (black) vertex of type a (a') and coordinates x (resp. x').

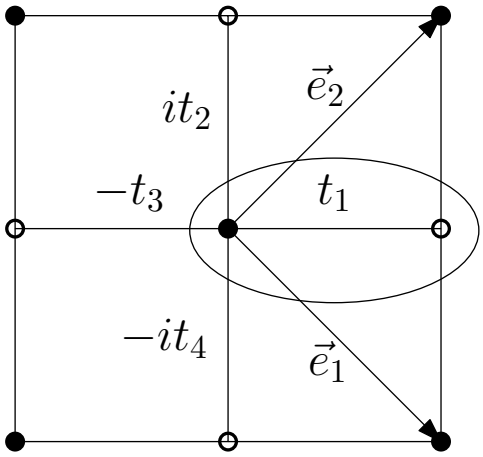

FIGURE 14. The square lattice with its fundamental domain (encircled; the black and white vertex in the same fundamental domain have the same coordinates $x = (x_1, x_2)$), the elementary vectors $\vec{e}_1, \vec{e}_2$ that serve for the definition of coordinates of vertices. Next to the four edges incident to the central black vertices are marked the values of the Kasteleyn matrix elements $K(b, w)$ (the Kasteleyn matrix is invariant under translations by $n\vec{e}_1 + m\vec{e}_2$.

3.6. **Phase structure: liquid, frozen and gaseous phases.** Once we get to (3.74), we would like to go back to Eqs. (2.14) and (2.38) and to repeat the arguments of Section 3.2 to compute the infinite volume of the free energy as

$$\lim_{L\to\infty} \frac{1}{L^2} \log Z_{G_L,\underline{t}} = \frac{1}{(2\pi)^2} \int_{[-\pi,\pi]^2} \log|\mu(k)|. \tag{3.81}$$

Also, one would like to prove that the analog of Theorem 3.8 holds. Namely, if the edge e_i has white vertex of coordinates x_i, a_i and black vertex of coordinates y_i, a_i' then the l.h.s. of (3.29) tends to

$$\pi_{G,\underline{t}}(e_1, \ldots, e_n \in M) := \prod_{j=1}^{n} K_{00}(e_j) \times \det\{K^{-1}_{a_i,a_j'}(x_i - y_j)_{i,j\le k}\} \tag{3.82}$$

where

$$K_{a,a'}(x - x') := \lim_{L\to\infty} K^{-1}_{\theta,\tau}(w, b) = \frac{1}{(2\pi)^2} \int_{[-\pi,\pi]^2} [M(k)]^{-1}_{a,a'} e^{-ik(x-x')}. \tag{3.83}$$

As we saw in Sections 3.2 and 3.3 for the hexagonal graph, for the proof of Theorems 3.3 and 3.8 and for the study of the long-distance behavior of K^{-1} (Proposition 3.15), it is of primary importance to know what are the zeros of $\mu(\cdot)$ and the asymptotic behavior of μ close to them. For the hexagonal graph with elementary fundamental domain it was immediate to study the zeros of μ, as a function of the weights A, B, C. Accordingly, we saw that the infinite-volume measure $\pi_{G,\underline{t}}$ obtained as the $L \to \infty$ limit of the measure on the torus of period L can be of the following two types:

- If A, B, C satisfy the triangle condition then μ has two simple zeros. All three types of dimers have a non-zero asymptotic density, and the inverse Kasteleyn matrix decays like the inverse of the distance. Then, $\pi_{G,\underline{t}}$ is called (with the nomenclature of [22]) a "liquid phase".
- If one of the three weights is strictly larger than the sum of the other two, then μ has no zeros and the dimers of the corresponding type have asymptotic density 1. Therefore, in the $L \to \infty$ limit there are no fluctuations in $\pi_{G,\underline{t}}$ and we have a so-called "frozen phase". In general, $\pi_{G,\underline{t}}$ is called "frozen" in [22] if there exist faces f, f' arbitrarily far away such that the difference $h(f) - h(f')$ is deterministic.
- we have not really proved this, but in the limit case where one of the weights equals the sum of the other two, then the two zeros coincide and at least one of the dimer densities is asymptotically zero (recall the geometric interpretation $\rho_A = \theta_A/\pi$ etc, and note that when the triangle degenerates to a segment, either one or two angles tend to zero; the other two tend to any two angles summing to π). In this case, fluctuations in $\pi_{G,\underline{t}}$ are non-zero, if two limit densities are both zero. However, $\pi_{G,\underline{t}}$ is still frozen, in the sense defined in the previous point. See Fig. 15.

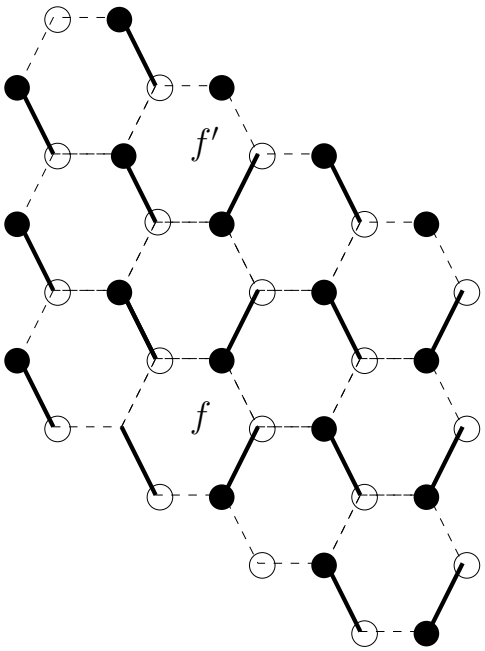

FIGURE 15. Assume that $B = C > 0$, while $A = 0$. Then, the densities of dimers of types of types B and C is $1/2$ and there are no dimers of type A, as in the picture. Then, it is clear that the height difference between two faces in the same vertical column (say, f and f') is deterministic (recall the definition of height function, as in Section 1.1).

In the general case of a graph G with elementary domain containing $n > 1$ black/white vertices, the study of the zeros of the so-called "characteristic polynomial"

$$P(z, w) := \mu(-i \log z, -i \log w)$$

on $\{(z, w) \in \mathbb{C}^2 : |z| = |w| = 1\}$ requires deeper algebraic ideas that we do not explain here, referring the interested reader to [22]. Let us mention here only the following (see also the discussion in Remark 4.7 below). In the general case, *three situations* can happen (excluding degenerate limit cases like that where $A = B + C$ in the hexagonal graph):

(1) μ has two simple zeros $p^+ = -p^-$ and $\alpha^{\pm}, \beta^{\pm}$, defined as in (3.52), satisfy $\alpha^{\omega}/\beta^{\omega} \notin \mathbb{R}$. Then the inverse Kasteleyn matrix decays like the inverse distance (liquid phase);
(2) μ has no zeros and the limit measure is frozen, in the sense above;
(3) μ has no zeros, the inverse Kasteleyn matrix decays *exponentially* with distance but there are no arbitrarily far away faces f, f' such that $h(f) - h(f')$ is deterministic. In the third kind of situation, the limit measure $\pi_{G,\underline{t}}$ is called a *gaseous phase* in [22]. This situation cannot happen on graphs with elementary fundamental domain ($n = 1$). Perhaps the simplest example where a gaseous phase occurs is the square-octagon graph of Fig 13 with constant weights. Note by the way that μ in (3.79) clearly has no zeros for $|z| = |w| = 1$. Another well-known case is that of the square grid with weights that are periodic only under even translations of the lattice, as in Fig. 12.

Given this input, it is not hard to check that the arguments of Sections 3.2 and 3.3 still work and one finds indeed (3.81) and (3.82).

Exercise 3.24. *Compute the characteristic polynomial for the graph 12 and verify that it has no zeros if $a \neq 1$.*

4. Surface tension, large deviation principle and Limit shape

4.1. **Height change and Legendre duality.** Suppose we have an infinite, periodic, planar weighted graph G (with weights $\underline{t}$) and we consider its L-periodization G_L, as in Section 2.4. We know how to compute its partition function $Z_{G_L,\underline{t}}$, via Theorem 2.14, and its infinite-volume free energy and infinite-volume correlations, as explained in Sections 3.2 and 3.3 in the case of the hexagonal graph with minimal fundamental domain consisting of just two vertices. Here, we will learn how to compute infinite-volume free energy and correlations for the model "with imposed slope".

Recall from Section 1.1 that on the toroidal (therefore not planar) graph G_L, the height function is not well-defined. However, given the oriented self-avoiding paths γ_1, γ_2 of Section 2.4 (recall that γ_1 has winding number $(1, 0)$ and γ_2 has winding $(0, 1)$) and given a dimer configuration $M \in \Omega_{G_L}$, let

$$\Delta_j(M) = \sum_{e \sim \gamma_j} \sigma_e (1_{e \in M} - 1_{e \in M_0}) = \sum_{e \sim \gamma_j} \sigma_e \, 1_{e \in M} + L\mathcal{K}_j, \qquad j = 1, 2, \tag{4.1}$$

where:

- the sum runs over the edges crossed by γ_j;
- σ_e, M_0 are as in (3.4); we assume that the reference configuration M_0 has been chosen to have the same periodicity as the weights $\underline{t}$;
- $\mathcal{K}_j = -(1/L)\sum_{e\sim\gamma_j}\sigma_e 1_{e\in M_0}$ is an L-independent constant that depends on the choice of M_0.

We will refer to $\Delta_j(M)$ as to the height change of M in direction j and to $\Delta_j(M)/L$ as to the j-th component of the slope of M. Given $\Delta \in \mathbb{Z}^2$, we let $\Omega_{G_L}(\Delta)$ be the subset of Ω_{G_L} consisting of configurations with height change equal $\Delta = (\Delta_1, \Delta_2)$. It is natural to introduce the probability measure

$$\pi_{L,\Delta}(M) := \frac{w_{\underline{t}}(M)}{Z_{L,\Delta}}, \qquad M \in \Omega_{G_L}(\Delta)$$

$$Z_{L,\Delta} := \sum_{M\in\Omega_{G_L}(\Delta)} w_{\underline{t}}(M) \quad (4.2)$$

where as usual $w_{\underline{t}}(M) := \prod_{e\in M} t_e$. (We drop the explicit dependence of the measure and of the partition function on the weights, to lighten notations). Note that $\pi_{L,\Delta}$ is just the restriction of $\pi_{G_L,\underline{t}}$ to $\Omega_{G_L}(\Delta)$.

We would like now to study the asymptotic (in L) properties of the measure $\pi_{L,\Delta}$, especially in terms of free energy. This will be a crucial ingredient for the variational problem that determines the "macroscopic shape", as discussed in next section. In principle, Kasteleyn's theory does not provide a formula to compute the partition function $Z_{L,\Delta}$ "with fixed height change". However, we can resort to a classical trick of statistical mechanics, known as "equivalence of ensembles" or "Legendre duality". In simple words, the idea is to replace the "micro-canonical" measure with constraint on the height change by a "gran-canonical" measure where the height change is not fixed, but a Lagrange multiplier suitably modifies the edge weights, thereby imposing the desired (average) height change. The fluctuations of Δ/L in the gran-canonical measure tend to zero as $L \to \infty$, so the micro-canonical and gran-canonical measures are actually close in a sense. This is a bit like imposing a given average magnetization for the Ising model by introducing a suitably chosen magnetic field instead of directly fixing the value of the sum of the spin variables.

Let us see how this works in practice. First of all it is easy to see, using the ideas of Section 1.1, that $\Delta_1(M)$ is unchanged if the path γ_1 is translated in direction $\vec{e}_2$, and similarly $\Delta_2(M)$ is unchanged if the path γ_2 is translated in direction $\vec{e}_1$. We let $\gamma_1^{(i)}, i = 1, \dots, L$ be the distinct translations of γ_1 in direction $\vec{e}_2$, and similarly for $\gamma_2^{(i)}, i = 1, \dots, L$. Next, given "magnetic fields"[4] $B = (B_1, B_2) \in \mathbb{R}^2$, we introduce modified edge weights simply by multiplying an edge weight t_e by $e^{B_j(N_j^+ - N_j^-)}$ where N_j^+ (resp. N_j^-) is the number of paths $\gamma_j^{(i)}, i \le L$ that cross e leaving its white endpoint on the right (resp. on the left). If the original edge weights were called $\underline{t}$, the new one will be denoted $\underline{t}^{(B)}$.

[4]The name "magnetic fields" is used due the above analogy with the Ising model

Example 4.1. *Suppose that the initial graph is the hexagonal lattice $\mathcal{H}$ with weights identically equal to 1. After the magnetic transformations, one obtains the usual A, B, C weighted hexagonal graph, with $A = 1, B = e^{-B_1}, C = e^{-B_2}$. Note that, by choosing a suitable $(B_1, B_2) \in \mathbb{R}^2$, we can obtain any weights A, B, C (recall that we can set one weights, say A, to one by multiplying all of them by a common prefactor).*

Example 4.2. *Suppose that the initial graph G is the square lattice $\mathbb{Z}^2$ with weghts identically equal to 1. A look at Figure 10 shows that the magnetic transformation turns the weights into those of Figure 14 with $t_1 = e^{-B_1}, t_2 = e^{-B_1-B_2}, t_3 = e^{-B_2}$ and $t_4 = 1$. Note that $t_1 t_3 = t_2$.*

The partition function of the dimer model with magnetically modified weights can be trivially written as

$$Z_{G_L,\underline{t}^{(B)}} = \sum_{\Delta} \sum_{M \in \Omega_{G_L}(\Delta)} w_{\underline{t}^{(B)}}(M), \tag{4.3}$$

where the sum runs over all values of Δ such that $\Omega_{G_L}(\Delta) \neq \emptyset$. Now, write

$$w_{\underline{t}^{(B)}}(M) = w_{\underline{t}}(M) e^{\sum_{e \in E} 1_{e \in M} \sum_{j=1,2} B_j \sum_{i=1}^{L} \chi_{i,j}(e)} \tag{4.4}$$

$$= w_{\underline{t}}(M) e^{\sum_{j=1}^{2} L B_j (\Delta_j(M) - L\mathcal{K}_j)} \tag{4.5}$$

where:

- in the exponent the first sum runs over all edges of G_L,
- $\chi_{i,j}(e)$ equals $+1$ (resp. -1) if path $\gamma_j^{(i)}$ crosses e leaving the white endpoint to the right (resp. left) and 0 if it does not cross it;
- $\mathcal{K}_j$ is the same constant as in (4.1).

As a consequence, the partition function (4.3) is given by a linear combination of the partition functions $Z_{L,\Delta}$, each of them weighted with the exponential of the height change multiplied by the magnetic field:

$$Z_{G_L,\underline{t}^{(B)}} = e^{-L^2 B \cdot \mathcal{K}} \sum_{\Delta} e^{LB \cdot \Delta} Z_{L,\Delta}. \tag{4.6}$$

The "magnetically modified" dimer model (on the torus) can be solved by discrete Fourier analysis, as we did in Sections 3.2 and 3.3 and, more generally, 3.5. We call $\mu^{(B)}(\cdot)$ its characteristic polynomial and we assume that its zeros are such that (3.81) holds (there is actually no loss of generality in this; see Remark 4.7 below). We let

$$F_{\underline{t}}(B) = \lim_{L \to \infty} \frac{1}{L^2} \log Z_{G_L,\underline{t}^{(B)}} = \frac{1}{(2\pi)^2} \int_{[-\pi,\pi]^2} dk \log |\mu^{(B)}(k)|.$$

On the other hand, setting

$$\rho = (\rho_1, \rho_2) := (\Delta_1/L, \Delta_2/L),$$

we see from (4.6) that this implies

$$\sigma^*_{\underline{t}}(B) := F_{\underline{t}}(B) + \mathcal{K} \cdot B = \lim_{L \to \infty} \frac{1}{L^2} \log \sum_{\rho} e^{L^2 \left[B \cdot \rho + \frac{1}{L^2} \log Z_{L,L\rho} \right]} \tag{4.7}$$

(the "$*$" in σ^* is because we are actually rather interested in the Legendre transform $\sigma(\cdot)$ of $\sigma^*(\cdot)$). Intuitively, the r.h.s. should be dominated by the

value of ρ such that the quantity in square parenthesis is maximal. The sum over ρ runs over the values such that $L\rho$ is in $\mathbb{Z}^2$.

Recall that $\sigma^*_{\underline{t}}(B)$ is a convex function of B (being the limit of convex functions). Define its Legendre transform as

$$\sigma_{\underline{t}}(\rho) := \sup_B \{\rho \cdot B - \sigma^*_{\underline{t}}(B)\}, \qquad \rho \in \mathbb{R}^2, \tag{4.8}$$

that is a convex function of ρ and as usual we have, by duality,

$$\sigma^*_{\underline{t}}(B) = \sup_\rho \{\rho \cdot B - \sigma_{\underline{t}}(\rho)\}. \tag{4.9}$$

Usually the function $\sigma_{\underline{t}}(\cdot)$ is referred to as "surface tension". We let N be the set where $\sigma_{\underline{t}}(\cdot)$ is strictly convex.

Remark 4.3. *In the hexagonal graph case with constant weights $\underline{t} \equiv 1$, the function $\sigma^*(\cdot)$ is strictly convex on the simply connected set*

$$\mathcal{B} = \{B = (B_1, B_2) : (1, e^{-B_1}, e^{-B_2}) \text{ satisfy the strict triangle condition}\} \tag{4.10}$$

while it is not strictly convex outside. In fact,

$$\partial_{B_j} F_{\underline{1}}(B) = -e^{-B_j} \partial_{e^{-B_j}} F_{\underline{1}}(B)$$

that, in view of Example 4.1, is minus *the density of north-west (if $j = 1$) or north-east oriented dimers. We know that if, say, $1 > e^{-B_1} + e^{-B_2}$ then north-east and north-west oriented dimers have both density zero, i.e. the gradient of $F_{\underline{1}}(\cdot)$ is constant, i.e. the function is not strictly convex. A similar argument holds when either e^{-B_1} or e^{-B_2} exceeds the sum of the oher two weights. Conversely, if $(1, e^{-B_1}, e^{-B_2})$ satisfy strictly the triangle condition, then the dimer densities (i.e. the angles of the triangle with those side-lengths) are bijectively determined by B. Therefore, the gradient of $F_{\underline{1}}(\cdot)$ cannot take the same value for two different values of B, and* (4.10) *follows.*

The surface tension $\sigma(\cdot)$ can be computed by evaluating explicitly the Legendre transform and the result is that, if $\rho = \mathcal{K} + r$, then [20]

$$\sigma(\rho) = \begin{cases} \frac{1}{\pi}\left[\Lambda(-\pi r_1) + \Lambda(-\pi r_2) + \Lambda(\pi(1 + r_1 + r_2))\right] \le 0, & r \in T \cup \partial T \\ +\infty & \text{otherwise} \end{cases} \tag{4.11}$$

with

$$\Lambda(\theta) = \int_0^\theta \ln(2\sin(t))dt$$

and T the triangle

$$T = \{\rho \in \mathbb{R}^2 : -1 < \rho_1 < 0, -1 < \rho_2 < 0, -1 < \rho_1 + \rho_2 < 0\}. \tag{4.12}$$

One can check that this expression vanishes for $r \in \partial T$ and is strictly negative, strictly convex and smooth for $r \in T$. In particular, $N = T + \mathcal{K}$

Similarly, for the dimer model on $\mathbb{Z}^2$ with constant weights $\underline{t} \equiv 1$, the function $\sigma^*_{\underline{t}}(\cdot)$ is strictly convex on a simply connected set $\mathcal{B}$ and its Legendre transform $\sigma_{\underline{t}}(\cdot)$ is strictly convex, smooth and strictly negative is the open square of vertices

$$(-1, 0), (0, 0), (0, -1), (-1, -1),$$

translated by $\mathcal{K}$, and $+\infty$ outside. Actually, for the surface tension one finds an expression similar to (4.11).

Exercise 4.4. *Compute $\mathcal{B}$ for the dimer model on $\mathbb{Z}^2$ and constant weights.*

Let us now go back to the problem of studying the constrained measure $\pi_{L,\Delta}$. We will deduce from (4.7) a Large Deviation Principle-type statement that essentially says that the probability that the height change Δ takes some value ρL, for the original model with unmodified weights $\underline{t}$, is at leading order

$$\exp[-L^2(F_{\underline{t}}(0)+\sigma_{\underline{t}}(\rho))]. \tag{4.13}$$

Note that $-\sigma_{\underline{t}}(\rho)$ has the interpretation of an entropy per unit volume, being approximately $1/L^2$ times the logarithm of the number of configurations with height change close to ρL. Note also that $F_{\underline{t}}(0)$ is just the infinite-volume free energy of the model with unmodfied weights $\underline{t}$.

The following is a rigorous version of the above statement:

Theorem 4.5. *For $\rho \in \mathbb{R}^2$ and $\epsilon > 0$, let $U(\rho,\epsilon)$ be the ball of $\mathbb{R}^2$ of radius ϵ centered at ρ.*

Let also

$$P_{L,\rho,\epsilon} := \pi_{G_L,\underline{t}}\left(\frac{\Delta(M)}{L} \in U(\rho,\epsilon)\right). \tag{4.14}$$

Then, for ρ in the interior of N (the set where the surface tension is strictly convex)

$$\limsup_{L\to\infty} \frac{1}{L^2}\log P_{L,\rho,\epsilon} \le -(F_{\underline{t}}(0)+\sigma_{\underline{t}}(\rho)) + O(\epsilon) \tag{4.15}$$

and

$$\liminf_{L\to\infty} \frac{1}{L^2}\log P_{L,\rho,\epsilon} \ge -(F_{\underline{t}}(0)+\sigma_{\underline{t}}(\rho)) - O(\epsilon). \tag{4.16}$$

Remark 4.6. *From Theorem 4.5 is is clear that $-\sigma_{\underline{t}}(\rho) \le F_{\underline{t}}(0)$ for every ρ, and actually the maximum of $-\sigma_{\underline{t}}(\cdot)$ equals $F_{\underline{t}}(0)$, as follows from (4.9). The value $\rho^*(0)$ of ρ that maximizes $-(F_{\underline{t}}(0)+\sigma(\rho))$, i.e. that minimizes $\sigma_{\underline{t}}(\rho)$, is therefore the typical value of $\Delta(M)/L$ under the measure $\pi_{G_L,\underline{t}}$. It is related to $\sigma^*_{\underline{t}}(\cdot)$ by*

$$\rho^*(0) = \nabla_B \sigma^*_{\underline{t}}(0), \tag{4.17}$$

assuming that $\sigma^*(\cdot)$ is differentiable at 0. More generally, under $\pi_{G_L,\underline{t}^{(B)}}$, the typical value of $\Delta(M)/L$ is

$$\rho^*(B) := \nabla_B \sigma^*(B). \tag{4.18}$$

Proof. For simplicity of exposition, we will suppose that the surface tension is not only strictly convex at ρ, but also differentiable (this is not really a loss of generality, since a convex function is differentiable almost everywhere and we can always slightly move the center of the ball $U(\rho,\epsilon)$).

Let

$$Z_{L,\rho,\epsilon} := \sum_{\Delta/L\in U(\rho,\epsilon)} Z_{L,\Delta}. \tag{4.19}$$

We have from (4.6), for any choice of B,

$$\frac{1}{L^2}\log Z_{L,\rho,\epsilon} \le B\cdot(\mathcal{K}-\rho) + \frac{1}{L^2}\log Z_{G_L,\underline{t}^{(B)}} + |B|\epsilon. \tag{4.20}$$

Choosing the value $B^*(\rho)$ that realizes the supremum in (4.8)[5], taking the limsup w.r.t. L and recalling that $\frac{1}{L^2}\log Z_{G_L,\underline{t}}$ converges to $F_{\underline{t}}(0)$ gives the upper bound (4.15).

As for the lower bound, assume by contradiction that

$$Z_{L,\rho,\epsilon} \le e^{-L^2[\sigma_{\underline{t}}(\rho)+M\epsilon]} \tag{4.21}$$

for some large M. Given $R>0$ define

$$\sigma^{(R)}_{\underline{t}}(\rho) := \sup_{|B|\le R}\big\{\rho\cdot B - \sigma^*_{\underline{t}}(B)\big\} \tag{4.22}$$

to be compared with (4.8). Note that its Legendre transform $\sigma^{*,(R)}_{\underline{t}}(B)$ coincides with $\sigma^*_{\underline{t}}(B)$ if $|B|<R$ and is $+\infty$ if $|B|>R$. The possible values of Δ/L are in a compact, L-independent set $\mathcal{C}$. Let ρ_i be a finite collection of $O(\delta^{-2})$ points of $\mathcal{C}\setminus U(\rho,\epsilon)$ such that the union of $U(\rho_i,\delta)$ covers $\mathcal{C}\setminus U(\rho,\epsilon)$. Using (4.21) we have

$$\begin{aligned} Z_{\underline{t}(B^*(\rho))}e^{L^2B^*(\rho)\cdot\mathcal{K}} &\le \sum_{\Delta/L\in U(\rho,\epsilon)} Z_{L,\Delta}e^{LB^*(\rho)\cdot\Delta} + \sum_i \sum_{\Delta/L\in U(\rho_i,\delta)} Z_{L,\Delta}e^{LB^*(\rho)\cdot\Delta} \\ &\le e^{L^2[B^*(\rho)\cdot\rho+\epsilon|B^*(\rho)|-\sigma_{\underline{t}}(\rho)-M\epsilon]} + e^{L^2\delta|B^*(\rho)|}\sum_i e^{L^2[B^*(\rho)\cdot\rho_i+\frac{1}{L^2}\log Z_{L,\rho_i,\delta}]} \\ &= e^{L^2[\sigma^*_{\underline{t}}(B^*(\rho))+\epsilon(|B^*(\rho)|-M)]} + e^{L^2\delta|B^*(\rho)|}\sum_i e^{L^2[B^*(\rho)\cdot\rho_i+\frac{1}{L^2}\log Z_{L,\rho_i,\delta}]}. \end{aligned} \tag{4.23}$$

Fix $R>0$ large enough (how large, not depending on ϵ or δ) so that $\sigma^{(R)}_{\underline{t}}(\cdot)$ defined in (4.22) coincides with $\sigma_{\underline{t}}(\cdot)$ in an ϵ-independent neighborhood of ρ. Given ρ_i, let B_i be a value of B that realizes the supremum in (4.22). From (4.20) we see that

$$\frac{1}{L^2}\log Z_{L,\rho_i,\delta} \le -B_i\cdot\rho_i + \sigma^*_{\underline{t}}(B_i) + o(1) + R\delta \tag{4.24}$$

$$= -\sigma^{(R)}_{\underline{t}}(\rho_i) + o(1) + R\delta, \tag{4.25}$$

where $o(1)$ vanishes as $L\to\infty$.

Since $\sigma_{\underline{t}}(\cdot)$ (and therefore $\sigma^{(R)}_{\underline{t}}(\cdot)$) is strictly convex around ρ, it follows that

$$\sup_{\rho'\notin U(\rho,\epsilon)}[B^*(\rho)\cdot\rho' - \sigma^{(R)}_{\underline{t}}(\rho')] \le \sigma^*_{\underline{t}}(B^*(\rho)) - c(\epsilon) \tag{4.26}$$

for some $c(\epsilon)>0$. In fact, if (4.26) fails then

$$B^*(\rho)\cdot(\rho'-\rho) = \nabla\sigma^{(R)}_{\underline{t}}(\rho)\cdot(\rho'-\rho) = \sigma^{(R)}_{\underline{t}}(\rho') - \sigma^{(R)}_{\underline{t}}(\rho) \tag{4.27}$$

for some $\rho'\neq\rho$, which contradicts strict convexity. Putting together (4.23), (4.24) and (4.26) we get

$$\begin{aligned} &e^{L^2[\sigma^*_{\underline{t}}(B^*(\rho))+o(1)]} = Z_{\underline{t}(B^*(\rho))}e^{L^2B^*(\rho)\cdot\mathcal{K}} \\ &\le e^{L^2[\sigma^*_{\underline{t}}(B)-\epsilon(M-|B^*(\rho)|)]} + e^{L^2[\sigma^*_{\underline{t}}(B^*(\rho))-c(\epsilon)+\delta R+o(1)]}\times\text{const}\times\delta^{-2}. \end{aligned} \tag{4.28}$$

[5] $B^*(\rho)$ is easily seen to be finite because ρ is in the interior of the region where $\sigma_{\underline{t}}(\cdot)$ is finite. Actually, $B^*(\rho)=\nabla\sigma_{\underline{t}}(\rho)$.

This is a contradiction if $M > |B^*(\rho)|$ and $R\delta < c(\epsilon)$.

■

Remark 4.7. *In general, the set $\mathcal{B}$ where $\sigma^*_{\underline{t}}$ is strictly convex is called, in the language of algebraic geometry, the "amoeba" of the characteristic polynomial $P(z,w) = \mu(-i\log z, -i\log w)$. For general bipartite, periodic weighted graphs, in general $\mathcal{B}$ is not simply connected. The complementary of $\mathcal{B}$ contains both unbounded components U_i and bounded connected components, or "holes", H_i. $\sigma^*_{\underline{t}}(\cdot)$ is affine on each component of the complementary of $\mathcal{B}$. The slope ρ_i of $\sigma^*_{\underline{t}}(\cdot)$ in H_i is in the interior of N (the set where the surface tension is strictly convex). On U_i, instead, the slope is "extremal" (i.e. it belongs to the boundary ∂N). This has interesting consequences, as discovered in* [22]. *Namely, each "hole" $H_1, \dots, H_k$ of $\mathcal{B}$ corresponds to a "gaseous phases" of the model. I.e., if B is in H_n, then the limit measure $\pi_{G,\underline{t}^{(B)}}$ is gaseous (in the sense of having exponentially decaying correlations, cf. Section 3.5) and it depends only on n (i.e. two values of B in the same hole yield the same infinite-volume measure). The typical slope under $\pi_{G,\underline{t}^{(B)}}$, for $B \in H_i$, equals ρ_i, the slope of the affine function $\sigma^*_{\underline{t}}(\cdot)$ in H_i. Correspondingly, $\sigma_{\underline{t}}$ has a conical singularity at ρ_i, where it is not C^1. The case where B belongs to an unbounded component U_i corresponds instead to the case where $\pi_{G,\underline{t}^{(B)}}$ is a "frozen phase" with extremal slope in ∂N.*

Finally, the generic point[6] *$B \in \mathcal{B}$ corresponds to a "liquid phase" $\pi_{G,\underline{t}^{(B)}}$, whose characteristic polynomial has two simple zeros.*

We refer to [22] *for a more complete discussion.*

4.2. **Macroscopic shape.** In many situations, a typical random dimer configuration M of a finite, large graph G has the following two features:

- if the lattice spacing is rescaled by $1/L$, with L the diameter of L, so that the diameter of the graph becomes of order one, then the height function $h_M(\cdot)$, also rescaled by the L, tends to a deterministic function $\bar{\phi}(\cdot)$ ("limit shape", or "macroscopic profile");
- the profile $\bar{\phi}(\cdot)$ is in general non linear and depends on the detail of the boundary of G.

See for instance Figures 4 and 16 in [20], in the case of the dimer model on the hexagonal graph.

The purpose of this section is two-fold:

- to explain how the limit shape is determined by a variational principle, involving the "surface tension" $\sigma_{\underline{t}}(\cdot)$ that we computed in the previous section;
- to explain how the same variational principle also gives the asymptotics of the free energy per unit volume of the dimer model in the large graph G.

Let us start by recalling (cf. also Remark 3.1) that for more conventional statistical mechanics models without "hard constraints", like the Ising

[6]For special choices of the weights, a hole H_i can shrink to a point. The corresponding value B would not be a "generic point B"

model, the free energy $\log Z_\Lambda$ in a large domain Λ is, at leading order, equal to

$$|\Lambda|(F + o(1))$$

where $|\Lambda|$ is the volume of Λ, F is a function that depends on temperature, magnetic field, etc. but not on the shape of Λ, and $o(1)$ vanishes as $|\Lambda| \to \infty$ (under the assumption that the surface area of the boundary of Λ is negligible w.r.t. to the volume). For the dimer model, this fails. I.e. the logarithm of the partition function, divided by the volume, turns out to depend non-trivially on the shape of the domain and actually on its fine details, even in the limit of infinite volume. See Fig. 16.

Let us make some more precise assumptions on the domains we are interested in. We start as usual from an infinite, weighted, periodic bipartite graph G. Recall from 1.3 that a dimer configuration M on G is in bijection with a tiling of the plane, where admissible tiles are unions of two adjacent faces of the dual graph G^*. Recall also that the height function $h_M(\cdot)$ is defined on faces of G, i.e. on vertices of G^*. A tileable domain U of the plane defines a finite graph $G(U)$ (see e.g. Figure 3: $G(U)$ is on the left and U is the union of rhombi on the right). Fixing by convention the height function to 0 at some vertex $v_0 \in G^*$ on the boundary of U, the height function on all other vertices of ∂U is the same for every tiling of U, i.e. for every dimer configuration on $G(U)$: the boundary height is independent of the configuration inside the domain. We call $h_{\partial U}$ the boundary height and we extend it to every point on ∂U (not just the vertices of the dual graph) by linear interpolation.

Let $\{U_L\}_{L\geq 1}$ be a sequence of finite and growing domains of the plane, formed of faces of G^*, that admit a tiling. We will make the following assumptions:

Assumption 4.8. *The sequence $\{U_L\}_L$ satisfies the following conditions:*

1. *U_L tends to the whole plane as $L \to \infty$, i.e. $G(U_L)$ tends to the whole graph G;*
2. *U_L is simply connected and the rescaled domain $\hat{U}_L := (1/L)U_L$ converges in Hausdorff distance to a domain U of $\mathbb{R}^2$, whose boundary is a simple, piecewise smooth curve;*
3. *the rescaled height function $u \in \partial\hat{U}_L \mapsto \hat{h}(u) := (1/L)h(uL)$ converges to a function $\phi_{\partial U} : \partial U \mapsto \mathbb{R}$ (such function is necessarily continuous and actually Lipshitz).*

Note that point (2) does not imply at all (3). Actually, one can easily construct two sequences such that $\hat{U}_L^{(1)}, \hat{U}_L^{(2)}$ tend to the same limit domain U but the boundary heights converge to two different boundary heights. See Figure 16.

Given the dimer model on the domain (or rather on the sequence of domains) $G(U_L)$, the probability measure $\pi_{G(U_L),\underline{t}}$ on dimer configurations M induces a probability measure on the height function h_M. We denote $\hat{h}_M$ the rescaled height function that to a vertex u in the rescaled graph $\hat{U}_L$ associates the value $(1/L)h_M(uL)$.

Further, we denote $\Phi_{U,\phi_{\partial U}}$ the set of Lipshitz functions $\phi : U \mapsto \mathbb{R}$ that coincide with $\phi_{\partial U}$ on ∂U (we will say that ϕ is compatible with the boundary

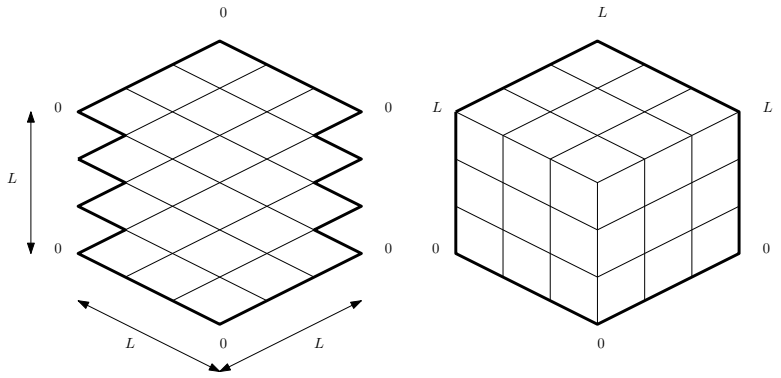

FIGURE 16. The two domains are approximately hexagons of side L, and once rescaled both tend to a unit hexagon. However, in the first case the boundary height is identically zero, while in the second it ranges from 0 to L and once rescaled tends to a non-trivial function. A possible tiling of the domains is drawn in order to visualize the boundary height more easily. Actually, the domain on the left has only one possible tiling, so that its free energy is zero, while the one on the right has exponentially many (in L) tilings, so its free energy is positive.

height). Given $\phi \in \Phi_{U,\phi_{\partial U}}$, we write $\hat{h}_M \overset{\delta}{\sim} \phi$ to mean that

$$|\hat{h}_M(u) - \phi(u)| \leq \delta \qquad \forall u \in \hat{U}_L.$$

In words, the rescaled height function $\hat{h}_M$ is everywhere within distance δ from the target profile ϕ. Finally, we let

$$\Sigma[\phi] := \int_U \sigma_{\underline{t}}(\nabla\phi)dx. \tag{4.29}$$

The following theorem has been proven in [4] for the dimer model on the square and hexagonal grids. The general case is treated in [23]:

Theorem 4.9. *There exists a unique minimizer $\bar{\phi}$ of $\Sigma[\phi]$ on $\Phi_{U,\phi_{\partial U}}$ and*

$$\lim_{L\to\infty} \frac{1}{|U_L|} \log Z_{G(U_L),\underline{t}} = -\Sigma[\bar{\phi}] \tag{4.30}$$

Moreover, given $\phi \in \Phi_{U,\phi_{\partial U}}$, we have

$$\lim_{\delta\to 0} \lim_{L\to\infty} \frac{1}{|U_L|} \log \left(\sum_{M:\hat{h}_M \overset{\delta}{\sim} \phi} w_{\underline{t}}(M) \right) = -\Sigma[\phi]. \tag{4.31}$$

For the statement about uniqueness of the minimizer, see also [6]. Note that the theorem in particular implies that, for large L, the height profile concentrates around the deterministic "macroscopic shape" $\bar{\phi}$. Determining

the actual macroscopic shape given a domain U and a boundary height $\phi_{\partial U}$ is in general not doable explicitly. For nice enough domains and boundary heights, the macroscopic shape can be actually worked out and it turns out to be parametrized by analytic functions [21]. In the regions where $\bar{\phi}$ is smooth and not "frozen"[7], it solves the non-linear, parabolic PDE

$$\mathrm{div}(\nabla\sigma_{\underline{t}} \circ \nabla\bar{\phi}) := \sum_{i,j=1}^{2} \sigma_{i,j}(\nabla\bar{\phi})\partial^2_{x_i x_j}\bar{\phi}_i(x) = 0 \tag{4.32}$$

where $\sigma_{i,j}(\rho) := \partial^2_{\rho_i\rho_j}\sigma_{\underline{t}}(\rho)$ and where the boundary value of $\bar{\phi}$ is fixed by the boundary height $\phi_{\partial U}$. Note that (4.32) is just the Euler-Lagrange equation associated to the variatonal problem of minimixing $\Sigma[\phi]$ over $\Phi_{U,\phi_{\partial U}}$.

For the proof of Theorem 4.9 we refer to [4, 23] but, in view of Theorem 4.5, claims (4.30) and (4.31) should look very natural to the reader.

5. Height flucuations in "liquid" phases, and massless Gaussian Field

One of the most interesting aspects of dimer models is the emergence, in the infinite-volume limit, of the so-called massless Gaussian field, also known as Gaussian Free Field (GFF). That is, the height function $h_M(\cdot)$, for a configuration sampled from the infinite-volume measure $\pi_{G,\underline{t}}$, tends on large scales to a log-correlated Gaussian field, provided that $\pi_{G,\underline{t}}$ is a liquid phase (in the sense of Section 3.5, i.e. if its characteristic polynomial μ has two simple zeros).

For definiteness, let us restrict once more to the case of the hexagonal lattice with elementary fundamental domain and we assume that the weights A, B, C satisfy the strict triangle condition, so that μ has simple zeros p^+ and $p^- = -p^+$ as in (3.19). We denote for lightness of notation the infinite-volume measure simply as Π. Given M sampled from Π, we denote $h_M(\cdot)$ the corresponding height function defined as in (1.2) on the faces of G, with the convention that the height is zero at some reference face, say the origin. Let us choose $c(e)$ in (1.2) to be equal to $\Pi(e \in M)$. This way, the gradients of $h_M(\cdot)$ are by construction centered. A first natural question is, how does the variance of the gradients behave at large distances. This is answered by the following (the proof is quite instructive and it is given later):

Theorem 5.1. *If A, B, C satisfy stricly the triangular inequality, then*

$$\mathrm{Var}_\Pi(h_M(f) - h_M(f')) \overset{|f-f'|\to\infty}{=} \frac{1}{\pi^2}\log|\phi(f) - \phi(f')| + O(1). \tag{5.1}$$

Note that the r.h.s. of (5.1) is also $1/(\pi^2)\log|f - f'| + O(1)$. It is a remarkable fact that the prefactor of the logarithm does not depend on the weights A, B, C. Actually, one finds the same prefactor $1/\pi^2$ on any bipartite periodic graph, provided the measure Π is a liquid phase! [22].

It is possible to prove also that the height gradient, suitably rescaled, tend to a Gaussian variable:

[7] It may happen that at some points $\nabla\bar{\phi}$ is an extremal slope, (i.e., for the model on the hexagonal graph, it is on the boundary of the triagle N of allowed slopes). In this case, the PDE (4.32) does not make sense literally, since $\sigma_{i,j}$ are not defined there.

Theorem 5.2. *Under the same assumptions as in the previous Theorem, one has the convergence in law*

$$\frac{h_M(f) - h_M(f')}{\sqrt{\mathrm{Var}_\Pi(h(f) - h(f'))}} \Rightarrow \mathcal{N}(0,1), \tag{5.2}$$

as $|f - f'| \to \infty$, *where* $\mathcal{N}(0,1)$ *is a standard Gaussian random variable.*

The proof of Theorem 5.2 is quite similar (just a bit more technical) to that of Theorem 5.1: one proves that all the moments of the random variable in the l.h.s. of (5.2) tend to the moments of $\mathcal{N}(0,1)$ (for Gaussian variables, moments determine the law). The way one estimates the n-th moment is very similar to the way one estimates the second moment in the proof of Theorem 5.1 (we will not give details for moments of order ≥ 3; see for instance [24, Th. 2.8] and see also the proof of Theorem 5.3 below).

We already see from Theorem 5.2 that on large scales height gradients become Gaussian, but we do not yet see a Gaussian "field". To see the emergenge of the GFF we should look not just at the law of the gradient $h_M(f) - h_M(f')$ but at the joint law of n such gradients. Then, we have:

Theorem 5.3. *Let* A, B, C *satisfy strictly the triangle inequality, let* $f_1, \dots, f_4$ *be four faces of the graph and let* $D(f_1, \dots, f_4)$ *denote the minimal distance between* $f_i, f_j, 1 \leq i \neq j \leq 4$. *Also, let*

$$\phi(f, f') := (\phi(f) - \phi(f'))\mathbf{1}_{f \neq f'} + \mathbf{1}_{f = f'}. \tag{5.3}$$

Then,

$$\begin{aligned} &\mathrm{Cov}_\Pi(h_M(f_2) - h_M(f_1), h_M(f_4) - h_M(f_3)) \\ &\qquad = \mathcal{G}(\{f_1, f_2\}, \{f_3, f_4\}) + O(1/D(f_1, f_2, f_3, f_4)) \\ &\qquad \mathcal{G}(\{f_1, f_2\}, \{f_3, f_4\}) := \frac{1}{2\pi^2} \Re \log \left(\frac{\phi(f_4, f_1)\phi(f_3, f_2)}{\phi(f_4, f_2)\phi(f_3, f_1)}. \right) \end{aligned} \tag{5.4}$$

Moreover, given $n \in \mathbb{N}$ *and faces* $f_i, f_i', i \leq n$ *we have*

$$\begin{aligned} \Pi \left[\prod_{i=1}^{n} (h_M(f_i) - h_M(f_i')) \right] = \sum_{\sigma \in M[n]} \prod_{i=1}^{n/2} \mathcal{G}(\{f_{\sigma(i)}, f'_{\sigma(i)}\}, \{f_{\sigma(i+1)}, f'_{\sigma(i+1)}\}) \\ + O(1/D(f_1, \dots, f_n')) \end{aligned} \tag{5.5}$$

where the sum runs over the matchings σ *of* $\{1, \dots, n\}$ *and is to be interpreted as zero if* n *is odd. Finally,* $D(f_1, \dots, f_n')$ *is the minimal distance among faces* $f_1, f_1', \dots, f_n, f_n'$.

The reason for distinguishing f equal or not equal f' in (**??**) is simply that we want (5.3) to make sense also when some faces coincide.

Recall the so-called Wick theorem for Gaussian processes: if $X_i, i \leq n$ are a jointly Gaussian family of centered random variables and $G_{i,j} := \mathbb{E}(X_i X_j)$, then the following identities hold (and actually they characterize Gaussian processes):

$$\mathbb{E}(\prod_{i \leq 2k} X_{j_i}) = \sum_{\sigma \in M[k]} \prod_{i \leq k} G_{j_{\sigma(i)}, j_{\sigma(i+1)}}$$

and is instead zero if the product contains an odd number of terms. Therefore, Theorem 5.3 is saying that, at large distances, height gradients form an approximately Gaussian process.

The following result is in the same spirit (and actually is an easy consequence of the previous one) and is another way of seeing that the height field tends to the GFF:

Theorem 5.4. *Given a smooth function $\psi : \mathbb{R}^2 \mapsto \psi(x) \in \mathbb{R}$ of compact support and average zero, $\int_{\mathbb{R}^2} \psi(x)dx = 0$, define*[8] *for $\epsilon > 0$*

$$h_\epsilon(\psi) := \epsilon^2 \sum_{f \in G^*} h_M(f)\psi(\epsilon f). \tag{5.6}$$

Then, for any finite family $\{\psi_i\}_{i\leq n}$ of such functions one has the convergence in law

$$\{h_\epsilon(\psi_i)\}_{i\leq n} \Rightarrow \{h(\psi)\}_{i\leq n} \tag{5.7}$$

as $\epsilon \to 0$, where $\{h(\psi_i)\}_{i\leq n}$ is the family of centered, jointly Gaussian variables of covariance

$$\mathbb{E}(h(\psi_i)h(\psi_j)) = \int_{\mathbb{R}^2} dx \int_{\mathbb{R}^2} dx' \psi_i(x)\psi_j(x')G(x-x'), \tag{5.8}$$

$$G(x) = -\frac{1}{2\pi^2} \log|\phi(x)|. \tag{5.9}$$

See later in this section for the proof (we will give details mostly for the covariance and just a few hints for higher moments).

Roughly speaking, Theorem 5.4 says that the random variable $h_\epsilon(\psi)$ tends to the integral

$$\int_{\mathbb{R}^2} \psi(x)X(x)dx,$$

with $X(\cdot)$ the centered Gaussian random function on $\mathbb{R}^2$ with covariance $G(\cdot)$. We say "roughly speaking" because $X(\cdot)$ is not really a function: since $G(0) = +\infty$, the random variable $X(x)$ is almost surely infinite for every x. Actually, one has to view $X(\cdot)$ not as a random function but as a random distribution. In fact, once it is integrated against a smooth function such as ψ, it gives a perfectly defined Gaussian variable $h(\psi)$ with finite variance (note that the r.h.s. of (5.7) is finite since ψ_i has finite support and the logarithmic singularity is integrable in two dimensions).

It might not be obvious that $G(\cdot)$ is actually a covariance, i.e. that for $i = j$ the r.h.s. of (5.1) is actually non-negative. Recall that

$$\phi(x) = -\beta^+ x_1 + \alpha^+ x_2 = z_1 + iz_2, \tag{5.10}$$

$$z_1 = -\Re(\beta^+)x_1 + \Re(\alpha^+)x_2, \tag{5.11}$$

$$z_2 = -\Im(\beta^+)x_1 + \Im(\alpha^+)x_2 \tag{5.12}$$

and recall that α^+, β^+ are not collinear (their ratio is not real), so that the linear map $x = Mz$, with M the 2×2 matrix defined by the above relation,

[8] In (5.5), $\psi(\epsilon f)$ can be taken to be the value of $\psi(\epsilon \cdot)$ at the center of the face f. This is not very important.

is a bijection of the plane. Rewrite the r.h.s. of (5.1) as

$$\int_{\mathbb{R}^2} dz \int_{\mathbb{R}^2} dz' \psi_1(Mz)\psi_2(Mz')G_0(z-z'), \tag{5.13}$$

$$G_0(z) = -\frac{1}{2\pi^2}\det(M)^2 \log|z|. \tag{5.14}$$

Note that $G_0(\cdot)$ is the Green's function (i.e. the inverse) of $-C\Delta$, with Δ the Laplacian on the plane and C a suitable positive constant. Since $-\Delta$ is a positive operator (its spectrum is $[0,\infty)$), its inverse (and therefore (5.1)) is also positive.

Therefore, formally, the centered Gaussian field $\hat{X}$ with covariance $G_0(\cdot)$ is the Gaussian field with probability density proportional to

$$e^{\frac{C}{2}\int dx \hat{X}(x)[\Delta\hat{X}](x)} d\hat{X} = e^{-\frac{C}{2}\int dx |\nabla \hat{X}(x)|^2} d\hat{X}. \tag{5.15}$$

The field $\hat{X}$ is called "massless" because there is no "mass-term $-m\hat{X}^2(x)$, $m>0$, in the exponent. The measure (5.14) makes no mathematical sense literally speaking, since the Lebesgue measure $d\hat{X}$ is not well defined, $\hat{X}$ is not differentiable and it is not a function anyway. The rigorous version of (5.14) is to view $\hat{X}$ as a random Gaussian distribution, as mentioned above. We refer to [26] for a a mathematician's introduction to the GFF.

Proof of Theorem 5.1. Recall that

$$h_M(f') - h_M(f) = \sum_{e \in \mathcal{C}_{f\to f'}} \sigma_e (1_{e\in M} - c(e)) \tag{5.16}$$

where $C_{f\to f'}$ is any nearest-neighbor path from f to f'. Therefore, one has

$$\begin{aligned}&\mathrm{Var}_\Pi(h_M(f) - h_M(f')) \\ &\quad = \sum_{e\in\mathcal{C}^{(1)}_{f\to f'}} \sum_{e'\in\mathcal{C}^{(2)}_{f\to f'}} \sigma_e\sigma_{e'}[\Pi(e,e'\in M) - \Pi(e\in M)\Pi(e'\in M)].\end{aligned} \tag{5.17}$$

In principle, the two paths $C^{(1,2)}_{f\to f'}$ can be the same, but we will see that it is convenient to choose them to be different. Next, we use (3.61) to express the dimer-dimer correlation and we are left with

$$\begin{aligned}&\mathrm{Var}_\Pi(h_M(f) - h_M(f')) \\ &= -\sum_{e\in\mathcal{C}^{(1)}_{f\to f'}} \sum_{e'\in\mathcal{C}^{(2)}_{f\to f'}} \sigma_e\sigma_{e'} K(e)K(e')K^{-1}(x_e - y_{e'})K^{-1}(x_{e'} - y_e),\end{aligned} \tag{5.18}$$

where x_e/y_e are the coordinates of the white/black vertex of e. It is out of question to perform the sum exactly, therefore the only hope is to be able to use simply the asymptotic behavior of K^{-1} at large distances, provided by Proposition 3.15. For this purpose, it is convenient to choose two paths $C^{(1,2)}_{f\to f'}$ that stay far from each other (except of course at the endpoints f, f' where they necessarily meet). More precisely, we will require that:

- the lengths of $\mathcal{C}^{(1,2)}_{f\to f'}$ are $O(|f-f'|)$;
- for some $c>0$, in a neighboorhood of size $c|f-f'|$ of f and of f' the two paths are straight and have distinct directions;

- outside these two neighborhoods, the two paths are at distance at least $c'|f-f'|$ for some $c'>0$.

See Fig. 17. Recall now the asymptotics of Proposition 3.15. First of all,

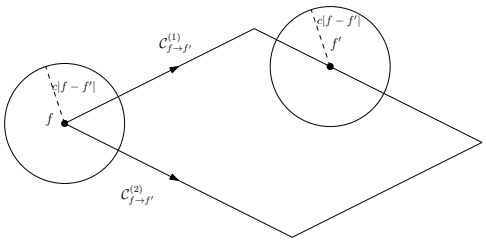

FIGURE 17. A schematic view of the paths $\mathcal{C}^{(1,2)}_{f\to f'}$

the contribution from the $O(|x|^{-2})$ term in (3.57) gives overall a quantity $O(1)$, uniformly in $|f-f'|$, when summed over e,e'.

Exercise 5.5. *Prove this.*

Therefore, we are left with

$$\begin{aligned}&\mathrm{Var}_\Pi(h_M(f)-h_M(f'))\\&=-\frac{1}{\pi^2}\sum_{e\in\mathcal{C}^{(1)}_{f\to f'}}\sum_{e'\in\mathcal{C}^{(2)}_{f\to f'}}\sigma_e\sigma_{e'}K(e)K(e')\Re\left[\frac{e^{-ip^+(x_e-y_{e'})}}{\phi(x_e-y_{e'})}\right]\Re\left[\frac{e^{-ip^+(x_{e'}-y_e)}}{\phi(x_{e'}-y_e)}\right]+O(1).\end{aligned}\tag{5.19}$$

Note first of all that

$$x_e=y_e+v_e,\tag{5.20}$$

where $v_e=(0,0)$ if the edge e is horizontal, $v_e=(-1,0)$ if it is north-east oriented and $v_e=(0,1)$ if it is north-west oriented. Then,

$$\begin{aligned}&-\frac{1}{\pi^2}\Re\left[\frac{e^{-ip^+(x_e-y_{e'})}}{\phi(x_e-y_{e'})}\right]\Re\left[\frac{e^{-ip^+(x_{e'}-y_e)}}{\phi(x_{e'}-y_e)}\right]\\&=\frac{1}{2\pi^2}\Re\left[\frac{e^{-ip^+(v_e+v_{e'})}}{\phi(x_e-x_{e'})^2}\right]+\frac{1}{2\pi^2}\Re\left[\frac{e^{-ip^+(x_e+y_e-x_{e'}-y_{e'})}}{|\phi(x_e-x_{e'})|^2}\right]+O(|x_e-x_{e'}|^{-3})\end{aligned}\tag{5.21}$$

where the error term comes from approximations of the type $\phi(x_e-y_{e'})=\phi(x_e-x_{e'})+O(1)$. Again, the error term can be neglected in the sum over e,e'. Next, note the following crucial identity:

$$\sigma_eK(e)e^{-ip^+v_e}=i\Delta_e\phi,\tag{5.22}$$

where $\Delta_e\phi$ is the change of ϕ (with ϕ as defined in (3.58)) when the edge e is crossed. To check (5.21), it is enough to recall the definition of v_e and of α^+,β^+ (cf. (3.52)) as well as the relation

$$A+Be^{-ip_2^+}+Ce^{ip_1^+}=0$$

that defines p^+. Then, the first term in the r.h.s. of (5.20), once summed over e, e', gives

$$-\sum_{e\in\mathcal{C}^{(1)}_{f\to f'}}\sum_{e'\in\mathcal{C}^{(2)}_{f\to f'}}\frac{1}{2\pi^2}\Re\left[\frac{\Delta_e\phi\Delta_{e'}\phi}{\phi(x_e-x_{e'})^2}\right]$$
$$=-\frac{1}{2\pi^2}\Re\int_{\phi(f)}^{\phi(f')}dz\int_{\phi(\tilde f)}^{\phi(\tilde f')}dz'\frac{1}{(z-z')^2}+O(1) \quad (5.23)$$

where:

- the integral performed on the complex plane $\mathbb{C}$;
- the integral over z runs over a contour from $\phi(f)$ to $\phi(f')$ (essentially this is the path obtained from (a continuous version of) the path $\mathcal{C}^{(1)}_{f\to f'}$, deformed through the application of the map $\phi(\cdot)$);
- the integral over z' runs over a contour from $\phi(\tilde f)$ to $\phi(\tilde f')$, with $\tilde f, \tilde f'$ two points at distance $O(1)$ from f, f';
- the two contours do not cross.

The reason why we take $f\neq\tilde f$ and $f'\neq\tilde f'$ is that otherwise the integral would be divergent, while the sum in the l.h.s. is not (for $x=0$, one can replace (3.57) by $K^{-1}(0)=O(1)$).

The integral 5.22 can be done exactly and the result is

$$\frac{1}{2\pi^2}\Re\log\left(\frac{(\phi(\tilde f')-\phi(f))(\phi(\tilde f)-\phi(f'))}{(\phi(\tilde f')-\phi(f'))(\phi(\tilde f)-\phi(f))}\right)=\frac{1}{\pi^2}\log|\phi(f)-\phi(f')|+O(1)$$
$$=\frac{1}{\pi^2}\log|f-f'|+O(1). \quad (5.24)$$

Finally, it remains only to prove that

$$\frac{1}{2\pi^2}\sum_{e\in\mathcal{C}^{(1)}_{f\to f'}}\sum_{e'\in\mathcal{C}^{(2)}_{f\to f'}}\Re\left[\frac{e^{-ip^+(x_e+y_e-x_{e'}-y_{e'})}}{|\phi(x_e-x_{e'})|^2}\right]=O(1). \quad (5.25)$$

Suppose (without loss of generality) that the paths $\mathcal{C}^{(1,2)}_{f\to f'}$ have been chosen so that they are made of a finite number of portions, each of them being a straight path of length $O(|f-f'|)$, in one of the three lattice directions (as is the case in Fig. 17). When, say, e runs over one such straight portion, the oscillating factor

$$e^{-ip^+(x_e+y_e)}$$

equals

$$e^{-2i(np^+\cdot w+const.)},$$

where n labels the edge e along the path and $w=(1,0)$ if the path is in direction $\hat e_1$, $w=(0,1)$ if it is in direction $\hat e_2$ and $w=(-1,-1)$ if it is vertical upward. In all cases, note from (3.19) that $p^+\cdot w$ is not a multiple of π (because A,B,C are assumed to satisfy strictly the triangle inequality), so that there is a non-trivial oscillation and, as a consequence,

$$r(N):=\sum_{n=1}^{N}e^{-2inp^+\cdot w} \quad (5.26)$$

is bounded uniformly in N by some constant K. Since $e^{-inp^+\cdot w}$ is the discrete gradient of $r(\cdot)$, one can perform a discrete summation by parts and transfer the discrete gradient (both for e and for e') to the function $1/|\phi(x_e - x_{e'})|^2$. Note that applying the discrete gradient twice to $1/|\phi(x_e - x_{e'})|^2$ gives a quantity that decays like $|x_e - x_{e'}|^{-4}$ at large distances. Altogether, we obtain that (5.24) is upper bounded by

$$const. \times \sum_{e \in \mathcal{C}^{(1)}_{f \to f'}} \sum_{e' \in \mathcal{C}^{(2)}_{f \to f'}} \frac{1}{|x_e - x_{e'}|^4} = O(1) \tag{5.27}$$

Exercise 5.6. *Fill in the details.*

■

Proof of Theorem 5.3. The proof of (5.3) is essentially identical to that of (5.1), so we give no further details.

As for (5.4) with $n > 2$, let us sketch the main idea. The analog of (5.17) is the following:

$$\begin{aligned} \Pi \left[\prod_{i=1}^{n} (h_M(f_i) - h_M(f_i')) \right] \\ = \sum_{e_1 \in \mathcal{C}_{f_1 \to f_1'}} \cdots \sum_{e_n \in \mathcal{C}_{f_n \to f_n'}} \sigma_{e_1} \dots \sigma_{e_n} K(e_1) \dots K(e_n) \\ \times \sum_{\rho \in \tilde{S}_n} (-1)^\rho \prod_{j=1}^{n} K^{-1}(x_i - y_{\rho(i)}), \end{aligned} \tag{5.28}$$

where:

- x_i/y_i are, as usual, the coordinates of the white/black vertex of e_i;
- the sum over ρ runs over all permutations of $\{1, \dots, n\}$ without fixed points[9] (i.e. such that $\rho(i) \neq i$ for every i);
- $(-1)^\rho$ is the signature of the permutation ρ.

For each occurrence of K^{-1} we use the decomposition

$$K^{-1}(x) = A(x) + \bar{A}(x) + R(x) \tag{5.29}$$

provided by Proposition 3.15, with $A(x) = e^{-ip^+ x}/(2\pi\phi(x))$ and $R(x)$ the error term $O(|x|^{-2})$. When we expand the product $\prod_{j=1}^{n} K^{-1}(x_i - y_{\rho(i)})$, we have three types of terms:

(1) those containing n terms of type $A(\cdot)$ or n terms of type $\bar{A}(\cdot)$;
(2) those containing both terms of type $A(\cdot)$ and of type $\bar{A}(\cdot)$ and no error term $R(\cdot)$;
(3) those containing at least one error term of type $R(\cdot)$.

The latter two types, once summed over $e_1, \dots, e_n$, contribute only to the error term $O(1/D(\dots))$ in (5.4). The argument to see this is the same as we gave for $n = 2$: the terms containing $R(\cdot)$ are small, while those

[9] The condition that the permutation ρ has no fixed point comes from the fact that the height gradients $h_M(f_i) - h_M(f_i')$ in (5.4) are centered. For the same reason, in (5.17), the term $K^{-1}(x_e - y_e)K^{-1}(x_{e'} - y_{e'})$ does not appear.

containing both A and $\bar{A}$ give a small contribution due to oscillations (recall the argument we used to estimate (5.24)).

It remains to look at the dominant terms, i.e. those of type (1). In turn, as far as the the permutation ρ is concerned, one should distinguish two cases:

(1) the permutation ρ is composed only of cycles of length 2 (there are no such permutations if n is odd);
(2) the permutation ρ contains at least a cycle of length ≥ 3.

Note that the permutations with no fixed points and only 2-cycles are in bijection with matchings of $\{1, \dots, n\}$ in (5.4). For instance, for $n = 4$, the permutation $\pi(1) = 4, \pi(2) = 3, \pi(3) = 2, \pi(4) = 1$ corresponds to the matching $(1, 4), (2, 3)$. Also, the signature of such permutations is constant and equals $(-1)^{n/2}$. Then, is not hard to realize that the permutations of type (1) produce (once one sums over the edges $e_1, \dots, e_n$) the r.h.s. of (5.4). The argument is essentially the same as for (5.22); the sign $(-1)^{n/2}$ compensate with a $(i^2)^{n/2}$, with each i coming from a $i\Delta_e$ as in (5.21)

As for the permutations with at least a cycle of length ≥ 3, a different argument is needed to see that their contribution is sub-dominant. The crucial point is the following combinatorial identity ([12, Eq. D.29]): if S_ℓ denotes the set of cyclic permutations of $\{1, \dots, \ell\}$ and $\ell \geq 3$, then

$$\sum_{\rho \in S_\ell} \prod_{i=1}^{\ell} \frac{1}{\phi(x_i - x_{\rho(i)})} = 0. \tag{5.30}$$

■

Proof of Theorem 5.4. Let us show first that the covariance of $h^\epsilon(\psi_1), h^\epsilon(\psi_2)$ is given as in (5.3) when $\epsilon \to 0$. Let g_1, g_2 be two faces of the graph, outside of the support of $\psi_i(\epsilon \cdot), i = 1, 2$, and at mutual distance of order $1/\epsilon$. We start by writing

$$\begin{aligned}\pi[h^\epsilon(\psi_1)h^\epsilon(\psi_2)] = \epsilon^4 \sum_{f,f'} \psi_1(\epsilon f)\psi_2(\epsilon f')\pi[(h_M(f) - h_M(g_1))h_M(f')] \\ + \epsilon^4 \sum_{f,f'} \psi_1(\epsilon f)\psi_2(\epsilon f')\pi[h_M(g_1)h_M(f')].\end{aligned} \tag{5.31}$$

Let us show first of all that the second line is $o(1)$. In fact, since ψ_1 is smooth and its integral is zero, we have that

$$\epsilon^2 \sum_f \psi_1(\epsilon f) = O(\epsilon) \tag{5.32}$$

and on the other hand $\pi[h_M(g_1)h_M(f')] = O(\log \epsilon^{-1})$ thanks to Theorem 5.1 and to Cauchy-Schwarz. Similarly, one can rewrite the covariance as

$$\begin{aligned}&\pi[h^\epsilon(\psi_1)h^\epsilon(\psi_2)] \\ &= \epsilon^4 \sum_{f,f'} \psi_1(\epsilon f)\psi_2(\epsilon f')\pi[(h_M(f) - h_M(g_1))(h_M(f') - h_M(g_2))] + o(1).\end{aligned} \tag{5.33}$$

At this point we can use Theorem 5.3, observing that $D(f_1, g_1, f_2, g_2) \geq c|f - f'|$ for some $c > 0$. I.e.,

$$\pi[h^\epsilon(\psi_1)h^\epsilon(\psi_2)] = \frac{\epsilon^4}{2\pi^2}\sum_{f,f'}\psi_1(\epsilon f)\psi_2(\epsilon f')\Re\log\left(\frac{(\phi(f')-\phi(g_1))(\phi(g_2)-\phi(f))}{(\phi(f')-\phi(f))(\phi(g_2)-\phi(g_1))}\right) + O\left(\|\psi_1\|_\infty\|\psi_2\|_\infty\epsilon^4\sum_{f\in S_1, f'\in S_2}\frac{1}{|f-f'|+1}\right), \quad (5.34)$$

where S_i is the support of the function $\psi_i(\epsilon\cdot)$. Since S_i has diameter $O(\epsilon^{-1})$, one sees easily that the latter sum is $O(\epsilon)$. As for the former sum, using (5.31) we are left with

$$\begin{aligned}\pi[h^\epsilon(\psi_1)h^\epsilon(\psi_2)] &= -\frac{\epsilon^4}{2\pi^2}\sum_{f,f'}\psi_1(\epsilon f)\psi_2(\epsilon f')\Re\log(\phi(f')-\phi(f)) + o(1)\\ &= -\frac{\epsilon^4}{2\pi^2}\sum_{f,f'}\psi_1(\epsilon f)\psi_2(\epsilon f')\Re\log(\epsilon\phi(f')-\epsilon\phi(f)) + o(1)\\ &= \frac{1}{2\pi^2}\int_{\mathbb{R}^2}dx\int_{\mathbb{R}^2}dx'\psi_1(x)\psi_2(x')G(x-x') + o(1). \quad (5.35)\end{aligned}$$

For higher moments, the proof is similar except that one uses (5.4) instead of (5.3).

Exercise 5.7. *Fill details.*

■

6. Related models: Ising, 6-vertex and spanning tree

In this section we show how two classical models of statistical mechanics, namely the Ising model and the six-vertex model, can be rewritten in terms of the dimer model. For the Ising model on a planar graph (for instance, the square grid) one gets the dimer model on a suitable planar, weighted, "decorated graph". The partition function can then be computed by Kasteleyn theory as explained in Section 2 and one recovers the celebrated Onsager solution for the free energy of the Ising model.

As for the six-vertex model, instead, one ends up with a model of "interacting dimers" that admits no simple solution any more. The mapping is anyway rich in consequences, that we will only hint to.

6.1. Ising and dimers.

Let Γ be a finite, planar graph. The Ising model with zero magnetic field on Γ is defined by assigning to every vertex i of Γ a spin variable $\sigma_i \in \{-1, +1\}$ and associating to the configuration $\sigma = \{\sigma_i\}_{i\in V(\Gamma)}$ the weight

$$\pi_\Gamma(\sigma) = \frac{e^{\sum_{(i,j)\in E(\Gamma)} J_{ij}(\sigma_i\sigma_j - 1)}}{Z_\Gamma^{Ising}}, \quad (6.1)$$

where:

- $E(\Gamma), V(\Gamma)$ denote the set of edges/vertices of Γ;

- $J_{ij} \geq 0$ is the positive "coupling" associated to edge (i, j) (the inverse temperature β has been absorbed into J_{ij});
- the partition function Z_Γ^{Ising} is

$$Z_\Gamma^{Ising} = \sum_{\sigma \in \{-1,+1\}^{V(\Gamma)}} e^{\sum_{(i,j)\in E(\Gamma)} J_{ij}(\sigma_i \sigma_j - 1)}.$$

The constant -1 in the exponent has been added for convenience; it does not modify the probability measure π_Γ.

Here we present a trick due to M. E. Fisher [11] that allows to rewrite Z_Γ^{Ising} as the partition function of a dimer model on a "decorated" graph G, that contains more vertices and edges than Γ. The point is that if Γ is planar then G is still planar, so that Kasteleyn theory can be applied.

The Ising-dimer mapping for the general planar case is explained in [11] (see also Remark 6.3 below); no spatial periodicity of the graph is assumed there. For simplicity of exposition, we will suppose here that Γ is a finite portion of the hexagonal lattice, that we draw in a sort of distorted way as in Fig. 18, for later convenience. We assign couplings J_1, J_2, J_3 to the three types of edges. Note that if, say, $J_3 \to +\infty$, then the model becomes equivalent to the Ising model on the square lattice with horizontal/vertical couplings J_1, J_2. This is because only configurations where all spins on J_3-type edges have the same sign have non-zero weight in the limit $J_3 \to +\infty$. Note also that vertices in the interior of the graph have degree 3, while vertices at the boundary have degree 1 or 2.

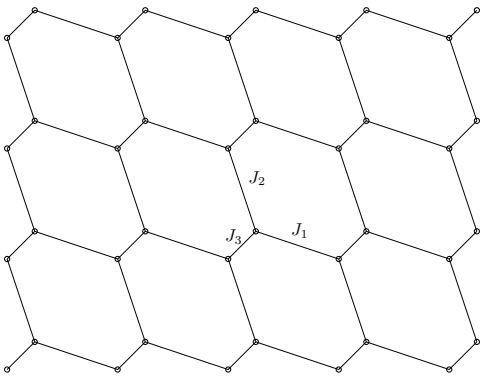

FIGURE 18. The planar graph Γ

The "decorated lattice" G is obtained by splitting every vertex of Γ as follows (see Fig. 19):

- vertices of degree 1 are left unchanged;
- vertices of degree 2 are split into a double vertex connected by a new edge;
- vertices of degree 3 are replaced by 3 vertices connected by new edges forming a triangle.

The newly added edges will be called "internal edges", while the edges already present in Γ are called "external edges" (the nomenclature is from [11]). (See [11] for the prescription to follow when Γ contains vertices of degree larger than 4). The resulting graph $G = G(\Gamma)$, when Γ is as in Fig.

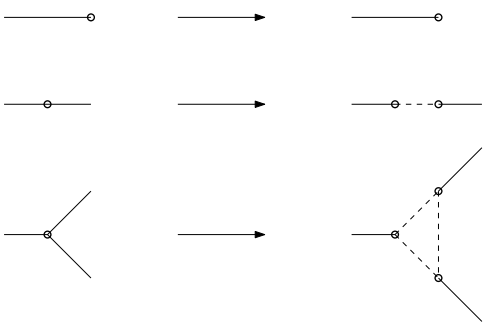

FIGURE 19. The rules of transformation of the vertices of degree $1, 2, 3$. Internal edges are drawn as dashed lines, external edges as solid lines.

18, is given in Fig. 20.

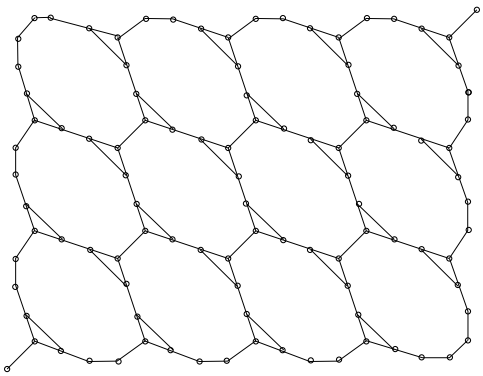

FIGURE 20. The planar graph G obtained from Γ. Both internal and external edges are drawn as solid lines here.

Given an external edge (i, j), let

$$v_{ij} := \tanh(J_{ij}).$$

Then, the following identity holds:

Theorem 6.1.

$$Z_{\Gamma}^{Ising} = 2^{|V(\Gamma)|} \left[\prod_{(i,j)\in E(\Gamma)} v_{ij}(1 + v_{ij})^{-1} \right] \times Z_{G}^{Dimer}, \tag{6.2}$$

where Z_G^{Dimer} is the partition function of the dimer model on G, with edge weights 1 on internal edges and $1/v_{ij}$ on external edges.

Proof. We start with the so-called "high-temperature expansion" of the Ising model. Namely, write

$$e^{J_{ij}(\sigma_i\sigma_j - 1)} = \cosh(J_{ij})e^{-J_{ij}}(1 + v_{ij}\sigma_i\sigma_j) \tag{6.3}$$

and apply such identity to all terms appearing in the Boltzmann weight (6.1). Then,

$$Z_\Gamma^{Ising} = \left[\prod_{(i,j)\in E(\Gamma)} \cosh(J_{ij}) e^{-J_{ij}}\right] \sum_\sigma \prod_{(i,j)\in E(\Gamma)} (1 + v_{ij}\sigma_i\sigma_j). \tag{6.4}$$

Then, expand the last product, choosing for every $(i,j) \in E(\Gamma)$ either 1 or $v_{ij}\sigma_i\sigma_j$. For each term of the expansion, mark the edge (i,j) if $v_{ij}\sigma_i\sigma_j$ has been chosen. When we sum over σ, since

$$\sum_{\sigma_i=\pm 1} \sigma_i^r = \begin{cases} 2 & \text{if} \quad r \text{ is even} \\ 0 & \text{if} \quad r \text{ is odd,} \end{cases} \tag{6.5}$$

we see that the non-zero terms in the expansions are those such that for every vertex $i \in V(\Gamma)$, there is an even number of marked edges containing it. Altogether, we see that

$$Z_\Gamma^{Ising} = \left[\prod_{(i,j)\in E(\Gamma)} \cosh(J_{ij}) e^{-J_{ij}}\right] \times \Upsilon_\Gamma$$

$$\Upsilon_\Gamma = \sum_{P\in\mathcal{P}(\Gamma)} \prod_{(i,j)\in P} v_{ij} \tag{6.6}$$

where $\mathcal{P}(\Gamma)$ is the set of all polygon configurations made of edges of Γ, such that at each vertex i an even number of sides meet.

At this point, the idea is to find a one-to-one mapping between polygons P on Γ and dimer configurations on G. Namely, we put a dimer on an external edge (i,j) of G iff the corresponding edge of Γ is *not* in P. Then, there is a unique way of putting dimers on the internal edges of G such that the resulting configuration is indeed a dimer covering of G. See Fig. 21. Since the edge weights of the dimer model on G are 1 on the internal edges and $1/v_{ij}$ on external edges, we see that the weight of a dimer configuration M on G is

$$\prod_{(i,j)\in M,(i,j) \text{ external}} \frac{1}{v_{ij}} = \frac{1}{\prod_{(i,j) \text{ external}} v_{ij}} \times \prod_{(i,j)\notin M,(i,j) \text{ external}} v_{ij}$$

$$= \frac{1}{\prod_{(i,j)\in E(\Gamma)} v_{ij}} \times \prod_{(i,j)\in E(\Gamma):(i,j)\in P} v_{ij}. \tag{6.7}$$

Therefore,

$$\Upsilon_\Gamma = Z_G^{Dimer} \times \prod_{(i,j)\in E(\Gamma)} v_{ij} \tag{6.8}$$

and the claim of the Theorem follows, observing that $e^{-J_{ij}} \sinh(J_{ij}) = v_{ij}/(1+v_{ij})$. ∎

Given Theorem 6.1, the computation of the infinite-volume free energy of the Ising model is done as in Sections 3.2 and 3.5, except that we have Pfaffians instead of determinants since the graph is non-bipartite. We give just the result and skip details. Note that, when the graph Γ of Fig. 18 tends to the infinite hexagonal graph, the decorated graph G tends to an infinite,

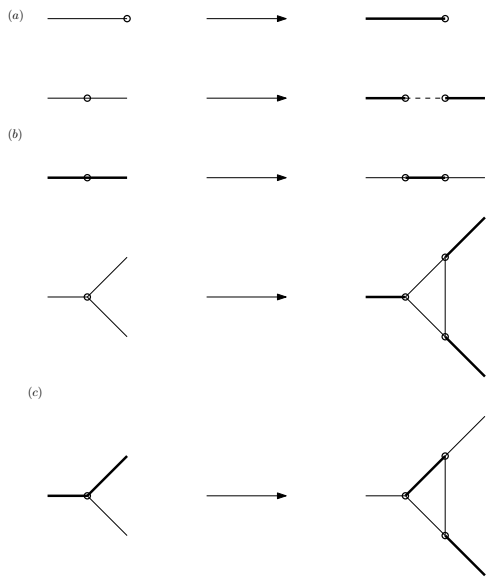

FIGURE 21. The local configurations of a polygon P (left) and of the dimer configuration (right). In drawing (a), the case of a degree-one edge, where there is no choice: it cannot belong to any polygon and the incident edge must be covered by a dimer. In drawing (b), the case of a degree-two vertex, where 2 cases are possible. In drawing (c) the case of a degree-three vertex: either it belongs to no polygon edge (first case) or to two of them (one of the three possibilities is given here)

periodic graph whose fundamental domain contains six vertices. See Fig. 22, where a Kasteleyn (clock-wise odd) orientation is given. The subsequent

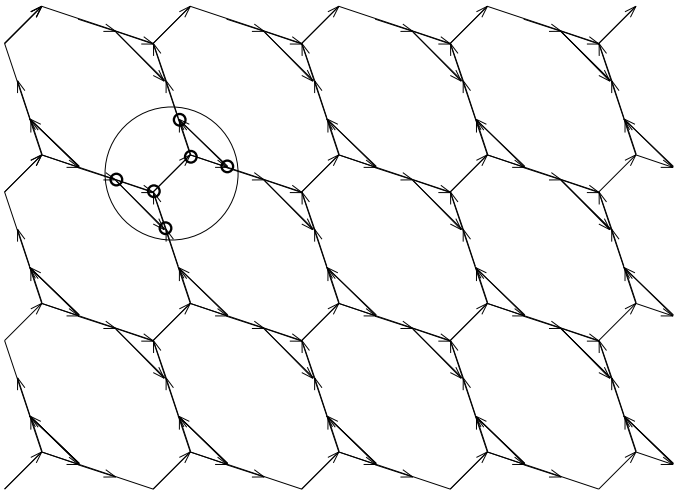

FIGURE 22. A finite portion of the infinite graph G obtained by decorating the infinite hexagonal graph, with its Kasteleyn orientation and its fundamental domain.

steps in the computation of the infinite-volume free energy are:

- periodizing the graph on a torus of size $L \times L$ to get a finite graph G_L;
- expressing the partition function as a sum of four Pfaffians;
- using discrete Fourier analysis as in Section 3.5 to reduce the four Kasteleyn matrices to a block-diagonal form, with 6×6 blocks;
- computing the determinant of each matrix, which turns out to be the product over the Fourier mode k of $\mu(k)$, the determinant of the 6×6 block;
- computing the four Pfaffians via the identity $\mathrm{Pf}(A)^2 = \det A$;
- taking the $L \to \infty$ limit of $(1/L^2) \log Z_{G_L}^{Dimer}$, whereby the sum over Fourier modes becomes an integral over $k \in [-\pi, \pi]^2$.

The final result is that the infinite-volume free energy of the Ising model on the hexagonal graph, with couplings J_1, J_2, J_3, is

$$\begin{aligned} F(J_1, J_2, J_3) &= 2\log 2 - \log(1+v_1)(1+v_2)(1+v_3) \\ &\quad + \frac{1}{2(2\pi)^2} \int_{[-\pi,\pi]^2} dk_1 dk_2 \log D(k_1, k_2, k_3), \\ D(k_1, k_2, k_3) &= 1 + v_1^2 v_2^2 + v_2^2 v_3^2 + v_3^2 v_1^2 - 2(1-v_1^2) v_2 v_3 \cos(k_1) \\ &\quad - 2(1-v_2^2) v_1 v_3 \cos(k_2) - 2(1-v_3^2) v_2 v_1 \cos(k_3) \end{aligned} \tag{6.9}$$

where $v_i = \tanh(J_i)$ and k_3 is defined by $k_1 + k_2 + k_3 = 2\pi$.

In particular, putting $J_3 = +\infty$ one recovers Onsager's formula [25] for the free energy of the Ising model on the square lattice and couplings J_1, J_2:

$$F(J_1, J_2) = \log 2 - (J_1 + J_2) \tag{6.10}$$

$$+ \frac{1}{2(2\pi)^2} \int_{[-\pi,\pi]^2} dk_1 dk_2 \log \left[\cosh(2J_1) \cosh(2J_2) \right. \tag{6.11}$$

$$\left. - \sinh(2J_1) \cos(k_1) - \sinh(2J_2) \cos(k_2) \right]. \tag{6.12}$$

From this expression one can recover the relation between J_1 and J_2 that determines the critical point, where F fails to be analytic.

Remark 6.2. *The Ising-to-dimers mapping given above fails if there is a magnetic field h, i.e. the Boltzmann weight* (6.1) *is modified by the exponential of h times the sum of σ_i for $i \in V(\Gamma)$.*

Remark 6.3. *Suppose that the planar graph Γ where the Ising model is defined contains vertices of degree ≥ 4. Then, one first changes Γ to $\tilde{\Gamma}$, where degrees are all ≤ 3. This is done by suitably "splitting vertices", as in Figure 23, and assinging a coupling J^* to the auxiliary edges: in the limit $J^* \to +\infty$, the partition function tends to that of the Ising model on the original graph.*

At this point, one applies the above Ising-to-dimer mapping to the Ising model on the graph $\tilde{\Gamma}$.

6.2. **6-vertex model and dimers.** TO BE DONE

6.3. **Temperley's bijection and spanning trees.** TO BE DONE

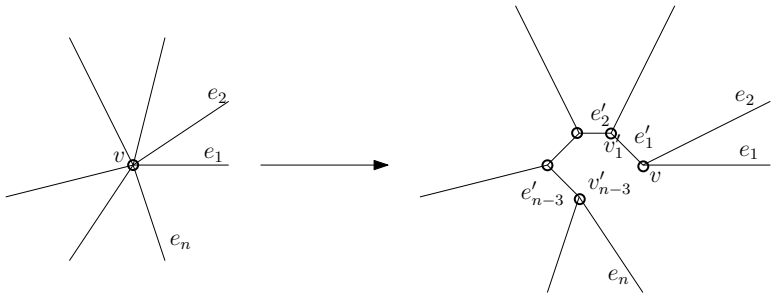

FIGURE 23. A vertex v of degree $n > 3$ is split into $n-2$ vertices, each of degree 3, and in the process $n-3$ "internal" edges $e'_1, \dots, e'_{n-3}$ are added.

APPENDIX A. A FEW TECHNICAL ESTIMATES

A.1. **Proof of Lemma 3.7.** Let Q_k the square of side $2\pi/L$ centered at k and with sides parallel to the horizontal/vertical axes of $\mathbb{R}^2$. Write

$$\left| \frac{1}{L^2} \sum_{k \neq k^{\pm}_{\theta,\tau}} \log|\mu(k)| - \frac{1}{(2\pi)^2} \sum_{k \neq k^{\pm}_{\theta,\tau}} \int_{Q_k} \log|\mu(q)| dq_1 dq_2 \right| \leq \frac{c}{L^3} \sum_{k \neq k^{\pm}_{\theta,\tau}} \max_{q \in Q_k} |\nabla_q \log|\mu(q)||. \quad \text{(A.1)}$$

Since $k = k^{\pm}_{\theta,\tau}$ is excluded from the sum and $\mu(\cdot)$ vanishes linearly at $p^{\pm}$, we can bound

$$\max_{q \in Q_k} |\nabla_q \log|\mu(q)|| \leq \frac{c'}{\min(\|k - p^+\|, \|k - p^-\|)}. \quad \text{(A.2)}$$

Note that this would be false for $k = k^{\pm}_{\theta,\tau}$, as the l.h.s. would be simply $+\infty$. It is standard to deduce that the r.h.s. of (A.1) is $O(1/L)$, because

$$\int_{\epsilon \leq |x| \leq 1} \frac{1}{|x|} dx_1 dx_2 \overset{\epsilon \to 0}{=} O(1). \quad \text{(A.3)}$$

Finally, one has

$$\int_{Q_{k^{\pm}_{\theta,\tau}}} \log|\mu(q)| dq_1 dq_2 \overset{L \to \infty}{=} o(1) \quad \text{(A.4)}$$

because $\log(\cdot)$ is integrable. Altogether, we have proven that

$$\left| \frac{1}{L^2} \sum_{k \neq k^{\pm}_{\theta,\tau}} \log|\mu(k)| - \frac{1}{(2\pi)^2} \int_{[-\pi,\pi]^2} \log|\mu(q)| dq_1 dq_2 \right| = o(1). \quad \text{(A.5)}$$

A.2. **Proof of** (3.33)**.** This is very similar to the proof of Lemma 3.7. One has

$$\left| \frac{1}{L^2} \sum_{k \neq k^{\pm}_{\theta,\tau}} \frac{e^{i(y-x)k}}{\mu(k)} - \frac{1}{(2\pi)^2} \sum_{k \neq k^{\pm}_{\theta,\tau}} \int_{Q_k} \frac{e^{i(y-x)q}}{\mu(q)} dq_1 dq_2 \right| \leq \frac{c}{L^3} \sum_{k \neq k^{\pm}_{\theta,\tau}} \max_{q \in Q_k} \left| \nabla_q \left\{ e^{i(y-x)q}/\mu(q) \right\} \right|. \quad \text{(A.6)}$$

The r.h.s. is $O(\log L/L)$ because

$$\int_{\epsilon \leq |x| \leq 1} \frac{1}{|x|^2} dx_1 dx_2 \stackrel{\epsilon \to 0}{=} O(|\log \epsilon|). \quad \text{(A.7)}$$

Also,

$$\int_{Q_{k^{\pm}_{\theta,\tau}}} \frac{e^{i(y-x)q}}{\mu(q)} dq_1 dq_2 \stackrel{L \to \infty}{=} o(1) \quad \text{(A.8)}$$

because $1/q$ is integrable close to zero, in two dimensions. The claim then follows.

A.3. **Proof of Lemma 3.9.** It is sufficient to prove the claim when $(\theta, \tau) - (\theta', \tau')$ equals either $(1, 0)$ or $(0, 1)$ and, for definiteness, assume we are in the former case. Also, without loss of generality, assume that $|k^{\pm}_{\theta,\tau} - p^{\pm}| \leq |k^{\pm}_{\theta',\tau'} - p^{\pm}|$. As was noted in Remark 3.6,

$$|k^{\pm}_{\theta',\tau'} - p^{\pm}| \geq \frac{\pi}{2L}, \quad \text{(A.9)}$$

while $|k^{\pm}_{\theta,\tau} - p^{\pm}|$ can be much smaller, possibly zero. Write the ratio of determinants as

$$\exp \left\{ \sum_{k \in \mathcal{D}_{\theta\tau}} \left(\log \mu(k) - \frac{1}{2} \log \mu(k^{\leftarrow}) - \frac{1}{2} \log \mu(k^{\rightarrow}) \right) \right\}, \quad \text{(A.10)}$$

with $k^{\rightarrow} = k + (\pi/L, 0) \in \mathcal{D}_{\theta',\tau'}$ and $k^{\leftarrow} = k - (\pi/L, 0) \in \mathcal{D}_{\theta',\tau'}$. Decompose $\mathcal{D}_{\theta,\tau}$ as the disjoint union $A \cup B$, with A containing the values of k at distance at most, say, $10/L$ from $p^{\pm}$, and B all the others. The cardinality of A is uniformly bounded as a function of L.

Note that for all $k \in A$, $|\mu(k^{\leftarrow})|$ and $|\mu(k^{\rightarrow})|$ are upper and lower bounded by positive constants times $1/L$, because μ vanishes linearly at $p^{\pm}$ and the values of $k^{\rightarrow}, k^{\leftarrow}$ are at distance of order $1/L$ from $p^{\pm}$ (cf. (A.9)). The same holds for $|\mu(k)|, k \in A$, except possibly for $k = k^{\pm}_{\theta,\tau}$. One has then

$$c_1 \leq \left| \frac{\mu(k^{+}_{\theta',\tau'}) \mu(k^{-}_{\theta',\tau'})}{\mu(k^{+}_{\theta,\tau}) \mu(k^{-}_{\theta,\tau})} \exp \left\{ \sum_{k \in A} \left(\log \mu(k) - \frac{1}{2} \log \mu(k^{\leftarrow}) - \frac{1}{2} \log \mu(k^{\rightarrow}) \right) \right\} \right| \leq c_2. \quad \text{(A.11)}$$

It remains to prove that (A.10), with k restricted to B, is upper and lower bounded (in absolute value) by L-independent positive constants. Write

$$\log \mu(k) - \frac{1}{2}\log \mu(k^{\leftarrow}) - \frac{1}{2}\log \mu(k^{\rightarrow}) \tag{A.12}$$

$$= -\frac{\pi^2}{L^2}\partial_{k_1}^2 \log \mu(k) - \frac{\pi^3}{6L^3}\partial_{k_1}^3 \log \mu(k)|_{k=k'} \tag{A.13}$$

where k' is a point in the segment joining $k^{\leftarrow}$ and $k^{\rightarrow}$. Since $\mu(\cdot)$ vanishes linearly at $p^{\pm}$,

$$|\partial_{k_1}^3 \log \mu(k')| = O((\min(|k-p^+|, |k-p^-|)^{-3}).$$

Here it is important that $k \in B$, since this means that $\partial_{k_1}^3 \log \mu(k')$, computed in the unknown point k', can be safely replacd by the derivative computed at k. Therefore,

$$\frac{1}{L^3}\sum_{k\in B} \partial_{k_1}^3 \log \mu(k') = O(1). \tag{A.14}$$

The sum of the term involving $\partial_{k_1}^2 \log \mu(k)$ requires more care since at first sight it diverges like $\log L$. However, write

$$\frac{\pi^2}{L^2}\partial_{k_1}^2 \log \mu(k) = \frac{1}{4}\int_{Q_k} \partial_{q_1}^2 \log \mu(q) dq_1 dq_2 + O(L^{-3}|\partial_{k_1}^3 \log \mu(k)|), \tag{A.15}$$

with Q_k the square of side $2\pi/L$ centered at k. Therefore, the ratio (A.10), with k restricted to B, times the exponential of

$$+\frac{1}{4}\int_{[-\pi,\pi]^2\setminus(N^+\cup N^-)} dk\ \partial_{k_1}^2 \log \mu(k), \tag{A.16}$$

with $N^{\pm}$ the neighborhood of radius $10/L$ around $p^{\pm}$, is upper and lower bounded in absolute value by positive constants.

The integral (A.16) has a finite limit as $L \to \infty$. Indeed, recalling from (3.52) that $\mu(p^\omega + k') = \alpha^\omega k_1' + \beta^\omega k_2' + O(|k'|^2)$, the possibly singular part of the integral is proportional to

$$\int \frac{dk}{(\alpha^\omega k_1 + \beta^\omega k_2)^2} \mathbf{1}_{\{(10/L)\le|k|\le 1\}}. \tag{A.17}$$

We make the change of variables $q_1 = \omega(\alpha_1 k_1 + \beta_1 k_2), q_2 = (\alpha_2 k_1 + \beta_2 k_2)$, where α_j, β_j were defined in (3.55). The Jacobian matrix A_ω has non-zero determinant (this is because, as observed in Remark 3.16, the ratio $\alpha^\omega/\beta^\omega$ is not real). Then, the integral becomes

$$\det(A_\omega)\int \frac{dq}{(q_1 + iq_2)^2} \mathbf{1}_{\{(10/L)\le|A_\omega q|\le 1\}} \tag{A.18}$$

$$= \det(A_\omega)\int \frac{dq}{(q_1 + iq_2)^2} \mathbf{1}_{\{(10/L)\le|q|\le 1\}} + O(1) = O(1). \tag{A.19}$$

In the first equality we used the fact that the symmetric difference between the balls of radius $10/L$ for q and for $A_\omega q$ has area of order L^{-2}, while the integrand is $O(L^2)$ there; in the second step, we noted that the integral is zero, using the symmetry $(q_1, q_2) \leftrightarrow (q_2, -q_1)$.

References

[1] J. Ben Hough, M. Krishnapur, Y. Peres, *Determinantal processes and independence*, Probability Surveys **3** (2006), 206-229.

[2] C. Boutillier and B. de Tilière, *Loops statistics in the honeycomb dimer model*, Ann. Probab. **37** (2009), 1747-1777.

[3] D. Cimasoni, *The geometry of dimer models*, arxiv.org/abs/1409.4631

[4] H. Cohn, R. Kenyon, J. Propp, *A variational principle for domino tilings*, J. Amer. Math. Soc. **14** (2000), 297-346.

[5] D. Cimasoni, N. Reshetikhin, *Dimers on Surface Graphs and Spin Structures. I*, Comm. Math. Phys. **275** (2007), 187-208.

[6] D. De Silva, O. Savin, *Minimizers of convex functionals arising in random surfaces*, Duke Math. J. **151** , Number 3 (2010), 487-532.

[7] B. de Tilière, *The dimer model in statistical mechanics*, http://perso.math.u-pem.fr/de_tiliere.beatrice/Cours/polycop_Dimeres.pdf

[8] J. Dubédat, R. Gheissari, *Asymptotics of Height Change on Toroidal Temperleyan Dimer Models*, Journal of Statistical Physics **159** (2015), 75-100.

[9] B. Farb, D. Margalit, *A Primer on Mapping Class Groups*, Princeton University Press, 2011.

[10] M. E. Fisher: *Statistical mechanics of dimers on a plane lattice*, Phys. Rev. **124**, 1664-1672 (1961).

[11] M. E. Fisher, *On the Dimer Solution of Planar Ising Models*, J. Math. Phys. **7** (1966), 1776-1781

[12] J. Fröhlich, R. Götschmann, P. A. Marchetti, *The effective gauge field action of a system of non-relativistic electrons*. Comm. Math. Phys. **173** (1995), 417-452.

[13] M. E. Fisher, J. Stephenson: *Statistical Mechanics of Dimers on a Plane Lattice. II. Dimer Correlations and Monomers*, Phys. Rev. **132**, 1411-1431 (1963).

[14] P. W. Kasteleyn: *The statistics of dimers on a lattice: I. The number of dimer arrangements on a quadratic lattice*, Physica **27**, 1209-1225 (1961); and J. Math. Phys. **4**, 287-293 (1963).

[15] P. W. Kasteleyn: *Dimer statistics and phase transitions*, J. Math. Phys. **4** (1963), 287-293.

[16] P. W. Kasteleyn, *Graph theory and crystal physics*, in "Graph Theory and Theoretical Physics" pp. 43110 Academic Press, London, 1967

[17] R. Kenyon: *Local statistics of lattice dimers*, Ann. Inst. H. Poincaré, Prob Stat. **33**, 591-618 (1997).

[18] R. Kenyon: *Conformal Invariance of Domino Tiling*, Ann. Probab. **28**, 759-795 (2000).

[19] R. Kenyon: *Dominos and the Gaussian Free Field*, Ann. Probab. **29**, 1128-1137 (2001).

[20] R. Kenyon: *Lectures on dimers*, Park City Math Institute Lectures, available at arXiv:0910.3129.

[21] R. Kenyon, A. Okounkov, *Limit shapes and the complex Burgers equation*, Acta Math. **199** (2007), 263–302

[22] R. Kenyon, A. Okounkov, S. Sheffield: *Dimers and amoebae*, Ann. Math. **163**, 1019-1056 (2006).

[23] N. Kuchumov, *Limit shapes for the dimer model*, arXiv:1712.08396

[24] B. Laslier, F. L. Toninelli: *How quickly can we sample a uniform domino tiling of the $2L \times 2L$ square via Glauber dynamics?*, Prob. Theory Rel. Fields **161**, 509-559 (2015).

[25] L. Onsager, *Crystal Statistics. I. A Two-Dimensional Model with an Order-Disorder Transition*, Phys. Rev. **65** (1944), 117.

[26] S. Sheffield: *Gaussian free fields for mathematicians*, Prob. Theory Rel. Fields **139**, 521-541 (2007).

[27] S. Sheffield,

[28] A. Soshnikov, *Determinantal random point fields*, Russian Mathematical Surveys, 2000, 55(5), 923-975.

[29] N.V. Temperley, M. E. Fisher: *Dimer problem in statistical mechanics-an exact result*, Philos. Mag. **6**, 1061-1063 (1961).
[30] G. Tesler, *Matchings in graphs on non-orientable surfaces*, J. Combin. Theory Ser. B, **78** (2000), 198231.

Université de Lyon, CNRS, Institut Camille Jordan, Université Claude Bernard Lyon 1, 43 bd du 11 novembre 1918, 69622 Villeurbanne cedex, France

E-mail address: toninelli@math.univ-lyon1.fr