THE PLANAR DIMER MODEL WITH BOUNDARY: A SURVEY.

RICHARD KENYON

1. Introduction

A **dimer covering** of a finite graph is a perfect matching of the graph, that is, a set of edges with the property that every vertex is contained in a unique edge. The **dimer model** is the statistical model dealing with the set of all dimer coverings of a graph.

Kasteleyn [13] and Temperley and Fisher [25] initiated the study of the dimer model by showing how to *count exactly* the number of perfect matchings of an $m \times n$ grid (when at least one of m and n is even). The number is

$$\left[\prod_{j=1}^{m}\prod_{k=1}^{n}\left(2\cos(\frac{\pi j}{m+1})+2i\cos(\frac{\pi k}{n+1})\right)\right]^{\frac{1}{2}}.$$

Later Kasteleyn [15] showed how to efficiently count the number of perfect matchings of any finite planar graph: the number is the square root of the determinant of a matrix closely related to the adjacency matrix of the graph.

Since that time a great deal of work has been done on the physical aspects of the dimer model [8, 12]. In the early 1990's some intriguing combinatorial results brought a renewed mathematical interest in the model. These results touched on the influence of the *boundary* on a random dimer covering of a bounded graph[7, 3, 4, 22, 6]. In this paper we would like to give a short survey of some of these new results.

I am writing this paper from the point of view of a mathematician; in particular I have avoided discussion of the large physics literature on the dimer model. This is partially because physics has a certain itinerary which does not always parallel that of mathematics, and partially due to the unsteadyness of my understanding of the physical point of view. The reader interested in the physical aspects of dimers should consult works of Kasteleyn [13, 14, 15], Fisher [9], Fan-Wu [8], McCoy-Wu [21], and more modern works of Henley [12], Richard et al [23] and others.

Figure 1 shows four planar tilings. The first is a tiling with **dominos**, that is, 1×2 rectangles in two orientations. The second is a tiling with lozenges, where a **lozenge** is a rhombus with a 60° angle (*Losange* is the French word for rhombus). Lozenges come in three orientations. The third example is a tiling with squares and isoceles right triangles, and the last is a tiling with "bibones" which are pairs of hexagons joined along an edge. Bibones also come in three orientations.

In the above tilings the tiles are required to meet edge-to-edge except in the case of the diabolo tilings, where the hypotenuse of a triangle may be across from two squares (or another entire hypotenuse, but no other combination).

Each of these tiling models is in fact a dimer model in disguise. See Figure 2. The reason for introducing them as tilings is simply a matter of taste. Apparently

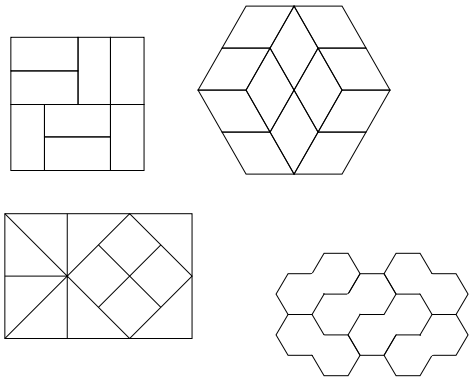

FIGURE 1. Four tiling problems.

mathematicians prefer to talk about tilings, physicists about dimers, and computer scientists about perfect matchings! The square-triangle tilings were called **diabolo tilings** by Jim Propp. Each tile is a union of two isosceles-right triangles in a certain isosceles-right triangulation of the plane. The word "bibone" comes from [16] where it was used in analogy with terminology of Thurston [27] who used the term "tribone" to denote a line of three adjacent vertices in the triangular lattice (dually, a line of three hexagons).

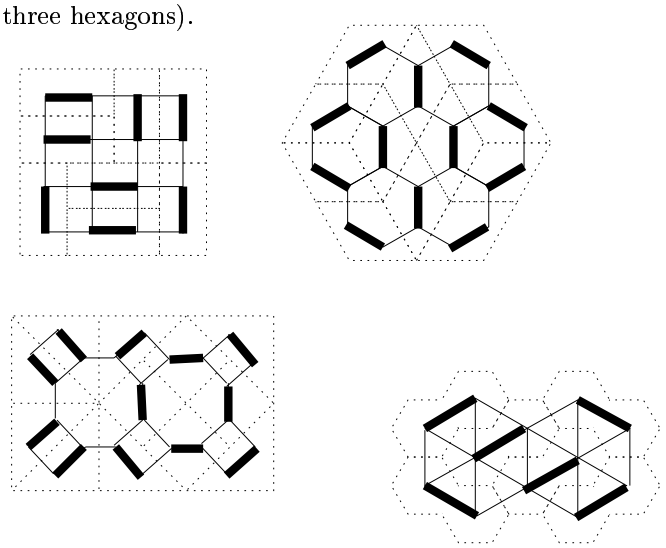

FIGURE 2. Matchings corresponding to the tilings of Figure 1.

Another well-known statistical model which is a planar dimer model in disguise is the Ising model (with no external field) on the square grid $\mathbb{Z}^2$. This fact was first made explicit in Fisher [9]. Recall that configurations in the Ising model are assignments of spins $\{+,-\}$ to points in $\mathbb{Z}^2$; each configuration has an **energy**

which we may take to be proportional to the number of pairs of adjacent vertices with differing spins. In Figure 3, we put the spins rather on the faces of the grid

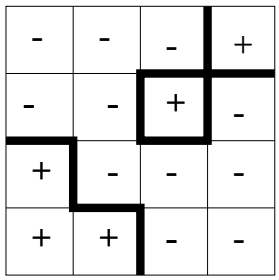

FIGURE 3. The Ising model.

so that for each pair of neighboring faces with differing spins there is a boldface edge; the set of boldface edges then forms a subgraph which has the property that at each vertex it has even degree. Conversely, any subgraph of the grid which has even degree at each vertex corresponds to exactly two different spin configurations (one may choose arbitrarily the spin of the face at the origin: then the remaining spins are determined). Figure 4 then shows the bijection between the Ising model and the dimer model on a related graph: one replaces each vertex in the Ising model with a butterfly graph. There is a way to assign energies to dimer configurations so that this bijection preserves energies: see section 1.1. The resulting planar graph is called the **Fisher lattice**.

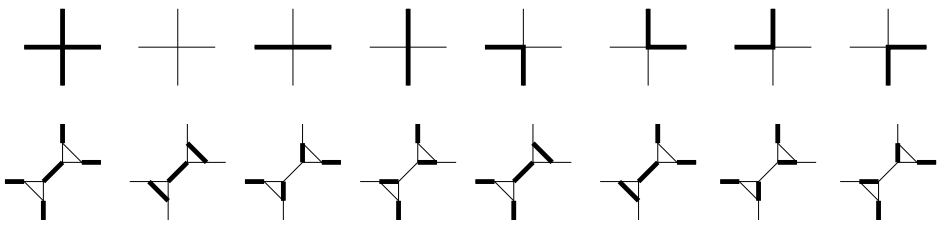

FIGURE 4. Matchings corresponding to the Ising model.

Other well-known statistical models such as the 6-vertex and 8-vertex model have similar planar dimer versions under certain restrictions on the vertex energies: see [8].

1.1. **Defining the measure.** For a finite planar graph G let $M(G)$ be the set of all its perfect matchings. We assume that the graph G comes equipped with an **energy** associated to each edge, which is the energy we must furnish to put a dimer on that edge. This energy may be positive or negative. A perfect matching then has associated to it an energy which is the sum of the energies for each of its *matched* edges.

Mathematically these energies can be thought of as simply a convenient device used to define the **measure** on $M(G)$. The measure is defined using *Gibbs' axiom* (hence it is called a Gibbs' measure), which says that the probability of a configuration is proportional to the exponential of its energy, that is, if a configuration C has energy $E(C)$ then its probability is $\frac{1}{Z}e^{-\beta E(C)}$, where Z is the constant of proportionality necessary to make the measure a probability, and β is a constant which in physics is equal to $\frac{1}{kT}$, k being a physical constant and T being the temperature. The function Z is called the **partition function** of the system and is by definition the sum over all configurations C of $e^{-\beta E(C)}$.

Writing $\beta = \frac{1}{kT}$ we have one measure for each value of T. As T decreases the configurations with lower energy have higher probability density. In fact in the limit $T = 0$ only the configurations of lowest energy have nonzero probability. These are called the **ground states**. On the other hand, in the limit as $T \to \infty$ all configurations have the same probability, so the problem just becomes one of counting the number of configurations. (In reality the situation is not quite so simple: we are usually interested in taking the size of the system to ∞ at the same time as we are taking the limit of the temperature...)

Often it is convenient to shortcut the exponentiation step and assign to an edge of energy E an **activity** equal to $e^{-\beta E}$. Then the probability of a configuration is proportional to the *product* of the activities of its matched edges.

In the Ising model each spin configuration has an energy equal to the number of neighbors of differing spins. This is the number of bold edges in Figure 3. In the corresponding dimer model we assign the energies of the edges in the butterfly to be zero, and the other edges energy 1. The corresponding Gibbs measures are now mapped to one another under the bijection.

1.2. **The boundary.** There are some natural questions to ask right from the start about these Gibbs' measures on $M(G)$. For each finite graph the measure is a sum of a finite number of point masses (since the number of configurations is finite). So perhaps the first question should be regarding the convergence of the measures when the graphs become large. Put differently, if we take a large graph G how does the measure depend on G?

In all the cases we are considering the graphs are "periodic": they are subgraphs of an infinite graph on which $\mathbb{Z}^2$ acts via translations (with finite quotient). So by "large" graph G we should mean one which contains a large neighborhood of the origin in $\mathbb{Z}^2$. Then the question really becomes one about the choice of boundary of G. Does the choice of boundary outside a large ball in G significantly affect what the measure looks like near the origin in $\mathbb{Z}^2$? Remember that we are only dealing with complete dimer coverings of G; we are not allowing the dimers to 'hang over' the boundary.

The answer is of course yes, the measure depends very importantly on how you choose boundary conditions. The most well-known example of this phenomenon is the existence of a phase transition in the Ising model. In the Ising model on a large

ball in $\mathbb{Z}^2$ and at low temperature, choose spins near the boundary of the ball to all be $+$. Then the probability that a spin near the origin will be $+$ is greater than $1/2$ by an amount which does not tend to zero as the size of the ball increases. On the other hand at high temperature this probability tends to $1/2$. We will see below examples of similar boundary influence in the tiling models as well, even in the case of *equal activities*.

This paper deals with the question of how the boundary influences the Gibbs' measure. The following sections each touch on one topic, and are ordered from elementary considerations to more complicated. The first section deals with purely combinatorial questions of the type: which boundary conditions allow the *existence* of a tiling? Section 3 discusses Kasteleyn's method of computing the partition function in each of our examples. In section 4 we discuss how Kasteleyn's method was extended to allow computation of *densities of local patterns*. In section 5 we discuss how the boundary of a region influences the *local entropies* and local densities in a region. In section 6 we discuss in the domino model how the boundary, even when it does not have an influence on the local statistics, can still affect the long-range properties of the limiting measure.

2. A Lipschitz criterion

What regions can be tiled with dominos? There is a surprisingly clean answer to this question [27, 10]. Figure 5 illustrates three elementary observations. The first region cannot be tiled because it has odd area: only even-area regions can be tiled since each tile has even area. The second region cannot be tiled despite having even area: if you color the unit squares black and white in a checkerboard coloring then the number of black and white squares differ, yet every domino covers exactly one square of each color. So if a region can be tiled then in a checkerboard coloring it has the same number of black squares as white squares. This condition is still not sufficient for existence of a tiling, as is illustrated by the third region in Figure 5. Here the principle is that if we cut the region into two parts then the excess of black squares over white squares in either part cannot exceed the length of the cut. This notion was made precise by Fournier [10] in an argument based on the **height function** of Thurston [27].

2.1. **Height function.** Given a domino tiling of a simply-connected region, the height function is an integer-valued function on the vertices of the dominos, defined as follows. Fix a checkerboard coloring of the whole plane. Pick some vertex v_0 (say on the boundary for simplicity) and define the height there to be zero. For any other vertex v take a lattice path from v_0 to v which does not pass through the middle of any domino. The height of v is defined to be the number of black squares adjacent to and to the left of the path, minus the number of white squares adjacent to and to the left of the path. In other words, on each step of the path the height increases by 1 if the square to its left is black and decreases by 1 if the square to its left is white. The value obtained at a vertex is independent of the path taken (because the height change going around a tile is zero), so the height is well-defined. See Figure 6. The height depends on the tiling and in fact determines the tiling: the tiles cross exactly those edges whose vertices have height difference 3. However the height on the *boundary* of the region is independent of the tiling; it depends only on the choice of starting point (a different starting point just shifts the height by a constant additive amount). So you don't need to know a tiling to

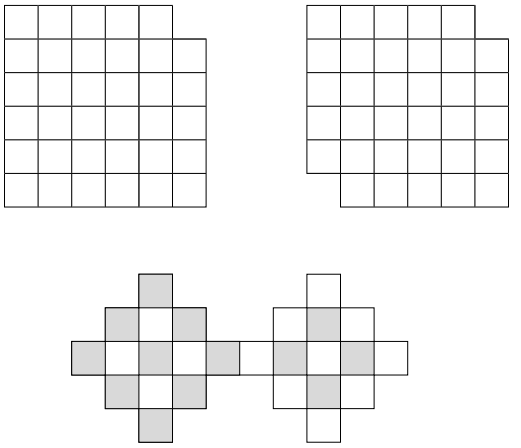

FIGURE 5. Three domino tiling problems

define the boundary height function; in fact the boundary height function gives a necessary and sufficient *condition for tilability*: a simply-connected region can be tiled if and only if between any two boundary points x, y the difference in the height functions $|h(x) - h(y)|$ is not too large compared to their distance in the graph. More precisely, for x, y vertices in U define $d(x, y)$ to be the shortest length of a path in U from x to y which has only black squares on its left. Then (see [10]) U is tilable if and only if for all $x, y \in \partial U$ we have $d(x, y) \geq |h(x) - h(y)|$.

2.2. **Tiling as interface.** The height function allows one to think of a tiling as a surface in $\mathbb{R}^3$: indeed, the graph of the height function is a surface spanning the graph of the boundary height function. See Figure 6. Thurston showed that the max of two height functions and the min of two height functions is again a height function. In particular the set of height functions under the natural partial order is a lattice and has a unique highest and lowest element.

A similar height function exists for both lozenge tilings and diabolo tilings. In the lozenge case use a black-and-white coloring of the underlying equilateral triangulation; the height along a lattice path increases by 1 on an edge with a black triangle to its left and otherwise decreases by 1. In the diabolo case use the black-and-white coloring of the underlying isoceles-right-triangulation. Indeed, for dimer coverings of any *bipartite* planar graph there is a height function with a similar definition and properties [2, 20]. In all these cases dimer coverings can be represented as interfaces.

2.3. **Nonbipartite case.** The non-bipartite case, e.g. for bibones or the Fisher lattice, seems to be genuinely different from the bipartite case. There is no locally-defined notion of long-range combinatorial order similar to the height function. In fact, regarding the existence of a tiling with given boundary, it may be that in all periodic examples any "sufficiently fat" region has a dimer covering. For example C. Kenyon and E. Remila [16] analyzed completely the bibone tiling problem, concluding that a "simply connected" subgraph of the triangular lattice with an even number of vertices has a perfect matching if it has no isthmus or degree-1 vertex.

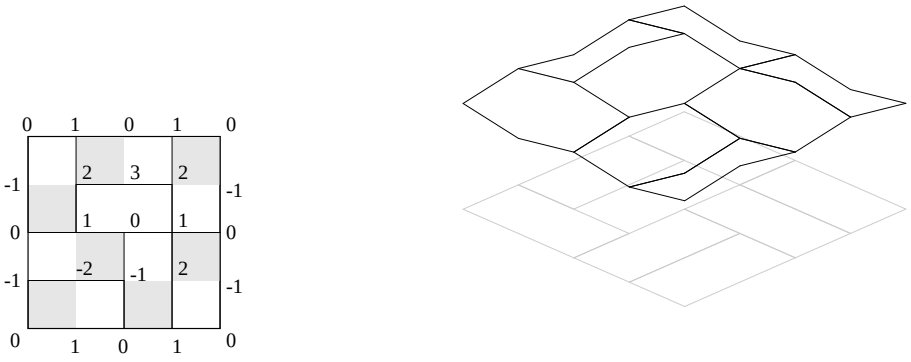

FIGURE 6. Heights and interface of a domino tiling.

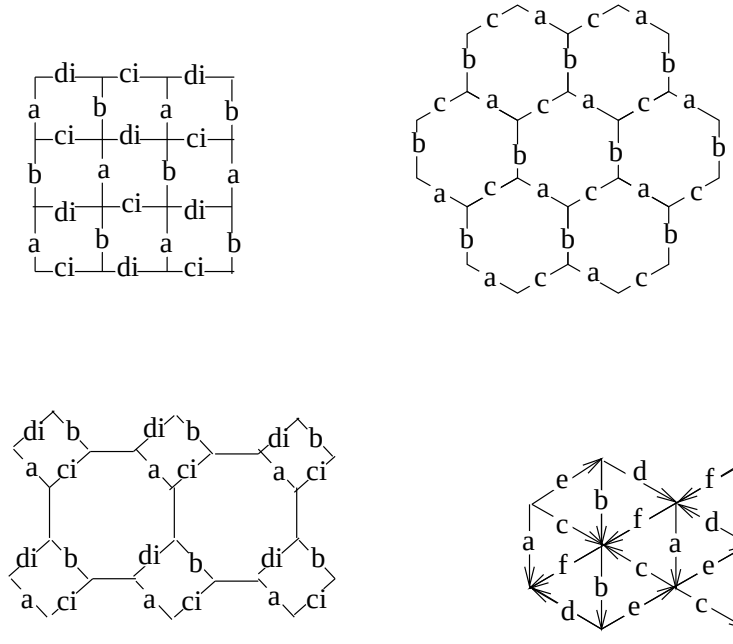

FIGURE 7. 'Natural' edge activities. The numbers a, b, c, d, e, f are positive reals and $i = \sqrt{-1}$.

3. KASTELEYN'S DETERMINANT

Let G be one of the weighted graphs of Figure 7 and $A = (A_{jk})$ be the associated adjacency matrix, that is, the matrix indexed by vertices of G, where A_{jk}

is 0 if vertices j and k are not adjacent, and otherwise A_{jk} is the weight on the corresponding edge, multiplied by -1 in case the edge is directed from k to j.

The matrices A are called **Kasteleyn matrices** for the corresponding graphs; Kasteleyn showed that the square root of the determinant of a Kasteleyn matrix is the partition function for the corresponding dimer model whose activities are given by the *absolute values* of the edge weights.

The reason for the strange choices of signs and directions is to make all the terms in the expansion of the determinants come out with the right sign. In particular, Kasteleyn showed that on any planar graph there is a way to direct the edges so that the associated "adjacency matrix" is a Kasteleyn matrix, i.e. has determinant whose square root is the partition function [15]: it suffices to direct the edges so that around any face the number of clockwise-directed edges is odd. There are many ways to direct the edges to make a Kasteleyn matrix, and moreover on bipartite graphs instead of directing the edges one can get by with putting unit complex number weights on the edges; this has a number of advantages over the direction scheme. In either case, a choice of signs or weights on edges of G is called **Kasteleyn flat** if it gives rise to a Kasteleyn matrix.

This explains (at least in a vague sense) the signs and directions in Figure 7; what about the activites? Surprisingly enough, in the bipartite cases when one is interested in *uniform* random dimer coverings of finite regions (i.e. even when all edge activities are 1) one is led naturally to consider the above activities. We'll see this below.

It turns out that in the lozenge case any two matchings have the same weight. One way to see this is to notice that, first, with these activities, the two possible matchings of a basic hexagon have the same weight (abc), and second, any two matchings can be obtained from one another by local rearrangements around a basic hexagon. This second fact follows from the lattice property of the height function: there is a lowest possible tiling and it has the property that it is the unique tiling whose height function has no "local maximum" in the interior. So one can move from any tiling to the lowest tiling by "rotating" basic hexagons downwards.

So in a sense these activities are superfluous for lozenges: if each matching has the same weight, we may as well set all the activites to be 1. The activities are *not* superfluous when the graph is not "simply connected", for example if we have periodic boundary conditions. For then it is not true that all matchings have the same energy. We'll see the effect of this in the next section.

In a similar fashion one can show (see e.g. [27]) that if we take $ab = cd$ in the domino case, or $ab = cd = 1$ in the diabolo case then any two matchings of a planar graph will have the same weight.

In the bibone case if $ab = cd = ef$ then all matchings will have the same weight (this is shown in [16]). We may as well set $ab = cd = ef = 1$ by a suitable normalization.

The property that all matchings of a planar region have the same weight, is a property of the measure called **conditional uniformity**. In probabilistic language it means that conditioning on the exterior of a region yields a uniform measure.

3.1. **Limiting partition functions.** It is natural to ask one's self what good these determinants are. For an arbitrary planar region we cannot hope to get a closed-form expression for the determinant of one of these huge matrices. So it is not

a surprise that people early in this subject started out by concentrating on very simple boundary conditions like periodic boundary conditions.

If we take a subgraph modelled on those of Figure 7 having *periodic* boundary conditions, we can compute in each case, using the method of Kasteleyn, the partition function Z_n. This is more complicated than indicated in the previous section since the underlying graph is not planar but toroidal: nonetheless Kasteleyn indicated how to compute the partition function of a graph on the torus as a sum of square-roots of four determinants of matrices derived from the Kasteleyn matrices [13, 26]. The advantage of this approach is that with periodic boundary conditions one can explicitly diagonalize the Kasteleyn matrices and thereby compute their determinants. Another advantage of periodic boundary conditions will become clear in section 5.2.

Let G_n be the toroidal graph obtained from the infinite graph by taking the quotient by the subgroup $n\mathbb{Z}^2$ of $\mathbb{Z}^2$. Let Z_n be the partition function of dimers on G_n.

To obtain a finite limit as $n \to \infty$ we define the **partition function per fundamental domain** to be $Z = Z_n^{1/n^2}$.

Values for $\log Z$ in each case have the impressive-looking formulas:

$$\log Z_{dominos} = \frac{1}{4\pi^2}\int_0^{2\pi}\int_0^{2\pi} \log\left(\frac{(a+be^{i\theta})^2}{e^{i\theta}} + \frac{(c+de^{i\phi})^2}{e^{i\phi}}\right) d\theta\, d\phi, \tag{3.1}$$

$$\log Z_{lozenges} = \frac{1}{4\pi^2}\int_0^{2\pi}\int_0^{2\pi} \log(a+be^{i\theta}+ce^{i\phi})\, d\theta\, d\phi, \tag{3.2}$$

$$\log Z_{diabolo} = \frac{1}{4\pi^2}\int_0^{2\pi}\int_0^{2\pi} \log\left(\frac{(2a+be^{i\theta})(a+2be^{i\theta})}{e^{i\theta}} + \frac{(2c+de^{i\phi})(c+2de^{i\phi})}{e^{i\phi}}\right) d\theta\, d\phi, \tag{3.3}$$

$$\log Z_{bibone} = \frac{1}{4\pi^2}\int_0^{2\pi}\int_0^{2\pi} \log(6-2\cos(\theta)-2\cos(\phi)+2\cos(\theta-\phi))\, d\theta\, d\phi \tag{3.4}$$

$$\log Z_{ising} = \frac{1}{4\pi^2}\int_0^{2\pi}\int_0^{2\pi} \log((t^2+1)^2)+2(t-t^3)(\cos(\theta)+\cos(\phi)))\, d\theta\, d\phi \tag{3.5}$$

In all case except the last the integral can be computed using elliptic functions.

Note that in the bibone case Z is independent of the weights! This is a consequence of the fact that even on the torus any two matchings can be obtained from one another using local transformations. The reason Z_{ising} depends on the temperature is that for the weights we chose the Gibbs' measure is not conditionally uniform.

3.2. **Phase transitions.** If $a > b + c + d$, then the integral in (3.1) is easy to evaluate and $\log Z_{dominos} = \log(a)$. It is *independent* of b, c and d! This implies that the expected number of 'a'-edges per fundamental domain is 1, and the expected number of non-'a'-edges is zero. So the system is "frozen" into a brickwork pattern of all 'a'-type edges. Similarly when any one of a, b, c, d is greater than the sum of the other three, Z is frozen into a brickwork pattern. On the other hand when each of a, b, c, d is less than the sum of the others, Z depends on all four variables and the probability of each edge type is nonzero.

Note that for a fixed temperature T, if a is the largest of the four activities then as T decreases there will be a point at which $a = e^{-\beta E_a}$ surpasses the sum of the other three activities. So this temperature is a critical temperature for the system: below this temperature the system is frozen.

In the lozenge case the same phenomenon occurs: the system is frozen whenever one of the three activities a, b, c is greater than the sum of the other two. We'll discuss the diabolo case in section 5.4.

The Ising model does not 'freeze' at any temperature except $T = 0$. However there is a phase transition of a less obvious type at $t = \sqrt{2} - 1$. This is the phase transition we discussed earlier. From (3.5) one can detect something interesting at this point because Z_{ising} has discontinuous derivative there. There is a lot of literature on the Ising model so we have chosen not to discuss this transition here: see [21].

4. Local statistics

Counting tilings is nice, but we'd really like to compute *densities* of patterns to understand the Gibbs' measure. In [17] we showed how the Kasteleyn matrix (or more precisely, its inverse) can be used to compute these densities in the planar dimer model.

Let G be a planar graph and K a Kasteleyn matrix for G. Let E be a set of k dimers covering $2k$ distinct vertices $v_1, \dots, v_{2k}$. The probability of all the dimers of E occurring in a μ-random matching is the square root of

$$a_E^2 \det \begin{pmatrix} A^{-1}(v_1, v_1) & \dots & A^{-1}(v_1, v_{2k}) \\ \vdots & \ddots & \vdots \\ A^{-1}(v_{2k}, v_1) & \dots & A^{-1}(v_{2k}, v_{2k}) \end{pmatrix}$$

where a_E is the product of the activities of the dimers E. (Actually, in [17] we only dealt with the bipartite case, but the general case follows from the same argument.)

So to compute the density of any local pattern we need "only" to compute the inverse Kasteleyn matrix, called the **coupling function**. Again this can be done explicitly only for simple regions like rectangles (for tori it can also be done using a sum of 4 inverses). Nonetheless with the help of a computer one can experiment on some reasonably-sized regions.

By way of example, in Figure 8 we computed for each $k \in [1, n]$ the probability of the vertical domino $\{(k, n/2), (k, n/2 + 1)\}$ occurring in a domino tiling of an $n \times n$ square. The probability is close to $1/4$ throughout most of the region. In fact one can show that as $n \to \infty$ the probability of any given edge e is $1/4 + O(1/d^2)$, where d is the combinatorial distance from e to the boundary. (On the contrary the probability of a boundary edge being covered is asymptotic to $\frac{1}{\pi}$.)

Using the coupling function (inverse Kasteleyn matrix) one can prove the following strong homogeneity property for uniform domino tilings of the square. If we tile the unit square $[0, 1] \times [0, 1]$ with $\epsilon \times 2\epsilon$ dominos, where $\epsilon = \frac{1}{2n}$, then at every interior point $(x, y) \in (0, 1)^2$, the density of any local pattern in a uniform random tiling converges as $n \to \infty$ to a constant independent of (x, y). In particular for large n if you zoom in to any interior point of the square then the measure that you see (as defined by its densities of patterns) is the same. The limiting measure (on tilings of $\mathbb{Z}^2$) is called the **Burton-Pemantle** measure μ_{BP}: they proved that μ_{BP} is the unique translation-invariant measure of maximal entropy on domino tilings of $\mathbb{Z}^2$.

So tilings of the square are very homogeneous. The situation is not at all the same for other kinds of regions. Figure 9 shows a random uniform domino tiling of a tilted square region called an *Aztec diamond*. Random domino tilings of an

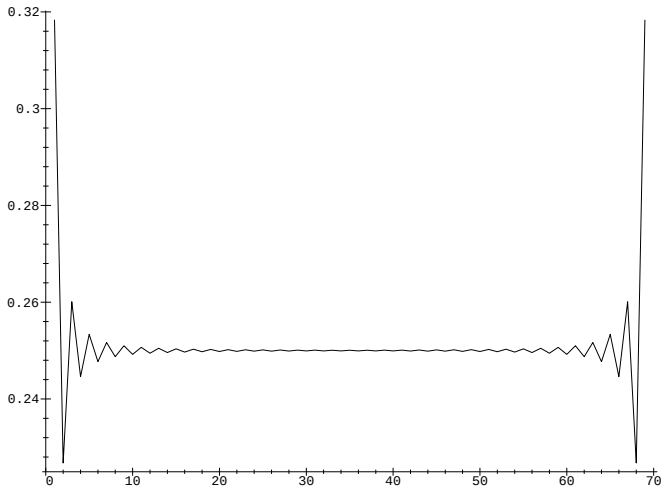

FIGURE 8. Vertical edge probabilities across the middle horizontal line of a square.

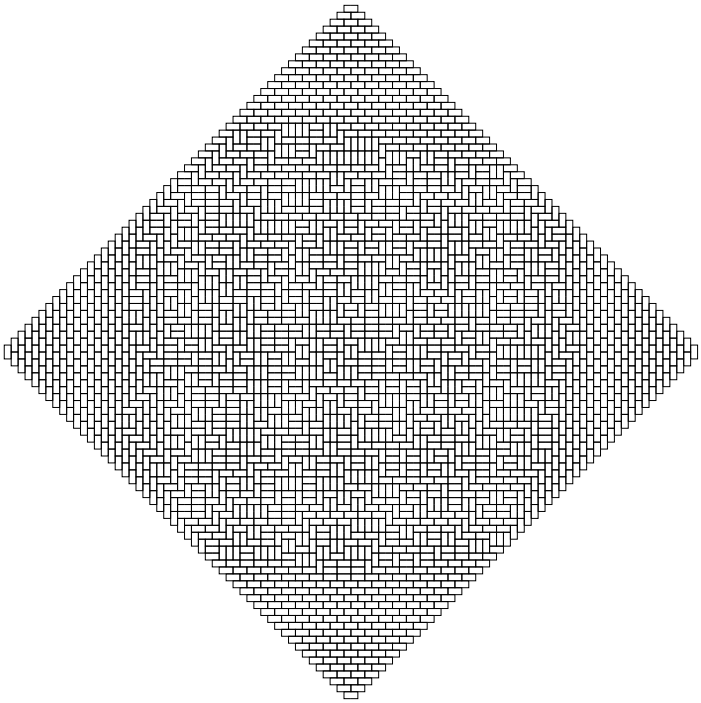

FIGURE 9. Random tiling of an Aztec diamond.

Aztec diamond are quite non-homogeneous, as can be seen from the figure. For example, Figure 10 shows the probability of a vertical domino on the horizontal bisector of the region. Not only is the probability distinctly different from $1/4$ except at the center (even when n tends to ∞ these probabilities do not approach $1/4$ except at the origin), but the probabilities at odd coordinates are completely different from those at even coordinates. In [3] the limiting probabilites of edges were computed. We rescale the coordinates so that the Aztec diamond is defined by $\{(x,y) : |x|+|y| \le 1\}$ and we are tiling with $\epsilon \times 2\epsilon$ dominos. Then as $\epsilon \to 0$ the probability of a vertical domino near coordinate $(x,0)$ converges to either

$$\begin{cases} 0 & \text{if } x < \frac{1}{\sqrt{2}} \\ \frac{1}{2} + \frac{1}{\pi}\mathrm{Tan}^{-1}\left(\frac{2x-1}{\sqrt{1-2x^2}}\right) & \text{if } -\frac{1}{\sqrt{2}} < x < \frac{1}{\sqrt{2}} \\ 1 & \text{if } \frac{1}{\sqrt{2}} < x, \end{cases}$$

or

$$\begin{cases} 1 & \text{if } x < \frac{1}{\sqrt{2}} \\ \frac{1}{2} + \frac{1}{\pi}\mathrm{Tan}^{-1}\left(\frac{-2x-1}{\sqrt{1-2x^2}}\right) & \text{if } -\frac{1}{\sqrt{2}} < x < \frac{1}{\sqrt{2}} \\ 0 & \text{if } \frac{1}{\sqrt{2}} < x, \end{cases}$$

depending on the parity of its distance (in multiples of ϵ) to the rightmost edge.

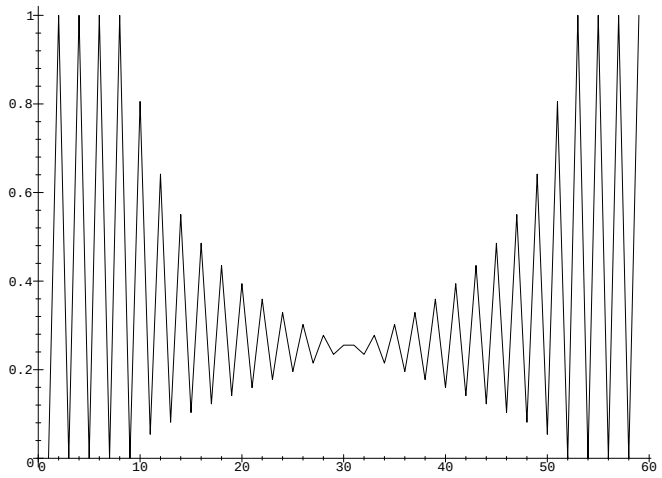

FIGURE 10. Vertical edge probabilities across the middle horizontal line of an Aztec diamond.

In [3], Cohn, Elkies and Propp actually compute the limiting probabilities of edges near *any* point (x,y), but were not able to compute the probabilities of other local patterns, only those of individual edges. Recently Helfgott [11] gave a formula for the inverse Kasteleyn matrix of the Aztec diamond, which would in principle allow one to compute densities, but its asymptotic form has not yet been worked out.

Note that when $|x| > \frac{1}{\sqrt{2}}$ the edge probabilites tend to 0 or 1. In fact [3] showed that outside the inscribed circle $x^2 + y^2 = \frac{1}{2}$, with probability tending to 1 the tiling is 'frozen' into one of the four brickwork patterns seen in the figure. One can think of this phenomenon as a "spatial" phase transition.

5. Boundary effects

It is clear from the previous section that the boundary has a long-range effect on the random dimer covering–at least in the domino tiling case. Quantifying this effect for general regions is a nontrivial problem. In the domino and lozenge case Cohn, Kenyon and Propp [4] gave the first treatment of this problem, computing the **average height function** and **entropy** of dominos or lozenges on large regions. The following sections 5.1 and 5.2 are essentially a rephrasing of the introduction of [4].

5.1. **The variational principle.** The starting point is the following result. If we take a large, tilable, fat region (say an approximate square) whose boundary is chosen in such a way that the height function on the boundary approximates a linear function of slope (s,t), then the number of tilings of the region grows like $\exp(Aent(s,t))$, where A is the area of the region and $ent(s,t)$ is a constant depending only on the slope (s,t).

In other words the *slope of the boundary height function determines the asymptotic number of tilings*, at least when the boundary is flat. What happens when the boundary is not approximately flat? Suppose we have a large tilable region U and we select a random tiling. This is like selecting a random height function with fixed boundary values. As we just noted, the number of tilings whose height function lies near a given one depends roughly on the average slope of the height function: there are many more ways to vary a tiling near a place where the height function is horizontal (slope $(0,0)$) than near places where the slope is nonzero. The steeper the slope, the fewer tilings there are lying close to that slope (indeed there is a maximal slope beyond which there are no tilings: one must have $|s| + |t| \leq 2$). So the idea is that the height function of the random tiling tries to be as horizontal as possible while maintaining its fixed boundary values. Let us make this more precise.

Since we are discussing limiting behavior, we need a growing sequence of tilable regions. Alternatively we can take a fixed domain and tile it with dominos on finer and finer lattices, as we did in the previous section. So, let $U \subset \mathbb{R}^2$ be a simply-connected domain and let $b\colon \partial U \to R$. For each $\epsilon > 0$ let U_ϵ be a tilable region in $\epsilon\mathbb{Z}^2$ approximating U in the Hausdorff topology, such that the scaled boundary height function ϵb_ϵ of U_ϵ approximates b after adding a constant (this presupposes that b has a nice Lipschitz extension to U). Note that we rescale b_ϵ at the same rate as the lattice. We are interested in the behavior of a uniform random tiling of U_ϵ.

The idea that the number of tilings lying close to a given one depends on the slope of the height function is quantified in the following way. The set of possible rescaled height functions of tilings of U_ϵ is converging to the set $L = L(U,b)$ of Lipschitz functions h on U with boundary values b, whose slope (h_x, h_y) satisfies $|h_x| + |h_y| \leq 2$. For each limiting height function $h \in L$ we can approximate the logarithm of the number of tilings of U_ϵ whose rescaled height lies close to h by an

integral

$$\frac{1}{\epsilon^2}\int_U ent(h_x, h_y)dx\,dy. \tag{5.1}$$

(Recall that a Lipschitz function is differentiable almost everywhere so this integral makes sense.)

Because L is compact and the "entropy" function $ent(s,t)$ is strictly concave one can show that there is a unique function $h_{max} \in L$ maximizing the integral (5.1). When $\epsilon \to 0$ the number of tilings whose height lies close to h_{max} overwhelmingly dominates all the other tilings. So we may conclude two things: First, the logarithm of the number of tilings of U_ϵ is given, up to lower order corrections, by the value of the integral (5.1) evaluated for h_{max}. Secondly, the average height function of a tiling of U_ϵ converges to h_{max}.

The function $ent(s,t)$ can be explicitly computed (see the next section) and the maximization property of h_{max} can be turned into an elliptic PDE: the function $h = h_{max}$ is the unique function with boundary values b satisfying

$$\left(2(1-R^2) - \sin^2(\frac{\pi h_x}{2})\right) h_{xx} + 2\sin(\frac{\pi h_x}{2})\sin(\frac{\pi h_y}{2})h_{xy} + \left(2(1-R^2) - \sin^2(\frac{\pi h_x^2}{2})\right) h_{yy} = 0,$$

where $R = \frac{1}{2}(\cos\frac{\pi h_x}{2} - \cos\frac{\pi h_y}{2})$. The function h may be only C^1, not C^2, in which case this equation holds in a distributional sense.

5.2. **The entropy as a function of slope.** The function $ent(s,t)$ for domino tilings is by definition the per-unit-area logarithm of the number of tilings of a big tilable region whose boundary height function has approximate slope (s,t).

It can be computed in a roundabout fashion using a graph on a torus. We use the activities of Figure 7, and the partition function for tilings with periodic boundary conditions (3.1). For this choice of weights, the average slope of a tiling of the torus is $(s,t) = (2(p_a - p_b), 2(p_c - p_d))$ where p_a, p_b, p_c, p_d are the edge probabilites of edges weighted a, b, c, d respectively (by average slope we mean simply the total height change as you go horizontally or vertically around the torus, divided by the length of the corresponding curves). These probabilities can be computed using derivatives of (3.1), e.g. $p_a = \frac{a}{Z}\frac{\partial Z}{\partial a}$. When none of a, b, c, d is greater than the sum of the others, the probability p_a turns out to be

$$p_a = \frac{1}{\pi}\mathrm{Sin}^{-1}\left(\frac{a\sqrt{(a+b+c-d)(a+b-c+d)(a-b+c+d)(-a+b+c+d)}}{(ab+cd)(ac+bd)(ad+bc)}\right), \tag{5.2}$$

and the other edge probabilities are symmetric. The average weight of a tiling is directly related to its average slope, since if a tiling has N_a edges of type a then $p_a = N_a/N$ where N is the total number of edges, and similarly for b, c, d.

Furthermore, and this is the key, almost all tilings for this Gibbs' measure have average slope close to this constant value (s,t). In fact there are so many with this average slope that we can ignore the remaining tilings and make the approximation that all tilings have the same weight. In particular after dividing by this weight factor, the corresponding measure is the same as the measure on the *unweighted* tilings whose average slope is restricted to be (s,t).

There is an amusing geometric interpretation of the above formula (5.2). Take a quadrilateral with edge lengths a, c, b, d in cyclic order, and which is cyclic, that is,

inscribed in a circle. There is a unique such quadrilateral up to congruence. Then $p_a = \frac{\theta_a}{2\pi}$ where θ_a is angle of arc cut off by the 'a' edge of the quadrilateral. There is an even more amusing geometric interpretation of the corresponding entropy $ent(a, b, c, d)$. In the upper half-space model of 3-dimensional hyperbolic space, take an ideal pyramid with a vertex at ∞ and remaining vertices on the four vertices of the above cyclic quadrilateral on the bounding plane. Then $1/\pi$ times the volume of this pyramid is the entropy $ent(s, t)$. So far we have no explanation for this fact except that the formulas agree.

5.3. **Conditional uniformity.** One may ask, since there is a two-parameter family of slopes, but a three-parameter family of activities (we are free to choose a normalization factor, so only the ratios $a : b : c : d$ count), which choices of activities give the same slope? This is an important question since it is not clear (and in fact not true) that activities which give the same average slope also give the same entropy. The answer is that for each given slope, the activities which maximize the entropy are those which satisfy $ab = cd$, that is, those which are conditionally uniform. Furthermore for each allowed slope (s, t) there is a unique (up to scale) choice of a, b, c, d whose Gibbs' measure satisfies $ab = cd$ and has average slope (s, t).

So we see that, when one considers the problem of understanding *uniform* tilings of general regions, that is the partition function with *constant* activities, one is naturally led to consider tilings with non-constant activities, but only those which satisfy conditional uniformity.

In [4] we conjectured that a complete desciption of the (asymptotic) local statistics on a tiling of a general region is obtained from the average height function h_{max}. That is, at a point where h_{max} has slope (s, t) the local statistics should be those of the unique conditionally uniform measure $\mu_{a,b,c,d}$ of the same average slope.

Very similar results hold for lozenge tilings. There is a similar PDE describing the asymptotic average height function. The edge probabilities satisfy $p_a = \theta_a/\pi$, where θ_a is the angle opposite the edge of length a in a triangle of sides a, b, c. (When one of a, b, c is greater than the sum of the other two then the edge probabilities are all 1 or 0). Here all Gibbs' measures defined using the weights in Figure 7 are conditionally uniform. (The entropy is similarly the volume of an ideal hyperbolic tetrahedron whose vertices are ∞ and the above triangle.)

For diabolo tilings there is an additional surprise.

5.4. **Diabolo tilings.** In the diabolo model, when the four activities a, b, c, d are nearly equal the edge probabilities are constant! In fact this system three phases, which we have denoted solid, liquid, and gaseous. (Actually the solid phases come in 4 types like the four brickwork patterns of dominos.) See Figure 11. We will restrict ourselves to the conditionally uniform measures: i.e. those which satisfy $ab = cd = 1$. We can then parametrize the phase space using positive reals a, c. The phase boundary between the 'gaseous' and 'liquid' phase is given by the equation

$$a^2 + a^{-2} + c^2 + c^{-2} = 5,$$

with $a^2 + a^{-2} + c^2 + c^{-2} < 5$ in the gaseous region. The boundary between the 'liquid' and 'solid' phases is given by

$$a^2 + a^{-2} - c^2 - c^{-2} = \pm 5.$$

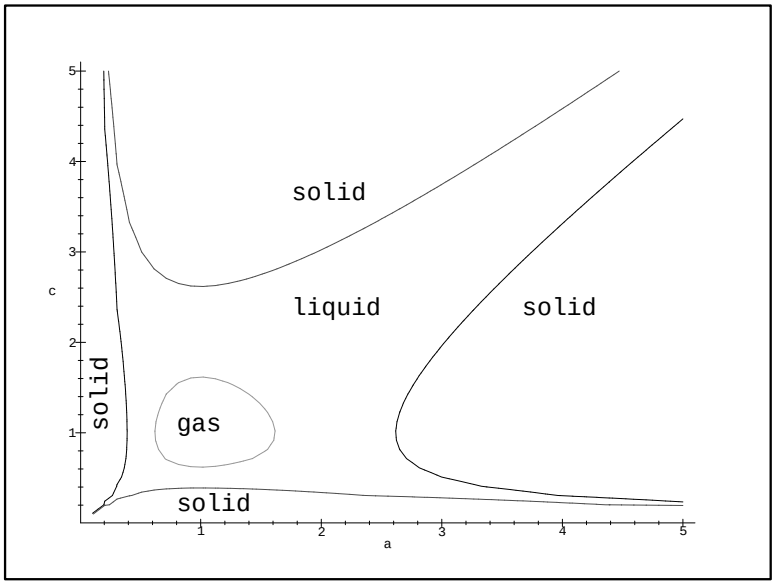

FIGURE 11. Phase boundaries in diabolo tilings.

We have not worked out all the details of this model as the computations are cumbersome. However there are some heuristic results: In the gaseous phase the height function has average slope zero. The measures $\mu_{a,c}$ in this phase are all identical. They appear to have exponential decay of correlations, that is, the probability of local patterns appearing at different points decays exponentially with the distance between the events. In the liquid phase, on the other hand, the measures $\mu_{a,c}$ are all distinct, the height function has non-zero slope and the correlations decay polynomially (probably quadratically). In each of the four solid phases the system is frozen, the entropy is zero and the correlations are infinite (distant events are not independent).

Figure 12 shows a *uniform* random diabolo tiling of a region called a **fortress**. All three phases can be seen to be present: near the corners the system is frozen; near the center the system is in a gaseous phase; in between is a liquid phase. Cohn, Pemantle and Propp [5] computed the asymptotic boundary curve between the three phases in the fortress; there is a single degree 8 algebraic equation (too long to give here) describing both boundaries simultaneously. At the centers of the sides of the diamond all three phases meet.

The existence of this gaseous phase in the center of Figure 12 is apparently due to the non-concavity of the function $ent(s,t)$ near $s = t = 0$. There seems to be a set of *disallowed slopes* in the model: any slope sufficiently close to zero but nonzero cannot occur naturally, due to entropy considerations. Put another way, if one attempted to tile an approximate square with boundary conditions approximating a linear function of small slope, the typical tiling would not be flat but rather have regions of average slope $(0,0)$ and regions of higher slope.

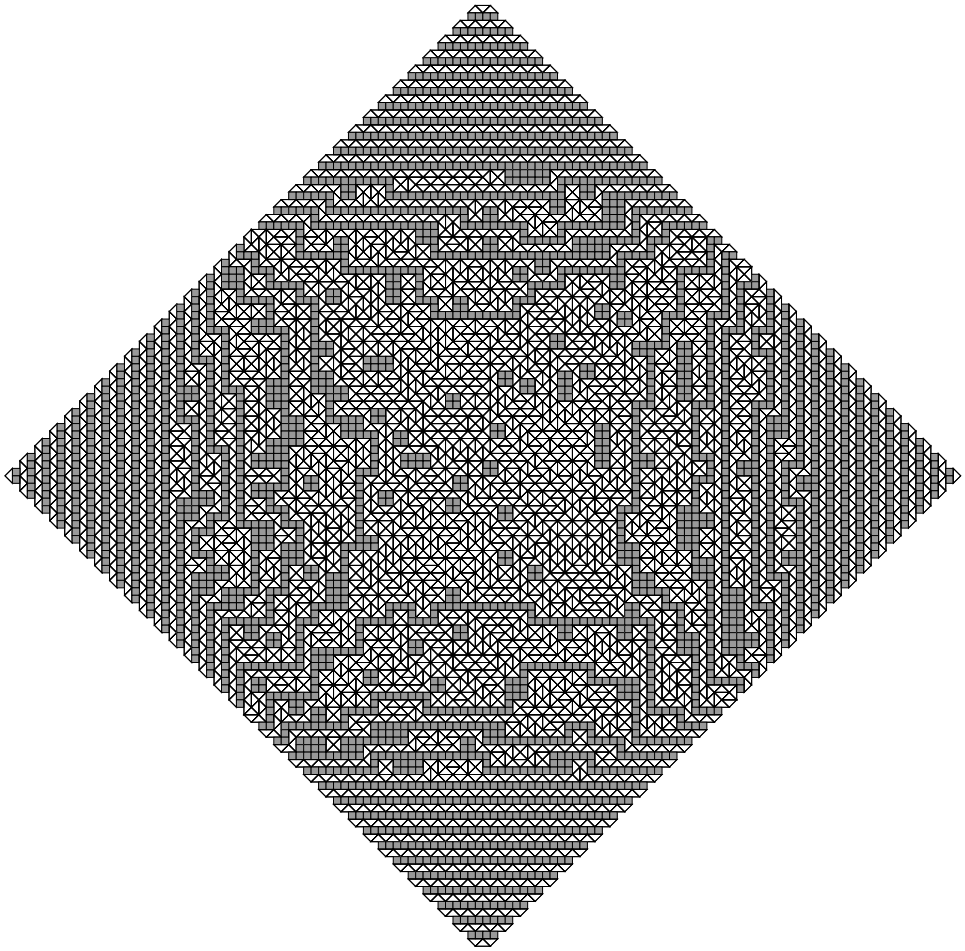

FIGURE 12. A random tiling of a "fortress" with diabolo tiles (courtesy Tilings Research Group, MIT).

6. Temperleyan boundary conditions

Finally we would like to return to take a closer look at a domino system with 'horizontal' boundary conditions, that is, boundary conditions in which the height function stays bounded as the system size grows. Again we concentrate on asymptotic properties, i.e. those which hold in a limiting sense as the lattice spacing ϵ tends to 0.

Unfortunately here again we do not have a complete picture for general boundary conditions. If we restrict ourselves to regions with sufficiently 'nice' horizontal boundary conditions, however, then we can say quite a lot. These special boundary conditions are those which arise from the **uniform spanning tree model** with free boundary conditions via Temperley's bijection (**Temperley's bijection** is a

bijection between the dimer model on the square grid and the uniform spanning model on an index-4 square grid: see [24]). These boundary conditions are therefore called **Temperleyan** boundary conditions. Rather than define precisely here what the bijection is, let us only mention that these boundary conditions are sufficiently rich to be able to approximate any planar domain.

One beautiful property of Temperleyan regions is that they exhibit **conformal invariance**, which for example implies that understanding (limiting properties of) dominos on a region allows one to understand (limiting properties of) dominos on any conformally equivalent region (two domains are conformally equivalent if there is a conformal bijection from one to the other). For example understanding domino tilings on a square allows one to understand dominos on any other simply connected domain.

For a Temperleyan region the boundary does not have a long-range effect on the local densities: the local statistics converge to the Burton-Pemantle measure μ_{BP} which we discussed earlier for the square. However the boundary does still have an effect on certain long-range properties of the measure, among them the height function. The boundary height function on a square is bounded and the average height function in the interior is as well (in fact the average height function is the harmonic function whose boundary values are the smoothing of the boundary height function). The average height function is not so important for our purposes since the *fluctuations* of the height function for a random tiling are unbounded. Indeed, in [19] it is shown that the random variable giving the height function at the center (or any other 'interior' point) of an $n \times n$ square is converging to a Gaussian with variance $\frac{1}{\pi^2} \log n + O(1)$. It is tempting to think of the height as a random interface as we did in section 5.1, but this would be incorrect. In section 5.1 we rescaled the height by ϵ, and the interface converged to a fixed surface. Here however the heights at distinct points in the square are essentially independent: the covariance $\mathbb{E}(h(v_1)h(v_2))$ remains bounded as $n \to \infty$. So the fluctuations are too wild to define a continuous (but nonzero) interface, even if one tries to rescale.

It might nonetheless be possible to average the height over small balls so as to make a smooth interface, and then try to understand what happens when the size of the balls shrinks. In fact one of the current challenges in this theory is to develop a "scaling limit", that is, a continuous object with an intrinsic, conformally invariant description, which has the properties of the limiting random tiling. Another challenge is to prove some kind of **universality**, that is, independence of the scaling limit from the local structure of the lattice. For example it would seem that both dominos and lozenges, when taken with Temperleyan boundary conditions, have the same limiting structure.

References

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

[2] Pavage par des dominos dans des graphes planaires bireguliers. *C. R. Acad. Sci. Paris* I **318** No. 6.(1994):591-594.

[3] H. Cohn, N. Elkies, J. Propp, Local statistics for random domino tilings of the Aztec diamond *Duke Math. J.* **85** No. 1 (1996):117-166.

[4] H. Cohn, R. Kenyon, J. Propp, A variational principle for domino tilings. preprint.

[5] H. Cohn, R. Pemantle, J. Propp, in preparation.

[6] N. Destainville, R. Mosseri, F. Bailly, Configurational entropy of codimension-one tilings and directed membranes, *J. Stat. Phys.***87**(1997),697-.

[7] N. Elkies, G. Kuperberg, M. Larsen, J. Propp, Alternating-sign matrices and domino tilings,I, *J. Alg. Comb.* **1** No 2. (1992):111-132.
[8] Chungpeng Fan and F. Y. Wu. General lattice model of phase transitions. *Physical Review B*, **2**(3) (1970):723–733.
[9] M. E. Fisher, On the dimer solution of planar Ising models. *J. Math. Phys* **7** No. 10 (1966):1776-1781.
[10] J.-C. Fournier, Pavage des figures planes sans trous par des dominos : fondement graphique de l'algorithme de Thurston et parallélisation, *Compte Rendus de L'Acad. des Sci.*, Serie I **320** (1995), 107–112.
[11] H. Helfgott, Edge effects on local statistics in lattice dimers: a study of the Aztec diamond (finite case), Thesis, Brandeis, 1998.
[12] C. L. Henley, Random tiling models, in *Quasicrystals: The state of the art*, ed. D. DiVincenzo and P. Steinhardt, World Scientific, 1991, 429-524.
[13] P. W. Kasteleyn, The statistics of dimers on a lattice. I. The number of dimer arrangements on a quadratic lattice, *Physica* **27** (1961), 1209–1225.
[14] P. W. Kasteleyn, Graph theory and crystal physics, in Graph Theory and Theoretical Physics, Academic Press, London 1967.
[15] P. W. Kasteleyn, *J. Math. Phys.* **4**(1963),287-
[16] C. Kenyon, E. Remila, Perfect matchings in the triangular lattice *J. Disc. Math.* **152** (1996)191-210.
[17] R. Kenyon, Local statistics of lattice dimers, *Ann. Inst. H. Poincaré, Prob. et Stat.* **33** (1997), 591-618.
[18] R. Kenyon, Conformal invariance of domino tiling. preprint.
[19] R. Kenyon, Conformal invariance of uniform spanning trees, in preparation.
[20] R. Kenyon, J. Propp, D. Wilson, Trees and matchings, preprint.
[21] B. McCoy, F. Wu, The two-dimensional Ising model, Harvard Univ. Press 1973.
[22] Boundary-dependent local behavior for 2-D dimer models, *Int'l. J. Modern Phys. B* **11** (1997),183-187.
[23] C. Richard, M. Höffe, J. Hermisson, M. Baake, Random Tilings: concepts and examples.
[24] H. Temperley, Combinatorics: Proceedings of the British Combinatorial Conference 1973, *London Math. Soc. Lecture Notes Series #13*, (1974),202–204.
[25] H. Temperley, M. Fisher, The dimer problem in statistical mechanics – an exact result. *Phil Mag.* **6**(1961):1061-1063.
[26] G. Tesler, Matchings in graphs on non-oriented surfaces, preprint.
[27] W. P. Thurston, Conway's tiling groups, *Amer. Math. Monthly* **97** (1990),757-773.

URAD1169 du CNRS, Bât 425, Université Paris-Sud, 91405 Orsay, France.