Classical and Quantum Dimer Models

Mike Gartner¹

¹Department of Physics, University of California at San Diego, La Jolla, CA 92093

We consider a classical discrete model defined on various types of graphs, and a class of quantum generalizations, the so called dimer models. In the classical setting, we discuss a closed formula for the partition function, interpretation of the model in terms of a random surface via the construction of height functions, and the relationship of the model to gauge theory. We then discuss the quantum dimer model, construct some ground states for particular instances of the model, and provide a completely superficial discussion of phase diagrams and critical points. The ultimate goal will be to provide an introduction to the theory accessible to someone who has never heard of a dimer. A secondary goal is to provide plausible evidence to support the notion that quantum dimer models exhibit complex phase structures, and have interesting applications in gauge theory.

INTRODUCTION

The dimer model arose initially as an attempt to describe the adsorption of diatomic molecules on the surface of crystals, and has had success in describing the behavior of partially dissolved crystals in equilibrium. It has since been applied as a model for many other physical systems, and garnered much interest from the mathematical community, due in part to the tractibility afforded by a closed form expression for the partition function in certain cases. This fact has allowed for a deep understanding of the classical model to be developed, including a beautiful characterization and classification of the phase diagram in terms of certain algebraic curves. The model also exhibits conformal invariance in the classical setting, and can be understood in the context of \mathbb{Z}_2 gauge theory. Quantum versions of the dimer model were proposed in studies of high temperature superconductors, specifically in the study of SU(2) singlet dominated phases in various spin models.

This review will be organized as follows: First, we will define the classical dimer model, consider its basic statisto-mechanical[7] properties, discuss height functions, and describe a gauge equivalence in the classical setting. We will then move on to defining the quantum model, construct many ground states, and review some of their features discovered in the literature. Finally, we will mention the quantum phase diagrams of the model.

We place a \star next to subsections of "The Classical Dimer Model" which may be be skipped without affecting the narrative of the "Quantum Dimer Models" section.

THE CLASSICAL DIMER MODEL

Here we will give a lightning review of the classical dimer model, and the progress that has been made in understanding its behavior. Naturally, such a short treatment does not do the theory justice, and many fundamental results have been left out. The goal is to provide some context before studying quantum variants of the model, and to display some of the interesting features that arise classically. In fact, we will find that the classical theory plays an important role in the study of ground states of certain quantum versions of the model. This will section will mostly consist of sketches of major ideas aimed at giving some flavor of the degree to which the model is understood, with a few of the technical details shunned and relegated to the depths of the appendix. The story is really quite fascinating, and the interested reader can consult [6] for a readable and thorough introduction.

What's a Dimer?

We consider a system which can be modeled as a graph G. States of this system are defined as follows:

Definition 1. Given a graph G, a **perfect matching** M of G, also called a **dimer configuration**, is a subset of edges of G such that every vertex is incident to exactly one edge.

Clearly not all graphs will allow for perfect matchings; for example having an even number of vertices is a necessary condition for existence. We will at times restrict our attention to various classes of graphs to provide enough structure to make analytical progress. In particular, we will consider graphs G adorned with such titles as planar, simple, \mathbb{Z}^2 -periodic, and bipartite. Appendix 1 provides definitions and simple examples for those unfamiliar with these properties. In Figure 1 below one example of a perfect matching is shown for a subgraph of the honeycomb lattice.

Fix some graph G, and consider any positive function on the edges of the graph $\nu : E \to \mathbb{R}$, called a weight function on G. The *energy* of a dimer configuration M will be defined as

$$\mathscr{E}(M) = -\sum\nolimits_{e \in M} \log \nu(e),$$

and the *Boltzmann weight* of the configuration will be

$$\nu(M) = e^{-\mathscr{E}(M)} = \prod_{e \in M} \nu(e)$$



FIG. 1: A finite graph, with a dimer configuration specified by bold edges. Source:[6]

Let $\mathscr{M}(G)$ be the set of dimer configurations on G, and let the partition function associated with G be given by $Z(G) = \sum_{M \in \mathscr{M}(G)} \nu(M)$. As usual, the Boltzmann weight divided by the partition function yields a Boltzmann measure on dimer configurations of G, and in the case where the weight function on edges is the constant function $\nu(e) = 1$ it merely counts the number of dimers. A first step towards understanding this model analytically is to find an explicit expression which one may use to compute the partition function. The following theorem gives us just that.

Theorem 1. Let G be a finite, bipartite, planar directed graph with an admissable orientation of its edges, let ν be a weight function, and let K be the Kastelyn matrix associated with G. Then Z(G) = |det(K)| [2]

The sketch of the proof is straightforward: the Kastelyn matrix is some (oriented and weighted) adjacency matrix, and by inspection it is clear that nonzero terms in its determinant are in one to one correspondence with dimer configurations. Furthermore, each nonzero term represents the weight of a corresponding dimer configuration, and thus all that must be verified is that every term in the determinant has the same sign. The rest of the argument consists of proving that this can always be accomplished with an "admissable orientation" of the edges of G, and that such an admissable orientation always exists. The part of the proof sketched here, and requisite definitions, are stated more formally in Appendices A and B.

Using this result, one can obtain an explicit form for the probability of obtaining a particular set of edges in a dimer configuration chosen with respect to the Boltzmann measure. This yields information about local statistics [4].

Ultimately, we would like to consider the model on an infinite graph, and we will now sketch how our results above can be carried over to this setting. The following sketch will apply to a bipartite, simple, planar, infinite, \mathbb{Z}^2 -periodic graph G. It turns out that we can consider a class of finite graphs $G_n \equiv G/n\mathbb{Z}^2$ on a torus, which in the appropriate limit allow us to extend our result to an

analagous result on the graph G. The graphs are defined on a torus so that we retain \mathbb{Z}^2 translation symmetry, which allows for Fourier techniques to carry through to the infinite setting. One finds that the computation of an analytic expression for the partition function of these graphs G_n is not precisely given by Theorem 1 (they are not planar, so there is no contradiction in this statement), but that a very similar theorem can be proven [6]. Once this is done, taking the limit provides a result for the partition function per fundamental domain of the graph of interest G.

Height Functions \star

Here we discuss a method whereby dimer configurations can be associated with objects called height functions. These play a crucial role in finding closed form expressions for the graphs G_n mentioned in the previous section, and also provide alternative descriptions of dimer phenomena in general. Intuitively, these will just be assignments of scalar height fields to faces of our graphs. An example is shown in Figure 2.



FIG. 2: A triangular lattice, and a particular choice of height function. Source: [6]

Let G be a bipartite graph. A flow ω on G is a real valued function defined on all oriented edges of G, ω : $\vec{E} \to \mathbb{R}$. Note that we have not specified an orientation for G here yet; rather, \vec{E} is the set of *all* possible oriented edges of G, so it contains two oppositely oriented copies of every edge of G. The **divergence** of a flow ω will be a real valued function on the vertices of G, $div(\omega): V \to \mathbb{R}$, which subtracts the total flow going into a vertex from that going out of it.

A dimer configuration M on G can be used to define a flow as follows. Color the vertices of G white and black, and then choose ω to be 1 for every directed edge \vec{e} such that $e \in M$ and \vec{e} begins on a white vertex. Likewise, choose ω to be -1 for every directed edge \vec{e} such that $e \in M$ and \vec{e} begins on a black vertex. It is easily verified that the difference in flows $\omega_2 - \omega_1$ constructed in this

way from two dimer configurations M_1, M_2 , must have no divergence. Now consider any two faces f_0, f_1 in G, and an arbitrary path γ in the dual graph G^* from f_0 to f_1 . The fact that $\omega_2 - \omega_1$ is divergence free can be shown to imply that the flux of $\omega_2 - \omega_1$ across γ is independent of the choice of γ . This means this quantity is a function of the face f_1 only, once we choose a reference face f_0 . If we further choose a fixed reference dimer configuration M_1 , then we see that we have constructed one such function on the faces of G for every dimer configuration M_2 . This is the **height function** associated to M_2 , and one can recast the theory of dimer models into a theory of appropriate height functions. This is one means by which dimer models can be used to understand the physics of partially dissolved crystals.



FIG. 3: A height function arising from dimers on the honeycomb lattice. One can imagine this as a 2d projection of the partially dissolved corner surface of a 3d crystal. On the right is a "Ronkin" function characterizing this height function; see [4] if you would like to know what this means. Source: [4]

Gauge Equivalence \star

Consider again the graphs G_n defined above, which are bipartite. We note that in this case one can choose a "preferred" orientation for G_n by labeling vertices white and black, and declaring that edges go from white to black. Then any function g on the (unoriented) edges can be associated with a 1-form, by which we mean a function f on oriented edges such that $f(-\vec{e}) = -f(\vec{e})$. This can be done by defining $f(\vec{e}) \equiv g(e)$ and $f(-\vec{e}) \equiv -g(e)$ when \vec{e} is oriented from white to black, and oppositely when it is oriented from black to white. Then if we are given two weight functions ν_1,ν_2 on G, the functions are said to be **gauge equivalent** if they differ by the differential of a one form:

$$\nu_1 = \nu_2 + df, for some f \in \Omega^0$$

Here Ω^0 is just the linear space of real functions on the vertices. More explicitly, this means:

$$\forall e = (w, b) \in G_n, \nu_1(e) = \nu_2(e) + f(b) - f(w)$$

Given two gauge equivalent weight functions, the difference in energies $\mathscr{E}_1(M) - \mathscr{E}_2(M)$ for some dimer configuration M of G_n will be a constant. Since this energy difference will be the same for all dimer configurations, the two weight functions will actually induce the same Boltzmann distributions, and thus the same physics.

If we choose some oriented cycle in our graph G_n , $\gamma = \{w_0, b_0, w_1, ..., b_{n-1}, w_n\}$, where $w_0 = w_n$, we can define the following quantity:

$$\int_{\gamma} \nu \equiv \sum_{i=1}^{n-1} [\nu(w_i, b_i) - \nu(w_{i+1}, b_i)]$$

which is called the **magnetic flux** through γ . It can be shown that ν_1 and ν_2 are gauge equivalent if and only if $\int_{\gamma} \nu_1 = \int_{\gamma} \nu_2$ for all cycles γ . These "integrals" also characterize the energy change under particular transformations of dimer configurations which are important in the quantum setting. Fix a weight function ν on G_n . If a dimer configuration M is such that every other edge of a cycle γ is included in M, then we can simply "rotate along γ " and leave the rest of the configuration unchanged to obtain a new dimer configuration M'. This type of operation is displayed in Figure 5 (b), where it is referred to as a loop flip. It is easily verified that

$$\mathscr{E}(M') = \mathscr{E}(M) \pm \int_{\gamma} \nu$$

under the action of rotation of M along γ . One can use this to show that knowledge of the magnetic fluxes through all cycles in G_n uniquely determines the relative Boltzmann weights of all dimer configurations [4]. We end this section by noting that the magnetic field coordinates $(B_x, B_y) \in \mathbb{R}^2$ are the notation used in the literature to describe the magnetic flux through two simple non-trivial horizontal and vertical cycles on the torus in which our graph is embedded (they look like $H_1(\mathbb{T}^2,\mathbb{Z})$) generators). They (partially) parametrize the gauge equivalence classes of weight functions^[4]. This will be useful in having some idea of how Figure 4 relates to anything mentioned in this review (One may recall that graphs embeddable on tori are crucial to extending the result of Thm 1 to the infinite case, so it is not so unbelievable that they will play a role in understanding the phase diagrams in the infinite case).

Phase Diagrams, Amoeba, Curves *

One can characterize phases of the classical dimer model by studying how the height functions behave. On a heuristic level, "frozen" phases are those in which there is some degree of determinism in height differences between different faces, "gaseous" phases are those in which height difference fluctuations have bounded variance, and "liquid" phases are those in which height differences grow universally as the logarithm of the distance between faces.

One can associate a curve known as the spectral curve to any dimer model on a graph G, and these curves have very nice algebraic properties, being what are known as plane algebraic Harnack curves. Furthermore, they allow detailed information about phase diagrams to be extracted relatively simply. One such example of this is that the genus of the spectral curve associated to a dimer model is the number of distinct "gaseous" phases that arise in the model. This story is quite involved; just note that the classical dimer model has surprising ties to algebraic geometry, and there exists powerful machinery that has allowed for the phases of the classical models to be understood on a very deep level.



FIG. 4: This "amoeba" describes the phases of a particular dimer model in magnetic field coordinate space. The take home message is that it looks cool. Source:[6]

QUANTUM DIMER MODELS

We begin by defining a Hilbert space from our classical dimer model.

Definition 2 (QDM Hilbert space). Given a graph G satisfying the conditions as in blah, the set of dimer configurations on G is an orthonormal basis for the quantum dimer Hilbert space \mathscr{H}_G corresponding to G.

We remark that the orthonomality of the basis specifies the inner product on \mathscr{H}_G , as is necessary in defining a Hilbert space. Note that even the most simple operators on \mathscr{H}_G must act on at least two edges in a dimer configuration in order to produce a new dimer configuration. Two such simple operators are "plaquette flips" and "loop flips", which act on just one plaquette or one loop of a dimer configuration M at a time. In a plaquette flip we rotate the edges around a particular plaquette in M to obtain a new configuration. When this is possible, aka when a plaquette is flippable, the operator does nothing to edges outside of the plaquette.

In a loop flip of a configuration M, we consider a closed loop of edges of G such that alternating edges in the loop are in M. We then define the loop flipped configuration to be that which is the same outside of the loop, and with the alternate choice of edges made inside the loop. These two moves are depicted in Figure 1 below for some particular graphs. These operations will not be precisely defined here, as it is unclear what the appropriate definition should be for a generic graph, but the intuitive idea is clear, and one may think of what follows as taking place on a square or triangular lattice for concreteness.



FIG. 5: In (a) plaquette flips for the honeycomb, triangular, and cuibc lattices are shown. In (b), the dotted line surrounds a flippable loop, and the action of a loop flip is shown. Source: [5]

Let \hat{F}_i be the plaquette flip operator which acts on a dimer configuration by flipping the i-th plaquette if it is flippable, and leaving all other edges in the configuration unchanged. In the case where the i-th plaquette is not flippable \hat{F}_i is zero. Let \hat{V}_i be the identity if the i-th plaquette is flippable, and zero otherwise. Then we can construct the following Hamiltonian for a quantum dimer model:

Definition 3. Let $H_{QDM} = \sum_{i \in I} -t\hat{F}_i + v\hat{V}_i$, which will be denoted the **Quantum Dimer Hamiltonian**. Here I is an indexing set for the elementary plaquettes of the graph G.

We begin by considering the QD hamiltonian when t=v, which is known as the *R*okhsar-Kivelson point. We note that in this case the hamiltonian is self adjoint, since it is everywhere defined and also symmetric. For this model, we can construct ground states explicitly. One can easily verify that

$$\hat{V}_i - \hat{F}_i = \frac{1}{2}(\hat{V}_i - \hat{F}_i)^2$$

so that

$$H_{RK} = \sum_{i \in I} \hat{V}_i - \hat{F}_i = \sum_{i \in I} \frac{1}{2} (\hat{V}_i - \hat{F}_i)^2 =$$

$$\sum_{i \in I} \frac{1}{2} (\hat{V}_i - \hat{F}_i)^{\dagger} (\hat{V}_i - \hat{F}_i)$$

This implies that the eigenvalues of H_{RK} are nonnegative, which in turn implies that any state annhilated by $\hat{Q}_i \equiv \hat{V}_i - \hat{F}_i \ \forall i \in I$ is a ground state of the system. Consider any state $\psi \in \mathscr{H}_G$; by definition $\psi = \sum_{M \in \mathbb{M}(G)} A_M | M \rangle$ for some set of amplitudes $\{A_M\}$. We note that

$$\hat{Q}_i(A_M|M\rangle) = \begin{cases} A_{M_i}|M_i\rangle - A_M|M\rangle & if flippable\\ 0 & else \end{cases}$$

Here flippable means that the ith plaquette of M is flippable, and M_i is the dimer configuration obtained from M by flipping the ith plaquette. With this understood, it is clear that \hat{Q}_i will annhiliate $|\psi\rangle$ so long as the following condition holds for $|\psi\rangle$:

If two dimer configuration states $|M\rangle$ and $|M'\rangle$ appear with nonzero amplitudes $A_M, A_{M'}$ in $|\psi\rangle$, and if M is related to M' via a single plaquette flip, then $A_M = A_{M'}$.

Thus the equal superposition state over all dimer configurations

$$|\psi_0\rangle = \frac{1}{\sqrt{Z}} \sum_{M \in \mathscr{M}(G)} |M\rangle \tag{1}$$

is a ground state of H_{RK} , and is normalized when Z is the equal weight classical partition function which counts the number of dimer configurations. This result can be used to find equal time correlation functions in this ground state, and one finds that in the planar case these equal time correlators are equivalent to those of the classical theory. So indeed, our classical results are significant in the quantum context. In particular, it has been shown that equal time correlators in the equal superposition ground state on a square lattice exhibit algebraic decay, and those on a triangular lattice exhibit exponential decay [1].

Topological Sectors and Topological Order

One may have thought that the condition sufficient to ensure a state is a ground state of H_{RK} mentioned above might allow for multiple ground states, and in this section we discuss one means by which this can be realized.

In both the bipartite and non-bipartite planar cases, there exist invariants associated with particular dimer configurations which cannot be changed via the action of \hat{F}_i . For example, in the simple case of a triangular lattice one can prove that given any non contractible path in the dual graph, and a dimer configuration M, the parity of the number of edges of M intersecting the path cannot be changed by \hat{F}_i [5]. This is shown in Figure 6 below.



FIG. 6: The dashed line is a non contractible curve extending infinitely in both directions on the triangular lattice. The parity of the number of dimer edges crossing the line cannot be altered by the action of plaquette flips. Source: [5]

On a torus a similar construction exists, yielding a two component invariant (W_x, W_y) . This is good, because as you may recall graphs embeddable on tori played a crucial role in extending our classical partition function result to the infinite setting. In fact this construction depends on topology and can be extended to a class of graphs embeddable on surfaces of arbitrary genus g, yielding 2g numbers when $g \ge 1$. The key point is that to each dimer configuration we can associate a set of numbers, called *winding numbers*, and that two dimer configurations can be related via \hat{F}_i if and only if they share the same set of winding numbers. These invariants motivate the following definition:

Definition 4. A topological sector of \mathscr{H}_G is a subspace spanned by the set T_G of all dimer configurations with some particular set of winding numbers.

We now note that in addition to (1), any equal superposition

$$|\psi\rangle = \frac{1}{\sqrt{Z}} \sum_{M \in T_G} |M\rangle \tag{2}$$

over all basis dimer states in a topological sector will satisfy the condition sufficient for $|\psi\rangle$ to be a ground state of H_{RK} . Thus if we have a graph embedded on a surface of genus g, we will have 2g winding numbers, each with two possible choices for their parity. This means there will be $2^{2g} = 4^g$ ground states given by the equal superposition states over topological sectors.

It turns out that there is an energy gap between these states and the lowest energy excited states in the QD model. Furthermore, while we have shown that at the Rokhsar-Kivelson point when t=v these states all have zero energy, for the general model given by H_{QDM} these states can be seperated by an energy gap. For example, on the triangular lattice different topological sectors have an energy gap which decays exponentially with the growth of the length of the system L, $\delta E = e^{-cL}$ [5]. This is an example of topological order that is ubiquitous in quantum dimer models; there is no local order parameter which allows one to distinguish ground states in these different topological sectors, as degeneracy depends on the topology of the graph in question.

Phases of the QDM

Here we merely remark that our quantum dimer model is really a family of models given by the value of the ratio t/v. The role of the Rokhsar-Kivelson point in general depends on the graph on which the model is being studied, but in many cases it plays a central role in describing phase transitions in the model.

For many planar non-bipartite lattices, the RK point is part of a \mathbb{Z}_2 RVB (Resonance Valence Bond) liquid phase, and exhibits \mathbb{Z}_2 topological order with four degenerate ground states arising from the topological sectors. They are "liquid" in the sense of exponential decay of dimer edge correlation functions [5]. If one allows our QDM Hilbert space to be enlarged, this phase also allows for deconfined monomer excitations, and vison excitations.

For many planar bipartite lattices, the RK point is a critical point between different "crystalline" phases of the model, given names such as the staggered, columnar, and plaquette phases. These appear, as one might expect, to be very ordered.

Needless to say, there are many more phases of the QDM and its variants that can arise, and a lot more can be said about the role of the RK point in this theory. These are all detailed points, about which I will say nothing more.

Relationship to Height Functions and Gauge Theory

Our discussion of height functions and gauge theory in the classical case can be extended to QDMs in several ways. The use of height functions to describe our QDMs can give rise to the construction of continuum QDM theories [5]. Often these models yield U(1) gauge theories in the continuum, though they can also result in descriptions in terms of \mathbb{Z}_2 gauge theories. Again, these points require more detail than I can afford in my remaining -3 pages, so I will say nothing more here.

Acknowledgements

I'd like to thank John McGreevy for an interesting class, and for allowing me to choose this topic despite my having no idea if or how it related to the course. I'd also like to thank anyone else who may have made it this far into the paper for taking the time to read it. I hope it was comprehensible.

- [2] P.W. Kastelyn, "Graph Theory and Crystal Physics," In Graph Theory and Theoretical Physics, pages 43110. Academic Press, London, 1967.
- [3] R. Kenyon, "Conformal Invariance of Domino Tiling," arXiv:math-ph/9910002
- [4] R. Kenyon, A. Okounkov and S. Sheffield, "Dimers and Amoebae," math-ph/0311005.
- [5] R. Moessner and K. S. Raman, "Quantum Dimer Models," arXiv:0809.3051 [cond-mat.str-el]
- [6] B. Tiliere, "The Dimer Model in Statistical Mechanics,"
- [7] This may or may not be a word

APPENDIX A: DEFINITIONS

Here we provide some definitions, in case they may be sought.

Definition 5. A graph G is said to be **planar** if it can be embedded in a plane (Shocking, I know!). This means that it can be drawn in a plane (e.g. \mathbb{R}^2) in such a way that edges interect only at vertices.



FIG. 7: Some examples and non-examples of planar graphs. Source: Wikipedia

Definition 6. A graph G is said to be **bipartite** if one can partition the vertices V of G into two sets $W \cup B = V$, such that every edge in G connects an element of W to an element of V.



FIG. 8: A bipartite graph. Unfortunately, what I have called W and B are labeled here as U and V. Source: Wikipedia

Definition 7. A graph G is said to be simple if it contains no graph loops (i.e. no edge begins and ends on the same vertex) and if it does not contain multiple edges between any two vertices.



FIG. 9: Some examples and non-examples of simple graphs. Source: Wolfram

Definition 8. A planar, bipartite graph G is said to be \mathbb{Z}^2 – **periodic** if it is embedded in the plane in such a way that translations in \mathbb{Z}^2 act by isomorphisms preserving the colorings of the vertices (i.e. black vertices go to black vertices, white vertices go to white vertices)

FIG. 10: The fundamental domain for the square octagon graph. This graph is \mathbb{Z}^2 -periodic. Source: [4]

Definition 9. The dual graph G^* of a planar graph G is the graph obtained by assigning a vertex to each face of G, and an edge for each edge separating two adjacent faces of G.

APPENDIX B: (PART OF) PROOF OF THEOREM 1

Let G be a finite, bipartite, planar directed graph with a weight function ν defined on its edges. Since G is bipartite, color the vertices black and white and label them as $w_i \in W$, $b_j \in B$, and any (unoriented) edge in G will be labeled by a pair (w_i, b_j) . The **Kastelyn matrix** associated to G is the following oriented weighted adjacency matrix:

$$K_{ij} = \begin{cases} \nu(w_i b_j) & if w_i \to b_j \\ -\nu(w_i b_j) & if b_j \to w_i \\ 0 & else \end{cases}$$

WLOG suppose G has 2n vertices, with n of these vertices white and n black. If this is not the case, there will be no dimer configurations on G. Consider the determinant of K:

$$det(K) \equiv \sum\nolimits_{\sigma \in S_n} Sign(\sigma) K_{1\sigma(1)} K_{2\sigma(2)} ... K_{n\sigma(n)}$$

By inspection, each nonzero term in the determinant corresponds to the Boltzmann weight of a dimer configuration, up to an overall sign. Furthermore, it is also clear that for every distinct dimer configuration M on G there exists a unique nonzero term in the determinant.

Thus if we can prove the existence of an "admissable" orientation on the edges of G such that the signs of all terms in the determinant are the same, the proof will be completed. This is not extremely difficult, but it is fairly involved and not the most enjoyable proof to work through. The interested reader can refer to [6] for this part of the proof. The treatment of this point there is understandable and explicit.