DIMER MODELS AND THEIR CHARACTERISTIC POLYGONS

Timothy Fong Nam Chan

Supervisor: Associate Professor Daniel Chan

School of Mathematics and Statistics
UNSW Australia

November 2016

Submitted in partial fulfillment of the requirements of the degree of
Bachelor of Science with Honours

I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, nor material which to a substantial extent has been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis.

I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.

Timothy Fong Nam Chan

Abstract

A dimer model is a bipartite graph embedded on the torus. First introduced to study crystal physics in the 1960s, the literature has recently been revived to study algebro-geometric aspects of string theory. In this field, a result by Broomhead tells us that sufficiently "nice" dimer models encode algebras that resolve toric singularities.

In this thesis we study dimer models from a purely combinatorial perspective. We make rigorous what is meant by the word "nice" by defining *geometric consistency*, and proceed to investigate properties of dimer models that meet this condition. We introduce the *characteristic polygon* of a dimer model, which is the key to putting Broomhead's result in context. Finally, we develop a graph algorithm that generates a geometrically consistent dimer model with any fixed characteristic polygon.

Acknowledgements

My greatest gratitude goes towards my supervisor this year, the one and only DChan[1], without whom I would have achieved very little over the past year. Even though I kept him in the dark about my progress many times when I should not have, his insights and his ability to see the forest through the trees often resolved problems that I had spent hours on. I know that he is greatly respected by students of all year-levels at UNSW, and I will strive to take his numerous lessons to heart as I move onto the next phase of my life.

Of course, the supportive atmosphere at the university is not made by any one individual. To my friends in the mathematics student cohort whom I have had the pleasure of getting to know this year, I thank you. To the support staff who have had to put up with me, I thank you. To the lecturers who inspire me with their dedication and enthusiasm, I thank you. In particular, I'd like to show my gratitude to Ian Doust and David Harvey for a most enjoyable introduction to higher-level mathematics, to Catherine Greenhill for the fantastic guidance she has given as honours coordinator, and to Thomas Britz for his ever-friendly encouragement. To my girlfriend, who had the nerve to ask to be included in the acknowledgements despite being the only influence to have made my thesis worse, I decidedly do not thank you. Lastly, I would like to recognize Daniel Gulotta at Princeton and Nathan Broomhead at Hannover for graciously allowing me to use their diagrams, and Tarig Abdelgadir at UNSW for his helpful correspondence.

And of course, a tip of the hat to you, the reader, for surely being one of the only people who will read my work, or at least make it past the abstract. As this thesis is a learning experience for me, any feedback is most appreciated and can be sent to `timothy.fn.chan@gmail.com`.

[1] Associate Professor Daniel Chan, School of Mathematics and Statistics, UNSW.

Contents

Chapter 1

Overview

1.1 Background and aim

Dimer models are embeddings of bipartite graphs on a surface, whose study lies in the intersection of mathematical physics, algebraic geometry, and combinatorics. First defined in 1967 by Kasteleyn, a physicist, he and others successfully used them as a mathematical model of molecular interactions in statistical mechanics. Of particular importance to their goals was the study of *perfect matchings*, from which one could define, for each dimer model, an object called the *characteristic polygon*.

The literature in statistical physics ran its course, but in recent years has been revived among the mathematical physics community after connections with string theory were discovered. In particular, dimer models shed light on the hugely influential AdS-CFT conjecture in theoretical physics. In this field, modern authors are mostly interested in graphs embedded on the torus (as opposed to the historical interest in planar embeddings).

String theory usually posits the existence of additional dimensions of spacetime that take the form of structure-rich manifolds, so the work of physicists has in turn drawn the attention of algebraic geometers. In both mathematics and physics, it turns out that only sufficiently "nice" dimer models are of use. Although there are many conflicting definitions on what constitutes a "nice" dimer model, we will stick to a definition known as *geometric consistency*. In 2010, Broomhead [1] showed that geometrically dimer models encode algebras that resolve *toric singularities*. Moreover, this toric singularity is uniquely determined by the dimer model's characteristic polygon. Hence, geometrically consistent dimer models can be used to study a particular infinite class of singular toric varieties called *Gorenstein affine toric threefolds*.

A natural question to ask here is "does every Gorenstein affine toric threefold admit a resolution that can be obtained from a dimer model?" In other words, *can we construct a geometrically consistent dimer model for an arbitrary characteristic polygon?*" The primary aim of the thesis is to answer this question using an algorithm discovered by Gulotta [2], although the road to get there is quite long.

A secondary aim of the thesis is to provide a coherent synthesis of existing research from mathematics and physics, communities in which the dimer model is presented quite differently in both content and writing style. One of the major differences is that algebraic geometers prefer to work not with dimer models *per se*, but rather quivers that are obtained from dimer models via a canonical dualisation process in graph theory. A by-product of this is that many results concerning dimer models have a corresponding dual result concerning quivers, but it takes a strong familiarity with the content to be able to translate between them. In order to provide a unified approach to the topic, in this thesis we present all results and proofs from the perspective of the dimer model, making no reference to the quiver except as part of the introductory material.

Lastly, particular emphasis is placed on providing a mathematically acceptable level of rigour for any material that is taken from the physics literature. This is most evident in Chapter 4, entire sections of which are devoted to filling in details omitted by the physicists.

1.2 Structure of the thesis

In Chapter 2, we introduce the basic definitions of a dimer model and its *universal cover*, along with some concepts in topology. We learn how to associate to each dimer model its *characteristic polygon*, which is constructed from the set of *perfect matchings* of the dimer model through a mechanism called the *height function*. Characteristic polygons will become an important object of study in the thesis, and we observe that they are uniquely determined by their *primitive side segments*. Lastly, we explain how to get from a dimer model to the quiver that algebraic geometers prefer to use, and present the result of Broomhead that gives the motivating question behind this thesis.

In Chapter 3, we define the notion of *geometric consistency*, a notion central to Broomhead's result, based on the intersection properties of paths on a dimer model called *zigzag paths*. After investigating some implications of geometric consistency, we associate to each vertex of a dimer model a *zigzag fan*, which encodes information about which zigzag paths pass through the vertex. A considerable amount of time is spent expounding on the properties of these fans. From this information, we are able to use zigzag paths to construct some perfect matchings of (geometrically consistent) dimer models, which gives bounds on their characteristic polygons. In fact, we construct enough perfect matchings and obtain enough bounds to completely determine the characteristic polygon of any geometrically consistent dimer model. Thus we are able to find an alternative description of characteristic polygons that relies on zigzag paths instead of perfect matchings. To be more specific, the concluding theorem of the chapter proves that the *winding numbers* of zigzag paths on a dimer model are simply related to the primitive side segments of the dimer model's characteristic polygon. We call this the *zigzag-PSS correspondence*.

In Chapter 4, we use this alternative description to answer our key question of how to find a geometrically consistent dimer model for any given characteristic polygon. We begin by defining the *zigzag path diagram* of a dimer model, an object that simplifies the structure of a dimer model's zigzag paths and puts the zigzag-PSS correspondence into a framework amenable to algorithmic methods. We also learn to invert this procedure; that is, to obtain dimer model from a zigzag path diagram. We are then able to define some *operations* that act on zigzag path diagrams (or, equivalently, dimer models) and have predictable effects on their characteristic polygons. Repeatedly applying these operations in a specific order defined by the *Stern-Brocot sequences*, we develop an entire algorithm that accepts a convex polygon X as its input and whose output is a dimer model whose characteristic polygon is X. We further prove that the output of the algorithm is geometrically consistent, providing a complete affirmative answer to our motivating question.

The main sources of reference are Broomhead [1] and Gulotta [2], although background, insights, and inspiration are drawn from a wide variety of authors. We assume that the reader has no more than a standard undergraduate education in mathematics, and no knowledge of algebraic geometry or physics is required at all. Original contributions include:

- Dualisation of the results of Broomhead from the quiver to the dimer case. This involves a number of entirely new proofs.
- A considerably longer and completely restructured presentation of the algorithm of Chapter 3. In particular, a much more detailed exploration of the zigzag path diagram and Stern-Brocot ordering.
- Filling in a number of deficiencies in the original proof of the algorithm.
- Extending the proof of the algorithm to show that it produces geometrically consistent dimer models (although this fact was observed by Broomhead).

Chapter 2

The dimer model

In this chapter we define the dimer model and its associated *universal cover*, which allows us visualise dimer models on the plane instead of on the torus. We show that for any dimer model, we can associate to it a *characteristic polygon*. We expound on some of the motivation alluded to in the introduction, and present the construction of the quiver of a dimer model, an object of central importance to the literature in toric geometry.

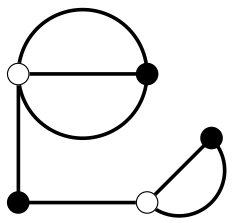
2.1 Basic definitions

Definition 2.1.1. A *graph* is an ordered triple $G = (V, E, I)$, where

- V is a finite set, called the *vertices* of G,
- E is a finite set, called the *edges* of G, and
- $I : E \to V \times V$ assigns to each edge an unordered pair of *endpoint* vertices.

A *bipartite graph* is a graph G whose vertex set can be partitioned into two sets (B, W) such every edge has an element of B and an element of W as an endpoint. We call B the *black* vertices and W the *white vertices* of G.

The reader is hopefully familiar with the standard way of visualising graphs; the figure below shows an example of a bipartite graph. Note that our definition of a graph differs from the standard definition since we allow multiple edges and loops (this type of structure is usually called a *multigraph* in the graph theory literature). For bipartite graphs, the endpoints of any edge are always distinct, so in particular, bipartite graphs do not have loops.



Definition 2.1.2. A *dimer model* (or simply *dimer*) D is a toric embedding of a bipartite graph that tiles the torus $\mathbb{T}$. In other words, it is a drawing of a graph on $\mathbb{T}$ such that:

1. Every edge is a simple curve of finite length.
2. No vertex lies on an edge unless it is an endpoint of that edge.
3. Edges do not cross.
4. D breaks the torus into regions homeomorphic to the unit ball of $\mathbb{R}^2$.

We use the following standard terminology throughout this thesis:
A *face* of D is one of the regions homeomorphic to the unit ball of $\mathbb{R}^2$ that D breaks the torus into.
The *boundary* of a face f is the set of edges that border f.
The *interior* of an edge e is the set of points t in $\mathbb{T}$ such that t lies on e but not one of its endpoints.
An edge incident with a vertex v (or simply an edge of v) is an edge that has v as an endpoint.
The *degree* of a vertex is the number of edges incident with it.
A *path* p on D is an alternating sequence of vertices and edges $(v_0, e_0, v_1, e_1, \ldots, v_{k-1}, e_{k-1}, v_k)$ that begins and ends in a vertex, such that the endpoints of e_i are v_i and v_{i+1}. Hence p "travels" along edges from one vertex to the next.[1]

We remind the reader of some basic topology, which can be found in most undergraduate texts such as [8]. The torus can be constructed by taking the unit square $[0,1] \times [0,1]$ and identifying opposite edges as in Figure 2.1. Equivalently, the torus can be described as the quotient of the plane $\mathbb{R}^2$ by the integer lattice $\mathbb{Z}^2$. We denote the quotient map by $\pi : \mathbb{R}^2 \to \mathbb{R}^2/\mathbb{Z}^2 = \mathbb{T}$, which is a map of oriented 2-manifolds.

Figure 2.1: Left: One way to think of the torus. Right: A dimer model represented in this way.

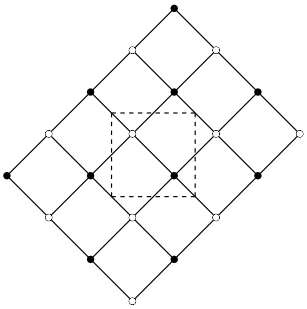
Hence, for any particular dimer model D, there is a unique doubly periodic planar graph (this time with infinitely many edges and vertices) $\widetilde{D}$ for which $D = \pi(\widetilde{D})$. This infinite graph

[1]Our definition of a *path* differs from the usual definition in graph theory, as we allow repeated edges and vertices (such an object is usually called a *walk*). In this sense, our definition matches the topological definition of a path moreso than the graph theoretic definition.

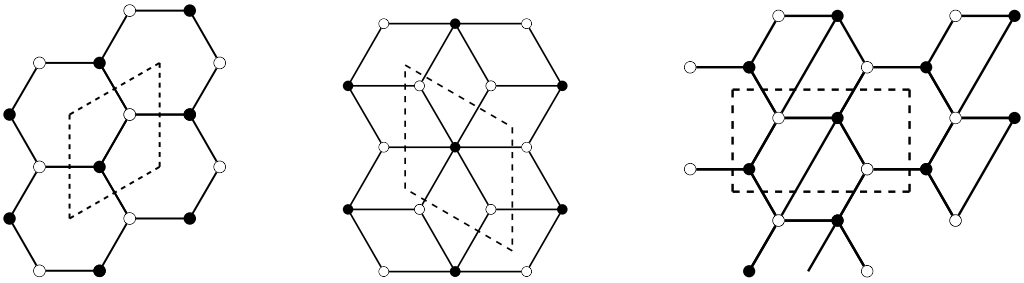
tiles the plane, and is called the *universal cover* of D. We call the unit square of this universal cover the *fundamental domain* of $\widetilde{D}$, as it captures all information about D and $\widetilde{D}$.

The universal cover provides a convenient way to not only visualise, but study dimer models as we shall see in Chapter 3. To go from the universal cover to the dimer model, we can *project* from $\mathbb{R}^2$ down to $\mathbb{T}$ using π. We can also *lift* a point $t \in D$ to any point $\widetilde{t}$ of $\widetilde{D}$ that satisfies $\pi(\widetilde{t}) = t$. Similarly, we can lift curves from D to $\widetilde{D}$. In particular, we can lift vertices, edges, and paths on D. It is easy to check that for any $(a, b) \in \mathbb{Z}^2$, the map $T : \mathbb{R}^2 \to \mathbb{R}^2$ corresponding to translation by (a, b) is homeomorphism such that $\pi(T(s)) = \pi(s)$ for any $s \in \mathbb{R}^2$. This translation defines a *deck transformation*, and allows us to move between different elements of $\pi^{-1}(u)$ for any $t \in \mathbb{T}$.

Example 2.1.3. The universal cover of the dimer model from Figure 2.1. We will use this dimer model (which we call the "square dimer") as an example that runs through the thesis, as it will turn out to be quite important in Chapter 4.



Some further examples of dimer models and their universal covers are shown below. In these examples, we have performed a change of basis in $\mathbb{Z}^2$, which is why our fundamental domains don't look like the unit interval.



2.1.1 Windings and crossings

We will only need a little bit more topology for our purposes. Again, details can be found in [8]. Recall that in informal terms, the *winding number* or simply *winding* of a curve measures the amount to which the curve wraps around an object. In the context of a oriented, closed curve γ on the torus, γ can be considered as an element of the fundamental group $\pi_1(\mathbb{T})$, and the winding number $[\gamma]$ is an element of $\mathbb{Z}^2$ that counts the number of times γ goes "around"

the two axes of the torus (with respect to a fixed orientation of $\mathbb{T}$). Any two oriented, closed curves that are homotopic (i.e. continuous deformations of each other) share the same winding number; in this sense, the winding number of γ is homotopy-invariant. We say that a winding number $(a, b) \in \mathbb{Z}^2$ is *primitive* if a and b are relatively prime.

Figure 2.2: Left: A path on $\mathbb{T}$ with winding $(1, 0)$. Right: A path on $\mathbb{T}$ with winding $(0, 1)$.

The following fact will not be proved.

Lemma 2.1.4 (Chapter 4, Lemma 2.1, [9])**.** *Let γ be an oriented, simple, closed curve on the torus with non-zero winding (a, b). Then its winding is primitive.*

Definition 2.1.5. Let γ_1 and γ_2 be closed curves on $\mathbb{T}$ with winding numbers (a, b) and (c, d) respectively. The *signed crossing number* $C_s : \pi_1(\mathbb{T}) \times \pi_1(\mathbb{T}) \to \mathbb{Z}$ is defined by $C_s(\gamma_1, \gamma_2) = ad - bc$.

The signed crossing number is antisymmetric and by definition, continuous deformations of γ_1 and γ_2 do not change the value of $C_s(\gamma_1, \gamma_2)$.

To give some intuition for the significance of the signed crossing number, suppose γ_1 and γ_2 are oriented, closed, simple paths that intersect transversally and only at a finite number of points. Define a point of intersection t to be *positively oriented* (resp. *negatively oriented*) of γ_1 and γ_2 if the tangent vector to γ_2 is between 0 and π radians anticlockwise (resp. clockwise) from the tangent to γ_1 at t. Then the physical interpretation of $C_s(\gamma_1, \gamma_2)$ is that is the number of positively oriented intersections of γ_1 and γ_2, minus the number of negatively oriented intersections.

In this case, we further define the *unsigned crossing number* $C(\gamma_1, \gamma_2)$ as the number of times γ_1 and γ_2 intersect (regardless of orientation). Note that this is *not* a homotopy-invariant quantity; Figure 2.3 shows how one may continuously deform two curves and reduce the number of unsigned crossings between them. Clearly, $|C_s(\gamma_1, \gamma_2)| \leq |C(\gamma_1, \gamma_2)|$, with equality holding if and only if every intersection of γ_2 with γ_1 has the same orientation.

Definition 2.1.6. If we have $|C_s(\gamma_1, \gamma_2)| < |C(\gamma_1, \gamma_2)|$, then we say that γ_1 and γ_2 have *extra crossings*.

Essentially, if γ_1 and γ_2 have extra crossings, then they intersect more than the minimal amount of times required by their winding numbers. A typical example of the sort of structure that gives rise to extra crossings is shown in the left of Figure 2.3.

Figure 2.3: Left: An example of extra crossings. Right: A continuous deformation of the curves that reduces the unsigned crossing number.

2.2 The characteristic polygon

A central object in this thesis is is the *characteristic polygon* of a dimer model, otherwise known in physics and mathematics literature as the *perfect matching polygon*, *toric diagram*, or *lattice polygon*. The definition requires a bit of prior machinery, which we develop first. The first piece of machinery is that of perfect matchings, which have always been important objects of study in graph theory, particularly for bipartite graphs like dimer models; although we will not be using it, Hall's Marriage Theorem [3] is a celebrated and foundational result in this area.

Definition 2.2.1. A perfect matching of a dimer D is a subset of the edges of D such that every of vertex of D is contained in exactly one those edges. The set of perfect matchings of D is naturally identified with the set of doubly periodic perfect matchings on $\widetilde{D}$.

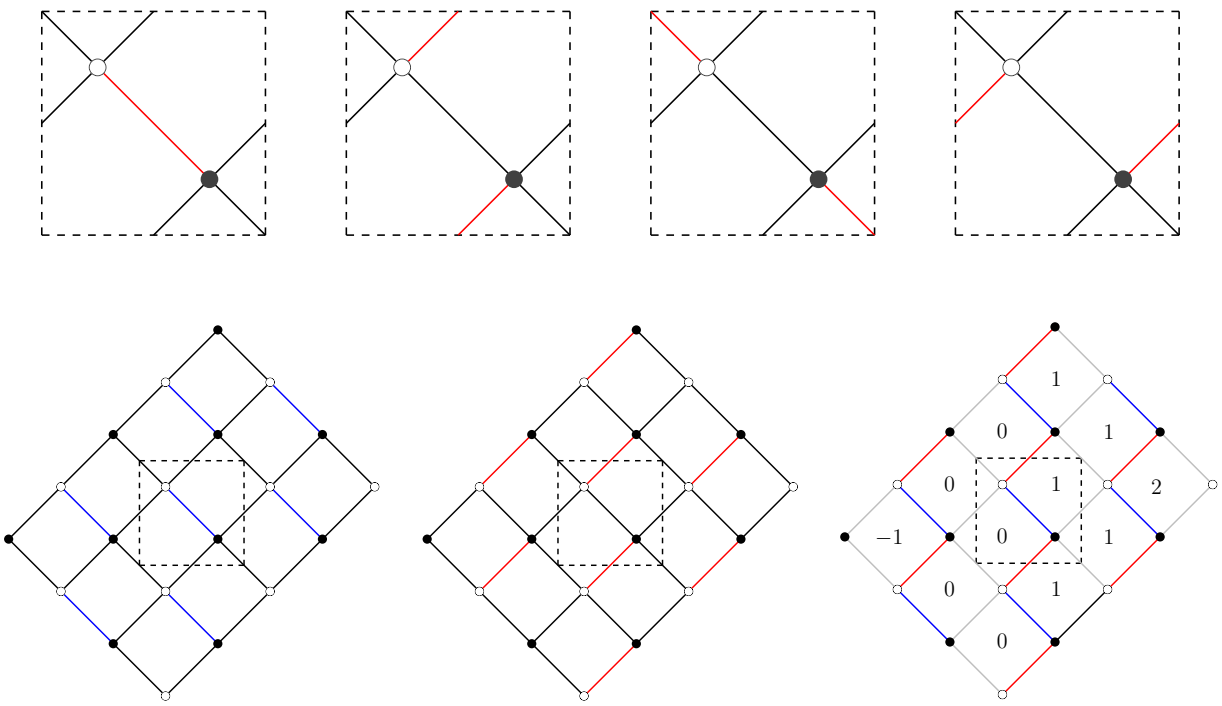
Fix a perfect matching M_0 and call it the *reference matching*. For any matching M, the symmetric difference (or simply the *difference*) $M \ominus M_0 = (M - M_0) \cup (M_0 - M)$ of M and M_0 partitions $\mathbb{R}^2 \backslash (M \ominus M_0)$ into a (possibly trivial, possibly infinite) number of regions that are disconnected from each other.

Definition 2.2.2. The height function $h_{M,M_0} : \mathbb{R}^2 \backslash (M \ominus M_0) \to \mathbb{Z}$ of M with respect to M_0 is a function that is constant on each connected region of $\mathbb{R}^2 \backslash (M \ominus M_0)$, and which increases (resp. decreases) by 1 when one crosses between regions through an edge $e \in M_0$ with a black (resp. white) vertex on the right, or when one crosses an edge $e' \in M$ with a black (resp. white) vertex on the left. This determines the height function up to addition of a constant.

We should check that this is well-defined (i.e. it does not matter which edge we choose to cross when moving between regions). Let e_0 and e_1 be consecutive edges in the boundary between the regions B_1 and B_2 of $\mathbb{R}^2 \backslash (M \ominus M_0)$. By the definition of a perfect matching, e_0 and e_1 cannot be part of the same perfect matching. Suppose without loss of generality that e_0 is in M_0 and e_1 is in M. If the crossing from B_1 to B_2 through e_0 has a black (resp. white) vertex on the right, then the crossing from B_1 to B_2 through e_1 has a black (resp. white) vertex on the left. Then the change in the height function is the same regardless of which edge is crossed. Applying this argument inductively, it does not matter at all which edge in the boundary of B_1 and B_2 we cross; we will always get the same change in the height function.

Example 2.2.3. The square dimer has exactly 4 perfect matchings, corresponding to the choice of a single edge of the dimer model. For the first two of these matchings, the associated perfect

matchings of the universal cover are shown. Fixing the reference matching to be the blue matching, we obtain their symmetric difference and height function.



Definition 2.2.4. Pick $t \in \mathbb{R}^2 \backslash (M \ominus M_0)$ and let $\widetilde{D}$ be the universal cover of a dimer model with the unit square $[0,1] \times [0,1]$ as its fundamental domain. Let $t_x = t + (1,0)$ and $t_y = t + (0,1)$. The *height change* $h(M, M_0) = (h_x(M, M_0), h_y(M, M_0)) \in \mathbb{Z}^2$ of M with respect to M_0 is defined by

$$h_x(M, M_0) = h_{M,M_0}(t_x) - h_{M,M_0}(t)$$
$$h_y(M, M_0) = h_{M,M_0}(t_y) - h_{M,M_0}(t)$$

It does not depend on the choice of t, and we can naturally consider height changes as an element of the lattice $\mathbb{Z}^2$.

To see that the height change is independent of our choice of t, let u be any other point in $M \ominus M_0$. As we did with t, let $u_x = u + (1,0)$ and $u_y = u + (0,1)$. Construct a curve p in $\mathbb{T}$ from t to u that crosses edges of $\widetilde{D}$ transversally. Then the height function at u differs from the height function at t by some integer C. Applying the deck transformation corresponding to translation by $(1,0)$ to p, we obtain a curve from t_x to u_x that crosses edges of M and M_0 in the same manner that p does. Hence the height function at u_x differs from the height function at t_x by the same integer C. Similarly, translating p by $(0,1)$ we find that the height function at u_y differs from the height function at t_y by C. Hence the height change in the same regardless of whether we compute it at t or u.

One other note is that since $\widetilde{D}$ has a periodic structure, an inductive argument shows that for $(a, b) \in \mathbb{Z}^2$,

$$h_{M,M_0}(p + (a, b)) - h_{M,M_0}(p) = ah_x(M, M_0) + bh_y(M, M_0)$$

The dependence of the height change on the choice of reference matching is given by the following lemma.

Lemma 2.2.5. *For any two perfect matchings M and M_1 combined with a reference matching M_0, we have the relationship*

$$h(M, M_1) = h(M, M_0) - h(M_1, M_0).$$

In other words, if we change our reference matching from M_0 to M', we simply shift all our height changes by a constant.

Proof. Simple case-by-case analysis of what happens to the height function when we cross an arbitrary edge $e \in D$. Suppose that we cross e with a *black* vertex on our right. If e is in M but not M_1 or M_0, then $h(M, M_1)$ decreases by one, $h(M, M_0)$ decreases by one, and $h(M_1, M_0)$ is unchanged. If e is in M and M_0 but not M_1, then $h(M, M_1)$ decreases by one, $h(M, M_0)$ is unchanged, and $h(M_1, M_0)$ increases by one. And so on, until we have exhausted all eight possibilities and find that the relationship holds for each. Then we exhaust the eight possibilities in the symmetric case that we cross e with a *white* vertex on our right. □

Definition 2.2.6. The *M_0-characteristic polygon* X_{M_0} of a dimer D with respect to a reference matching M_0 is the convex hull of (i.e. the smallest convex polygon containing) the set

$$\{h(M, M_0) : M \text{ is a matching of } D.\}$$

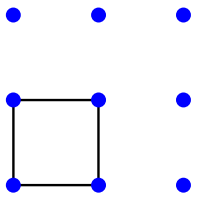
As noted in Lemma 2.2.5, a different choice of M_0 causes a translation of the resulting polygon, so hereafter we will only talk about X, the characteristic polygon of a dimer model, defined up to translation.

Since our polygon no longer depends (up to translation) on M_0, we will at times hide the reliance of $h(M, M_0)$ on M_0 by simply writing $h(M)$. Of course, we must remember that when we compute h for varying M, we must always use the same reference matching. We may also use M to mean either a perfect matching or its associated point in $\mathbb{Z}^2$, making statements such as "M lies in the characteristic polygon"; the context will always make it clear which we are referring to.

In the case of Example 2.2.3, for our chosen blue and red perfect matchings we obtain a height change of $(1, 0)$. Computing the other height changes with respect to the same blue perfect matching, we get a set of height changes

$$\{(0, 0), (1, 0), (0, 1), (1, 1)\},$$

from which we obtain a 1×1 square as the characteristic polygon. The blue dots denote lattice points of $\mathbb{Z}^2$.



The characteristic polygon X of a dimer model has *corners* and *sides* in the usual sense. An *extremal* perfect matching is one that lies on the boundary of X, and a *corner* perfect matching is one that lies on a corner of X. The *multiplicity* of point p in $\mathbb{Z}^2$ is the number of perfect matchings whose height change equals p. By definition of the characteristic polygon, it is nonzero only when p lies in X.

Definition 2.2.7. A *primitive side segment* (*PSS*) of a convex polygon X is an interval between two adjacent lattice points on the boundary of X. By giving the sides of the polygon an anticlockwise orientation, primitive side segments become vectors in $\mathbb{Z}^2$. The PSS set of X is the multiset formed by these vectors.

For brevity, if a PSS has vector (a, b), then we call it an (a, b)-PSS. An important observation is that for a convex polygon X, convexity forces the slopes of its sides to appear in a certain order around the polygon. In particular, if we normalise an (a, b)-PSS to a unit length vector

$$\frac{(a, b)}{\sqrt{a^2 + b^2}}$$

and consider it as an element of the unit circle, then the cyclic order of the PSSes around the unit circle must match their order as unnormalised vectors around X (up to cyclic permutation). We summarise this as follows:

Lemma 2.2.8. *A convex polygon is uniquely defined by its PSS set. The order that the PSSes appear around the polygon matches matches their cyclic order around the unit circle.*

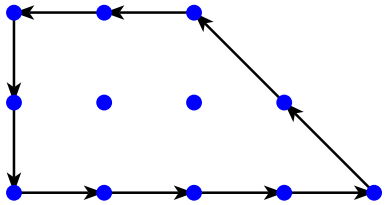


Figure 2.4: This convex polygon is uniquely defined by its ten PSSes: four with vector $(1, 0)$, two with vector $(-1, 1)$, two with vector $(-1, 0)$, and two with vector $(0, -1)$.

2.3 Motivation

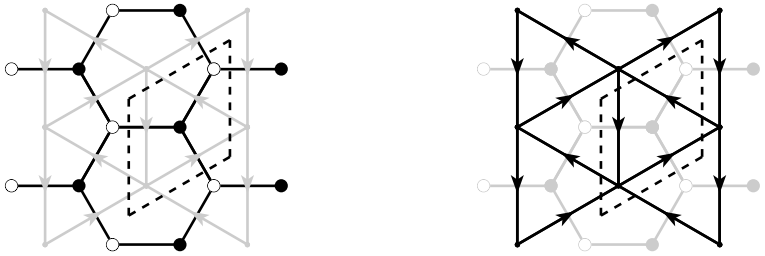
Why should the modern physicist or mathematician be interested in dimer models and characteristic polygons? As is often the case, they are examples of combinatorial objects that succinctly encode information about algebraic objects. For example, in theoretical physics, they are a convenient way of describing *conformal field theories*, which are frameworks through which one can model quantum mechanical behaviour. Using these dimer models, mathematical physicists can shed light on the conjectured anti-de Sitter/conformal field theory (AdS/CFT) correspondence, a highly influential conjecture in the field of high energy physics concerning the relationship between string theory and quantum field theory. To give an idea as to its importance, the article in which the conjecture first appeared ([20]) currently stands at over 12000 citations, making it the most cited article in high energy physics of all time.

The study of the AdS/CFT correspondence (and of string theory in general) requires a large amount of high-level mathematics. For example, modern formulations of string theory posit the existence of additional "folded up" dimensions beyond the familiar temporal and three spatial dimensions. This gives some intuition for why theoretical physicists need to work with manifolds that have a rich structure. These manifolds are described in algebraic geometry, which has drawn mathematicians to the study of dimer models.

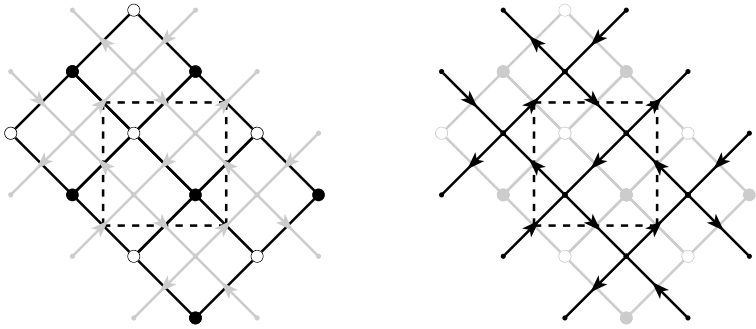
Actually, algebraic geometers do not usually work directly with a dimer model, but rather on a quiver that is *dual* to a dimer. Here, we briefly explain this construction, firstly so that we can give some context to an important motivating result of Broomhead's, and secondly to provide a more complete introduction to the literature.

Given a dimer model D, we can consider the canonical *dual graph* Q, which has a vertex dual to every face of D and a face dual to every vertex of D. Two vertices of Q are connected by an edge if and only if the vertices, considered as faces of D, were separated from each other by an edge of D. Hence the dual graph has an edge for every edge of D. A vertex of D with degree n becomes a face of Q with n edges in its boundary, and vice versa.

The dual graph Q is not bipartite, but the bipartite nature of D *does* allow us to impose additional structure on Q. More specifically, we can naturally impose a consistent orientation of the edges of Q to make Q into a quiver i.e. a graph with *directed* edges. To every face f



(a) Left: The "hexagonal" dimer. Right: The associated quiver.



(b) Left: The "square" dimer. Right: The associated quiver.

Figure 2.5: Two examples of dimer models are their associated quivers. Diagrams courtesy of Broomhead [1].

of Q, if f is dual to a black vertex of D, then we call f a *black face* and orient the edges in the boundary of f anticlockwise around the face. Conversely, if f is dual to a white vertex of D, then we call f a *white face* and orient the edges in the boundary of f clockwise around the face. Figure 2.5 gives the quivers of two important dimer models in the literature.

From a quiver Q associated to a particular dimer model, one can obtain a *superpotential algebra* A. We do not describe the construction here, but it is a quotient of the path algebra of Q. In [1], Broomhead proves that for dimer models meeting a certain condition called *geometric consistency* (which we will define in the next chapter), the algebra A can be used to *resolve* a particular sort of toric variety.

Theorem 2.3.1 (Theorem 1.4, [1])**.** *For a geometrically consistent dimer model, the superpotential algebra A obtained from it is a noncommutative crepant resolution (NCCR) of a Gorenstein affine toric threefold.*

One should not worry too much about the exact meanings of all the words here; in essence, the theorem tells us that dimer models can be used to study the features and behaviour of a class of toric singularities. *Threefolds*, referring to three-dimensional complex manifolds, are of interest because the number of additional (real) dimensions in many versions of string theory is six (in other words, three additional dimensions over $\mathbb{C}$).

Not only does Broomhead tell us that A is a resolution of *some* toric variety, he actually tells us *which one in particular*. In fact, this variety is *uniquely determined* by the characteristic

polygon of the dimer model.[2] Along with the work of Gulotta [2] and Stienstra [14], Broomhead is able to conclude the following:

Theorem 2.3.2 (Theorem 1.5, [1])**.** *Every Gorenstein affine toric threefold R is associated to a convex polygon X. R admits an NCCR that is the superpotential algebra of a geometrically consistent dimer model with characteristic polygon X.*

This naturally leads us to the main question that we ask in this thesis:

How can we construct geometrically consistent dimer models? More specifically, can we construct a geometrically consistent dimer model with a given characteristic polygon?

[2]This is due to something called the Calabi-Yau condition, see Broomhead's paper for details.

Chapter 3

Geometric consistency and the correspondence

Unfortunately, the question of whether we can construct a dimer model for any given characteristic polygon is not so easily answered. A big reason for this is that the set of perfect matchings is dependent on the global structure of the dimer model. In other words, there is no simple way of *finding* perfect matchings.

However, there is hope. In this chapter, we prove that for dimer models meeting the *geometric consistency* condition of Broomhead in [1], we *can* find a number of perfect matchings. Further, these perfect matchings lie on the corners of the characteristic polygon, giving rise to a correspondence between objects called the *zigzag paths* of a dimer model, and the PSSes of the associated characteristic polygon. We call this the *zigzag-PSS correspondence*. This will give us an alternative construction of the characteristic polygon that is more tractable than perfect matchings and the height function.

Much work is necessary to prove this correspondence, and we warn that the proofs in this chapter can be quite technical. A number of the proofs in this chapter are original due to the decision to present all content from the perspective of dimer models instead of quivers. It turns out that in many cases, direct dualisation of the proofs is not the most efficient course of action, as in some sense the dimer model is easier to work with. The reader is encouraged to refer to figures where available, or to draw their own. Although we develop some interesting theory in this chapter, the main takeaway for the purposes of answering our key question is the aforementioned zigzag-PSS correspondence (Theorem 3.4.9).

3.1 Zigzag paths & flows

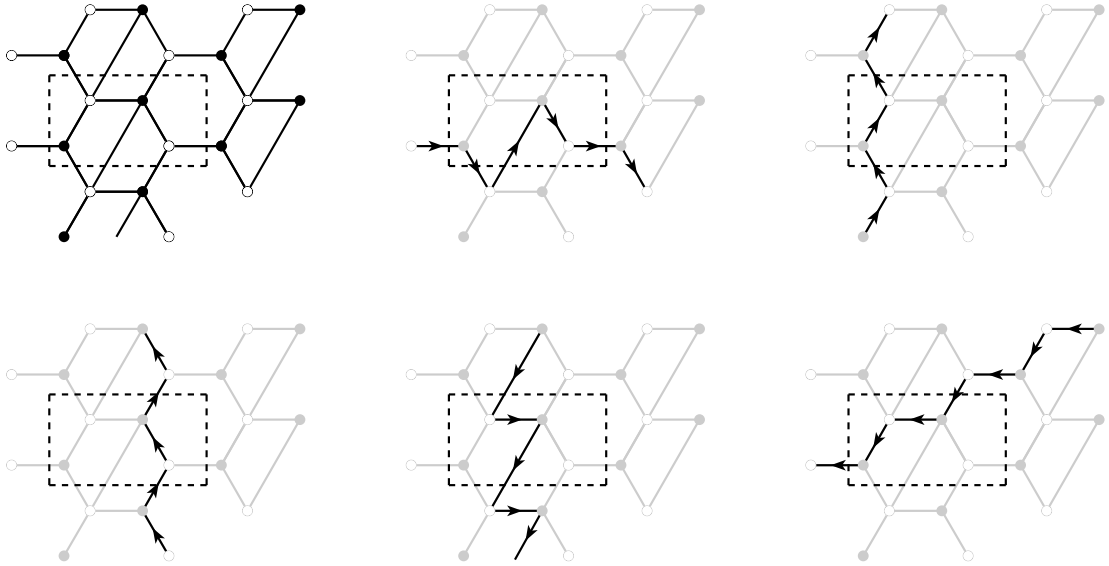
Definition 3.1.1. A *zigzag path* η is an (oriented, closed) path on D that turns maximally right at a black vertex, and maximally left at a white vertex. A *zigzag flow* $\widetilde{\eta}$ is a lifting of η

to $\widetilde{D}$. We say that an (oriented) edge e of a zigzag path or flow is a *zig* if it goes from a white to a black vertex, and a *zag* if it goes from a black to white vertex.

Here, we know that η is closed because there are only a finite number of edges in D for η to travel along. It is clear from the definition that a zigzag path (resp. flow) is completely defined by the choice of a single edge in the path (resp. flow) and a direction. The term *zig-zag pair* will denote a sequence (e_1, e_2) of two consecutive edges in a zigzag path or flow, where e_1 is a zig and e_2 is a zag of that path or flow. When the distinction between η and $\widetilde{\eta}$ is not important, we may simply refer to them collectively as a zigzag.

Since every zigzag path η is closed, it has a winding number $[\eta] \in \mathbb{Z}^2$. Of course, multiple zigzag paths can have the same winding number. We say that two zigzag flows $\widetilde{\eta}, \widetilde{\zeta}$ (or zigzag paths η, ζ) are *parallel* if $[\eta]$ is a positive scalar multiple of $[\zeta]$, and *anti-parallel* if it is a negative scalar multiple. In a slight abuse of notation, we may at times write $[\widetilde{\eta}]$ to mean $[\eta]$.

Example 3.1.2. The following dimer has 5 zigzag paths: one with winding number $(1, 0)$, two with winding number $(0, 1)$, one with winding number $(0, -1)$, and one with winding number $(-1, -1)$. For each of these zigzag paths, we illustrate a lifting of the path into $\widetilde{D}$, which forms a zigzag flow.



Remark 3.1.3. As explored in Section 2.3, in some areas it is preferred to work mainly with the dual of a dimer model. Dualising the edges of a zigzag flows $\widetilde{\eta}$ on $\widetilde{D}$, we get a a curve $\widetilde{\eta}^*$ on $\widetilde{Q}$ that alternates between going anticlockwise around a black face and clockwise around a white face. This defines a *quiver zigzag flow*. We also have $[\widetilde{\eta}] = [\widetilde{\eta}^*]$, since both paths are homotopic to the path composed of line segments connecting $\widetilde{\eta} \cap \widetilde{\eta}^*$. Almost all statements regarding dimer zigzag flows dualise very easily to the quiver case, although the proofs can require more thought.

Definition 3.1.4. Two (distinct) paths γ_1, γ_2 in $\widetilde{D}$ *intersect in an edge* if they share an edge. They *intersect in a vertex* (or simply *intersect*) if they share a vertex. Clearly, the former implies the latter.

In the case that the two paths are not distinct, of course γ_1 and γ_2 share an edge, but this does not capture any useful information about the topology of γ_1. In this case, we must take a bit more care and replace the concept of intersection in an edge with that of *self-intersection in an edge*, which amounts to saying that γ_1 goes through the same edge on two separate occasions.[1] Extending this, γ_1 *self-intersects in a vertex* if and only if γ_1 is not simple). We can provide analogous definitions of intersection and self-intersection in the case that we are dealing with paths of D instead of $\widetilde{D}$.

Lemma 3.1.5. *Let p be a finite, simple, closed path in $\widetilde{D}$. Any zigzag flow $\widetilde{\eta}$ that intersects p in an edge must intersect it in at least two edges.*

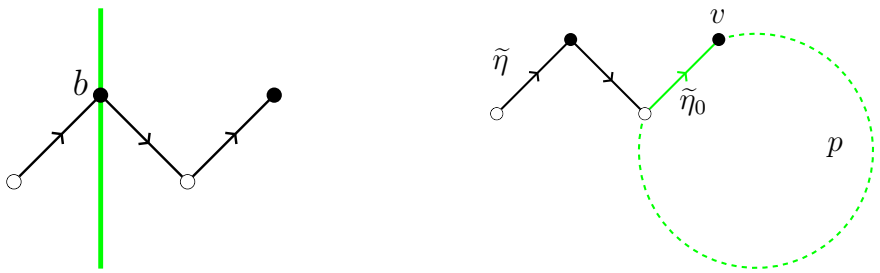
Proof. By the Jordan curve theorem, p breaks the plane into a region *interior* to p and a region *exterior* to p. The key observation is that $\widetilde{\eta}$ cannot cross through p without using an edge of p. This is because zigzag flows always make maximal turns in each direction (see Figure 3.1a).

Suppose $\widetilde{\eta}$ intersects p in at least one edge $\widetilde{\eta}_0$. Further suppose that $\widetilde{\eta}_0$ is a zig of $\widetilde{\eta}$, and goes clockwise around p (see Figure 3.1b). What is the edge $\widetilde{\eta}_1$ of $\widetilde{\eta}$ that is immediately after $\widetilde{\eta}_0$? We know that $\widetilde{\eta}_0$ intersects $\widetilde{\eta}_1$ at a black vertex v since $\widetilde{\eta}_0$ is a zig of $\widetilde{\eta}$. By definition, $\widetilde{\eta}$ makes a maximal right turn from $\widetilde{\eta}_0$ to $\widetilde{\eta}_1$ at v. Certainly, p intersects v in two edges, one of which is $\widetilde{\eta}_0$, the other of which we call e. If $\widetilde{\eta}_1$ is e, then $\widetilde{\eta}$ intersects p in at least two edges, and we are done. If it is not, then $\widetilde{\eta}$ must turn into the interior of p. There are only a finite number of edges in this region, so condition (a) of geometric consistency forces $\widetilde{\eta}$ to cross p a second time. Again, by our key observation above, this crossing must use an edge of p. Thus $\widetilde{\eta}$ crosses p at least twice.

The case that $\widetilde{\eta}_0$ is a zag that goes anticlockwise around p follows by a symmetric argument. The remaining cases (namely that $\widetilde{\eta}_0$ is a zig that goes anticlockwise around p, or that it is a zag that goes clockwise around p) follow by considering the edge $\widetilde{\eta}_{-1}$ that *precedes* $\widetilde{\eta}_0$ in $\widetilde{\eta}$. □

Actually, with a bit more thought, it is realised that not only does $\widetilde{\eta}$ intersect p in at least two edges, it must intersect p in an even number of edges in total. This is because intersections occur in pairs; every time $\widetilde{\eta}$ enters the interior of p, it must leave at some later stage. We will use this fact in a later proof.

[1] For a more formal definition of self-intersection, one way to think about γ_1 is as a map from a (not necessarily finite) subset of $\mathbb{Z}$ to the set of edges of $\widetilde{D}$. Then $\widetilde{\eta}$ self-intersects in an edge if the map is not injective.



(a) (b)

Figure 3.1: Left: No zigzag flow can cross the green path p without using an edge of $\widetilde{\eta}$, because otherwise $\widetilde{\eta}$ would by definition have turned onto p at the vertex b. Right: The edge of $\widetilde{\eta}$ following $\widetilde{\eta}_0$ must either turn into the interior of p (if such an edge exists), or continue to follow p.

Lemma 3.1.6. *Let $W = (e_1, e_2, \ldots, e_n)$ be a finite path in $\widetilde{D}$ from v to w. Then there exists a (finite) simple path P from v to w that is a subsequence of W.*

Proof. This is a standard trick in graph theory that deletes unnecessary loops, and will be proved by induction. If n is 0 or 1, then W is already simple. Now let n be arbitrary and suppose the claim is true for all paths with length less than n. Consider $W = (e_1, \ldots, e_n)$. If W does not self-intersect in a vertex, then W is already simple, and we are done. Otherwise, W self-intersects in a vertex x. Let W' be the subsequence of W formed by deleting the sequence of edges between the two repetitions of x in W. It is a path from v to w with length less than n. By the inductive assumption, W' has a subsequence P that forms a simple path from v to w. P is also a subsequence of W, so it meets the desired properties. □

3.2 Geometric consistency

Recall that in Chapter 1, we presented Theorem 2.3.1, which required that a dimer model be *geometrically consistent* in order to be useful. Now that we have introduced zigzag paths and flows, we can provide an exact definition of geometric consistency, which was introduced by Broomhead based on the work of Kenyon in [4].

Definition 3.2.1 (Dual of Definition 3.12, [1])**.** We say that a dimer model is *geometrically consistent* if the following conditions hold:

(a) No zigzag flow self-intersects in an edge.

(b) If $\widetilde{\eta}$ and $\widetilde{\zeta}$ are (distinct) zigzag flows with $\{[\eta], [\zeta]\}$ linearly dependent, then they do not intersect in an edge.

(c) If $\widetilde{\eta}$ and $\widetilde{\zeta}$ are (distinct) zigzag flows with $\{[\eta], [\zeta]\}$ linearly independent, then they intersect in exactly one edge.

Here, it is indeed important to distinguish between between zigzag paths and zigzag flows; for example, for condition (c), a path with winding $(2, 1)$ will always intersect a path with winding $(0, 1)$ at least twice on a torus, but their associated flows may only intersect once on the universal cover.

In essence, geometric consistency requires that zigzag flows behave like the straight lines of ordinary Euclidean geometry. The conditions of the definition match the notions that straight lines do not intersect themselves (condition a), parallel lines do not meet (condition b), and non-parallel lines intersect exactly once (condition c).

Remark 3.2.2. Many different alternative characterisations of *consistency* exist, for example the isoradiality condition first introduced by Duffin [21] and Merkat [22]. Broomhead [1] alone introduces 3 characterisations (consistency, algebraic consistency, and geometric consistency). See [12] for more examples. When reading the literature, one should take care not to confuse these conditions, especially because different conditions are often given the same name!

From this point onwards in the chapter, we assume that a dimer D is geometrically consistent unless otherwise stated.

Condition (a) of geometric consistency gives us a decent idea of how zigzag paths behave. They cannot self-intersect in an edge, but they *might* self-intersect in a vertex. Our first order of business is to *close this loophole* with a lemma that tells us that zigzag flows are simple curves.

Lemma 3.2.3. *Suppose v is a vertex of $\widetilde{D}$. If $\widetilde{\eta}$ is a zigzag flow that intersects v, then it does so in exactly one zig-zag pair.*

Proof. The following proof is original. Suppose v is black, as the general result follows from symmetry. By definition of zigzag flows, if a flow $\widetilde{\eta}$ intersects v then it does so in at least one zig-zag pair. Suppose $\widetilde{\eta}$ intersects v in two distinct zigzag pairs. By condition (a) of geometric consistency, these pairs must be disjoint. Give the edges incident with v an ordering $e_1, e_2, \ldots, e_n$ according to their cyclic order around v. Without loss of generality, let the first zigzag pair in which $\widetilde{\eta}$ intersects v be (e_1, e_2), and the second be (e_k, e_{k+1}), where $k > 2$ from disjointness. We can further choose these zigzag pairs such that $\widetilde{\eta}$ does not intersect v between e_2 and e_k. Now, e_2 is a zag of $\widetilde{\eta}$, but it is also a zig of a second flow $\widetilde{\eta}'$. This flow is distinct from $\widetilde{\eta}$ by condition (b) of geometric consistency.

Let v' be the endpoint of e_2 that is not v, and v'' be the endpoint of e_k that is not v. Let p' be the finite path in $\widetilde{D}$ that starts at v' and follows $\widetilde{\eta}$ until it meets v''. By Lemma 3.1.6, we can simplify p' to a finite simple path p from v' to v''. Then (unoriented) path $P = e_2 \circ p \circ e_k$ is a finite simple closed path. We know $\widetilde{\eta}'$ intersects P at e_2, so Lemma 3.1.5 tells us that it intersects P in a second edge. But P is a subset of $\widetilde{\eta}$, so $\widetilde{\eta}'$ must intersect $\widetilde{\eta}$ in at least two edges, contradicting geometric consistency.

□

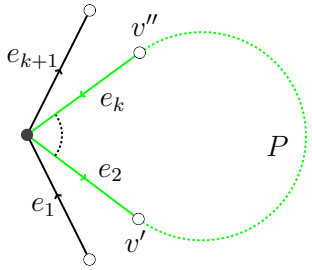


Figure 3.2: The green path P is finite, closed, and simple. It intersects $\widetilde{\eta}'$ at e_2.

A useful notion that geometric consistency gives us is that every zigzag flow $\widetilde{\eta}$ has a globally defined left side and right side. More precisely, we have a partition of the set of vertices of $\widetilde{D}$ (including the vertices of $\widetilde{\eta}$) into two subsets as follows.

Definition 3.2.4. The *left* (resp. right) of $\widetilde{\eta}$ consists of vertices that are connected to a *black* (resp. white) vertex of $\widetilde{\eta}$ via a finite path in $\widetilde{D}$ that does not intersect $\widetilde{\eta}$ in an edge. Our definitions of the left and right of $\widetilde{\eta}$ matches our intuitive notion of left and right for an observer travelling along the flow (see figure 3.3a).

We check that every vertex $v \in \widetilde{D}$ does indeed lie in one of these two sets. The condition that $\widetilde{D}$ tiles the plane forces the existence a finite path p from v to b for any black vertex of b of $\widetilde{\eta}$. Either this path does not intersect $\widetilde{\eta}$, or it does. If it does not, then v is on the left of $\widetilde{\eta}$ and we are done. Otherwise, consider the segment of p between v and the first vertex of p that lies in $\widetilde{\eta}$. Then this is a finite path that does not intersect $\widetilde{\eta}$ in an edge so v lies on either the left or right of $\widetilde{\eta}$.

The following proof is original. To see that the left and right of $\widetilde{\eta}$ are disjoint sets, suppose for the sake of contradiction that v is on both the left and right sides of $\widetilde{\eta}$. Then v is connected to a black vertex b of $\widetilde{\eta}$ via a finite path p_1 that does not intersect $\widetilde{\eta}$ in an edge. Similarly, v is connected to a white vertex w of $\widetilde{\eta}$ via a finite path p_2 that does not intersect $\widetilde{\eta}$ in an edge. Now p_1 and p_2 intersect in finitely many vertices, and the last of these vertices is on the left and right of $\widetilde{\eta}$ if and only if v is, so we can without loss of generality suppose that p_1 and p_2 do not intersect in an vertex other than v.

We can, moreover, suppose without loss of generality that p_1 does not contain any black vertex b' of $\widetilde{\eta}$ that is distinct from b (otherwise, we could simply set $b = b'$ to obtain a shorter path), and that it does not contain any white vertex of w' of $\widetilde{\eta}$ other than possibly v (otherwise, we could set $v = b$ and $w = w'$ to change p_1 to a trivial path). Similarly, we can suppose that p_2 does not contain any vertex of $\widetilde{\eta}$ other than w and possibly v. Using Lemma 3.1.6, we can still further suppose without loss of generality that p_1 and p_2 are finite and simple.

Finally, let p_3 be path along $\widetilde{\eta}$ between b and w. It is simple by Lemma 3.2.3, and is composed of an odd number of edges since it goes from a black to a white vertex. Consider the (unoriented) path $P = p_1 \circ p_3 \circ p_2$. As discussed above, the only intersections between the p_i are at their starting and ending points, so P is finite, simple, and closed. By the discussion

following Lemma 3.1.5, the number of edges in which $\widetilde{\eta}$ and P intersect is even. An odd number of these intersections are just the edges of e_3, but this means that there is at least one more edge of $\widetilde{\eta}$ in which $\widetilde{\eta}$ intersects P. This edge must be in either p_1 or p_2, contradicting our original assumption that they do not intersect $\widetilde{\eta}$ in an edge. Hence the left and right of $\widetilde{\eta}$ are disjoint.

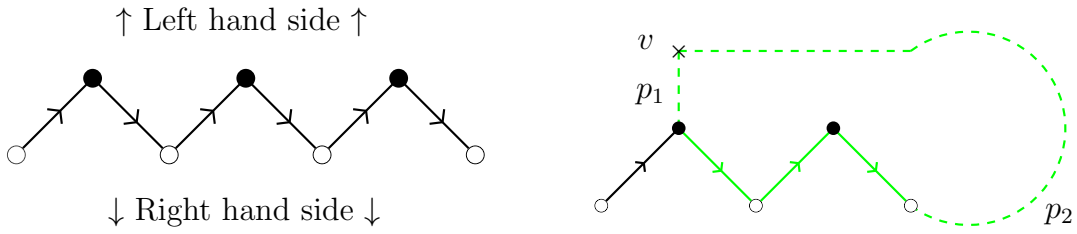


Figure 3.3: Left: Physical meaning of the left and right of a zigzag flow. We consider the black vertices to be part of the left, and the white vertices to be part of the right. Right: The green path P must intersect $\widetilde{\eta}$ at least one more time than is illustrated.

3.3 Zigzag fans

We mentioned at the beginning of the chapter that we will prove a correspondence between the zigzag paths of a dimer model, and the PSSes of the associated characteristic polygon. While this might at first sound unintuitive, the key observation as follows.

Suppose $\widetilde{\eta}$ is a zigzag flow. It is within the realm of possibility that there exists a perfect matching of $\widetilde{D}$ that uses every second edge of $\widetilde{\eta}$. If such a perfect matching M_0 exists, then how does it affect the height function? Suppose $[\eta] = (1, 0)$, as illustrated in Figure 3.4a. Consider the path p in $\mathbb{R}^2$ immediately to the left of $\widetilde{\eta}$, as in Figure 3.4b. Then, going along p from some point $t \in p \backslash \widetilde{D}$ to $t + (1, 0)$, we never cross an edge of M_0, and every edge we *do* cross has a black vertex on its right. Hence for any other perfect matching M, the height function h_{M,M_0} can only decrease as we go along p. This forces $h_x(M, M_0)$ to be negative, creating a bound on the characteristic polygon that we will make explicit in Section 3.4.

Figure 3.4: Left: A zigzag flow $\widetilde{\eta}$, highlighted in red. Right: A perfect matching (highlighted in blue) that uses every second edge of $\widetilde{\eta}$, and the path in $\mathbb{R}^2$ immediately to the left of $\widetilde{\eta}$ (highlighted in green).

First, though, we have to figure out how to construct such a perfect matching M_0. The problem remains that perfect matchings are *global* objects, and structures in one area of a graph can influence perfect matchings on a global scale. This motivates us to the development of *zigzag fans*, which are objects that encode the information about the zigzag paths that pass through a particular vertex of a dimer model. The goal here is to find a way to construct perfect matchings that does not require any knowledge of the global structure of D beyond geometric consistency, i.e. to reduce a *global* problem to a *local* one.

Definition 3.3.1. Let v be a vertex in $\widetilde{D}$. The *local zigzag set* of v is

$$\chi(v) = \{\widetilde{\eta} : v \text{ is a vertex of } \widetilde{\eta}\}.$$

The *local zigzag fan* of v is

$$\xi(v) = \{[\eta] : \widetilde{\eta} \in \chi(v)\} \subset \mathbb{R}^2.$$

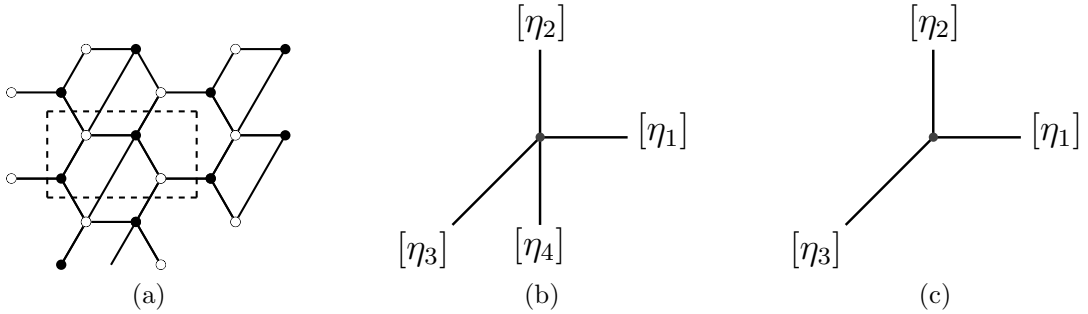
The *global zigzag fan* of D is

$$\Xi = \{[\eta] : \widetilde{\eta} \text{ is a zigzag path}\} \subset \mathbb{R}^2.$$

Equivalently, it is the union of the $\xi(v)$ over all $v \in D$. Clearly, $\xi(v) \subseteq \Xi$ for any v.

Every vertex v is the endpoint of finitely many edges, so $\chi(v), \xi(v)$, and Ξ are all finite sets. Considering a (local or global) zigzag fan as a set of vectors in $\mathbb{R}^2$, we have a notion of the angle between two elements of a fan. We prove later that if $\{[\eta], [\zeta]\}$ is linearly dependent, then $[\eta] = \pm[\zeta]$. Although for now we are forced to work in the more general case, keep in mind that some of the results of this section can, in retrospect, be stated in simpler terms.

The property that $[\eta] = \pm[\zeta]$ if $\{[\eta], [\zeta]\}$ is linearly dependent will further give us the ability to give the vectors of a zigzag fan an unambiguous natural cyclic order around the origin. Figure 3.5 illustrates examples of a global zigzag fan and a local one.



(a) (b) (c)

Figure 3.5: Left: The dimer model from Example 3.1.2. Middle: The global zigzag fan. Note that the elements are numbered according to their anticlockwise order around the origin. Right: The local zigzag fan of the black vertex at the bottom-left of the fundamental domain.

The following lemma crystallises the relationships between some of the concepts that we have just encountered.

Lemma 3.3.2. *Let $\widetilde{\eta}$ and $\widetilde{\zeta}$ be zigzag flows with and let η and ζ be their associated zigzag paths. The following are equivalent:*

1. *$[\zeta]$ is (strictly) between 0 and π radians anticlockwise from $[\eta]$.*
2. *$C_s(\eta, \zeta)$ is positive.*
3. *$\widetilde{\zeta}$ crosses $\widetilde{\eta}$ from right to left.*
4. *$\widetilde{\eta}$ and $\widetilde{\zeta}$ cross in an edge that is a zig of $\widetilde{\zeta}$ and a zag of $\widetilde{\eta}$.*

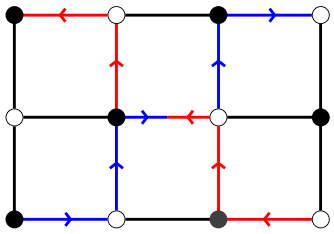


Figure 3.6: The blue zigzag path crosses the red one from left to right. Conversely, the red zigzag path crosses the blue one from right to left.

Proof. If $\widetilde{\zeta}$ crosses $\widetilde{\eta}$ from right to left, then $\widetilde{\zeta}$ passes from a white vertex to a black one as it crosses $\widetilde{\eta}$. Hence the edge is a zig of $\widetilde{\zeta}$, and therefore a zag of $\widetilde{\eta}$. This argument is reversible, so statements (3) and (4) are equivalent.

For the equivalence of (1) and (2), suppose that the signed crossing number $C_s(\eta, \zeta)$ is positive. In other words, $ad - bc > 0$, where $(a, b) = [\eta]$ and $(c, d) = [\zeta]$. Rewriting in terms of matrices:

$$\begin{pmatrix} a \\ b \end{pmatrix} \cdot \left[\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} c \\ d \end{pmatrix} \right] > 0$$

which is precisely what it means for $[\zeta]$ to be at an angle less than π anticlockwise from $[\eta]$, since $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ is an clockwise rotation matrix by $\pi/2$. This argument is reversible.

Now focus on (1) and (3). Suppose $[\zeta]$ is (strictly) between 0 and π radians anticlockwise from $[\eta]$. Pick a vertex v of $\widetilde{\eta}$ and a vertex v' of $\widetilde{\zeta}$. By repeatedly applying to v the deck transformations corresponding to translation by $[\eta]$, we obtain a sequence of vertices of $\widetilde{\eta}$ that lie on a straight line $l \subset \mathbb{R}^2$ with gradient $[\eta]$. Doing the same for v', we obtain a line l' with gradient $[\zeta]$ that, by our assumption, must cross l from right to left. But $\widetilde{\eta}$ lies within a bounded region of l and $\widetilde{\zeta}$ lies within a bounded region of l'. Combining with the fact that $\widetilde{\eta}$ and $\widetilde{\zeta}$ intersect in an edge exactly once, we see that $\widetilde{\zeta}$ crosses $\widetilde{\eta}$ from right to left. Hence (1) implies (3). This argument is reversible so (3) implies (1) as well, and we are done. □

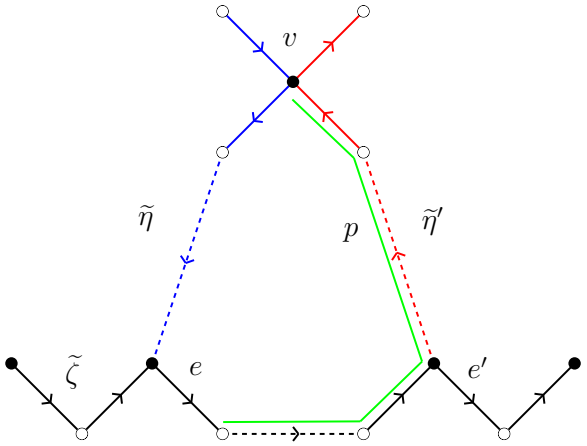


Figure 3.7: Artistic rendition of the proof of Lemma 3.3.5. The green path p connects a white vertex of $\widetilde{\eta}$ to a black one without using any edges of $\widetilde{\eta}$. This contradicts the fact that the left and right of a zigzag flow are distinct equivalence classes.

Lemma 3.3.3. *Given a vertex $v \in \widetilde{D}$ and a zigzag flow $\widetilde{\eta} \in \chi(v)$, there exists a $\widetilde{\zeta} \in \chi(v)$ such $[\zeta]$ is strictly between 0 and π radians in an anticlockwise direction from $[\eta]$.*

Proof. Since $\widetilde{\eta} \in \chi(v)$, by Lemma 3.2.3 it intersects v in precisely a zig-zag pair. Call the zag e. But e is also a zig of a (distinct) flow $\widetilde{\zeta} \in \chi(v)$. By the equivalence of (4) and (1) in the previous lemma, this means $[\zeta]$ lies at an angle strictly between 0 and π radians anticlockwise from $[\eta]$. □

Corollary 3.3.4. *Every (local or global) zigzag fan contains at least 3 elements that are pairwise linearly independent.*

Lemma 3.3.5 (Dual of Lemma 4.13, [1])**.** *If $\widetilde{\eta}$ and $\widetilde{\eta}'$ are distinct parallel zigzag flows, then they do not intersect in a vertex.*

Proof. By contradiction. There are many ways to construct a contradiction, but we will use an original argument that is simpler than the one presented in [1].

Assume v be a black vertex, as the general result follows from symmetry, and suppose that $\widetilde{\eta}, \widetilde{\eta}' \in \chi(v)$ are parallel. By Lemma 3.2.3, both $\widetilde{\eta}$ and $\widetilde{\eta}'$ intersect v in a zig-zag pair, and by condition (b) of geometric consistency, these pairs are disjoint. By Corollary 3.3.4, we can find a zigzag flow $\widetilde{\zeta} \in \chi(v)$ which is not parallel to $\widetilde{\eta}$ (and $\widetilde{\eta}'$). From geometric consistency, $\widetilde{\zeta}$ intersects each of $\widetilde{\eta}$ and $\widetilde{\eta}'$ in precisely one edge, which we respectively denote by e and e'. Because $\widetilde{\eta}$ and $\widetilde{\eta}'$ are parallel, $C_s(\eta, \zeta)$ and $C_s(\eta', \zeta)$ share the same sign, so Lemma 3.3.2 tells us that e and e' are either both zigs of $\widetilde{\zeta}$, or both zags. An example of this is given in Figure 3.7.

By geometric consistency, $\widetilde{\eta}$ and $\widetilde{\eta}'$ do not intersect in an edge, so e and e' are distinct. Consider the path p_1 along $\widetilde{\zeta}$ that is strictly between e and e'. Since e and e' are either both zigs or both zags of $\widetilde{\zeta}$, p_1 must be composed of an odd number of edges. So p_1 starts at a white

vertex and ends at a black one, or vice-versa. Suppose without loss of generality that it starts at a white vertex (otherwise, by swapping the names of $\widetilde{\eta}$ and $\widetilde{\eta}'$, we can relabel e and e', and thus reverse the orientation of p). Then by construction, p_1 is a finite path from a white vertex of $\widetilde{\eta}$ to a black vertex of $\widetilde{\eta}'$. Further, p_1 contains no edges from $\widetilde{\eta}$ since p_1 is a subset of $\widetilde{\zeta}$, which by geometric consistency intersects $\widetilde{\eta}$ in an edge at e and only e.

Now consider the path p_2 that follows $\widetilde{\eta}'$ from the end of p_1, all the way to v. By geometric consistency, p_2 does not contain any edges of $\widetilde{\eta}$. Composing p_1 and p_2, we obtain a finite path from a white vertex of $\widetilde{\eta}$ to v, which is a black vertex of $\widetilde{\eta}$ that does not intersect $\widetilde{\eta}'$ in an edge. Hence, straight from Definition 3.2.4, we are forced to conclude that v is on both the right and left sides of $\widetilde{\eta}$, a contradiction. □

Proposition 3.3.6. *Every zigzag path η is an oriented, simple, closed curve, and has primitive winding number.*

Proof. This proof was omitted by Broomhead. By definition, zigzag paths are oriented. They are closed because a dimer model has only a finite number of edges for a zigzag path to travel along. The primitivity of the winding number will come from applying Lemma 2.1.4 once we have established that η is simple.

To see that η is simple, it suffices to show that η does not intersect itself in a vertex. Let w be an arbitrary vertex of η, and let $v, v' \in \pi^{-1}(w)$ be vertices of $\widetilde{D}$. Then $v' = Tv$, where T is some deck transformation. Let $\widetilde{\eta}$ be a zigzag flow that passes through v and projects down to η. By Lemma 3.2.3, $\widetilde{\eta}$ intersects v in exactly one zig-zag pair (e_0, e_1).

By applying T to $\widetilde{\eta}$, one obtains a second zigzag flow $\widetilde{\eta}' \in \pi^{-1}(\eta)$ that passes through v' in the unique zigzag pair (Te_0, Te_1). Of course, $(\pi(Te_0), \pi(Te_1)) = (\pi(e_0), \pi(e_1))$, and by the previous lemma, $\widetilde{\eta}'$ is the only zigzag flow that passes through v' and is parallel to $\widetilde{\eta}$.

This shows that for any vertex in $\pi^{-1}(w)$, exactly one zigzag flow passing through it is parallel to $\widetilde{\eta}$. Thus these flows must comprise the entirety of $\pi^{-1}(\eta)$. Moreover, we learn that $(\pi(e_0), \pi(e_1))$ is the *unique* zig-zag pair in which η intersects w, so η does not intersect itself at w. Since we can choose w to be any vertex of η, η must not intersect itself at all. Hence η is simple. Primitivity of the winding number follows from Lemma 2.1.4. □

Corollary 3.3.7. *Every element of a zigzag fan is primitive. In particular, if $[\eta]$ and $[\zeta]$ are linearly dependent elements of Ξ, then $[\eta] = \pm[\zeta]$.*

Corollary 3.3.8. *For a (local or global) zigzag fan ξ we have an injective map $\phi : \xi \hookrightarrow \Xi \to \mathfrak{R}$, where $\mathfrak{R}$ is the set of rays in $\mathbb{R}^2$ that start at the origin. Explicitly, an element $[\eta] \in \xi$ maps to the ray $\{\lambda[\eta] : \lambda \in \mathbb{R}_{>0}\}$.*

Proof. The map ϕ is a composition of injective maps: The inclusion from ξ to Ξ is surely injective, and the previous corollary tells us that the map from Ξ to $\mathfrak{R}$ is injective as well. □

The above corollary greatly simplifies our understanding of zigzag fans, allowing us to visualise elements as rays. Thinking of Ξ as a set of rays in $\mathbb{R}^2$, the elements of Ξ have a well-defined (anticlockwise) cyclic order around the origin (see Figure 3.5). **Henceforth, we shall treat the elements of zigzag fans as rays instead of vectors.**

Now that the rays of Ξ have been given a cyclic order, we can discover what the physical meaning of this ordering is.

Lemma 3.3.9. *Every ray $R \in \xi(v)$ is represented by exactly one zigzag flow $\widetilde{\eta} \in \chi(v)$. This flow intersects v in a single zig-zag pair.*

Proof. By definition of zigzag fans, R must be represented by at least one zigzag flow $\widetilde{\eta} \in \chi(v)$. If it were to correspond to a distinct second flow $\widetilde{\zeta} \in \chi(v)$, then $\widetilde{\eta}$ and $\widetilde{\zeta}$ would be distinct parallel flows that intersect at v, contradicting Lemma 3.3.5. The final claim is a direct application of Lemma 3.2.3. □

Proposition 3.3.10 (Dual of Proposition 4.15, [1])**.** *Let R^- and R^+ be rays in $\xi(v)$, and let $\widetilde{\eta}^-$ and $\widetilde{\eta}^+$ respectively be their unique representatives in $\chi(v)$. Then $\widetilde{\eta}^-$ and $\widetilde{\eta}^+$ intersect each other in an edge of v if and only if R^- and R^+ are consecutive in $\xi(v)$.*

Proof. Uniqueness of the representatives $\widetilde{\eta}^-$ and $\widetilde{\eta}^+$ comes from Lemma 3.3.9. Assume without loss of generality that v is black, as the general result follows from symmetry. Similarly to the proof of the previous lemma, we construct an original argument that is simpler than the one in [1].

We prove the forward implication first. Suppose that $\widetilde{\eta}^+$ and $\widetilde{\eta}^-$ are zigzag flows that intersect in an edge a incident with v, and let R^+ and R^- be their respective rays in $\xi(v)$. By geometric consistency, R^+ is not proportional to R^-, and we can without loss of generality assume that R^+ is between *less* 0 and π radians anticlockwise from R^-.

Suppose for the sake of contradiction that there is a ray $R \in \xi(v)$ that lies strictly between R^- and R^+ (that is, R is anticlockwise from R^- and clockwise from R^+). Then there exists a zigzag flow $\widetilde{\eta}$ that has winding number R and intersects v. Let b and c respectively denote the unique edges at which $\widetilde{\eta}^-$ intersects $\widetilde{\eta}$ and $\widetilde{\eta}$ intersects $\widetilde{\eta}^+$. By Lemma 3.3.2, $\widetilde{\eta}$ crosses $\widetilde{\eta}^-$ from right to left and $\widetilde{\eta}^+$ from left to right. But v is a black vertex of both $\widetilde{\eta}^-$ and $\widetilde{\eta}^+$, so by definition v is on the left side of both flows. Hence at v, $\widetilde{\eta}$ has already crossed $\widetilde{\eta}^-$, but is yet to cross $\widetilde{\eta}^+$. In other words, with respect to $\widetilde{\eta}$, v lies between b and c. Figure 3.8 illustrates an example.

Consider the path p_1 along $\widetilde{\eta}^-$ that lies strictly between a and b. By geometric consistency, a is the unique edge in which $\widetilde{\eta}^-$ intersects $\widetilde{\eta}^+$, so p_1 does not intersect $\widetilde{\eta}^+$ in an edge. Then consider p_2, the path along $\widetilde{\eta}$ that starts where p_1 ends, and ends at v. As shown above, v lies between b and c, and by geometric consistency, c is the unique edge in which $\widetilde{\eta}$ intersects $\widetilde{\eta}^+$. Hence p_2 does not intersect $\widetilde{\eta}^+$ in an edge. Composing p_1 with p_2, we obtain a finite path p between the two endpoints of a that does not intersect $\widetilde{\eta}^+$. However, a is an edge of $\widetilde{\eta}^+$, so

the endpoints are a white and black vertex of $\widetilde{\eta}^+$. Using p and Definition 3.2.4, we find that these vertices are on both the left and right sides of $\widetilde{\eta}^+$, a contradiction. So no such $R \in \xi(v)$ between R^- and R^+ exists, which means that R^- and R^+ are consecutive in $\xi(v)$ (with R^+ anticlockwise from R^-).

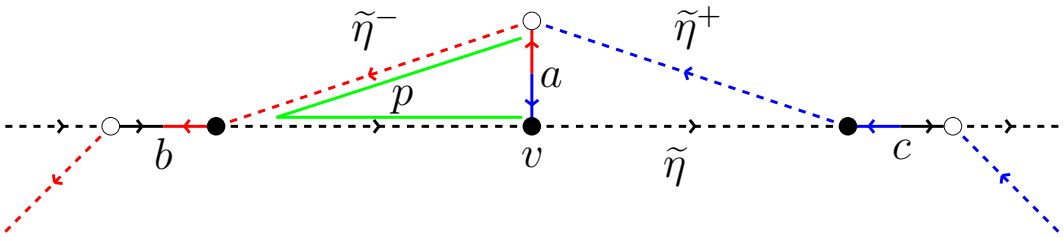


Figure 3.8: Yet another artistic masterpiece, this time illustrating the proof of Proposition 3.3.10. The green path p connects a white vertex of $\widetilde{\eta}^+$ to a black one without using any edges of $\widetilde{\eta}^+$. This contradicts the fact that the left and right of a zigzag flow are distinct equivalence classes.

Now we prove the reverse implication. Suppose R^+ and R^- are consecutive in $\xi(v)$. Using Lemma 3.3.3, assume without loss of generality that R^+ is at an angle between 0 and π in an anticlockwise direction from R^-. Let $\widetilde{\eta}^+$ and $\widetilde{\eta}^-$ respectively be their unique representatives in $\chi(v)$. By Lemma 3.3.9, $\widetilde{\eta}^-$ intersects v in a single zig-zag pair $(\widetilde{\eta}_0^-, \widetilde{\eta}_1^-)$. Now $\widetilde{\eta}_1^-$ is a zag of $\widetilde{\eta}^-$, but it is also a zig of a (distinct) flow $\widetilde{\zeta} \in \chi(v)$. By Lemma 3.3.2, $[\widetilde{\zeta}]$ lies between 0 and π radians anticlockwise from $[\widetilde{\eta}^-] = R^-$. Further, by the forward direction of the proof of this proposition, $[\widetilde{\zeta}]$ is consecutive to R^-. So it must be that $[\widetilde{\zeta}] = R^+$. By uniqueness of representatives $\widetilde{\zeta} = \widetilde{\eta}^+$, and therefore $\widetilde{\eta}^-$ and $\widetilde{\eta}^+$ intersect in the edge $\widetilde{\eta}_1^-$, which is incident with v. □

This result has a neat visual interpretation, which rounds out what we have learned in this section.

Corollary 3.3.11. *The zigzag flows that intersect a vertex $v \in \widetilde{D}$ do so in single zig-zag pairs. The cyclic order of these pairs around v is the same as the cyclic order of the corresponding rays in $\xi(v)$.*

Proof. The first statement comes from geometric consistency. Suppose v is black and let $e_1, e_2, \ldots, e_n$ be an anticlockwise cyclic ordering of the edges around v. Working modulo n, each of these edges e_i defines a zig of zigzag flow, with its corresponding zag being e_{i+1}. Respectively denote these zigzag flows as $\eta_1, \eta_2, \ldots, \eta_n$. These η_i are distinct by the first statement, and comprise all the zigzag flows incident with v. Since the zig of e_{i+1} is a zag of e_i, η_{i+1} intersects η_i in an edge of v, so by Proposition 3.3.10, $[\eta_{i+1}]$ is consecutive to $[\eta_i]$ in $\xi(v)$. A symmetric argument works when v is white. □

Remark 3.3.12. Using Proposition 3.3.6, a bit of thought reveals that Corollary 3.3.11 works even if we set make $v \in D$ and replace the phrase "zigzag flows" with "zigzag paths". A dimer model satisfying such a property is defined by Gulotta in [2] to be *properly ordered*, which serves as his characterisation of consistency. In the same paper, he formulates the algorithm that will be the focus of Chapter 4, proving that the algorithm outputs properly ordered dimer models. We will be proving these dimer models are not only properly ordered, but also geometrically consistent.

3.4 The zigzag-PSS correspondence

Now we are in a position to construct perfect matchings of the form conjectured at the beginning of Section 3.3. In particular, the matchings constructed are the ones that correspond to the corners of the characteristic polygon of D. In this section, the distinction between zigzag paths and zigzag flows is much less important because we have just showed that many of the properties of zigzag flows are shared by zigzag paths (see Proposition 3.3.6 and Remark 3.3.12).

Let σ be a ray of $\mathbb{R}^2$ that is not a ray of Ξ. For any vertex $v \in \widetilde{D}$, find the rays R^-, R^+ of $\xi(v)$ that make the smallest possible anticlockwise and clockwise angles with σ, respectively. These are called the *neighbours* of σ in $\xi(v)$. Now consider the set of zigzag flow passing through v that have winding R^- and R^+. Lemma 3.3.5 tells us that there is exactly one zigzag flow through v with winding R^-, and one with winding R^+. Call these flows $\widetilde{\eta}^-$ and $\widetilde{\eta}^+$. By Proposition 3.3.10, these two flows share a (unique) edge that has v as an endpoint. We name this edge $e_\sigma(v)$.

Lemma 3.4.1. *For $v \in \widetilde{D}$ and rays σ, σ' not in Ξ, $e_\sigma(v) = e_{\sigma'}(v)$ if and only σ and σ' have the same neighbours in $\xi(v)$.*

Proof. Suppose σ and σ' have the same neighbours R^-, R^+ in $\xi(v)$. Let $\widetilde{\eta}^-$ be the unique zigzag flow through v with winding R^-, and $\widetilde{\eta}^+$ be the unique zigzag flow through v with winding R^+. Then by definition, $e_\sigma(v)$ and $e_{\sigma'}(v)$ are the same, as they are both the unique edge in which $\widetilde{\eta}^-$ intersects $\widetilde{\eta}^+$.

Conversely, suppose $e_\sigma(v) = e_{\sigma'}(v)$. Recalling that a zigzag flow is completely defined by the choice of an edge and an orientation, $e_\sigma(v)$ is contained in exactly two (distinct) zigzag flows $\widetilde{\eta}^-, \widetilde{\eta}^+$. So e_σ must have been defined as the unique edge of v in which $\widetilde{\eta}^-$ intersects $\widetilde{\eta}^+$. Then $[\widetilde{\eta}^-]$ and $[\widetilde{\eta}^+]$ must be the neighbours of σ in $\xi(v)$. (Note that $\widetilde{\eta}^-$ and $\widetilde{\eta}^+$ intersect at e, so by geometric consistency $[\widetilde{\eta}^-] \neq [\widetilde{\eta}^+]$.) Applying the same argument to σ' instead of σ, we get the same two neighbours $[\widetilde{\eta}^-], [\widetilde{\eta}^+]$. Hence σ and σ' have the same neighbours in $\xi(v)$. □

Proposition 3.4.2. *For any ray $\sigma \notin \Xi$, the set*

$$M_\sigma = \{e_\sigma(v) : v \in \widetilde{D}\}$$

is a perfect matching.

Proof. Certainly, every vertex v is an endpoint of $e_\sigma(v) \in M_\sigma$, so we need only prove that no vertex is an endpoint of two edges of M_σ. This is the same as proving that $e_\sigma(v) = e_\sigma(v')$ for every v, where v' is the other endpoint of $e_\sigma(v)$.

Define $e = e_\sigma(v)$ for brevity, and let v' be the endpoint of e that is not v. By definition, e is the unique edge in which two zigzag flows $\widetilde{\eta}^-$ and $\widetilde{\eta}^+$ intersect, where $[\eta^-]$ and $[\eta^+]$ are the neighbours of σ in $\xi(v)$. Since $\widetilde{\eta}^+$ and $\widetilde{\eta}^-$ share e, they must also share v'. Hence $\widetilde{\eta}^+$ and $\widetilde{\eta}^-$ are in $\chi(v')$ and intersect in an edge of v'. Applying Proposition 3.3.10 to v', we find that these $[\eta^-]$ and $[\eta^+]$ are consecutive in $\xi(v')$. Since they are consecutive and σ lies between them, they must be the neighbours of σ in $\xi(v')$. Then, by definition, $e_\sigma(v')$ is the unique edge incident with v' in which $\widetilde{\eta}^-$ intersects $\widetilde{\eta}^+$, which of course is e. Thus $e_\sigma(v) = e = e_\sigma(v')$. □

Corollary 3.4.3. *Two rays produce the same perfect matching if and only if they have the same neighbours in Ξ. Hence the procedure defined above produces exactly $|\Xi|$ perfect matchings.*

Proof. The following proof is original. It is clear from the definition that $M_\sigma = M_{\sigma'}$ if σ and σ' have the same neighbours in Ξ.

It is fairly intuitive that the reverse implication follows essentially from Corollary 3.3.4 and Lemma 3.4.1, although the proof is a bit more long-winded. Suppose that σ and σ' do not have the same neighbours in Ξ. Without loss of generality, we may assume that σ' lies no more than π radians anticlockwise of σ. Travelling anticlockwise from σ to σ', one must encounter at least one ray $Q \in \Xi$, and by definition of Ξ this ray must be in $\xi(v)$ for some $v \in D$. From here, we want to prove that σ and σ' do not have the same neighbours in $\xi(v)$.

Certainly, Q is a ray that separates σ from σ' when travelling anticlockwise from σ to σ'. Because σ' is no less than π radians *clockwise* of σ, Lemma 3.3.3 tells us that when travelling clockwise from σ to σ', one must also encounter at least one ray of $\xi(v)$. So σ is separated from σ' in either direction by at least one ray of $\xi(v)$. By Corollary 3.3.4 there are at least 3 rays in $\xi(v)$, so by the pigeonhole principle, σ is separated from σ' by in one direction by at least two rays of $\xi(v)$, and in the other direction by at least one. Then it must be that σ and σ' do not have the same neighbours in $\xi(v)$, so by Lemma 3.4.1 $e_\sigma(v) \neq e_{\sigma'}(v)$, and therefore $M_\sigma \neq M_{\sigma'}$. □

Remark 3.4.4. Since rays that are not separated by an element of Ξ produce the same perfect matching, the use of σ to denote rays that are not in Ξ draws deliberate comparisons to the standard notation for *cones* used in (toric) geometry.

We have just constructed some perfect matchings, but we have not yet shown that they are of the form conjectured at the beginning of Section 3.3. This is what we do now.

Corollary 3.4.5. *Let $\widetilde{\eta}$ be a zigzag flow, and let σ (resp. σ') be rays not in Ξ whose anticlockwise (resp. clockwise) neighbour in Ξ is $[\eta]$. Then M_σ contains all zags of $\widetilde{\eta}$, and $M_{\sigma'}$ contains all zigs.*

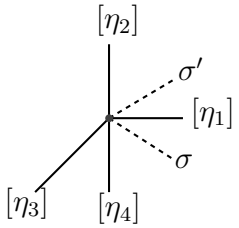


Figure 3.9: Here, $[\eta_1]$ is the anticlockwise neighbour of σ, and the clockwise neighbour of σ' in Ξ.

Proof. Suppose that the anticlockwise neighbour of σ in Ξ is $[\eta]$. For every vertex $v \in \widetilde{\eta}$, $[\eta]$ appears in the local zigzag fan $\xi(v)$, so by the construction of M_σ, $e_\sigma(v)$ is an edge of $\widetilde{\eta}$. Thus every vertex of $\widetilde{\eta}$ is matched by an edge of $\widetilde{\eta}$. The only way this is possible is if M_σ uses every second edge of $\widetilde{\eta}$.

Further, for every $v \in \widetilde{\eta}$, $e_\sigma(v)$ is the unique edge in which $\widetilde{\eta}$ intersects the flow $\widetilde{\zeta}$ through v corresponding to the other neighbour of σ in Ξ. By our original assumption, $\widetilde{\zeta}$ is anticlockwise from $\widetilde{\eta}$, so Lemma 3.3.2 tells us that $e_\sigma(v)$ is a zag of $\widetilde{\eta}$. Hence M_σ contains every zag of $\widetilde{\eta}$. A similar argument shows that $M_{\sigma'}$ contains every zig of $\widetilde{\eta}$. □

Lemma 3.4.6. *Let $\widetilde{\eta}$ be a zigzag flow, and let M_0 be a perfect matching that intersects half of the edges of $\widetilde{\eta}$. Then the height change of any perfect matching M with respect to M_0 in the direction of $\widetilde{\eta}$ is non-positive. More precisely,*

$$\langle h(M, M_0), [\eta] \rangle \leq 0.$$

This defines a half-plane in which the characteristic polygon must live (see Figure 3.10). Note that setting $M = M_0$ shows that M_0 lies on the boundary line of this half-plane.

Proof. Consider the path P on the torus that lies immediately to the left of $\widetilde{\eta}$ (see Figure 3.4b). Since M_0 intersects half of the edges of $\widetilde{\eta}$, none of the edges that P crosses are contained in M_0. Additionally, every edge that P *does* cross has a black vertex on its right, so each of these crossings will either have no effect on the height function or decrease it by one, depending on whether or not the crossed edge is in M.

If the winding of η is denoted (a, b) and our fundamental domain is the unit square, then we have just deduced that

$$\begin{aligned} 0 &\geq h_{M,M_0}(p + (a, b)) - h_{M,M_0}(p) \\ &= ah_x(M, M_0) + bh_y(M, M_0) \\ &= \langle h(M, M_0), [\eta] \rangle \end{aligned}$$

The physical interpretation of this is that if l is the line through $h(M)$ normal to $[\eta]$, then the characteristic polygon lies on the side of l opposite the direction of $[\eta]$. □

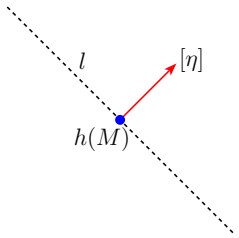


Figure 3.10: The characteristic polygon is bounded by l and must lie on the side of l opposite the direction of $[\eta]$, which in this case is the half-plane in the bottom left of the diagram.

Proposition 3.4.7. *The matchings of the form M_σ for $\sigma \notin \Xi$ correspond to the corners of the characteristic polygon X. The order of the corners around the polygon is the same as the order of the ray directions.*

Proof. Let $\sigma \notin \Xi$, and let $[\eta^-]$ and $[\eta^+]$ be its neighbours in Ξ. Corollary 3.4.5 tells us that M_σ uses half of the edges of $\widetilde{\eta}^-$, so by Lemma 3.4.6, $h(M_\sigma)$ is a point of X that lies on a line that bounds X. Hence $h(M_\sigma)$ must lie on a side of the characteristic polygon. Further, this side is normal to $[\eta^-]$. Similarly, using $\widetilde{\eta}^+$ instead of $\widetilde{\eta}^-$, we find that $h(M_\sigma)$ lies on a side of the characteristic polygon that is normal to $[\eta^+]$. These two sides are distinct because $[\eta^-] \neq [\eta^+]$, so $h(M_\sigma)$ must in fact lie on a corner of the characteristic polygon.

Now to show that these corners comprise *all* the corners of the characteristic polygon. First of all, X, being the convex hull of *all* matchings, certainly contains the convex hull of the M_σ's. Meanwhile, the characteristic polygon must lie inside the intersection of all half-planes described in Lemma 3.4.6, which is the convex hull of the $h(M_\sigma)$'s. Hence the two convex hulls are the same, so every corner of the characteristic polygon corresponds M_σ for some $\sigma \notin \Xi$.

Finally, we prove that the cyclic order of the corners and ray directions coincide. If σ and σ' are separated by only one element $[\eta]$ of Ξ, then M_σ and $M_{\sigma'}$ both satisfy the inequality of Lemma 3.4.6:

$$\langle h(M, M_\sigma), [\eta] \rangle \leq 0 \text{ and } \langle h(M, M_{\sigma'}), [\eta] \rangle \leq 0.$$

In words, the characteristic polygon lies on the same side of two parallel lines, and there exists a point of the polygon on each of the lines. The only way that this can be possible is if the two

lines are the same, i.e. $h(M_\sigma)$ and $h(M_{\sigma'})$ share a side. Hence M_σ and $M_{\sigma'}$ lie on consecutive corners of the characteristic polygon. □

Corollary 3.4.8. *The characteristic polygon has $|\Xi|$ corners.*

Theorem 3.4.9 (Adapted from Theorem 9.3, [14])**.** *In a geometrically consistent dimer D, the zigzag paths are in bijective correspondence with the primitive side segments of the characteristic polygon. Specifically, an zigzag path with winding (a, b) corresponds to a $(-b, a)$ PSS, and vice versa (see Figure 3.11).*

Proof. Let $[\eta] \in \Xi$, and let n be the number of zigzag paths in D that have winding $[\eta]$. Let σ and σ' be rays not in Ξ that are separated only by $[\eta]$, as in Corollary 3.4.5. By Proposition 3.4.7, M_σ and $M_{\sigma'}$ lie on consecutive corners of the characteristic polygon. Without loss of generality, suppose that the corner corresponding to $M_{\sigma'}$ is anticlockwise from the corner corresponding to M_σ. Then by the proposition σ' is anticlockwise from σ, with $[\eta]$ separating them.

Let e be an arbitrary edge of the dimer, and let v be one of its endpoints. Suppose that e does not belong to any of the zigzag flows of winding $[\eta]$. If $[\eta] \notin \xi(v)$, then σ and σ' have the same neighbours in $\xi(v)$. Then $e_\sigma(v) = e_{\sigma'}(v)$, so e is in either both or neither of M_σ and $M_{\sigma'}$. Meanwhile, if $[\eta]$ *is* in $\xi(v)$, then $e_\sigma(v)$ and $e_{\sigma'}(v)$ are both edges of $\widetilde{\eta}$ that intersect v. So $e \neq e_\sigma(v)$ and $e \neq e_{\sigma'}(v)$. So regardless of whether or not $[\eta]$ is in $\xi(v)$, e is in either both or neither of M_σ and $M_{\sigma'}$.

Now suppose that e *does* belong to one of the zigzag paths of winding $[\eta]$. By Corollary 3.4.5, e either lies in M_σ or $M_{\sigma'}$, but not both. Therefore, the symmetric difference $M_\sigma \ominus M_{\sigma'}$ is simply the union of the zigzag flows with winding $[\eta]$. More specifically, since $[\eta]$ is the anticlockwise neighbour of σ and the clockwise neighbour of σ' in Ξ, so by Corollary 3.4.5, $e_\sigma(v)$ is a zag of $\widetilde{\eta}$ and $e_{\sigma'}(v)$ is a zig.

Suppose $[\eta] = (1, 0)$. If we cross η from below to above using an edge $e \in \eta$, then by the above argument, e is in M_σ if e is a zag, and in $M_{\sigma'}$ if it is a zig. Further, if it is a zag then e has a white vertex on the right as we cross it, and if it is a zig then e has a black vertex on the right. Then the height function of $M_{\sigma'}$ with respect to M_σ increases by one when crossing $\widetilde{\eta}$ from below to above. If $[\eta] = (0, 1)$, then the height function decreases by one when crossing $\widetilde{\eta}$ from left to right.

More generally, for n zigzag paths, each with (primitive) winding number (a, b), the height change $h(M_\sigma, M_{\sigma'})$ is $n(-b, a)$. Geometrically, $h(M_{\sigma'}, M_\sigma)$ is the vector from $h(M_\sigma)$ to $h(M_{\sigma'})$, so we have our correspondence: every (a, b) zigzag path creates a $(-b, a)$ primitive side segment in the characteristic polygon (i.e. a clockwise rotation of the PSS by $\pi/2$ radians). □

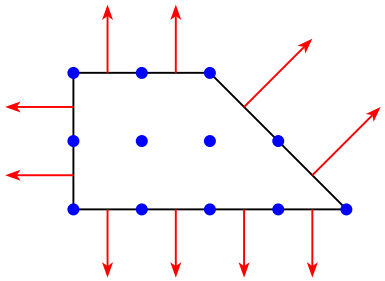


Figure 3.11: A visual representation of the correspondence between zigzag paths and primitive side segments, as stated in Theorem 3.4.9. For each red vector, there is a zigzag path on the dimer model with winding number equal to that vector. Therefore, any geometrically consistent dimer model with the above characteristic polygon must have exactly two zigzag paths with winding $(-1, 0)$, four with winding $(0, -1)$, two with winding $(1, 1)$, two with winding $(0, 1)$, and no others.

Chapter 4

From polygons to dimers: An inverse algorithm

Recall the key question that we asked in Chapter 1: Given a convex polygon M in $\mathbb{Z}^2$, is it possible to construct a dimer model with M as its characteristic polygon? Furthermore, can we construct the dimer model in such a way that it is geometrically consistent?

The zigzag-PSS correspondence of Section 3.4 gives us an alternative understanding of the characteristic polygon, one that does not rely explicitly on the height function, or on perfect matchings. The big benefit is that zigzag paths are more amenable to an algorithmic methods. In particular, the correspondence suggests that if we develop methods to manipulate zigzag paths, then we will also attain the capability to manipulate characteristic polygons.

For the remainder of this thesis, we will draw characteristic polygons rotated by $\pi/2$ radians clockwise from usual. This makes it so that for geometrically consistent dimer models, the windings of the zig-zag paths are *equal* to rather than *perpendicular* to the PSSes, visually simplifying the correspondence that was explored in Section 3.4.

A particular algorithm by Gulotta [2] will be the focus of this chapter. The main idea behind the algorithm is to start with a large characteristic polygon X_0 for which a dimer model is already known, and to make changes to the dimer model in such a way that we chisel away a corner of X_0. Applying this procedure many times, we slowly sculpt X_0 into any smaller convex polygon that we wish. The "chiselling" procedure is achieved by utilising the zigzag-PSS correspondence; instead of trying to change the set of perfect matchings on the dimer model, we try to change the set of zigzag paths. However, much work yet is needed before we are ready to present the algorithm.

We should mention that there exist alternative procedures to associate dimer models to convex polygons, proposed by authors both prior and subsequent to Gulotta's work. For example, Stienstra [14] uses entirely different methods, but his approach requires much trial-and-error, and is not guaranteed to find a solution. It sometimes, however, produces multiple valid solutions. Ishii & Ueda [11] generalise Gulotta's construction, although the resulting dimer may

not be geometrically consistent. Lastly, Stienstra has rewritten Gulotta's construction into a programmable algorithm using matrices in [15].

This chapter represents a complete overhaul of the relevant material in [2]. Sections 4.1 and 4.2 are additions that properly lay the groundwork for the algorithm, while Sections 4.3 and 4.4 are heavily restructured to increase the level of mathematical rigour. One last note is that although much of the content in this chapter is original, a majority of the diagrams are courtesy of Gulotta. Some of these diagrams have been colourised.

4.1 The Stern—Brocot Sequences

Before we jump into the meat of Gulotta's construction, we take a short interlude (and hopefully a welcome break) to investigate the *Stern-Brocot ordering*. The results in this section elucidate nontrivial results that Gulotta used (sometimes implicitly) but did not prove or even formally state in the original proof of the algorithm. At the end of the day, they are just some properties of numbers that, although interesting, are only stepping stones to help us achieve our end goal. Therefore, we will try to get through this section quickly and with minimal discussion. The important results to remember here are the neighbours property (Lemma 4.1.5) and covering property (Lemma 4.1.9).

Definition 4.1.1. The Stern-Brocot sequence S_n of order $n \geq 1$ is defined inductively as follows: For each pair of two consecutive values $\left(\frac{b}{a}, \frac{d}{c}\right)$ (called *parents*) in S_{n-1}, compute the *mediant* value $\frac{b+d}{a+c}$. Place all parents and mediants in order of size to form S_n. The base sequence, S_0, is defined as $\left(\frac{0}{1}, \frac{1}{0}\right)$.

Some elementary consequences of the definition are the following:

Lemma 4.1.2. *Any element $\frac{b}{a}$ of $S_n - S_{n-1}$ is the mediant of two parents in S_{n-1}, one of which is in $S_{n-1} - S_{n-2}$. (Otherwise, $\frac{b}{a}$ would have appeared in an earlier sequence.)*

Lemma 4.1.3 (Mediant Property)**.** *For $\frac{b}{a} < \frac{d}{c}$, their mediant $\frac{b+d}{a+c}$ lies (strictly) between them.*

Proof. Suppose $\frac{b}{a} < \frac{d}{c}$, where a, b, c, d are non-negative integers and a, c are nonzero. By cross-multiplying, we obtain $bc < ad$. Then

$$\frac{b+d}{a+c} - \frac{b}{a} = \frac{a(b+d) - b(a+c)}{a(a+c)} = \frac{ad - bc}{a(a+c)} > 0.$$

Similarly,

$$\frac{b+d}{a+c} - \frac{d}{c} = \frac{c(b+d) - d(a+c)}{c(a+c)} = \frac{bc - ad}{c(a+c)} < 0.$$

Hence

$$\frac{b}{a} < \frac{b+d}{a+c} < \frac{d}{c}.$$

□

The first few Stern-Brocot sequences are shown below:

$$S_0 = \left(\frac{0}{1}, \frac{1}{0}\right)$$
$$S_1 = \left(\frac{0}{1}, \frac{1}{1}, \frac{1}{0}\right)$$
$$S_2 = \left(\frac{0}{1}, \frac{1}{2}, \frac{1}{1}, \frac{2}{1}, \frac{1}{0}\right)$$
$$S_3 = \left(\frac{0}{1}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{1}{1}, \frac{3}{2}, \frac{2}{1}, \frac{3}{1}, \frac{1}{0}\right)$$
$$S_4 = \left(\frac{0}{1}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{1}{1}, \frac{4}{3}, \frac{3}{2}, \frac{5}{3}, \frac{2}{1}, \frac{5}{2}, \frac{3}{1}, \frac{4}{1}, \frac{1}{0}\right)$$

Note that the presence of the parent pairs $\left\{\frac{0}{1}, \frac{1}{1}\right\}$ and $\left\{\frac{1}{1}, \frac{1}{0}\right\}$ in S_1 induces a symmetry in the sequences, ensuring that the reciprocal of an element of S_n is also in S_n. It may help to think of the sequences via a tree as in Figure 4.1.

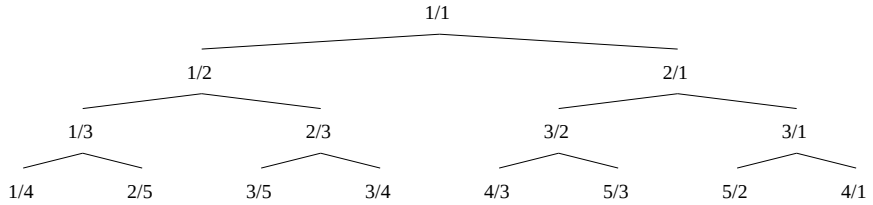


Figure 4.1: The Stern-Brocot tree. The n-th level of the tree displays only the elements of $S_n - S_{n-1}$. Be careful that this tree hides the fact that each element has two parents, not just one.

Definition 4.1.4. The Stern-Brocot ordering S is the union of the Stern-Brocot sequences, disallowing repeated elements. That is,

$$S = S_0 \cup S_1 \cup S_2 \cup \cdots = \left(\frac{0}{1}, \frac{1}{0}, \frac{1}{1}, \frac{1}{2}, \frac{2}{1}, \frac{1}{3}, \frac{2}{3}, \frac{3}{2}, \frac{3}{1}, \frac{1}{4}, \dots\right).$$

We denote by $S(k)$ the k-th element of the Stern-Brocot ordering.

Lemma 4.1.5 (Neighbours Property)**.** *If $\frac{b}{a} < \frac{d}{c}$ are consecutive in some Stern-Brocot sequence, then $ad - bc = 1$.*

Proof. By induction. For $S_0 = \left(\frac{0}{1}, \frac{1}{0}\right)$ we certainly have $ad - bc = 1 - 0 = 1$. Now suppose that the result holds for S_n, and consider consecutive terms in S_{n+1}. If two fractions are neighbours in S_n, then the mediant property ensures they are separated by and only by their mediant in S_{n+1}. So any pair of consecutive terms in S_{n+1} must consist of an element of $S_{n+1} - S_n$ and one of its parents. In other words, $F_{n+1} = \left(\dots, \frac{b}{a}, \frac{b+d}{a+c}, \frac{d}{c}, \dots\right)$, where $\frac{b}{a}$ and $\frac{d}{c}$ are the appropriate parents, and neighbours in S_{n+1} are of the form $\left(\frac{b}{a}, \frac{b+d}{a+c}\right)$ or $\left(\frac{b+d}{a+c}, \frac{d}{c}\right)$. Now $a(b+d) - b(a+c) = ad - bc = 1$

from the inductive hypothesis, and similarly $(a+c)d-(b+d)c = ad-bc = 1$. So the result is proved for all pairs of neighbours. □

Corollary 4.1.6. *Every member of a Stern-Brocot sequence is an irreducible fraction.*

Proof. Let $\frac{b}{a}$ be a fraction in a Stern-Brocot sequence. If $\frac{b}{a}$ is $\frac{1}{0}$, then it is irreducible. Otherwise, it has a neighbour $\frac{d}{c} > \frac{b}{a}$. By the previous lemma, $ad - bc = 1$. Bezout's identity tells us that the greatest common divisor of b and a is 1, i.e. $\frac{b}{a}$ is irreducible. □

Remark 4.1.7. The Stern-Brocot sequence S_n is intimately connected to the well-known Farey sequence F_n, defined as the sequence of completely reduced fractions between 0 and 1, ordered by size. Incredibly, Lemma 4.1.5 (and even its converse!) holds for the Farey sequences, although we will not be needing this result. See [6] for a creative and geometric proof.

Lemma 4.1.8. *For any element of $S_n - S_{n-1}$, the sum of its numerator and denominator (nd-sum) is at least $n+1$.*

Proof. Induction: For $n = 1$, the minimum nd-sum is certainly 2. By lemma 4.1.2, any fraction in $S_n - S_{n-1}$ is the mediant of a fraction $\frac{b}{a}$ in $S_{n-1} - S_{n-2}$ and a fraction $\frac{d}{c}$ in S_{n-1}. The nd-sum of $\frac{d}{c}$ is at least 1, and by our inductive assumption, the nd-sum of $\frac{b}{a}$ is at least n. So the nd-sum of $\frac{a+d}{b+c}$ is at least $n+1$. □

Lemma 4.1.9 (Covering property)**.** *Every positive irreducible fraction lies in a Stern-Brocot sequence.*

Proof. Certainly $\frac{0}{1}$ and $\frac{1}{0}$ lie in a Stern-Brocot sequence. All remaining cases will be proved by contradiction. Let $\frac{p}{q} \in \mathbb{Q}$ be an irreducible fraction with $p, q \in \mathbb{N}^+$, and suppose that it is not equal to any fraction in the Stern-Brocot sequences. Then $\frac{p}{q}$ lies strictly between two neighbours of S_{p+q}, which we call $\frac{b}{a}$ and $\frac{d}{c}$. The mediant property combined with the inductive definition of the Stern-Brocot sequences implies that one of these elements is in $S_{p+q} - S_{p+q-1}$. We can rearrange the equation $\frac{b}{a} < \frac{p}{q} < \frac{d}{c}$ into the following:

$$ap - bq > 0 \text{ and } dq - cp > 0.$$

Since a, b, c, d, p, q are all integers, this is equivalent to saying that $ap - bq \geq 1$ and $dq - cp \geq 1$. Multiplying by appropriate factors and combining the two inequalities gives

$$(c+d)(ap-bq) + (a+b)(dq-cp) \geq a+b+c+d.$$

Cancellation simplifies the left hand side to $adp - bcq + adq - bcp$, which factorises into $(ad-bc)p + (ad-bc)q$. But $\frac{b}{a}$ and $\frac{d}{c}$ were neighbours in S_{p+q}, so by the neighbours property, $ad - bc = 1$. This leaves us with the inequality

$$p + q \geq a + b + c + d.$$

However, recall that at least one of $\frac{b}{a}$ and $\frac{d}{c}$ was in $S_{p+q} - S_{p+q-1}$. Suppose without loss of generality that it is $\frac{b}{a}$. By lemma 4.1.8, $b + a \geq p + q + 1$, leading to the conclusion that $p + q \geq a + b + c + d \geq p + q + 1$, a contradiction. □

Corollary 4.1.10. *Every positive irreducible fraction appears exactly once in the Stern-Brocot ordering.*

Proof. The previous lemma shows that every positive rational number appears at least once in the ordering. The mediant property ensures that the same value will never be computed twice. □

4.2 Zigzag path diagrams

A key point about the zigzag-PSS correspondence is that it is not so much the zigzag paths themselves that are important, but rather their winding numbers. Of course, we must also remember that the correspondence only holds if our dimer model is geometrically consistent, a condition that concerns the way in which zigzag paths intersect. To this end, it would be convenient to be able to visualise zigzag paths in a way that suppresses their zigzagging structure and retains only the core information about their windings and intersections with each other.

To do this, we introduce the concept of a zigzag path diagram, which is based on the work of [4] and [5]. In [4], Kenyon introduces the idea of *train tracks* on planar embeddings of graphs, where every face is a quadrilateral. Hanany and Vegh [5] were the first to realise that from a dimer model, one could construct a graph (known to them as a *rhombus lattice*) whose *train tracks* were in some sense equal to the zigzag paths of the dimer. For more information about these rhombus lattices, see [1].

We will choose not go into excruciating topological detail in this section; it is important to keep in mind that it is purely a *tool* to help us encode data about zigzag paths, as opposed to an important object in its own right. Nevertheless, we explain much more about the properties of zigzag path diagrams than the two paragraphs Gulotta writes in [2], where the majority of these properties are left unstated.

Definition 4.2.1. For every edge e of D, fix a point $\dot{e}$ somewhere in the interior of that edge, called the vertebra of e. For each zigzag path η of D, the sequence of edges of η defines a sequence of vertebrae $(\dot{e}_1, \dot{e}_2, \ldots, \dot{e}_n, \dot{e}_1)$. Use lines to connect this sequence together in the manner suggested by η to form an oriented closed path $\dot{\eta}$, which we call the *spine* of η. The collection of all spines, embedded in $\mathbb{T}$, is the *zigzag path diagram* Z of D.

In such a way, we collapse each zigzag path η into a less complex object (one could say it is a *barebones* description of η), as illustrated in Figure 4.2. For brevity, we call a spine with winding (a, b) an (a, b)-spine.

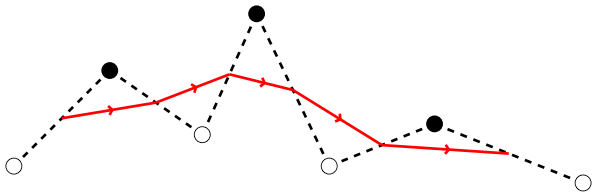


Figure 4.2: The dotted zigzag path η, and the associated path (in red) on the zigzag path diagram.

Since both η and $\dot{\eta}$ meet at each $\dot{e}_i$, it is clear from the definition that η and $\dot{\eta}$ are homotopic, and hence $[\eta] = [\dot{\eta}]$. So the zigzag path diagram does indeed capture the desired information about the windings of the zigzag paths of D.

What about the intersection properties of the zigzag paths? By construction, each vertebra of Z represents an edge of D, so two zigzag paths η and ζ intersect in an edge if and only if $\dot{\eta}$ and $\dot{\zeta}$ intersect in a vertebra. It is less simple (but still possible) to understand the analogue of intersection in a vertex, but we need not worry about this because the definition of geometric consistency only concerns itself with edge-intersections.

Now that we have motivated the definition of the zigzag path diagram, we should probably develop a basic intuition of what a zigzag path diagram actually looks like. For example, what does the zigzag path diagram look like around a particular vertex? Since zigzag paths intersect vertices in zig-zag pairs and turn maximally right at black vertices, a black vertex v of D becomes surrounded by edges of the zigzag path diagram, all of which point anticlockwise around v. Hence v corresponds to an anticlockwise-oriented face $\dot{v}$ in Z. Similarly, a white vertex of D corresponds to a clockwise-oriented face in Z. As shown in Figure 4.4, the remaining, unoriented faces of Z correspond to faces of D. This essentially comes from the fact that the boundary of any face of D alternates between meeting black vertices and white vertices.

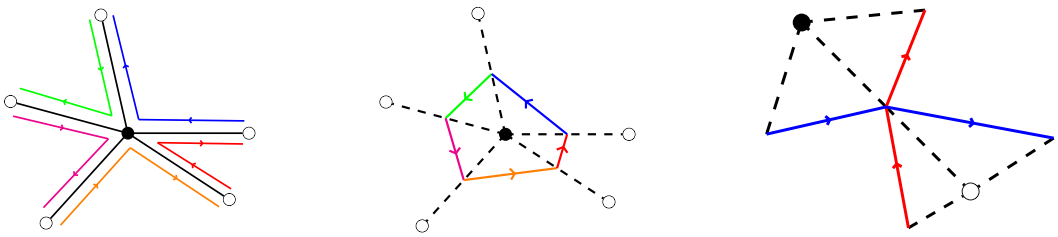


Figure 4.3: Left: The set of five zigzag paths around a black vertex b in D. Middle: The local structure of the zigzag path diagram near b; b 'becomes' a anticlockwise-oriented face with five sides. Right: Spines must cross transversally at $\dot{e}$ if the endpoints of e have degree at least three, as is the case for a geometrically consistent dimer model.

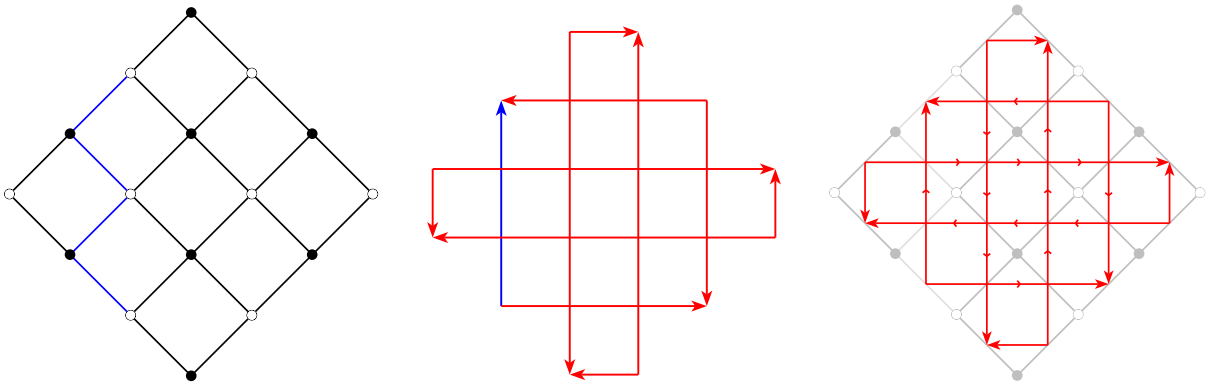


Figure 4.4: Left: A portion of the universal cover of the square dimer, where we have distinguished the zigzag path in blue. Middle: The zigzag path diagram Z. Right: Overlay, demonstrating the orientations of the faces of Z around vertices of D.

Now we know how to get from a dimer model to a zigzag path diagram. Actually, we are also interested in the reverse process. That is, we want to obtain a dimer $D(Z)$ from a collection of spines Z on the torus such that the winding numbers of the zigzag paths of $D(Z)$ are reflective of the winding numbers of the spines of Z. If we can do this, then dimer models and zigzag path diagrams become equivalent in a sense, and developing techniques that modify the winding numbers of zigzag paths becomes the same as developing techniques that modify the winding numbers of spines. Hopefully, the latter task is much simpler due to the suppression of the zigzag structure. Further, if the intersection properties of the spines of Z are reflective of the edge-intersection properties of $D(Z)$, then we can easily test whether $D(Z)$ is geometrically consistent.

It turns out that we can go from a collection of spines to a dimer model, provided that the spines meet certain properties outlined below.

Definition 4.2.2. A good configuration of curves Z is a finite collection of closed, simple curves (called *spines*) embedded on the torus that satisfy the following properties:

1. Every point of $\mathbb{T}$ lies on at most two spines.

2. Each spine intersects other spines at a finite number of points, and all of these intersections are transverse.

3. The intersections of each spines with other spines alternate between being positively and negatively oriented (i.e. the spine is crossed from right-to-left, then from left-to-right, then from right-to-left, etc.). See Figure 4.5.

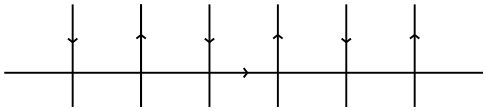
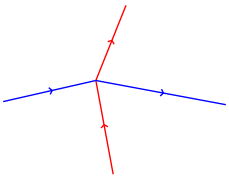


Figure 4.5: The horizontal path is intersected by the vertical ones first from left-to-right, then right-to-left, then left-to-right, etc. Note that because we are on a torus, we require that the orientation of the last intersection must be opposite of the orientation of the left intersection.

The procedure to get from a good configuration of curves Z to a dimer model $D(Z)$ is fairly simple. Let $V(Z)$ (the vertebrae of Z) be the set of points in $\mathbb{T}$ at which two distinct spine of Z intersect. The spines of Z cut the torus into a number of regions. Regions that are homeomorphic to the unit ball of $\mathbb{R}^2$ are called *faces* of Z. For every face of Z that is consistently anticlockwise-oriented, place a black vertex of $D(Z)$ somewhere in the interior. For every face of Z that is consistently clockwise-oriented, place a white vertex of $D(Z)$ somewhere in the interior. If two vertices in $D(Z)$ have corresponding faces in Z intersect at an element $t \in V(Z)$, then connect those two vertices with an edge that passes through t. Roughly speaking, this is saying that two vertices of $D(Z)$ are connected by an edge if their corresponding faces in Z share a corner.

For this construction to be useful, we desire three things: Firstly, that $D(Z)$ is a dimer model. Secondly, that each spine of Z corresponds to a zigzag path of $D(Z)$, and that their windings are the same. Lastly, that two spines intersect if and only if the zigzag paths they represent intersect in an edge.

The first question asks whether we get a bipartite graph from this process at all. The answer is yes, because the essential structure of Z is forced by conditions 1, 2, and 3. Let $t \in V(Z)$.[1] By condition 1, it is the intersection of *exactly* two spines of Z, and by condition 2, this intersection is transverse. Then the intersection must look something like the diagram below:



The point is that t is a corner of four faces. For the above diagram, just considering t alone, we see that the top right and bottom left faces are already inconsistently oriented, while the top left face is potentially consistently anticlockwise-oriented, and the bottom right face is potentially consistently clockwise-oriented. But actually, condition 3 forces them to the consistently oriented! In the above diagram the red spine is crossed from left to right. By

[1] This implicitly assumes that there is at least one intersection between spines of Z. We follow the example of Gulotta by not formally placing this in our definition because it is a degenerate case that we will never actually encounter, but that greatly complicates our proofs. It is degenerate in the sense that the resulting graph has no vertices and no edges.

condition 3, it must next be crossed from right to left. So the face at the top left of the diagram will continue to be oriented anticlockwise (see diagram below). Inductively, we realise that the entire face must be consistently oriented anticlockwise. A similar story goes for the face on the bottom right, except that it will be oriented clockwise. The moral here is that every $t \in V(Z)$ is a corner of exactly four faces, one of which is anticlockwise, and one of which is clockwise. Hence two anticlockwise-oriented faces can never meet at an element of $V(Z)$, so no black vertices will be joined by an edge. The same reasoning applies for white vertices, so we see that $D(Z)$ is bipartite. The proof that $D(Z)$ tiles the torus is quite topologically arduous, so we do not present it here.[2]

The bigger question that we asked above still looms: given a good configuration of curves Z, do the zigzag paths of $D(Z)$ have the same winding numbers as the spines of Z? The answer is also yes. As we have just shown, for every vertebra t of Z, one of the faces of which it is a corner corresponds to a black vertex, and another corresponds to a white vertex. By definition, these two vertices are connected by an edge that passes through t. Hence there is a bijective correspondence between the vertebrae of Z and the edges of $D(Z)$.

Suppose we have a spine $C \in Z$, and let (t_1, t_2) be consecutive vertebrae of C. Suppose that at t_1, C is crossed from right to left (the general result will follow from symmetry). Then at t_2, C is crossed from left to right. This is the case for the first two intersections of the red spine in the previous diagram. Looking at the diagram, we see that t_1 and t_2 are consecutive corners in a face of that corresponds to a white vertex w of $D(Z)$. Because they are *consecutive* corners, the edges of $D(Z)$ that pass through t_1 and t_2 must be consecutive around w. Hence these edges form a zag-zig pair of a zigzag path. Continuing inductively, C must form an entire zigzag path. To see that *every* zigzag path corresponds to a spine, notice that every vertebra in Z is crossed by two spines. Hence the corresponding zigzag paths cross every possible edge of $D(Z)$ in every possible direction. Thus these zigzag paths must comprise *all* the zigzag paths of $D(Z)$. Again, each zigzag path is homotopic to its spine, so their winding numbers are equal.

To see that two spines intersect (in a vertebra) if and only if the zigzag paths they represent intersect in an edge, suppose that $t \in V(Z)$. Then this vertebra lies at the corner of an anticlockwise-oriented face and a clockwise-oriented face, and hence t becomes an edge of $D(Z)$. The two spines that intersect at t become zigzag paths in $D(Z)$ that intersect at this edge.

[2]This requires the weak but additional constraint that the spines of Z have windings that span $\mathbb{Z}^2$. Again, this is to deal with degenerate cases that we will never encounter in practice, so like other authors, we will not go into detail.

Having a good configuration of curves allows us to retrieve a dimer model, but we will need two more conditions to be able to construct Gulotta's algorithm.

Definition 4.2.3. We declare that a pair of spines forms an *opposite pair* if they have opposite winding numbers, do not intersect, and if one of the two regions that they bound on the torus contains no crossings. We further define an opposite pair to be *left-handed* if the region lies on the left of an observer travelling along one of the spines, and *right-handed* if the region lies on the right (see Figure 4.6). If the windings of an opposite pair are given to be (e, f) and $(-e, -f)$, we may refer to the pair as an (e, f)-$(-e - f)$ pair.

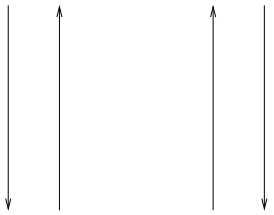


Figure 4.6: Pairs of spines that are left-handed and right-handed, respectively. It is possible for a pair of spines to be both left- and right-handed, although this is not particularly important.

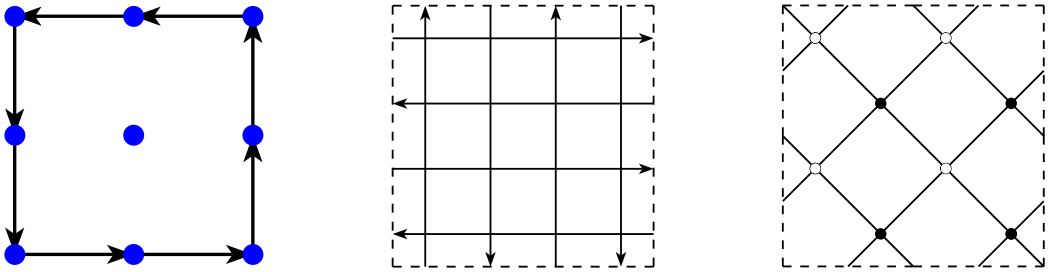
Definition 4.2.4. A good configuration of curves Z is said to be *very good* if it additionally satisfies the following property.

4. For any winding number $(e, f) \in \mathbb{Z}^2$, if there are n spines with winding (e, f) and m spines with winding $(-e, -f)$, then there are $\min(m, n)$ right-handed pairs of (a, b)- and $(-a, -b)$-spines.

5. There are no extra crossings between any pair of spines.

Condition 4 requires that anti-parallel spines must pair up whenever possible. Although this will seem like a cryptic requirement for now, it is important because it will ensure that we can always perform the next step of Gulotta's algorithm. Meanwhile, Condition 5 is essentially what tells us that the dimer model associated to a very good configuration of curves is geometrically consistent.

Example 4.2.5. Here we construct a dimer model with a 2×2 square X as its characteristic polygon. Finding the PSS set of X and applying the zigzag-PSS correspondence, we know that any geometrically consistent dimer model with characteristic polygon X must have exactly eight zigzag paths: two each with windings $(1, 0)$, $(0, 1)$, $(-1, 0)$, and $(0, -1)$. Instead of trying to directly construct such a dimer model, we try to arrange a good configuration of curves Z whose spines have these windings. This turns out to be very easy, as shown below. We then reverse engineer a dimer model $D(Z)$ whose zigzag paths have the desired windings.



Note that this dimer model is very similar to one that we have seen before; it is actually the quotient of the universal cover of the square dimer (Example 2.1.3) by $2\mathbb{Z} \times 2\mathbb{Z}$ (in other words, setting the fundamental domain to be $[0,2] \times [0,2]$ instead of $[0,1] \times [0,1]$). It is also easy to verify that this dimer model is geometrically consistent.

In fact, we can generalise this construction. If we want a very good configuration of curves with n $(1,0)$-spines, n $(-1,0)$-spines, m $(0,1)$-spines, and m $(0,-1)$-spines, we can do this by arranging spines in the same cross-hatch pattern as in the middle figure above. These windings generate (through the zigzag-PSS correspondence) a rectangular polygon with horizontal and vertical sides. In this way, all such rectangles can be realised as the characteristic polygon of some quotient of the universal cover of the square dimer. Because all of these dimer models share the same universal cover, they are all geometrically consistent.

4.3 The operations

The previous example gives us a partial answer to the main question of the thesis. However, we can do better. Suppose we want to find a dimer model whose characteristic polygon is the polygon X on the right of the below figure. To do this, we will start with a larger polygon with a known geometrically consistent dimer model. In this case, we use the 2×2 square, whose dimer model D and zigzag path diagram Z we found in the previous section. Because a convex polygon is uniquely defined by its PSS set (Lemma 2.2.8), to get from the square to X it suffices to remove a $(1,0)$-PSS and a $(0,1)$-PSS, and introduce a $(1,1)$-PSS. By the zigzag-PSS correspondence, we can do this by manipulating the spines of Z in such a way that we change their windings.

Of course, this is far from all that we want to do. In the first place, the correspondence only holds if our manipulation preserves geometric consistency. Further, we must ensure that, after

manipulation, the spines define a good configuration of curves so that we can turn it back into a dimer model. Both of these requirements come down to the intersection properties of spines, so considerable care must be taken in how we perform our manipulations. Gulotta's insight is that certain sorts of manipulation keep a vast majority of intersection properties unchanged.

We begin by looking at the simplest possible case—the case where we start with $(1,0)$-, $(0,1)$-, $(-1,0)$-, and $(0,-1)$-spines, and try to create $(1,1)$- and $(-1,-1)$-spines from them. Although this may at first appear to be an investigation with limited scope, we will see later that the techniques developed are far more general.

Let Y be a very good configuration of curves with at least n $(1,0)$-spines, n $(0,1)$-spines, m $(-1,0)$-spines, and m $(0,-1)$-spines. Suppose that we want to turn n $(1,0)$-spines, n $(0,1)$-spines, m $(-1,0)$-spines, and m $(0,-1)$-spines into n $(1,1)$-spines and m $(-1,-1)$-spines, where n and m are not both 0. We can write this process in shorthand as

$$n(1,0) + n(0,1) + m(-1,0) + m(0,-1) \to n(1,1) + m(-1,-1).$$

We can break this problem into three cases depending on the values on n and m, and each case will require a slightly different approach to solve:

- **Operation I**: $m = 0$ or $n = 0$.
- **Operation II**: $n = m$.
- **Operation III**: $0 \neq n \neq m$.

We now proceed to present the operations themselves.

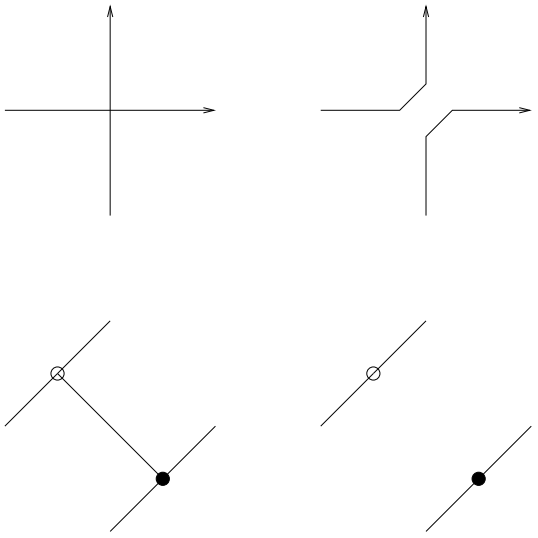
4.3.1 Operation I

Suppose we want to perform the merging

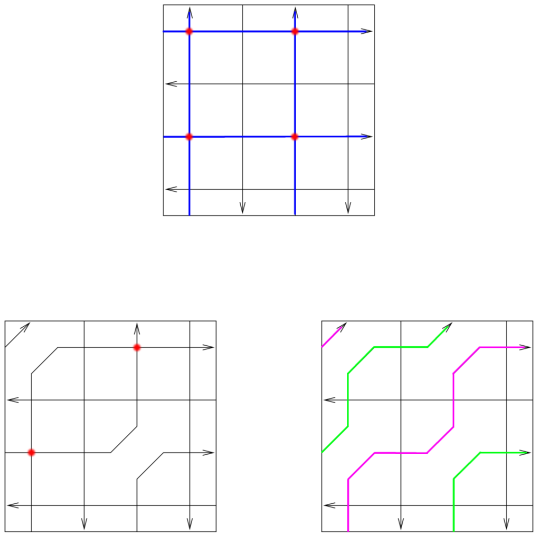
$$n(1,0) + n(0,1) \to n(1,1).$$

First, pick out n $(1,0)$- and n $(0,1)$-spines. As much as possible, choose these spines to be ones that do not form part of a right-handed pair. The absence of extra crossings in Y (condition 5 of Definition 4.2.4) means that each $(1,0)$-spine and $(0,1)$-spine intersects exactly once. Hence there are a total of n^2 intersections between the chosen $(1,0)$- and $(0,1)$-spines.

If $n = 1$, then merge our chosen spines by diverting their (unique) intersection, as illustrated in Figure 4.7a. Note that there is only one way to perform this merging that gives the resulting spine a well-defined orientation. Of course, we can perform this diversion procedure inside as small a region of the torus as we wish. Since there are only a finite number of crossings in our zigzag path diagram, we can, in particular, ensure that our region of diversion be so small that it contains no crossings other than the one being diverted.



(a) Operation I for $n = 1$. Top: Before/after of the operation on the zigzag path diagram. Bottom: The effect on the dimer.



(b) Operation I: $n = 2$. Top: The original zigzag path diagram. Bottom left: Diagram obtained by diverting two intersections. Bottom right: Diagram obtained by diverting all four intersections.

Figure 4.7: Summary of Operation I.

For $n > 2$, we generalise our construction from the $n = 1$ case by diverting all n^2 intersections between chosen spines. The result is n $(1, 1)$ paths that do not intersect each other. The proper application of Operation I in the case that $n = 2$ is illustrated in the bottom right of Figure 4.7b.

Similarly, one can define a merging procedure for

$$m(-1, 0) + m(0, -1) \to m(-1, -1).$$

4.3.2 Operation II

Suppose we want to perform the merging

$$n(1, 0) + n(0, 1) + n(-1, 0) + n(0, -1) \to n(1, 1) + n(-1, -1).$$

We begin by picking out n spines from each of the windings to be merged in such a way that we obtain n right-handed $(1, 0)$-$(-1, 0)$ pairs and n right-handed $(0, 1)$-$(0, -1)$ pairs. We know that there are enough right-handed pairs for us to do this because Y is a very good configuration of curves. Due to the absence of extra crossings in Y, the chosen $(1, 0)$- and $(0, 1)$-spines intersect at n^2 points, and similarly the $(-1, 0)$- and $(0, -1)$-spines intersect at n^2 points. Divert the n^2 crossings between the $(1, 0)$- and $(0, 1)$-spines as in Operation I, and then separately divert the n^2 crossings between the $(-1, 0)$- and $(0, -1)$-spines. Again following the example of Operation

I, we choose our regions of diversion to be small enough that they contains no other crossings in the zigzag path diagram. The result of this (for $n = 1$) is shown in the top right of Figure 4.8.

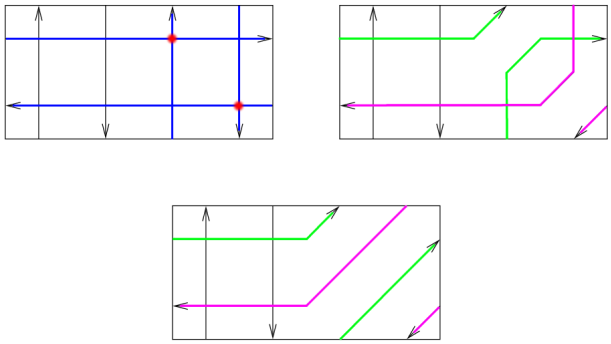


Figure 4.8: Operation II for $n = 1$. Top right: the zigzag path diagram after diverting the red intersections indicated in the top left diagram. Bottom: The diagram after moving the spines past each other; they no longer intersect.

However, as can be seen in the figure, the diversions raise a problem: we now have spines intersecting each other *twice*. Looking ahead, this will mean that the associated dimer model will have zigzag paths that intersect each other in two edges, and thus will not be geometrically consistent. Constructing geometrically consistent dimer models is our ultimate goal, so we pre-empt the issue by moving the spines past each other, as shown in the bottom of Figure 4.8. See Figure 4.9 for Operation II at work for $n = 2$.

Remark 4.3.1. This process of moving spines past each other was included by Gulotta in the original version of his algorithm even though he had no conception of geometric consistency. He writes "there is nothing inherently wrong with extra crossings, but we may find it desirable to produce diagrams without them". It is unclear if he suspected that it may become relevant to the work of others. If we do not perform the process, we obtain a zigzag path diagram whose associated dimer model is a vertex of degree two. Recalling that such a vertex becomes a face with two sides (i.e. a *di-gon*) in the associated quiver, the author feels compelled to relay a joke of Broomhead's that we should not let di-gons be di-gons!

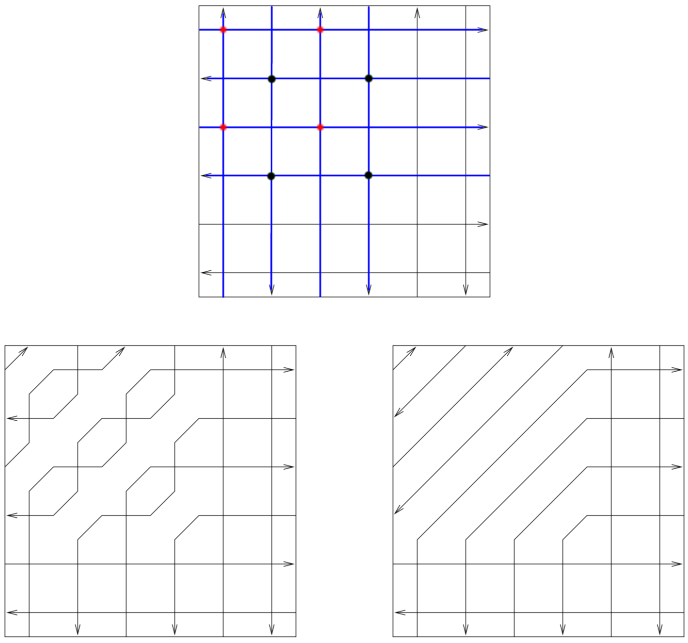


Figure 4.9: Operation II for $n = 2$. We first divert the four intersections (in red) between chosen $(1, 0)$- and $(0, 1)$-spines, and the four intersections (in black) between chosen $(-1, 0)$- and $(0, -1)$-spines. Then we push the resulting spines past each other to remove unnecessary intersections.

4.3.3 Operation III

Finally, suppose we want to perform the merging

$$n(1,0) + n(0,1) + m(-1,0) + m(0,-1) \to n(1,1) + m(-1,-1),$$

where $n \neq m$. Assume $n > m$. First, make $n-m$ $(1, 1)$-spines via Operation I. Then, completely remove m right-handed $(1, 0)$-$(-1, 0)$ pairs, and similarly remove m right-handed $(0, 1)$-$(0, -1)$ pairs. At this stage, we have performed an operation

$$\begin{aligned} n(1,0) + n(0,1) + m(-1,0) + m(0,-1) &\to m(1,0) + m(0,1) + m(-1,0) + m(0,-1) + (n-m)(1,1) \\ &\to (n-m)(1,1). \end{aligned}$$

It remains to insert m pairs of $(1, 1)$- and $(-1, -1)$-spines. We can do this by drawing new right-handed $(1, 1)$-$(-1, -1)$ pairs directly to the side of one of the $m - n$ $(1, 1)$-spines that we created earlier (see Figure 4.10, bottom right). We require that these "clone" spines follow the original closely enough that they intersect other spines in the same manner as the original does. We further require that the spine closest to the original be a $(-1, -1)$-spine, so that we preserve the property that spines are intersected with alternating orientations (Definition 4.2.2). Figure 4.10 provides a step-by-step illustration of Operation III. We can define an analogous operation if $m > n$.

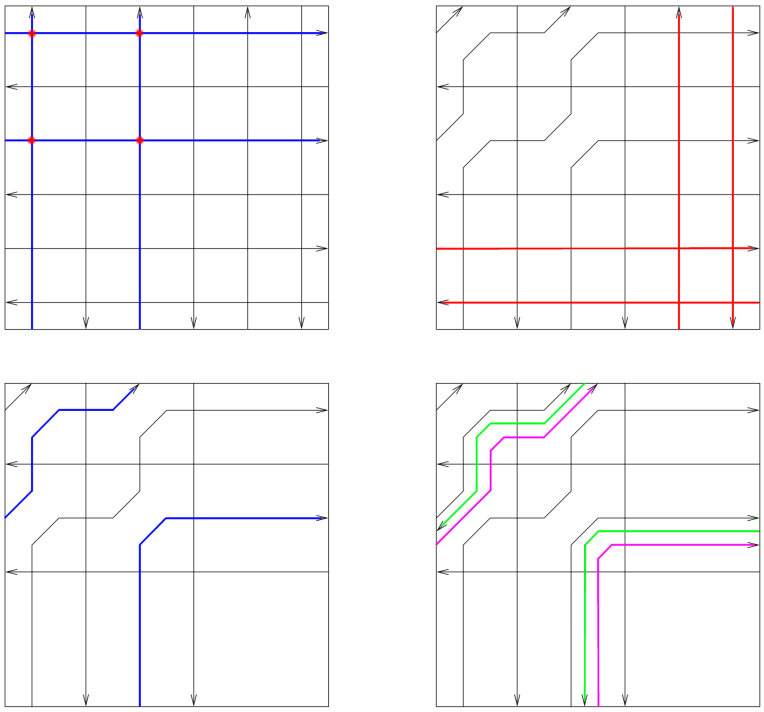


Figure 4.10: Operation III for $n = 3$, $m = 1$. We first make two $(1,1)$-spines via Operation I, and then remove some existing pairs of spines. Then we identify one of the $(1,1)$-spines we have made and introduce right-handed pairs of "clone" spines.

4.3.4 A more general case

Obviously, the above operations aren't very useful unless we can extend them a bit. Luckily, they turn out to be easy to extend, at least in the following case:

Suppose we want to perform the merging

$$n(a,b) + n(c,d) + m(-a,-b) + m(-c,-d) \to n(a+c,b+d) + m(-a-c,-b-d),$$

where (a,b) and (c,d) are neighbours in some Stern-Brocot sequence. We have not yet investigated how to do this beyond the simplest case where $(a,b) = (1,0)$ and $(c,d) = (0,1)$. Consider the automorphism of $\mathbb{R}^2$ defined by the action of the matrix

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix}^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -c \\ -b & a \end{pmatrix} = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}),$$

where $ad - bc = 1$ because of Lemma 4.1.5. Since this matrix has integer entries, it is also an automorphism of $\mathbb{Z}^2$. Hence it induces an automorphism of the torus $\mathbb{T} = \mathbb{R}^2/\mathbb{Z}^2$. Applying this change of basis to our zigzag path diagram, all (a,b) spines are sent to $(1,0)$ spines, $(-a,-b)$ to $(-1,0)$, (c,d) to $(0,1)$, and $(-c,-d)$ to $(0,-1)$.

We can now apply our three operations to create n $(1,1)$-spines and m $(-1,-1)$-spines, and then revert the change of basis using the matrix $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$ to find that we have created n $(a+c, b+d)$-spines and m $(-a-c, -b-d)$-spines while leaving unmerged spines unchanged. Thus we have reduced the problem of making $(a+c, b+d)$ and $(-a-c, -b-d)$ spines to the problem of making $(1,1)$ and $(-1,-1)$ spines, which we have already solved. This is a trick we call $\mathrm{SL}_2(\mathbb{Z})$ equivalence.

Recall that the slope of a closed curve on the torus with winding (e, f) is defined to be f/e. The following lemma proves that, under some constraints, the three operations preserve the property of being a very good configuration of curves.

Lemma 4.3.2. *Let Y be a very good configuration of curves and suppose that we perform a merging of spines*

$$n(a,b) + n(c,d) + m(-a,-b) + m(-c,-d) \to n(a+c, b+d) + m(-a-c, -b-d)$$

via one of the operations to obtain a collection of spines Z. If we assume that all spines of Y have slopes that appear before $(b+d)/(a+c)$ in the Stern-Brocot ordering, then Z is a very good configuration of curves.

We begin by reminding the reader what we have to prove here.

1. Every point of $\mathbb{T}$ lies on at most two spines.
2. Each spine intersects other spines at a finite number of points, and all of these intersections are transverse.
3. The intersections of each spines with other spines alternate between being positively and negatively oriented.
4. For any winding number $(e, f) \in \mathbb{Z}^2$, if there are j spines with winding (e, f) and k spines with winding $(-e, -f)$, then there are $\min(j, k)$ right-handed pairs of (e, f)- and $(-e, -f)$-spines.
5. There are no extra crossings between any pair of spines.

Proof. First we prove conditions 1, 2, and 3. The essential observation is that the three operations involve only the following actions: diversions of intersections, deletions of spines, and insertions of spines. Further, when doing these actions, we require that they occur in a small enough region that no intersection properties are affected. The details are not difficult to work out from here, but are so lengthy to comprehensively explain that one would risk losing sight of the main objective, so we will only provide brief explanations here.

Diversions of intersections are *removals* of intersections, so certainly conditions 1 and 2 hold. For any spine that is not one of the newly created $(a+c, b+d)$- or $(-a-c, -d-b)$-spines, condition 3 holds because absolutely nothing changes about the locations and orientations of its intersections. For any spine C that *is* one of the newly created spines, by considering the orientations of the spines of Y that created C, condition 3 is easily verified (refer to Figure 4.7b for inspiration). Meanwhile, deletions of spines certainly preserve conditions 1 and 2. Condition 3 holds because spines are removed in opposite pairs. Lastly, insertions of spines meet condition 1 by construction. The inserted spines are clones of a spine that is made through diversion of intersections. The clone spines share the same intersection properties as the original spine, and we have proved above that spines made by diversions of intersections meet conditions 2 and 3. Hence the cloned paths must do so as well.

For condition 4, we need to check the possibilities for (e, f) case-by-case and operation-by-operation. Because Y is already a very good configuration of curves, we need only check the case where (e, f) is one of the winding numbers involved in the merging process. That is, it suffices to consider the cases $(e, f) = (a, b)$, $(e, f) = (c, d)$, and $(e, f) = (a + c, b + d)$.

Let $(e, f) = (a, b)$ or $(e, f) = (c, d)$. Operation I preserves condition 4 because we always merge unpaired paths where possible. The same condition is preserved by Operation II because we only merge right-handed pairs. Operation III involves an application of Operation I, followed by the deletion and insertion of right-handed pairs. Hence it too preserves condition 4.

Now let $(e, f) = (a + c, b + d)$. By our assumption in the statement of the lemma, there are no $(a + b, c + d)$- or $(-a - c, -b - d)$-spines in Y, so $\min(j, k) = \min(m, n)$ in Z. Operation I does not create any $(-a - c, -b - d)$-spines, so $\min(m, n) = 0$ and we get condition 4 for free. For Operation II, $\min(m, n) = n$. Figure 4.11 demonstrates that Operation II always creates spines that form right-handed pairs, so we do indeed get n right-handed pairs from the operation. Lastly, Operation III inserts m right-handed pairs of spines in its final step, and $\min(m, n) = m$. So all three operations preserve condition 4.

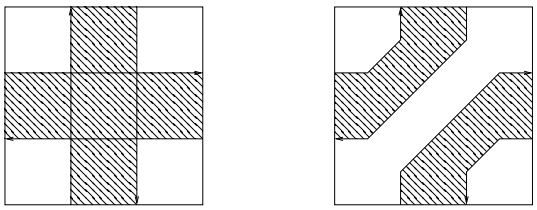


Figure 4.11: Left: We start out with two right handed pairs of spines. The shaded regions are free of crossings. Right: The regions formed by the merged pairs are still free of crossings.

The proof of condition 5 is somewhat lengthy. We remind the reader that two spines γ_1 and γ_2 are said to have extra crossings if $C(\gamma_1, \gamma_2) > |C_s(\gamma_1, \gamma_2)|$. Since we only merge spines that

are neighbours in some Stern-Brocot sequence, the neighbours property (Lemma 4.1.5) tells us that $ad - bc = 1$. For the purposes of this proof, we will express our operation as

$$\alpha_1 + \alpha_2 + \cdots + \alpha_{n+m} + \alpha'_1 + \alpha'_2 + \cdots + \alpha'_{n+m} \to \beta_1 + \beta_2 + \cdots + \beta_{n+m},$$

where the α_i denote the (a, b)- and $(-a, -b)$-spines, α'_i denote the (c, d)- and $(-c, -d)$-spines, and β_i denote the new $(a + b, c + d)$- or $(-a - b, -c - d)$-spines.

Let γ_1 and γ_2 be spines on Z. We check for extra crossings in three cases which, together with the antisymmetric nature of signed crossing number, together exhaust all possibilities for γ_1 and γ_2. First, note that every spine on Z either existed in Y and was unchanged by $\mathcal{O}$, or is one of the β_i created by $\mathcal{O}$.

Our first case is when γ_1 and γ_2 are spines that existed in Y. Then by our assumption that Y is a very good configuration of curves, there are no extra crossings between γ_1 and γ_2. Next, the case when γ_1 and γ_2 are both newly created spines (i.e. one of the β_i). By construction, the newly created spines do not intersect each other, so $C(\gamma_1, \gamma_2) = 0$. Then $C_s(\gamma_1, \gamma_2)$ must also be 0, so there are no extra crossings between γ_1 and γ_2.

The final case is trickier. Let γ be any spine that existed in Y, and let (e, f) denote the winding number of γ. By lemma 2.1.5, we know that for any i,

$$\begin{aligned} |C_s(\beta_i, \gamma)| &= |(a + c)f - (b + d)e| \\ &= |(af - be) + (cf - de)| \end{aligned}$$

Note that the presence of the absolute value sign makes the above identity hold regardless of whether β_i is a $(a + b, c + d)$- or $(-a - b, -c - d)$-spine.

Because (e, f) is the winding of γ, which existed in Y, our assumption in the statement of the lemma requires that f/e appears in S before $(b + d)/(a + c)$ does. The mediant property ensures that $(b+d)/(a+c)$ is the first element of S that lies between $\frac{b}{a}$ and $\frac{d}{c}$, so f/e does not lie between $\frac{b}{a}$ and $\frac{d}{c}$. It is easily verified that implies that $(af - be)$ and $(cf - de)$ share the same sign, which in turn implies that $|(af - be) + (cf - de)| = |af - be| + |cf - de|$. So the above equation becomes

$$\begin{aligned} |C_s(\beta_i, \gamma)| &= |af - be| + |cf - de| \\ &= |C_s(\alpha_i, \gamma)| + |C_s(\alpha'_i, \gamma)|, \end{aligned}$$

or, summing over all i,

$$\sum_{i=1}^{n+m} |C_s(\beta_i, \gamma)| = \sum_{i=1}^{n+m} |C_s(\alpha_i, \gamma)| + \sum_{i=1}^{n+m} |C_s(\alpha'_i, \gamma)|.$$

Meanwhile, the total number of unsigned crossings between the newly created spines and γ is much easier to calculate; recall that when performing Operation I, we require that the diversion be performed in a region with no crossings, so the intersection properties of γ are unchanged by the operation. More precisely,

$$\sum_{i=1}^{n+m} C(\beta_i, \gamma) = \sum_{i=1}^{n+m} C(\alpha_i, \gamma) + \sum_{i=1}^{n+m} C(\alpha_i', \gamma).$$

Our inductive hypothesis tells us that there are no extra crossings between the α_i and γ, nor between the α_i' and γ, so the right hand sides of the above two equations are equal. So

$$\sum_{i=1}^{n+m} |C_s(\beta_i, \gamma)| = \sum_{i=1}^{n+m} C(\beta_i, \gamma).$$

But of course, we have the individual bound $|C_s(\beta_i, \gamma)| < C(\beta_i, \gamma)$ for all i. Since all crossing numbers are finite in the zigzag path diagrams of our algorithm, we are forced us to conclude that

$$|C_s(\beta_i, \gamma)| = C(\beta_i, \gamma) \text{ for all } i.$$

So there are no extra crossings between the newly created spines and the pre-existing ones. Thus there are no extra crossings between any pair of spines in Z, and Z is a very good configuration of curves. □

4.4 The algorithm

Now that we know how to merge certain spines together, we can apply these operations in a clever way to get very good configurations of curves whose spines have arbitrary winding numbers, thus allowing us to construct a dimer model with a given convex polygon as its characteristic polygon. It turns out that one must go about this task in small steps because $\mathrm{SL}_2(\mathbb{Z})$ equivalence, while decently general, does not allow us to merge spines with arbitrary winding numbers. The "clever way" is suggested by Lemma 4.3.2, and this is where the results we have developed about the Stern-Brocot tree will come in handy. The description of algorithm will be given in its entirety before it is proved. Our goal here is to be much clearer than Gulotta is in [2] about exactly what we are doing (and that we can, in fact, do it at all).

Definition 4.4.1. A *tangent line* to a convex polygon X in $\mathbb{Z}^2$ is a line l such that $l \cap X \subseteq \partial X$ and $l \cap X \neq \varnothing$. Here, we allow the slopes $0 = \frac{0}{1}$ and $\infty = \frac{1}{0}$. Note that a convex polygon has exactly two tangent lines with a given slope (see Figure 4.12). Denote by $B(s) \subset \mathbb{R}^2$ the closure of the region bounded by the two tangent lines of X with slope s.

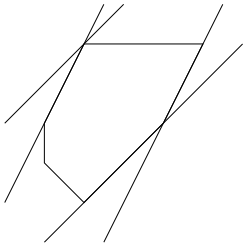


Figure 4.12: The tangent lines to a convex polygon for slopes 1/1 and 2/1.

Let X be a characteristic polygon for which we would like to construct a dimer (remember that X is defined only up to translation in $\mathbb{Z}^2$). Let X_0^+ be the smallest rectangle in $\mathbb{Z}^2$ with horizontal and vertical sides that contains X. By Example 4.2.5, we know a very good configuration of curves whose associated dimer model is geometrically consistent and has characteristic polygon X_0^+; this dimer model is some quotient of the universal cover of the square dimer.

For $n \geq 1$, inductively define

$$X_n^+ = X_{n-1}^+ \cap B(S(n)),$$

where $S(n)$ is the n-th element of the Stern-Brocot ordering (Definition 4.1.4). The physical meaning of X_n^+ is that it the subset of X_0^+ that been "cut" by the tangent lines of X that correspond to the first n elements of the Stern-Brocot ordering. Our algorithm will work by modifying the zigzag path diagram such that the characteristic polygon changes from X_n^+ to X_{n+1}^+. In such a way, we start with X_0 and progressively chip away at the polygon like a sculpture (see Figure 4.13), eventually arriving (after a finite number of steps) at a polygon X^+ that has been cut along *all* tangent lines of positive slope.

Figure 4.13: A before/after of the sort of cut that we will perform to get from X_n^+ to X_{n+1}^+. Here, we have not illustrated the other tangent line of slope $S(n+1)$, which lies towards the opposite side of the convex polygon.

We will then use a very similar procedure to cut X^+ along the tangent lines of X that have *negative* slope. Set $X_0^- = X^+$ and inductively define, for $n \geq 1$,

$$X_n^- = X_{n-1}^- \cap B(S^-(n)),$$

where $S^-(n) = -S(n)$. We will give a method for changing X_n^- to X_{n+1}^-. Once we have cut along all of these lines (which again is a finite task), the resulting convex polygon has been cut along all tangent lines of X, so it must in fact be identical to X, and the algorithm concludes.

Now we describe the specifics of the modifications we need to make to the zigzag path diagram in order to perform these cuts. We aid the description with an explicit and nontrivial example of the algorithm in practice (Example 4.4.2).

To get X_{n+1}^+ from X_n^+ we need to cut X_n^+ along the tangent lines of X with slope $B(S(n+1))$. As previously discussed, a characteristic polygon is uniquely defined by its set of primitive side segments, so it suffices to consider how the set of PSSes changes when we perform the cut. The same is true when we want to change X_n^- into X_{n+1}^-.

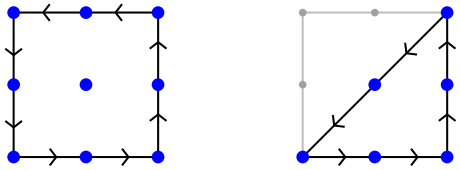


Figure 4.14: To get from the characteristic polygon on the left to the one on the right, it suffices to change the PSS set by removing two $(-1, 0)$ and two $(0, -1)$ PSSes, and adding two $(-1, -1)$ PSSes.

Suppose $n = 0$. Then $S(n + 1) = 0/1$. We desire to cut X_0^+ so that it lies inside $B(0/1)$. By our choice of X_0^+ as the *smallest* rectangle containing X, the slope $0/1$ tangent lines to X are exactly the lines defined by the horizontal sides X_0^+. So no changes are necessary to the PSSes are necessary, and we obtain a polygon X_1^+ that is identical to X_0^+. Similarly, no changes need to be made to get X_2^+ from X_1^+ since $S(2) = 1/0$ and the slope $1/0$ tangent lines to X are exactly the lines defined by the *vertical* sides of X_0^+. Hence $X_2^+ = X_0^+$. (Since this argument is true for arbitrary X, there is in future no need to consider the first two elements of the Stern-Brocot ordering when applying Gulotta's algorithm.)

For $n \geq 2$, $S(n + 1)$ is the mediant of two elements b/a and d/c of S that precede $S(n + 1)$ in the Stern-Brocot ordering, and we can write $S(n + 1) = (b + d)/(a + c)$. Further, $S(n + 1)$ is the first element of S that lies between b/a and d/c, so X_n^+ does not contain any PSSes that lie between (a, b) and (c, d) in the circle-ordering of Lemma 2.2.8. Hence the lemma tells us that the sides defined by the (a, b) and (c, d) PSSes are consecutive in the boundary of X_n^+. Similarly, the sides defined by the $(-a, -b)$ and $(-c, -d)$ PSSes are consecutive in the boundary of X_n^+ as well. The cut from X_n^+ to X_{n+1}^+ consists of the creation two new sides (possibly of length 0, if the cut is trivial) to the polygon of slope $(b + d)/(a + c)$. This can only be done by introducing $(a + c, b + d)$ and $(-a - c, -b - d)$ PSSes.

Focus on the first of these sides for now, which is comprised of $(a+c, b+d)$ PSSes. Again by Lemma 2.2.8, this new side will lie between the sides defined by the (a, b) and (c, d) PSSes at the expense of shortening the original sides (see Figure 4.13). In terms of the PSS set, we are

replacing some number n of (a, b) and (c, d) PSSes with n $(a + c, b + d)$ PSSes. The immediate question to ask is whether there are indeed at least n (a, b) and (c, d) PSSes in X_n^+. The answer here is that if n were greater than the number of (a, b) PSSes or the number of (c, d) PSSes, then the resulting polygon X_{n+1}^+ would no longer meet one of the tangent lines to X of slope b/a or d/c (see Figure 4.15). The final polygon X is a subset of X_{n+1}^+, so it would not meet this line either, contradicting the definition of a tangent line. Hence we do have n (a, b)- and (c, d)-PSSes.

Similarly, the second side created by the cut is comprised of $(-a - c, -b - d)$ PSSes, and will lie between the sides defined by the $(-a, -c)$ and $(-b, -d)$ PSSes. In terms of the PSS set, we are replacing some number m of $(-a, -b)$ and $(-c, -d)$ PSSes with m $(-a - c, -b - d)$ PSSes.

Figure 4.15: The black tangent line with slope $S(n+1)$ is not allowed to progress any further into the pictured polygon X_n^+. Otherwise, X would not intersect the red tangent line l, contradicting the requirement that $l \cap X \neq \varnothing$.

The bottom line here is that to perform the cut from X_n^+ to X_{n+1}^+, we need to remove n (a, b) and (c, d) PSSes, remove m $(-a, -b)$ and $(-c, -d)$ PSSes, introduce n $(a + c, b + d)$ PSSes, and introduce m $(-a - c, -b - d)$ PSSes. According to Theorem 3.4.9, we can do this by performing a merging of spines

$$n(a, b) + n(c, d) + m(-a, -b) + m(-c, -d) \to n(a + c, b + d) + m(-a - c, -b - d)$$

on the zigzag path diagram. Recalling that (a, b) and (c, d) are neighbours in some Stern-Brocot sequence, this merging is something that we can accomplish using the $\mathrm{SL}_2(\mathbb{Z})$ equivalence of Subsection 4.3.4.

Since X has finitely many sides, at some stage we will have cut X_0^+ along all tangent lines to X of positive slope. Now we repeat the procedure for the tangent lines of negative slope via an analogous procedure, eventually reaching X itself and concluding the algorithm.

Example 4.4.2. We first remind the reader of the first few terms of the Stern-Brocot ordering:

$$S = \left(\frac{0}{1}, \frac{1}{0}, \frac{1}{1}, \frac{1}{2}, \frac{2}{1}, \frac{1}{3}, \frac{2}{3}, \frac{3}{2}, \frac{3}{1}, \frac{1}{4}, \dots\right)$$

Suppose we want to find a dimer model with the characteristic polygon X pictured at the bottom of Figure 4.16. The smallest rectangle with vertical and horizontal sides that contains X is a 4×4 square at the top left of the figure, and the associated zigzag path diagram is drawn beside it. We begin our algorithm at the first nontrivial step, that of going from X_2^+ to X_3^+. This corresponds to cutting along X_2^+ along the tangent lines of X with slope $S(3) = 1/1$. To do this, we merge the appropriate number of $(1,0)$-, $(0,1)$-, $(-1,0)$-, and $(0,-1)$-spines into $(1,1)$- and $(-1,-1)$-spines via the proper choice of operation (refer to Section 4.3). For Figure 4.16, this involves the merging

$$3(1,0) + 3(0,1) + (-1,0) + (0,-1) \to 3(1,1) + (-1,-1)$$

via Operation III. The modified zigzag path diagram is shown in the second row of the figure, next to X_3^+.

The next step of the algorithm brings us from X_3^+ to X_4^+, where $S(4) = 1/2$. First, we identify the parents of $1/2$ in the Stern-Brocot sequence, namely $0/1$ and $1/1$. We identify the slope $1/2$ tangent lines to X and cut X_2^+ along those lines by merging $(1,0)$-, $(1,1)$-, $(-1,0)$-, $(-1,-1)$-spines into $(2,1)$- and $(-2,-1)$-spines. In our case, the quantities we need are

$$(1,0) + (1,1) + (-1,0) + (-1,-1) \to (2,1) + (-2,-1),$$

which we can perform using $\mathrm{SL}_2(\mathbb{Z})$ equivalence. With this, we have performed all nontrivial cuts of positive slope, and we set $X_0^- = X_4^+$.

It remains to perform this procedure on the negative slopes, i.e. using

$$S^- = \left(-\frac{0}{1}, -\frac{1}{0}, -\frac{1}{1}, -\frac{1}{2}, -\frac{2}{1}, -\frac{1}{3}, -\frac{2}{3}, -\frac{3}{2}, -\frac{3}{1}, -\frac{1}{4}, \dots\right).$$

Again, we know how to do this because of $\mathrm{SL}_2(\mathbb{Z})$ equivalence. In the case of our example, we need only a single cut of slope -1, with the associated merging $(1,0) + (0,-1) \to (1,-1)$ via Operation I. The resulting convex polygon X_3^- is identical to X. All we need to do is convert our zigzag path diagram into a dimer model, and we are done. The final dimer obtained from our example is shown in Figure 4.17.

Theorem 4.4.3 (Extension of Theorem 6.1, [2])**.** *At each step of the algorithm, we have the following:*

(i) We have a very good configuration of curves.

(ii) The resulting dimer model is geometrically consistent.

Proof. By induction. It is easy to check that these conditions hold for our initial dimer X_0^+. Because we create spines in the same order as the Stern-Brocot ordering, Lemma 4.3.2 immediately gives us (i). It remains to prove geometric consistency. Let Z be the zigzag path diagram obtained after some number of steps in the algorithm, and let D be the associated dimer model.

For condition (a) of geometric consistency, if a zigzag flow $\widetilde{\eta}$ self-intersects in an edge, then its projection η must self-intersect in an edge too. This is equivalent to saying that $\dot{\eta}$ self-intersects. But by construction, we never create spines on Z that self-intersect. Hence condition (a) holds.

For condition (b) of geometric consistency, suppose that $\widetilde{\eta}$ and $\widetilde{\zeta}$ are distinct zigzag flows on $\widetilde{D}$ that intersect in an edge e. Since e is a zig of one of these flows and a zag of the other, $\widetilde{\eta}$ and $\widetilde{\zeta}$ project down to distinct zigzag paths η and ζ. These zigzag paths intersect at $\pi(e)$. Suppose for the sake of contradiction that $\{[\eta], [\zeta]\}$ is linearly dependent, and consider the spines of η and ζ in Z. Since our algorithm only produces spines of primitive winding we must have $[\dot{\eta}] = \pm[\dot{\zeta}]$. Then $C_s(\dot{\eta}, \dot{\zeta}) = 0$, and by the extra crossings condition of Lemma 4.3.2, $\dot{\eta}$ does not intersect $\dot{\zeta}$. In other words, η does not intersect ζ in an edge. This contradicts the existence of $\pi(e)$. Hence if $\{[\eta], [\zeta]\}$ is linearly dependent, then $\widetilde{\eta}$ and $\widetilde{\zeta}$ do not intersect in an edge.

Finally, for condition (c) of geometric consistency, let $\widetilde{\eta}$ and $\widetilde{\zeta}$ are distinct zigzag flows on $\widetilde{D}$ with $\{[\eta], [\zeta]\}$ linearly independent. Then their projections η and ζ are distinct zigzag paths on D. Considering their associated spines in Z, by linear independence we know $C_s(\dot{\eta}, \dot{\zeta}) \neq 0$. Again by the extra crossings condition of Lemma 4.3.2, $C(\dot{\eta}, \dot{\zeta}) \neq 0$. That is, $\dot{\eta}$ and $\dot{\zeta}$ intersect, and therefore η and ζ intersect in at least one edge of D.

Suppose for the sake of contradiction that $\widetilde{\eta}$ and $\widetilde{\zeta}$ intersect two or more times. Then, picking any two consecutive intersections e_1 and e_2, we must have that one of these intersections is positively oriented, while the other is negatively oriented. Since the quotient map $\mathbb{R}^2 \to \mathbb{T}$ is a map of oriented manifolds, we know that $\pi(\eta)$ and $\pi(\zeta)$ intersect in the edges $\pi(e_1)$ and $\pi(e_2)$, one of which is positively oriented, and the other of which is negatively oriented (this also tells us that $\pi(e_1)$ and $\pi(e_2)$ are distinct). Then $\dot{\eta}$ and $\dot{\zeta}$ intersect at the vertices of the zigzag path diagram that correspond to $\pi(e_1)$ and $\pi(e_2)$. These intersections contribute $+2$ to $C(\dot{\eta}, \dot{\zeta})$ and 0 to $C_s(\dot{\eta}, \dot{\zeta})$. Then we cannot possibly have $C(\dot{\eta}, \dot{\zeta}) = |C_s(\eta, \zeta)|$, which contradicts the extra crossings condition of Lemma 4.3.2. Hence $\widetilde{\eta}$ and $\widetilde{\zeta}$ intersect in exactly one edge of $\widetilde{D}$, completing our proof of the theorem. □

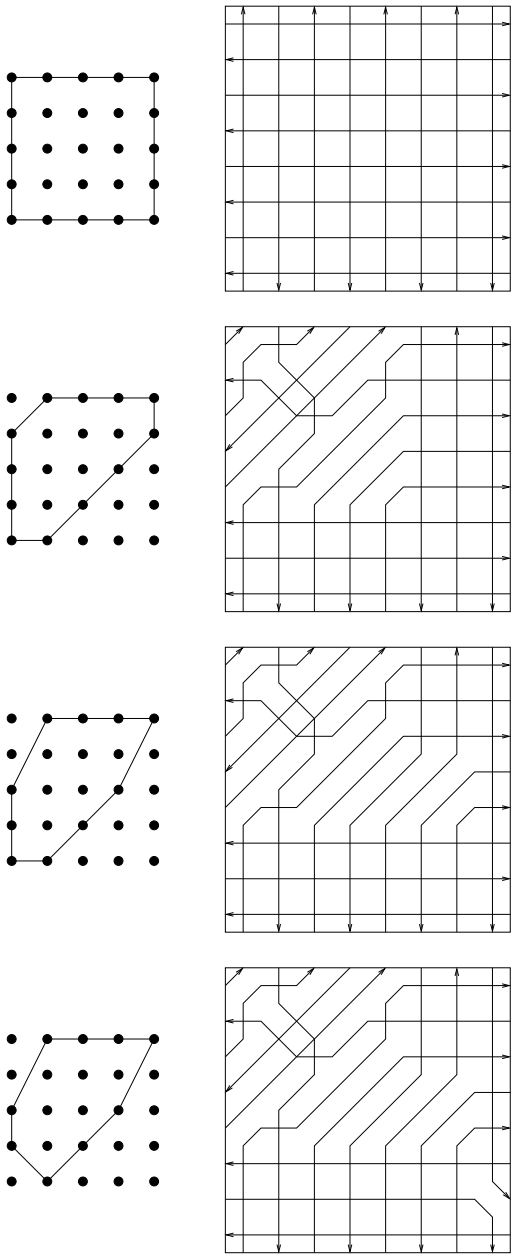


Figure 4.16: An example of the algorithm in action. The associated dimer model is shown in Figure 4.17.

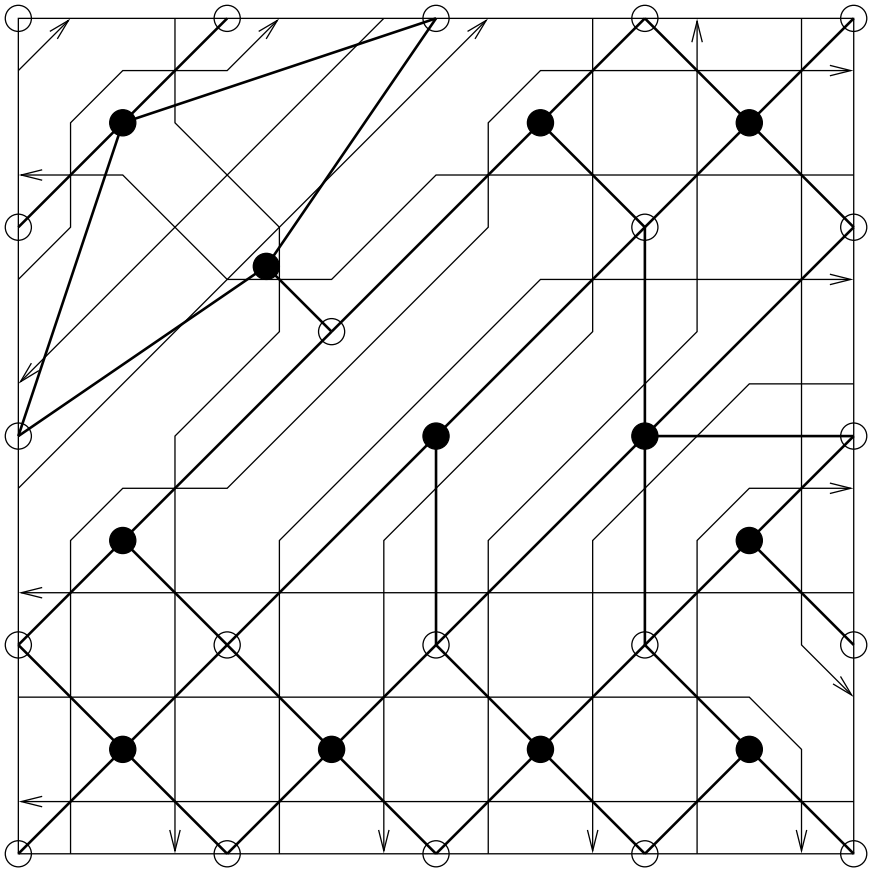


Figure 4.17: The dimer corresponding to the final zigzag path diagram in Figure 4.16.

Chapter 5

Conclusion

With Theorem 4.4.3, we are able to answer our original question of whether there exists a geometrically consistent dimer model with any given characteristic polygon. The answer is a resounding yes, and moreover, our proof is constructive. So not only do we conclude that every Gorenstein affine toric threefold admits a resolution, we can actually *find* what the resolution is.

While it is hoped that this thesis is an adequate starting point for the interested reader to continue learning about dimer models, the reality of the matter is that the literature on dimer models is extensive and split among different fields in mathematics and physics. I take this opportunity to remark about some of the research that I was *not* able to cover in this thesis. Perhaps most importantly, I was not able to explore the details behind the AdS-CFT correspondence in [20] that underlies modern interest in the dimer model. Further, in choosing to focus on modern applications, the older literature on the dimer model was largely ignored; Kasteleyn's innovative ideas touches on many interesting aspects of physics and combinatorics, such as the *enumeration* of perfect matchings. Indeed, it is not hard to count the number of perfect matchings that correspond to a particular lattice point in the boundary of the characteristic polygon, although this result is not important to the thesis.

Even within the scope of my thesis, I did not explain exactly *how* every convex polygon is associated to a toric singularity, as this requires the development of toric geometry. An introduction to this area can be found in [24], while readers with a strong background in algebra or theoretical physics are likely to prefer [23], which distils the relevant mathematics and additionally provides some context in string theory.

On the other hand, I feel that I have at least understood Gulotta's algorithm well and elucidated some of its subtleties. I have worked to fill in many missing details in the algorithm, and had the opportunity to do a number of exciting things, including formulating some original proofs and extending Gulotta's result, albeit slightly.

I conclude by saying that I have learned a lot in the course of producing this thesis. Sometimes delightful and sometimes frustrating, many hours have been put into understanding a

tiny piece of mathematical knowledge. Further, I had never imagined that the act of *writing* a thesis could be so difficult. Many times I would attempt to commit a concept to paper, only to struggle to find a concise explanation, or worse, realise that the concept itself was ill-formed in my mind. As in thinking, so too in writing—one must strive to be absolutely clear. Hopefully the lessons learned over the last year will have a lasting impact. Again, I would like to express my gratitude to Daniel Chan for his supervision, to Nathan Broomhead and Daniel Gulotta for allowing me to use their diagrams, and to you, if you have somehow made it all the way through.[1]

[1] And to my girlfriend, because apparently one mention isn't good enough.

Bibliography

[1] N. Broomhead, "Dimer models and Calabi-Yau algebras", Ph.D. Thesis, University of Bath, 2009.

[2] D. R. Gulotta "Properly ordered dimers, R-charges, and an efficient inverse algorithm", *Journal of High Energy Physics* 10 (2008), arXiv:0807.3012.

[3] P. Hall "On Representatives of Subsets", *J. London Math. Soc.*, **10** (1): 26-30.

[4] R. Kenyon and J.M. Schlenker, "Rhombic embeddings of planar quad-graphs with faces of degree 4", *Transactions of the AMS*, vol. 357, no. 9, pp. 3443-3458, 2004.

[5] A. Hanany and D. Vegh, "Quivers, Tilings, Branes and Rhombi", *Journal of High Energy Physics* 10 (2007) 029, arXiv:hep-th/0511063.

[6] R. Honsberger, *Ingenuity in Mathematics*, New York: Random House/Singer School Division, 1970.

[7] R. Diestel, *Graph Thoery*, 5th ed., Springer, 2016.

[8] J. Munkres *Topology*, 2nd ed., Pearson, 2000.

[9] R. J. Daverman and R. B. Sher, *Handbook of geometric topology*, Elsevier, 2002.

[10] K. Ueda and M. Yamazaki, "A Note on Dimer Models and McKay Quivers", *Communications in Mathematical Physics*, vol. 301, no. 3, pp. 723747, 2010, arXiv:math/0605780.

[11] A. Ishii and K. Ueda, "Dimer models and the special McKay correspondence", *Geometry & Topology*, vol. 19, no. 6, pp. 3405-3466, 2015, arXiv:0905.0059.

[12] A. Ishii and K.Ueda, "A note on consistency conditions on dimer models", 2010, arXiv:1012.5449.

[13] P. Kasteleyn, "Graph theory and crystal physics", *Graph Theory and Theoretical Physics*, pp. 43-110, Academic Press, London, 1967.

[14] J. Stienstra, "Hypergeometric systems in two variables, quivers, dimers and dessins denfants", *Modular Forms and String Duality*, pp. 125–161, 2008, arXiv:0711.0464.

[15] J. Stienstra, "Computation of Principal A-determinants through Dimer Dynamics", arXiv:0901.3681.

[16] A. Hanany and K. D. Kennaway, "Dimer models and toric diagrams", arXiv:hep-th/0503149

[17] K. D. Kennaway, "Brane Tilings", *International Journal of Modern Physics A*, **22**:18, pp. 2977-3038, 2007.

[18] S. Franco, A. Hanany, D. Martelli, J. Sparks, D. Vegh, and B. Wecht, "Gauge theories from toric geometry and brane tilings", *Journal of High Energy Physics*, (1):128, pp. 40 (electronic), 2006.

[19] S. Franco, A. Hanany, D. Martelli, J. Sparks, D. Vegh, and B. Wecht, "Brane dimers and quiver gauge theories", *Journal of High Energy Physics*, (1):128, pp. 40 (electronic), 2006.

[20] J. M. Maldacena, "The Large N limit of superconformal field theories and supergravity", *International Journal of Theoretical Physics*, vol. 38, pp. 1113-1133, 1999.

[21] R. J. Duffin, "Potential theory on a rhombic lattice", *J. Combinatorial Theory*, 5:258272, 1968.

[22] C. Merkat, "Discrete Riemann surfaces and the Ising model", *Comm. Math. Phys.*, 218(1):177216, 2001.

[23] C. Closset, "Toric geometry and local Calabi-Yau varieties: An introduction to toric geometry (for physicists)", 2009, arXiv:0901.3695.

[24] J. Maclaurin, "The Resolution of Toric Singularities", Honours Thesis, University of New South Wales, 2006.