Research Experience for Undergraduates
Research Reports
Indiana University, Bloomington
Summer 2010

October 27, 2010

Contents

Preface

During the summer of 2010 eight students participated in the Research Experiences for Undergraduates program in Mathematics at Indiana University. The program ran for eight weeks, from June 21 through August 13, 2010. Six faculty served as research advisers. Two faculty members oversaw a pair of related projects; all other faculty advised one student each.

The program opened with an introductory pizza party. On the following morning, students began meeting with their faculty mentors; these meetings continued regularly throughout the first few weeks. During week one, there were short presentations by faculty mentors briefly introducing the problem to be investigated. The week concluded with a formal reception at the IMU for REU students and mentors from around the campus. Several other IU faculty gave talks on their favorite topics during the first half of the program. Students also received orientations to the mathematics library and to our computing facilities. In week three, students gave short, informal presentations to each other on the status of work on the project. Brief training sessions on using LaTeXwere given during week four; the campus SUR program also hosted a two-day GRE preparation workshop, in which five of our students participated. Housed in a common dorm, the students themselves organized a barbecue and a trip to the Slocum puzzle collection. In week five, they received a personal tour of the cyclotron facility. During week six, we hosted the Indiana Mathematics Undergraduate Research conference, which featured 25 lectures by 44 students from Rose-Hulman Institute of Technology, Wabash College, Purdue University, and Indiana University, and ended with an hourlong panel discussion on graduate school. The next week ended with a picnic at Bryan Park. The program concluded with the students giving formal, hourlong presentations to the REU students and faculty, and the submission of final reports, contained in this volume.

It took the help and support of many different groups and individuals to make the program a success.

We thank the National Science Foundation for major financial support through the REU program through grants DMS-0453309 and DMS-0851852. We also thank Indiana University's Vice President for Diversity, Equity, and Multicultural Affairs, Dr. Edwin Marshall, and Vicki Roberts, Associate Vice President for Administration and Culture, for additional funding. We thank Indiana University for the use of facilities, including library, computers, and recreational facilities. We thank the staff of the Department of Mathematics for support, especially Mandie McCarty for coordinating the complex logistical arrangments (housing, paychecks, information packets, meal plans, frequent shopping for snacks and gluten-free items) and Cheryl Miller for her assistance in coordinating the application process. We thank Indiana graduate student Brent Stephens for serving as LaTeXconsultant and for compiling this volume.

Thanks to mathematics faculty Russell Lyons, Allan Edmonds, Michael Jolly, Elizabeth Housworth, and Bruce Solomon for giving lectures to our group. We also thank Professor Paul Kirk and Professor Carl Cowen (of Purdue Uni-

versity and IUPUI) and graduate students Kate Kearney and Steven Morrow for serving on the graduate school conference panel. Thanks to David Baxter of the Center for Exploration of Energy and Matter (nee IU cyclotron facility) for his personal tour.

This program could not exist without the faculty mentors, whose expertise and generous donation of time and energy enabled our participants to have a truly exceptional experience. A special thanks to the professors who led research projects: from Mathematics, Hari Bercovici, Chris Connell (two students), Larry Moss, Kevin Pilgrim, and Matthias Weber. This year marked the second in a row with projects in biology. We thank Distinguished Professor of Biology Michael Lynch (mentor for two students) and math graduate student JuanJuan Chai for serving as a consultant.

KMP
August, 2010

Figure 1: REU Participants, from left to right:Garrett Proffitt, Hayley Miles-Leighton, Michael Anselmi, Nathan Dowlin, Hamza Ghadyali, Linh Truong, Elizabeth Kammer, Komi Messan

Evolution of mutation rates in a population

Michael Anselmi

1 Introduction

Population genetics is a field of biology that concerns itself with the genetic basis of evolution, and differs from many other biological fields in that most of its insights are theoretical and neither observational nor experimental. [1] Just as different branches of mathematics have different central objects of study—homeomorphisms in topology; groups, rings and fields in abstract algebra—so too do the various fields of biology. In population genetics, the primary objects of study are the frequencies and fitnesses of genotypes in populations.

As far as population genetics is concerned, evolution is the change in genotype frequencies in a population over time. Unfortunately, it is impossible to observe these changes directly as the time scale of evolution can be on the order of millions of years. Consequently, we may observe the state of a population at any given point in time, but we cannot directly track how a population evolves. Fortunately all hope is not lost. In order to gain insight into how a population evolves, one constructs mathematical models and studies their behavior, checking to see if the model predicts the state or states in which we expect the population to be under certain circumstances.

In the context of population genetics, the question "how does evolution proceed?" would be expressed as "how do allele frequencies change over time?" In this paper, we attempt to model the interaction of deleterious mutations and variable mutation rates in an idealized population. We would like to gain insight into how mutational forces can steer evolution over large time scales. To this end we begin to develop an iterative model that, when supplied with biologically meaningful parameters, will yield data that reveal the distribution of mutation rates throughout any and all of a population's equilibrium states. Additionally, we would like this data to suggest a closed-form description of those equilibria.

Why is such a pursuit of interest to us? An example: many strains of bacteria are harmless to humans; in fact, the human gut is teeming with bacteria. However, for certain species of bacteria it is believed that strains which accrue mutations more quickly than the wild-type strain are largely responsible for the emergence of antibiotic resistance in the species. [2] Understanding the conditions necessary for the development of mutator strains may augment our understanding of such emergences.

Figure 1: A snippet of DNA with no base-substitution mutations

2 Background

Mutation is the ultimate source of genetic variation. Therefore, understanding the mechanisms driving evolution requires understanding the mechanisms driving mutation. We now make precise what we mean by "mutation" in this paper.

Throughout DNA we find the traditional Watson-Crick bases: adenine, thymine, cytosine, and guanine. Normally, as is depicted in Figure 1, adenine is paired with thymine and cytosine is paired with guanine. However, when we encounter what is known as a single base-substitution (henceforth referred to as a mutation), a single base mutates to another base, resulting in a mismatch. Here we assume that any given mutation leads to exactly one of two possible outcomes. Suppose the thymine in Figure 1 mutates to the guanine in Figure 2. This base pair mismatch may alter the individual's fitness (reproductive viability). The individual's fitness could increase, decrease, or remain unchanged. However, since most mutations that do alter fitness do so deleteriously, we assume that each mutation to fitness results in a reduction of fitness. We call such a mutation a type 1 mutation.

DNA-based life has a genetic repair system called mismatch repair that corrects errors like those caused by type 1 and type 2 mutations. However, if we develop a mutation at a locus responsible for the proper functioning of mismatch repair, the performance and reliability of mismatch repair is expected to decrease. This does not directly decrease an individual's fitness, but it leads to an increased rate of accrual of both type 1 and type 2 mutations in the future. A mutation of this kind is a type 2 mutation. We assume that each type 2 mutation increases an individual's mutation rate.

In short, a type 1 mutation decreases an individual's fitness but does not alter his mutation rate. A type 2 mutation does not decrease an individual's fitness but instead increases his mutation rate.

3 Model

Before we can begin developing our model, we must explicitly define our population to be modeled. We must make a number of assumptions about our

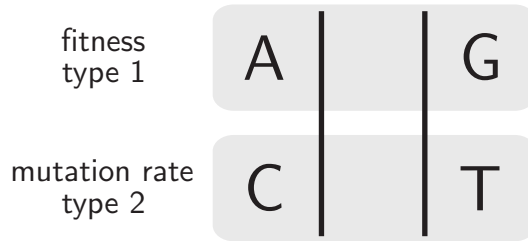


Figure 2: One of each type of mutation has occurred

population in order to develop a reasonably simple model. First, we assume that our population is isolated and is not affected by any forces not accounted for by our model. Second, our population is asexual in the sense that offspring are the product of exactly one parent. Additionally, every individual produces exactly one child and generations do not overlap. Consequently the population size remains fixed. We envision N individuals reproducing, generating N offspring, and then dying immediately afterward. Finally we assume that the population size is infinite for reasons that will become apparent later, particularly in section 4.

Every individual in our population is completely specified by two parameters x and y, which denote the number of type 1 mutations and the number of type 2 mutations an individual has, respectively. We denote the density of individuals at time t (generation t) with x number of type 1 mutations and y number of type 2 mutations by $D(t)_{x,y}$. Additionally, we say that an (x, y) individual has fitness $w(x)$ and genomic mutation rate $U(y)$, where w is a strictly decreasing function of x and U is a strictly increasing function of y.

An example fitness function w is $w(x) = (1 - s)^x$, where w maps $\mathbb{Z}^{\geq}$ (the nonnegative integers) to the half-open interval $(0, 1]$ and s is fixed in $(0, 1)$. Here, 1 is chosen to be peak fitness, with fitness decreasing toward zero as additional type 1 mutations are accumulated. In population genetics, s is referred to as the selection coefficient, a value which quantifies the deleterious effect (selective disadvantage) of a single type 1 mutation. This particular fitness function implies multiplicative epistasis of type 1 mutations, by which we mean the following: if two type 1 mutations taken individually each confer a 1/2 probability of survival, then taken together they confer a $(1/2) \cdot (1/2) = 1/4$ probability of survival. That is to say, the effects on the probability of survival of individual type 1 mutations are independent. [1] This function can be modified to be more biologically accurate by mapping to zero after sufficiently many type 1 mutations have been accumulated.

The mutation rate function U should map $\mathbb{Z}^{\geq}$ to the interval $[U_{\min}, \infty)$ or $[U_{\min}, U_{\max}]$ and should be strictly increasing. We are not yet sure what U should be, but for the time being we have assumed an exponential growth pattern and chosen U to be defined by $U(y) = e^{ky} + c$, with $k, c > 0$ fixed. Here we have a minimum mutation rate of $U_{\min} = 1 + c$. Biologically speaking, a

positive minimum mutation rate is necessary because no life has or can have a mutation rate of zero. Mathematically speaking, if $U_{\min}$ were to equal zero, then $(x, 0)$ would be an absorbing state for all x.

We wish to calculate $D(t+1)_{x',y'}$ given that $D(t)_{x,y}$ is known for all x and y. $D(t+1)_{x',y'}$ is determined in three consecutive steps:

$$D'(t)_{x,y} = D(t)_{x,y} \cdot \frac{w(x)}{\bar{w}}, \tag{1}$$

$$D(t+1)_{x',y'|x,y} = D'(t)_{x,y} \cdot T_{x,y,x',y'}, \tag{2}$$

$$D(t+1)_{x',y'} = \sum_x \sum_y D(t+1)_{x',y'|x,y}. \tag{3}$$

Steps (1) and (2) correspond to natural selection and mutation, respectively. Before a generation reproduces, we must adjust the densities of individuals according to their fitnesses and then normalize by that generation's mean fitness. In the above notation, $D(t)_{x,y}$ is adjusted to $D'(t)_{x,y}$ according to fitness. Next, we determine the density of (x, y) individuals whose offspring are of type (x', y'). That is, we calculate $D(t+1)_{x',y'|x,y}$ from $D'(t)_{x,y}$ and a four-dimensional array T of transition probabilities, where $T_{x,y,x',y'}$ denotes the probability that an (x, y) parent produces an (x', y') child. Finally, in step (3) we calculate $D(t+1)_{x',y'}$ by summing over the conditional densities of type (x', y').

We describe steps (1) and (2) in detail, beginning with step (1). We start at time 0 and initialize the population by defining various $D(0)_{x,y}$ as we please. For example, we typically set $D(0)_{0,0} = 1$ and $D(0)_{x,y} = 0$ for $x > 0$ and $y > 0$. The term $w(x)/\bar{w}$ accounts for the effect of type 1 mutations on reproductive output, decreasing $D(t)_{x,y}$ if $w(x) < \bar{w}$ and increasing $D(t)_{x,y}$ if $w(x) > \bar{w}$.

Step (2) is easily the most complicated iteration step, with the complexity lying entirely with $T_{x,y,x',y'}$. In probability notation, we have that

$$T_{x,y,x',y'} = \Pr\{\text{child } (x', y') | \text{ parent } (x, y)\}. \tag{4}$$

Assuming an infinite genome size, x' and y' become mutually independent, so we rewrite

$$= \Pr\{\text{child } x' | \text{ parent } (x, y)\} \cdot \Pr\{\text{child } y' | \text{ parent } (x, y)\}. \tag{5}$$

However, the number of and the rate of arrival of additional type 2 mutations are independent of type 1 mutations, so finally we have

$$T_{x,y,x',y'} = \underbrace{\Pr\{\text{child } x' | \text{ parent } (x, y)\}}_{P_1} \cdot \underbrace{\Pr\{\text{child } y' | \text{ parent } y\}}_{P_2}, \tag{6}$$

denoting the first and second terms on the right-hand side of (6) by P_1 and P_2, respectively. It is here that the assumption of infinite population size is necessary; (2) would not be valid otherwise, as this process would lose its determinicity and become stochastic due to binomial sampling. All that remains

to be done is model P_1 and P_2. Before continuing, we note that type 1 mutations are assumed not to back-mutate, while type 2 mutations are allowed to back-mutate with a rate independent of the forward mutation rate.

An (x', y') child of an (x, y) parent develops exactly $x' - x$ type 1 mutations not found in the parent, and inherits the remaining x type 1 mutations from the parent. Let $k = x' - x$ and let W be a random variable such that $\Pr\{W = k\} = \Pr\{\text{child } x' | \text{ parent } (x, y)\} = P_1$. Because mutations are rare events and since we are assuming an infinite genome size, W is well-approximated by a Poisson random variable with parameter λ. Here, $\lambda = U(y) \cdot f$, where f is fixed in $(0, 1)$ and denotes the proportion of incoming mutations U that will be of type 1. (Consequently, $(1 - f)$ denotes the proportion of incoming mutational events of type 2.) λ was so chosen because we require that $\mathrm{E}[W] = U(y) \cdot f$, and the expected value of any Poisson random variable is its λ parameter. We now have that

$$P_1 = m(k) = \Pr\{W = k\} = \begin{cases} \frac{\lambda^k}{k!} \cdot e^{-\lambda}, & k = 0, 1, 2, \cdots \\ 0, & k = -1, -2, -3, \cdots . \end{cases} \tag{7}$$

Modeling P_2 is a more difficult task. A y' child of a y parent accrues i additional type 2 mutations absent from the parent, and $j \leq y$ of the parent's type 2 mutations back-mutate in the child such that $i - j = y' - y$. Let X and Y be random variables denoting the incoming number of forward and backward type 2 mutations, respectively. We have

$$\mathrm{E}[X + Y] = U(y) \cdot (1 - f), \tag{8}$$

where $(1 - f)$ is the proportion of incoming mutations of type 2 (including both forward and back mutation). Thus the random variable $Z := X - Y$ is such that $\Pr\{Z = k\} = P_2$, where $k = y' - y$. Unfortunately it is not possible to let X be a Poisson random variable and Y be a binomial random variable, for then (8) forces the forward and backward type 2 mutation rates to be mutually dependent. We have not yet identified a model for P_2 suitable for the infinite genome size model. However, should a suitable probability mass function $n(k)$ of Z be found, then our model would be complete.

Finally, we halt the iteration of the system once an equilibrium state is encountered, which is when the population's various x and y densities no longer change.

4 Future work

After choosing a satisfactory model for P_2, we will then relax the assumption of infinite population size. A finite population size introduces a stochastic element called genetic drift, an unbiased (with respect to selection) dispersive evolutionary force that removes genetic variation from the population; it is mutation's counter-force. [1]

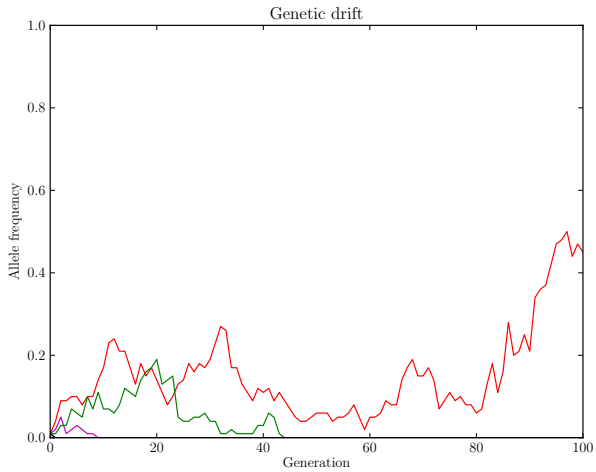


Figure 3: Allele frequency fluctuations due to genetic drift

To conceptualize how genetic drift acts on a population, consider a randomly mating diploid population of N individuals. For the purposes of genetic drift, it is equivalent to consider this population as a population of $2N$ alleles of a certain gene. Assuming no evolutionary forces other than genetic drift, the reproductive process is the following:

1. Randomly select one of the $2N$ alleles from the parent generation.
2. Duplicate the selected allele.
3. Place the duplicate in the new generation. [1]

Suppose our gene under consideration has two mutually exclusive alleles: A and B. Further suppose that our population consists of 100 alleles and that the initial frequencies of A and B are 99/100 and 1/100, respectively. Then there will be a $(1 - 1/100)^{100} \approx 37\%$ chance that the B allele with vanish from the population in only one generation. However it is also possible that genetic drift can elevate a new allele to fixation. In Figure 3 we see the effects of drift on the B allele frequency in several populations, each of which consists of 100 alleles and has B's initial frequency set to 1/100.

5 Acknowledgments

I thank my advisor, Professor Michael Lynch, for his advice and guidance this summer, along with attempting to teach me a semester's worth of population genetics in only a few weeks' time. I also thank Matthew Ackerman for similar

reasons. Finally, many thanks to Kevin Pilgrim and the National Science Foundation for making this REU possible, and to Amanda McCarty for the many delicious snacks provided throughout the summer.

References

[1] John H. Gillespie. *Population Genetics: A Concise Guide*. The Johns Hopkins University Press, 2nd edition, 2004.

[2] Emmanuel Tannenbaum, Eric Deeds, and Eugene I. Shakhnovich. Equilibrium distribution of mutators in the single fitness peak model. *Physical Review Letters*, 91:138105, 2003.

Torsion Subgroups of CAT(0) Groups

Nathan Dowlin *Hamza Ghadyali*

Abstract

Given a CAT(0) group G acting geometrically on a proper CAT(0) space, we attempt to demonstrate that any torsion subgroup of G has finite cardinality.

1 Introduction

Euclid's fifth postulate, commonly referred to as the parallel postulate, states that given a line ℓ and a point p not on ℓ there is exactly one line through p that does not intersect ℓ. When this postulate is removed, we allow for a much larger class of metric spaces which can have peculiar properties. For example, in hyperbolic geometry there are infinitely many lines through p that do not intersect ℓ. Furthermore, in elliptic geometry no such line exists since every line through p will intersect ℓ. Indeed, these are salient features of $\mathbb{H}^n$, hyperbolic n-space, and $\mathbb{S}^n$, the n-sphere, respectively. Together, $\mathbb{H}^n$, $\mathbb{S}^n$, and $\mathbb{E}^n$ (Euclidean n-space), are the model spaces of constant negative, positive, and 0 curvature, respectively.

CAT(0) spaces are of particular interest because while they are incredibly general, the lack of positive curvature provides us with many powerful tools. In this paper, we are primarily concerned with CAT(0) groups, or groups which act geometrically on proper CAT(0) spaces. More precisely, a CAT(0) group is a subgroup of the group of isometries of a proper CAT(0) space that acts properly discontinuously and cocompactly on the space. In studying these groups, one natural question that arises is whether there can exist an infinite subgroup consisting entirely of torsion (finite order) elements. Not only is this question interesting in its own right, but an important cut point theorem by Eric Swenson requires this to be true in order to complete his proof. For $\mathbb{E}^n$, Bieberbach has shown that no such infinite torsion subgroup can exist, and the same result has been shown for δ-hyperbolic groups; we studied Olshanskii's proof of this theorem in great depth. CAT(0) spaces lie somewhere in between these two, and a similar result is expected for CAT(0) groups.

We use various geometric approaches to arrive at a proof of this result. At first, since the CAT(0) case most closely resembles the situation in δ-hyperbolic groups, we attempted to follow Olshanskii's proof substituting CAT(0) groups for δ-hyperbolic groups. However, without word-hyperbolicity, word-processing in the CAT(0) group became incredibly complicated very quickly. Instead, we resorted to the geometry of CAT(0) spaces and examined how it must restrict

the groups acting on them. Because Swenson's proof requires only that torsion subgroups fixing a point in the boundary at infinity to be finite, we assume that our torsion subgroups do in fact have a fixed point in the boundary. Using this assumption, we have derived many characteristics regarding the manner in which torsion groups act on horospheres. While we have been unable to prove the theorem in general, these characteristics have allowed us to show something weaker about a torsion group T fixing a point c in the boundary at infinity. We proved that if the boundary at infinity of the horosphere at c consists only of the point c, the torsion subgroup can only be infinite if there exists an infinite torsion group with a finite number of conjugacy classes, which is currently a major open question in group theory.

2 Background

We provide the basic definitions that recur throughout this paper here.

Definition 2.1. Let (X, d_x) and (Y, d_y) be a metric spaces. An *isometry* between X and Y is a map $\phi\colon X \to Y$ such that $\forall x, x' \in X$, $d_x(x, x') = d_y(\phi(x), \phi(x'))$. A *(unit speed) geodesic* in X is an isometry $\gamma\colon I \to X$ where I is a connected subset of $\mathbb{R}$. If $I = [a, b]$, $I = [a, \infty)$, or $I = \mathbb{R}$, then γ is a *geodesic segment, ray, or line*, respectively.

Definition 2.2. Let X be a metric space. We call X a *geodesic metric space* if $\forall x, x' \in X$ there is a geodesic segment between x and x', (ie $\exists\gamma\colon [0, a] \to X$ such that $\gamma(0) = x$ and $\gamma(a) = x'$). If $\forall x, x' \in X$ this geodesic segment is unique, X is a *unique geodesic metric space.*

The definition below for a CAT(0) space comes from [Ru].

Definition 2.3. Let (X, d) be a metric space. X is *proper* if closed metric balls are compact. Let (X, d) be a proper complete geodesic metric space. If $\triangle abc$ is a geodesic triangle in X, then we consider $\triangle \bar{a}\bar{b}\bar{c}$ in $\mathbb{E}^2$, a triangle with the same side lengths, and call this a *comparison triangle.* Let $\triangle abc$ be a geodesic triangle in X. Then $\triangle abc$ satisfies the CAT(0) *inequality* if for any comparison triangle and any two points p, q on $\triangle abc$, the corresponding points $\bar{p}, \bar{q}$ on the comparison triangle satisfy

$$d(p, q) \leq d_{\mathbb{E}^2}(\bar{p}, \bar{q})$$

If every geodesic triangle in X satisfies the CAT(0) inequality, then we say X is a CAT(0) space.

Remark 2.4. If (X, d) is a CAT(0) space, then the following hold:

(a) The distance function $d\colon X \times X \to \mathbb{R}$ is convex.

(b) X is a unique geodesic metric space.

(c) X is contractible.

Definition 2.5. G is a CAT(0) group if there exists a CAT(0) space X such that G is a subgroup of the group of isometries of X and G acts both properly discontinuously and cocompactly on X. (G acts cocompactly if the fundamental domain is compact.)

Example 2.6. A simple example of a CAT(0) group is $\mathbb{Z}^2$ with the action on $\mathbb{R}^2$ by translations.

Definition 2.7. Let G be a CAT(0) group, and let L be a proper subgroup. Let S be a set of generators such that $< S >= L$. If $g \in G \setminus L$, then the group $< S \cup g >$ will be called the *extension* of L by g. For notational convenience, this group will be written as $Ext(L, g)$.

Definition 2.8. Let G be a CAT(0) group, and let T be a torsion subgroup. We say that T is a *maximal torsion subgroup* of G if $\forall g \in G \setminus T$, the group $Ext(T, g)$ is not a torsion group (i.e. it contains a translation).

Definition 2.9. Two geodesic rays $\gamma, \gamma' \colon [0, \infty) \to X$ are said to be *asymptotic* if there exists a constant K such that $d(\gamma(t), \gamma'(t)) \leq K$ for all $t \geq 0$. This gives an equivalence relation. The set of all equivalence classes forms the *boundary at infinity of* X, denoted $\partial_\infty X$. If γ is in the equivalence class c, we say γ *goes to* c or γ *approaches* c.

The action of an isometry on X has a natural extension to $\partial_\infty X$. Let $g \in Isom(X)$, where $Isom(X)$ is the group of all isometries from the metric space X to itself. Given two geodesic rays $\gamma, \gamma' \colon [0, \infty) \to X$, if there exists K such that $d(\gamma(t), \gamma'(t)) \leq K$ for all $t \geq 0$, then $d(g(\gamma(t)), g(\gamma'(t))) \leq K$ for all $t \geq 0$. Therefore, $g \circ \gamma$ and $g \circ \gamma'$ are in the same equivalence class, so g preserves equivalence classes of geodesic rays and hence extends to a map on the boundary.

Example 2.10. It is easy to see that both $\partial_\infty \mathbb{E}^2$ and $\partial_\infty \mathbb{H}^2$ is $\mathbb{S}^1$.

Definition 2.11. Let $B(p, r)$ denote the ball centered at p of radius r. Let $\gamma \colon [0, \infty) \to X$ with $\gamma(0) = x_0$ be a geodesic ray emanating from x_0 in the equivalence class c. We define the *horoball at* c *determined by* γ to be

$$\mathcal{B}_{c,\gamma} = \bigcup_{t \in [0,\infty)} B(\gamma(t), d(\gamma(t), x_0))$$

Since $\gamma(0) = x_0$ and γ a geodesic, $d(\gamma(t), x_0)) = t$. The definition can then be written more concisely as

$$\mathcal{B}_{c,\gamma} = \bigcup_{t \in [0,\infty)} B(\gamma(t), t)$$

The corresponding *horosphere at* c *determined by* γ denoted by $\mathcal{H}_{c,\gamma}$ is the boundary of the horoball $\mathcal{B}_{c,\gamma}$. With this notation, the geodesic ray γ is required to be in the equivalence class c. When the context is clear or we are referring to any horosphere at c, we will drop the second subscript.

Definition 2.12. Equivalently, we can define the horosphere using the Busemann function. Let γ be a geodesic ray defined as in the previous definition. Given a point $y \in X$ we define the *Busemann function* with respect to γ

$$b_\gamma(y) = \lim_{t\to\infty}(d(\gamma(t), y) - t)$$

The horosphere is then the level set $\mathcal{H}_{c,\gamma} = b_\gamma^{-1}(0)$, and similarly the horoball is $\mathcal{B}_{c,\gamma} = b_\gamma^{-1}((-\infty, 0])$.

Definition 2.13. We extend the notion of boundary at infinity to horoballs and horospheres as follows. If $\mathcal{B}_c$ is a horoball at c and $\mathcal{H}_c$ is the corresponding horosphere, then we will denote the boundary at infinity of these sets by $\partial_\infty \mathcal{B}_c$ and $\partial_\infty \mathcal{H}_c$ respectively.

$\partial_\infty \mathcal{B}_c$ is defined to be the set of equivalence classes of all geodesic rays γ which, if $\gamma(0) \in \mathcal{B}_c$, then $\gamma(t) \in \mathcal{B}_c$ for all $t \in [0, \infty)$. $\partial_\infty \mathcal{H}_c$ is naturally the boundary of $\partial_\infty \mathcal{B}_c$ as a subset of $\partial_\infty X$.

3 Results

The main theorems we use regarding properties of CAT(0) groups and their torsion subgroups can be found in [BH]. These results will be stated without proof. Throughout this section, G refers to a CAT(0) group acting on a CAT(0) space X properly discontinuously and cocompactly, and T denotes a subgroup of G consisting of only torsion elements. We will make the assumption throughout that T fixes a point in the boundary at infinity, and we will call this point c.

Lemma 3.1. *Every torsion element t fixes a point $p_t \in X$*

Proof. Let $t \in G$ be a torsion element and define $F =< t >$. Let $x \in X$ and let Fx be the orbit of x under F. Since F is finite, this is a bounded set. Therefore, $\mathcal{C}(Fx)$ is a compact, convex set, where $\mathcal{C}(Fx)$ denotes the convex hull of the set Fx. Moreover, $t(Fx) = Fx$, so $t(\mathcal{C}(Fx)) = \mathcal{C}(Fx)$. Applying the Schauder Fixed Point Theorem, t fixes a point in $\mathcal{C}(Fx)$. □

Remark 3.2. Since each torsion element t fixes a point $p_t \in X$, it also fixes the geodesic ray emanating from p_t in the equivalence class c.

Lemma 3.3. *CAT(0) groups have finitely many conjugacy classes of finite subgroups.*

Lemma 3.3 implies directly that CAT(0) groups have finitely many conjugacy classes of finite order elements. Therefore, given a torsion group $T \subseteq G$, the equivalence relation $t_1 \sim t_2$ if $\exists g \in G$ such that $t_1 = gt_2g^{-1}$ divides T into a finite number of equivalence classes. We will call these equivalence classes *conjugacy classes of T over G.*

Lemma 3.4. *Let $\mathcal{H}_{c,\gamma}$ be a horosphere at c. If $g \in T$, then g stabilizes $\mathcal{H}_{c,\gamma}$.*

Proof. Let $\mathcal{H}_{c,\gamma}$ be the horosphere at c determined by the geodesic ray γ. Define $x \in \mathcal{H}_{c,\gamma}$ by $x = \gamma(0)$. Let $g \in T$, and let $y \in X$ be a fixed point for g. Then g fixes α, where α is the geodesic ray from y to c. Let $a \in \mathcal{H}_{c,\alpha}$.

$$b_\alpha(ga) = \lim_{t\to\infty}(d(\alpha(t), ga) - t)$$

Since g^{-1} is an isometry,

$$d(\alpha(t), ga) = (d(g^{-1}(\alpha(t)), a)$$

and

$$b_\alpha(ga) = \lim_{t\to\infty}(d(g^{-1}(\alpha(t)), a) - t)$$

But for all t, $g^{-1}(\alpha(t)) = \alpha(t)$, so $b_\alpha(ga) = \lim_{t\to\infty}(d(\alpha(t), a) - t) = b_\alpha(a) = 0$. This is true for all a in $\mathcal{H}'_c$, so $g(\mathcal{H}_{c,\alpha}) = \mathcal{H}_{c,\alpha}$. The geodesic ray α intersects all horospheres inside $\mathcal{H}_{c,\alpha}$, and this intersection is fixed by g, so the same argument applies to all horospheres inside $\mathcal{H}'_c$. Therefore, if $\mathcal{H}_c \subset \mathcal{B}_{c,\alpha}$, then $\mathcal{H}_{c,\gamma}$ is stabilized by g.

Assume $\mathcal{H}_{c,\alpha}$ lies outside $\mathcal{H}_{c,\alpha}$. Then γ intersects $\mathcal{H}_{c,\alpha}$; let $z = \gamma(t_0)$ be the point at which they intersect. Note that $d(x, z) = t_0$, so $\mathcal{H}_{c,\gamma}$ is the unique horosphere a distance t_0 outside of $\mathcal{H}_{c,\alpha}$.

Let $b \in \mathcal{H}_c$. As before,

$$b_\gamma(gb) = \lim_{t\to\infty}(d(\gamma(t), gb) - t) = \lim_{t\to\infty}(d(g^{-1}(\gamma(t)), b) - t)$$

Since g^{-1} fixes c, $g^{-1}(\gamma)$ is also a geodesic ray going to c. Therefore, g maps $\mathcal{H}_c$ to another horosphere at c. Additionally, we have

$$d(x, z) = t_0 = d(g^{-1}(x), g^{-1}(z))$$

and g^{-1} stabilizes $\mathcal{H}'_c$, so $g^{-1}(z)) \in \mathcal{H}'_c$. Combining this with the fact that $g^{-1}(x)$ and $g^{-1}(z)$ lie on a geodesic ray going to c, we see that $g^{-1}(\mathcal{H}_c)$ is either the horosphere t_0 inside $\mathcal{H}'_c$ or the horosphere t_0 outside $\mathcal{H}'_c$. But if $g^{-1}(\mathcal{H}_c)$ is inside $\mathcal{H}'_c$, then so is $g(g^{-1}(\mathcal{H}_c))$, which contradicts our assumption that $\mathcal{H}_c$ lies outside $\mathcal{H}'_c$. $g^{-1}(\mathcal{H}_c)$ must then be the horosphere t_0 outside $\mathcal{H}'_c$, which is $\mathcal{H}_c$. Hence, g stabilizes the horosphere $\mathcal{H}_c$.

□

Lemma 3.5. *Given a horosphere $\mathcal{H}_c$, if there exists a geodesic line γ such that $\gamma \subseteq H_c$, then $\partial_\infty \mathcal{H}_c$ contains at least two points.*

Proof. Define $\alpha, \beta\colon [0, \infty) \to X$ by $\alpha(t) = \gamma(t)$ and $\beta(t) = \gamma(-t)$. Then α and β are geodesic rays in $\mathcal{H}_c$. Assume $\partial_\infty \mathcal{H}_c$ consists of only c. Then α and β are geodesic rays from $\gamma(0)$ to c, so by uniqueness of geodesics, $\alpha = \beta$. But if $\gamma(-1) = \gamma(1)$, then the geodesic segment $\gamma([-1, 1])$ is a loop, which is a contradiction. Therefore, $\partial_\infty \mathcal{H}_c$ contains at least two points. □

Lemma 3.6. *Let $x \in X$ be a translation, and let $a, b \in \partial_\infty X$ be endpoints of the axis of x. If there exists $c \in \partial_\infty X$ such that $xc = c$ and $a \neq c \neq b$, then the axis of x is contained in a horosphere at c.*

Proof. Let $\gamma\colon [0, \infty) \to X$ denote the axis of h, and let r be the distance that h translates points on the axis. Note that since γ is the axis of a translation, it is also a geodesic line. Assume γ is not contained in any horosphere at c.

Let $\mathcal{H}_c$ be a horosphere at c such that $\mathcal{H}_c \cap \gamma$ consists of exactly two points at least r apart (we can do this because horospheres are convex; we just have to choose one large enough). Call these two points u and v, with

$$d(u, v) = d(xu, v) + r$$

Then hv and $h^{-1}u$ are inside the horosphere $\mathcal{H}_c$ and hu and $h^{-1}v$ lie outside $\mathcal{H}_c$.

Let $\alpha\colon [0, \infty) \to X$ be the geodesic ray from u to c. Then $\mathcal{H}_c = b_\alpha^{-1}(0)$. h fixes c, so $h(\alpha)$ and $h^{-1}(\alpha)$ are geodesics going to c. Let $x \in \mathcal{H}_c$. Then

$$b_\alpha(hx) = \lim_{t\to\infty} (d(\alpha(t), hx) - t) = \lim_{t\to\infty} (d(h^{-1}(\alpha(t)), x) - t) = b_{h^{-1}\alpha}(x)$$

This is the same for any x in $\mathcal{H}_c$, so h maps $\mathcal{H}_c$ to another horosphere at c. Call this horosphere $\mathcal{H}'_c$.

So we have that $h(\mathcal{H}_c) = \mathcal{H}'_c$. But $h(u)$ lies outside the horosphere $\mathcal{H}_c$ and $h(v)$ lies inside, so they can not lie on the same horosphere at c, so we have reached a contradiction. Therefore, γ is contained in some horosphere at c. $\square$

Proposition 3.7. *Let $\mathcal{H}_c$ be a horosphere at c, where $c \in \partial_\infty X$. If there exists $c' \in \partial_\infty X$ and a horosphere $\mathcal{H}_{c'}$ at c' such that $c \in \partial_\infty \mathcal{H}_{c'}$, then $c' \in \partial_\infty \mathcal{H}_c$.*

Lemma 3.8. *Let C be an infinite conjugacy class of T (over G), and let $m \in C$. Then $C = \{x_i m x_i^{-1}\}_{i=1}^\infty$, where $x_i \in G$. Suppose each x_i a translation. Let $a_i \in \partial_\infty X$ be the point farthest from c which is fixed by $x_i m x_i^{-1}$ and let c_i be the point at infinity towards which x_i translates. If $\lim_{i\to\infty} x_i c = \lim_{i\to\infty} a_i = c$ and $\lim_{i\to\infty} c_i = c$, then $\lim_{i\to\infty} | < x_i m x_i^{-1} m^{-1} > | = \infty$. In particular, if x is a translation towards c, then $| < x m x^{-1} m^{-1} > | = \infty$.*

Proof. Sketch of proof.

We will begin with the particular case, when x is a translation towards c. Let $\gamma\colon (-\infty, \infty) \to X$ be the axis of x oriented such that $\gamma(\infty) = c$, and r be the distance that x maps points on the axis. Define $\alpha\colon [0, \infty) \to X$ by $\alpha(t) = \gamma(t)$. Then x maps the horosphere $\mathcal{H}_{c,\alpha}$ to the horosphere $\mathcal{H}_{c,\beta}$, where $\beta(t) = \gamma(t+r)$. Geometrically, x maps a horosphere at c to the unique horosphere a distance r inside it.

Now examine xmx^{-1}. Since x^{-1} maps horospheres at c *outward* by r, it acts as a contraction on geodesics going to c centered at γ, so it can be thought of as contracting the horosphere as it moves it outwards. The torsion element m acts as a rotation or reflection about some fixed set in the horosphere, so the contraction has very little effect on the action of m except that it can translate

this fixed set, so m is rotating or reflecting about a different point or set. x then maps this horosphere back to the original horosphere by a dilation about γ. The element xmx^{-1} therefore acts in the same way as m, but centered at a different point on the horosphere. We believe that this implies $xmx^{-1}m^{-1}$ is a translation, and we get our intuition from the manifold case. A simple example is rotations in the Euclidean plane: if a is a rotation by θ centered at p and b is a rotation by θ centered at q, then ab^{-1} is a translation.

In the more more general part of our lemma, we have an infinite conjugacy class $C = \{x_i m x_i^{-1}\}_{i=1}^{\infty}$. Let x_i and t_i have the properties assumed in the lemma. Then the sequence $\lim_{i\to\infty} x_i m x_i^{-1} m^{-1}$ is limiting to the above case, so it is limiting to a translation. Therefore, the order must approach infinity. If we were to compare this to Euclidean case mentioned previously, this sequence is analogous to a sequence ab_i^{-1}. As before, a is a rotation by θ centered at p, and we define b_i to be a rotation by θ_i centered at qi. If $\lim_{i\to\infty} \theta_i = \theta$ and if $\lim_{i\to\infty} q_i = q$, then the element ab_i^{-1} is a rotation by an angle ϕ_i, with $\lim_{i\to\infty} \phi_i = 0$. Thus, $\lim_{i\to\infty} | < ab_i^{-1} > | = \infty$.

□

Theorem 3.9 (Main Theorem)**.** *Let T be a maximal torsion subgroup of G, and let $c \in \partial_\infty X$ be fixed by T. If $\partial_\infty H_c = \{c\}$, then T has a finite number of conjugacy classes.*

Proof. We have that $\forall t \in T$, $tc = c$. Let $W \subseteq T$ be the subset consisting of all elements that fix at least one other point in the boundary (i.e. $t \in W$ if $t \in T$ and $\exists c_t \in \partial_\infty X$ such that $c_t \neq c$ and $tc_t = c_t$).

Assume W is a finite set. Recall that G has a finite number of conjugacy classes of finite order elements, so T has a finite number of conjugacy classes over G. Let C be one such conjugacy class, and let $m \in C$. Each $t \in C$ can then be written as xmx^{-1} for some $x \in G$. But $t(xc) = xmx^{-1}(xc) = xm(c)$. Since $m \in T$, $mc = c$, and $t(xc) = xc$. Therefore, xc is a fixed point for t. We assumed that W is finite, so this tells us that if $G_c \subset G$ is the stabilizer subgroup of G fixing c, then T has a finite number of conjugacy classes over G_c. Let $C_1, C_2, ...C_n$ be the conjugacy classes of T over G_c, and let $m_i \in C_i$. Each element $t \in T$ can be written xm_jx^{-1}, where $t \in C_j$ and $x \in G_c$.

Case 1: $\forall t \in T, \exists x \in G_c$ such that x is torsion and $t = xm_jx^{-1}$ for some j. All such x's fix c, and T is maximal, so either every such x is an element of T, or $\exists x$ such that the group extension of T by x is not a torsion group.

In the latter case, there must be some element $z \in Ext(T, x)$ such that $| < z > | = \infty$. But T and x both stabilize horospheres at c, so z must as well. Let a and b be the endpoints of the axis of z. Let $\mathcal{H}_c$ be a horosphere at c, and let $p \in \mathcal{H}_c$. Then $\forall n \in \mathbb{Z}, z^n(p) \in \mathcal{H}_c$, so $\lim_{n\to\infty} z^n(p) = a \in \partial_\infty \mathcal{H}_c$ and $\lim_{n\to-\infty} z^n(p) = b \in \partial_\infty \mathcal{H}_c$. Since $a \neq b$, this violates $\partial_\infty \mathcal{H}_c = \{c\}$.

Therefore, every such x is an element of T, so T has a finite number of conjugacy classes.

Case 2: There exists $t \in C_j$ such that $t = xm_jx^{-1}$ and x a translation fixing c. If c is an endpoint for the axis of x, then by Lemma 3.8 $| < tm^{-1} > | = \infty$,

which violates T being a torsion group. If c is not an endpoint for the axis of x, then by Lemma 3.6 the axis of x is contained in a horosphere at c. The axis is a geodesic, so Lemma 3.5 tells us that $\partial_\infty \mathcal{H}_c$ contains at least two points, so $\partial_\infty \mathcal{H}_c \neq \{c\}$.

Therefore, whenever W is finite, T has a finite number of conjugacy classes.

Assume now that W is infinite. Since T has a finite number of conjugacy classes over G, at least one of these conjugacy classes must contain an infinite subset of W. Let C be one such conjugacy class, and let $m \in W \cap C$. Let $S \subset \partial_\infty X$ be the set of points at infinity fixed by m. Since $m \in W$, S consists of at least two points.

Each element $t \in C$ can be written as xmx^{-1} for some x in G. Let $a \in S$.

$$t(xa) = (xmx^{-1})(xa) = xm(a) = xa$$

Therefore, if $t = xmx^{-1}$, then t fixes the set xS. We will now enumerate elements of C as $\{x_i m x_i^{-1}\}_{i=1}^{\infty}$, and define $t_i = x_i m x_i^{-1}$. Each t_i fixes $x_i S$; let c_i be the point in $x_i S$ farthest from c. Each t_i fixes both c and c_i, so applying Lemma 3.4, t_i stabilizes $\mathcal{H}_c \cap \mathcal{H}_{c_i}$, where $\mathcal{H}_c$ and $\mathcal{H}_{c_i}$ are any two horospheres at c and c_i, respectively.

Pick a point p in X. Let γ be the geodesic ray from p to c, and let α_i be the geodesic ray from p to c_i. The sequence $\{t_i p\}_{i=1}^{\infty}$ is infinite, so by discreteness it must approach infinity. For all integers i, $p \in \mathcal{H}_{c,\gamma} \cap \mathcal{H}_{c_i,\alpha_i}$, so $t_i p \in \mathcal{H}_{c,\gamma} \cap \mathcal{H}_{c_i,\alpha_i}$. Therefore, in the limit as i goes to infinity, the intersection $\mathcal{H}_{c,\gamma} \cap \mathcal{H}_{c_i,\alpha_i}$ must also approach infinity.

Now let us examine the sequence $\{c_i\}_{i=1}^{\infty}$. Since $\partial_\infty X$ is sequentially compact, there is a convergent subsequence $\{c_{i_j}\}_{j=1}^{\infty}$.

Case 1: There exists a convergent subsequence $\{c_{i_j}\}_{j=1}^{\infty}$ that converges to a point $a \neq c$. Let β be the geodesic ray from p to a. The sequence of intersections $\mathcal{H}_{c,\gamma} \cap \mathcal{H}_{c_{i_j},\alpha_{i_j}}$ must then approach c, so c must be in the boundary at infinity of $\lim_{j\to\infty} \mathcal{H}_{c_{i_j},\alpha_{i_j}}$ which equals $\mathcal{H}_{a,\beta}$. Therefore, $c \in \partial_\infty H_{a,\beta}$. By Proposition 3.7, $a \in \partial_\infty \mathcal{H}_{c,\gamma}$. Since $a \neq c$, this contradicts our assumption that $\partial_\infty \mathcal{H}_c = \{c\}$.

Case 2: Every subsequence $\{c_{i_j}\}_{j=1}^{\infty}$ converges to c. Then the sequence itself must converge to c.

$$\lim_{i\to\infty} c_i = c$$

Recall that c_i is the farthest fixed point from c for a given t_i. Therefore, $\lim_{i\to\infty} x_i S = c$. Since S contains at least two points, it has non-zero diameter. Hence, the elements x_i must be acting as contractions on $\partial_\infty X$ near c, with the fixed point of these contractions limiting to c. The only isometries that behave in this way on the boundary are translations towards points approaching c. Therefore, we can apply Lemma 3.8 to derive our contradiction.

$$lim_{i\to\infty} | < t_i m^{-1} > | = \infty$$

This implies that either there exists an infinite order element, in which case T is not torsion, or there is no bound on the order of elements in T, which violates Lemma 3.3.

This completes the proof. □

4 Conclusion and Future Research

We have yet to prove the general theorem that torsion subgroups of CAT(0) groups are finite. However, we have shown that in a particular case, the question can be reduced to a major open question regarding torsion groups. Many believe that infinite torsion groups with a finite number of conjugacy classes do not exist, and if this turns out to be true, then our case would be proved.

The one problem with our proof is that we have been unable to rigorously prove Lemma 3.9, which is essential for our proof. We will keep working on this part of the paper, and if a proof continues to elude us then we will try other approaches for the subcases requiring this lemma. Once this section is complete, we have a few ideas which may be applied to proving the theorem in general. In the case we proved, we assumed that $\mathcal{H}_c$ has only one point in its boundary at infinity, and we often derived our contradiction by showing that there was a geodesic in the horosphere $\mathcal{H}_c$. Even without our assumption, this information may be useful. A geodesic line in the horosphere $\mathcal{H}_c$ may imply that the CAT(0) space X is equal to the product $\mathbb{R} \times Y$, where Y is a CAT(0) space. If the CAT(0) group G acting on X has an infinite torsion subgroup, then perhaps we can construct a CAT(0) group G' which acts on Y geometrically and contains an infinite torsion subgroup T'. By induction, this would imply that X is infinite dimensional, so it cannot be proper. These are the potential directions in which we may take this research project in the coming months.

5 Acknowledgments

We are deeply grateful to our advisor, Chris Connell, who contaminated us with his endless enthusiasm and passion for the project. He was our indefatigable tour guide through the universe of negatively curved spaces, and we cannot thank him enough for all that we learnt with him this summer. We would also like to thank Kevin Pilgrim, for his excellent direction of the program, the NSF for providing this invaluable opportunity for undergraduates, and our good friend Linh Truong who provided thoughtful insights at many stages along the way.

References

[BH] M.R. Bridson, A. Haefliger: Metric Spaces of Non-positive Curvature, Grundl. Math. Wiss., Vol. 319, Springer, Berlin 1999.

[IO] S. Ivanov and A. Yu. Olshanskii. Hyperbolic groups and their quotients of bounded exponents, Trans. Amer. Mat. Soc. 348 (1996), 2091-2138.

[Ru] K. Ruane. CAT(0) Boundaries of Truncated Hyperbolic Space, Proc. of Topology, 29 (2005), no.1, 317-331.

[Sw] E. Swenson. A cut point theorem for CAT(0) groups, J. Differential Geom. 53:2 (1999), 327-358.

Systems of Natural Logic with Adjectives and Their Completeness

Elizabeth Kammer

Abstract

In natural logic, the goal is to create a system of logic that is as similar to natural language as possible. In order to build a natural logic, simple sentence forms are considered, slowly incorporating more language throughout time. As background, we first consider the fragments of the form *All* X *are* Y and then add *No* X *are* Y, where X and Y are nouns. We will look at the rules of logic and completeness of their proof systems. Next, we will introduce the idea of intersecting adjectives which are adjectives with meaning separate from the noun being modifying. For example, the plural noun red cars means the intersection of everything red and everything that is a car. This paper will then present versions of the simple systems that contain intersecting adjectives and will discuss their completeness.

1 Introduction

Natural logic is concerned with fragments of natural language and the logic derived from systems using only these sentences. In particular, we are interested in the completeness of these systems of logic. We will begin with a language consisting of one sentence and slowly adding more phrases and looking at their completeness theorems as background. Then we will move on to systems including intersecting adjectives and discuss the completeness of their system.

Historically, logic was largely a study of syllogisms. The classic example of a syllogism is as follows:

$$\begin{array}{l} \textit{All men are mortal.} \\ \textit{Socrates is a man.} \\ \hline \textit{Socrates is mortal.} \end{array}$$

In this syllogism, whatever makes the two statements above the line true will also make the statement below the line true. It is also important to notice that if we have any three sentences of these particular forms, models which make the first two true will also make the third sentence true. We can show this as:

$$\begin{array}{l} \textit{All } m \textit{ are } n. \\ S \textit{ is a } m. \\ \hline S \textit{ is } n. \end{array}$$

We explore this type of idea within our systems of natural logic. We look at simple sentence forms with variables as placeholders for nouns. We create a system of rules based on what can be derived using only these sentences, and study the completeness of these rules.

First we look at the simple systems as done by Larry Moss[1]. These systems initially contain only the fragments of the form *All* X *are* Y and then add those like *No* X *are* Y, where X and Y are nouns. We consider at the rules of logic and completeness of their proof systems. The following systems of logic involve intersecting adjectives. This type of adjective includes those which have meanings separate from the nouns they modify. For example, it Jane is a female student, and Jane is an athlete, we can deduce that Jane is a female athlete. Intersecting adjectives include things such as colors, but does not include adjectives such as tall or short. For intersecting adjectives, we assume that the adjectives and nouns are sets of things described by that particular word, and an intersecting adjectives is the set of things in the intersection of the adjectives and nouns. Intersecting adjectives include things such as colors, but does not include adjectives such as tall or short. If an object is a red car, we take the intersection of things that are red and things that are cars; however, if something is a short child, this is not the intersection of all short things and children because a short five year old is very different from a short twelve year old. Also note that we can add adjectives productively, meaning that a noun could have multiple adjectives modifying it. An example of this would be that many countries have red, white, and blue flags. We will allow finitely many adjectives to modify a noun in our system. After looking at some systems for background, we will look at those including adjectives and their completeness.

2 Background

Before looking at a particular system, there are two concepts, syntax and semantics, which we have in all systems. The work presented in the background is not original, but will help develop an understanding of the systems which we are considering. We will discuss ideas within each of these that we will use throughout the rest of the paper. Then we will begin to look at systems done by Larry Moss[1], building up from the simplest to the more complicated systems.

For the first idea, syntax, we have narrow forms that sentences can take. When working with basic nouns we use $X, Y, ...$ to represent nouns and when dealing with adjectives we denote the intersecting adjectives either by red or when used productively as $c_1, c_2, ..., c_k$. When representing nouns, possibly with or without adjectives, attached we use the variables m, n. In any of our systems Γ denotes a sets of statements in that particular logic that can take the form of the sentences in our syntax. In the syntax for our system, we also deal with sentences of the forms *All* $X(n)$ *are* $Y(m)$, *No* $X(n)$ *are* $Y(m)$, and *Some* $X(n)$ *are* $Y(m)$, using only some of these sentences depending on our particular system. In the syntax, $\Gamma \vdash S$ means that there is a proof tree with leaves from Γ and has the root S that only follows the rules of our system.

When looking at the semantics of a system, we first create a model $\mathcal{M}(M, [\![\ \ ]\!])$, which contains a set M, and a subset $[\![X]\!] \subseteq M$ for each variable X and $[\![red]\!] \subseteq M$ for each adjective. In semantics, we get the following:

$$\begin{array}{lll} \mathcal{M} \models \textit{All } X \textit{ are } Y & \textit{iff} & [\![X]\!] \subseteq [\![Y]\!] \\ \mathcal{M} \models \textit{Some } X \textit{ are } Y & \textit{iff} & [\![X]\!] \cap [\![Y]\!] \neq \emptyset \\ \mathcal{M} \models \textit{No } X \textit{ are } Y & \textit{iff} & [\![X]\!] \cap [\![Y]\!] = \emptyset \end{array}$$

We allow $[\![X]\!]$ to be empty, and in this case, $\mathcal{M} \models \textit{All } X \textit{ are } Y$ vacuously. For a Γ, we use $\mathcal{M} \models \Gamma$ to mean that $\mathcal{M} \models S$ for all $S \in \Gamma$. Additionally, we have the following semantic meaning in our systems with adjectives:

$$[\![c_1 c_2 ... c_k X]\!] = [\![c_1]\!] \cap [\![c_2]\!] \cap ... \cap [\![c_k]\!] \cap [\![X]\!]$$

For our final semantic definition, we write $\Gamma \models S$ to mean that every model with makes all sentences in Γ true also make S true.

Next we present the idea of the Soundness of our systems.

Lemma 2.1 (Soundness)**.** *If $\Gamma \vdash S$, then $\Gamma \models S$.*

For a proof, see work by Larry Moss[1]. That all of our systems are sound tells us that anything we can prove in our systems, will also be true semantically. This guarantees that our systems do not produce nonsensical results. We would also like the converse of this statement to be true.

Theorem 2.2 (Completeness)**.** *If $\Gamma \models S$, then $\Gamma \vdash S$.*

Completeness guarantees that any result we can prove semantically, will also be provable in our system. We leave the proofs of completeness to the individual systems.

2.1 All

We begin with the simplest system, $\mathcal{L}(all)$, which utilizes only sentences of the form *All X are Y*. The rules of logic for this system are shown in Figure 1. And present a proof of completeness. First, we present an example to display the difference between semantics and syntactical proofs.

Example 2.3. Let $\Gamma =$ All A are B, All B are C, All C are D.

Claim 1: $\Gamma \models$ All A are D.

From Γ, we have the following

$$[\![A]\!] \subseteq [\![B]\!], [\![B]\!] \subseteq [\![C]\!], [\![C]\!] \subseteq [\![D]\!].$$

So, we have

$$[\![A]\!] \subseteq [\![D]\!].$$

Claim 2: $\Gamma \vdash$ All A are D.

$$\dfrac{\dfrac{\text{All } A \text{ are } B \quad \text{All } B \text{ are } C}{\text{All } A \text{ are } C} \qquad \text{All } C \text{ are } D}{\text{All } A \text{ are } D}$$

$$\frac{}{\textit{All } X \textit{ are } Y} \qquad \frac{\textit{All } Y \textit{ are } Z \quad \textit{All } X \textit{ are } Z}{\textit{All } X \textit{ are } Y}$$

Figure 1: The Rules of Logic for $\mathcal{L}(all)$.

And now we turn to the completeness of $\mathcal{L}(all)$.

Theorem 2.4. *The logic of $\mathcal{L}(all)$ is complete.*

Proof. Suppose $\Gamma \models S$ and let S be All X are Y. We begin by making a model $\mathcal{M}(\Gamma)$. Let

$$M = \text{all variables in } \Gamma$$

and we set the semantics of any variable be

$$[\![V]\!] = \{W : \Gamma \vdash \text{All } W \text{ are } V\}$$

Claim: $\mathcal{M} \models \Gamma$ Let All A are $B \in \Gamma$. We must show $[\![A]\!] \subseteq [\![C]\!]$. Let $P \in [\![A]\!]$. Then we have that $\Gamma \vdash$All P are A. And we can get the following proof tree

$$\frac{\begin{array}{c}\vdots\\ \text{All } P \text{ are } A\end{array} \quad \text{All } A \text{ are } B}{\text{All } P \text{ are } B}$$

Since $\Gamma \vdash$ All P are B, $B \subseteq [\![{}^*B^*]\!]$. From this we can conclude that $[\![A]\!] \subseteq [\![C]\!]$.

Since $\mathcal{M} \models \Gamma$ and $\Gamma \models$ All X are Y, $\mathcal{M} \models$ All X are Y. Therefore, we have $[\![X]\!] \subseteq [\![Y]\!]$. Since we have

$$\frac{}{\text{All } X \text{ are } Y}$$

we know that $X \in [\![X]\!]$ and $X \in [\![Y]\!]$. Therefore, $\Gamma \vdash$ All X are Y. □

Thus we have the completeness of $\mathcal{L}(all)$ of our simplest proof system. From here, we move on to systems that add complexity to this system.

2.2 All and No

We expand our language to also contain sentences of the form *No X are Y*. Note that *No X are X* means that there are no X. And, in addition to the rules of $\mathcal{L}(all)$, the system $\mathcal{L}(all, no)$ also contains the rules listed in Figure 2. Again, we are interested in the completeness of the system.

Theorem 2.5. *The logic of $\mathcal{L}(all, no)$ is complete.*

Proof. Let Γ be a set of sentences in $\mathcal{L}(all, no)$ Suppose $\Gamma \models S$. We consider the model M(Γ) where M = set of sets a such that the following are true

if $V \in a$ and $\Gamma \vdash$ All V are W, then $W \in a$
if $V, W \in a$, then $\Gamma \nvdash$ No V are W

we set the semantics of any variable to be

$$[\![Z]\!] = \{a \in \mathcal{M} : Z \in a\}$$

We claim that $\mathcal{M} \models \Gamma$. By our first condition, we have that if *All* X *are* Y belongs to Γ, then $[\![X]\!] \subseteq [\![Y]\!]$. If *No* X *are* Y belongs to Γ, then let $a \in [\![X]\!]$, so $X \in a$. By the second condition, $Y \notin a$ and therefore, $a \notin [\![Y]\!]$. Which shows that $[\![X]\!] \cap [\![Y]\!] = \emptyset$. So we have that $\mathcal{M} \models \Gamma$.

Since we have that $\mathcal{M} \models \Gamma$ and $\Gamma \models S$, we have $\mathcal{M} \models S$. First we consider the case where S is *All* X *are* Y. Let

$$a = \{V : \Gamma \vdash \text{All } X \text{ are } V\}$$

Case I: $a \notin \mathcal{M}$ Then there must be some $A, B \in a$ such that $\Gamma \vdash$ *No* A *are* B. Then we get the following proof tree

$$\dfrac{\dfrac{\begin{array}{c}\vdots\\ \text{All } X \text{ are } B\end{array} \qquad \dfrac{\dfrac{\begin{array}{c}\vdots\\ \text{All } X \text{ are } A\end{array} \quad \begin{array}{c}\vdots\\ \text{No } A \text{ are } B\end{array}}{\text{No } X \text{ are } B}}{\text{No } B \text{ are } X}}{\text{No } X \text{ are } X}}{\text{All } X \text{ are } Y}$$

Case II: $a \in \mathcal{M}$. Then since $a \in [\![X]\!]$, we have $a \in [\![Y]\!]$. Therefore, $Y \in a$, and $\Gamma \vdash$ All X are Y.

Next we consider the case when S is *No* X *are* Y. Here, we let

$$a = \{V : \Gamma \vdash \text{All } X \text{ are } V or \Gamma \vdash \text{All } Y \text{ are } V\}$$

Notice that $X, Y \in a$. We claim that $a \notin \mathcal{M}$. To see this notice that if $a \in \mathcal{M}$, we have that $a \in [\![X]\!] \cap [\![Y]\!]$. Which implies that $[\![X]\!] \cap [\![Y]\!] \neq \emptyset$ which contradicts $\mathcal{M} \models$ *No* X *are* Y. So we do in fact have that $a \notin \mathcal{M}$, implying that there are some $V, W \in a$ such that $\Gamma \vdash$ *No* V *are* W. There are four possible cases, depending on if $\Gamma \vdash$ *All* X *are* V or $\Gamma \vdash$ *All* Y *are* V and $\Gamma \vdash$ *All* X *are* W or $\Gamma \vdash$ *All* Y *are* W. We first explore the case where $\Gamma \vdash$ *All* Y *are* V and $\Gamma \vdash$ *All* X *are* W.

$$\dfrac{\dfrac{\begin{array}{c}\vdots\\ \text{All } X \text{ are } W\end{array} \qquad \dfrac{\dfrac{\begin{array}{c}\vdots\\ \text{All } Y \text{ are } V\end{array} \quad \begin{array}{c}\vdots\\ \text{No } V \text{ are } W\end{array}}{\text{No } Y \text{ are } W}}{\text{No } W \text{ are } Y}}{\text{No } X \text{ are } Y}$$

The proof tree for $\Gamma \vdash$ *All* X *are* V and $\Gamma \vdash$ *All* Y *are* W is similar. Next we consider the case where $\Gamma \vdash$ *All* X *are* V and $\Gamma \vdash$ *All* X *are* W.

$$\frac{\textit{All } X \textit{ are } Y \quad \textit{No } Y \textit{ are } Z}{\textit{No } X \textit{ are } Z} \qquad \frac{\textit{No } X \textit{ are } Y}{\textit{No } Y \textit{ are } X} \qquad \frac{\textit{No } X \textit{ are } X}{\textit{All } X \textit{ are } Y}$$

Figure 2: The rules logic of $\mathcal{L}(all, no)$ when combined with the rules of logic for $\mathcal{L}(all)$.

$$\dfrac{\dfrac{\dfrac{\vdots}{\text{All } X \text{ are } W} \quad \dfrac{\dfrac{\dfrac{\vdots}{\text{All } X \text{ are } V} \quad \dfrac{\vdots}{\text{No } V \text{ are } W}}{\text{No } X \text{ are } W}}{\text{No } W \text{ are } X}}{\text{No } X \text{ are } X}}{\text{All } X \text{ are } Y} \quad \dfrac{\vdots}{\text{No } X \text{ are } X}}{\text{No } X \text{ are } Y}$$

The case where $\Gamma \vdash$ *All* Y *are* V and $\Gamma \vdash$ *All* Y *are* W follows similarly. Therefore, we have that $\Gamma \vdash S$. $\square$

We also have know that $\mathcal{L}(all, no)$ is complete.

2.3 All and Intersecting Adjectives

In this section, we look at the simplest logic including adjectives. There are many similarities between this case and $\mathcal{L}(all)$; however, we add a few rules concerning adjectives that are not derivable from the our original logic. Additionally, the proof of completeness of $\mathcal{L}(all, adjectives)$ is similar to that of our first completeness proof.

Theorem 2.6. *The logic of $\mathcal{L}(all, adjectives)$ is complete.*

Proof. Suppose $\Gamma \models S$ and let S be All n are m. We begin by making a model $\mathcal{M}(\Gamma)$. Let

$$M = \text{all nouns in } \Gamma$$

and we set the semantics of any variable be

$$[\![V]\!] = \{n : \Gamma \vdash \text{All } n \text{ are } V\}$$

$$[\![red]\!] = \{m : \Gamma \vdash \text{All } n \text{ are } red\ p \text{ for some } p \in M\}$$

Claim: $\mathcal{M} \models \Gamma$ Let All n are $m \in \Gamma$. There are a number of cases, dependent on n, m, but we will only explore a few as others are similar.

Case I: $n = X, m = red\ Y$ We must show $[\![X]\!] \subseteq [\![red]\!] \cap [\![Y]\!]$. Let $p \in [\![X]\!]$. Then we have that $\Gamma \vdash$All p are X. And we can get the following proof trees

$$\frac{\dfrac{\vdots}{\text{All } p \text{ are } X} \quad \text{All } A \text{ are } red\ Y}{\text{All } p \text{ are } red\ Y}$$

$$\frac{}{\textit{All } n \textit{ are } m} \qquad \frac{\textit{All } m \textit{ are } l \quad \textit{All } n \textit{ are } l}{\textit{All } n \textit{ are } m}$$

$$\frac{}{\textit{All red } n \textit{ are } n} \qquad \frac{\textit{All } n \textit{ are } m \quad \textit{All } n \textit{ are red } m}{\textit{All } n \textit{ are red } l}$$

Figure 3: The rules of logic for $\mathcal{L}(all, adjectives)$.

$$\frac{\begin{array}{c}\vdots\\ \text{All } p \text{ are } red\ Y\end{array} \quad \overline{\text{All } red\ Y \text{ are } Y}}{\text{All } p \text{ are } Y}$$

It follows that $p \in [\![red]\!]$ and $p \in [\![Y]\!]$. So we have that $[\![X]\!] \subseteq [\![red]\!] \cap [\![Y]\!]$.

Case II: $n = red\ blue\ X, m = green\ Y$ We must show $[\![red]\!] \cap [\![blue]\!] \cap [\![X]\!] \subseteq [\![green]\!] \cap [\![Y]\!]$. Let $p \in [\![red]\!] \cap [\![blue]\!] \cap [\![X]\!]$, then $p \in [\![red]\!], p \in [\![blue]\!]$, and $p \in [\![X]\!]$, so $\Gamma \vdash$ All p are $red\ a$, $\Gamma \vdash$ All p are $blue\ b$, and $\Gamma \vdash$ All p are X. And get the following trees

$$\frac{\frac{\begin{array}{c}\vdots\\ \text{All } p \text{ are } red\ a\end{array} \quad \frac{\begin{array}{c}\vdots\\ \text{All } p \text{ are } blue\ b\end{array} \quad \begin{array}{c}\vdots\\ \text{All } p \text{ are } X\end{array}}{\text{All } p \text{ are } blue\ X}}{\text{All } p \text{ are } red\ blue\ X} \quad \text{All } red\ blue\ X \text{ are } green\ Y}{\text{All } p \text{ are } green\ Y}$$

$$\frac{\begin{array}{c}\vdots\\ \text{All } p \text{ are } green\ Y\end{array} \quad \overline{\text{All } green\ Y \text{ are } Y}}{\text{All } p \text{ are } Y}$$

Other cases are done similarly. It follows that $\mathcal{M} \models \Gamma$.

Since $\mathcal{M} \models \Gamma$ and $\Gamma \models$ All n are m, $\mathcal{M} \models$ All n are m. Here again, we have many cases, but will look at one to understand the general method.

$n = red\ X, m = blue\ green\ Y$ Therefore we have $[\![red]\!] \cap [\![X]\!] \subseteq [\![blue]\!] \cap [\![green]\!] \cap [\![Y]\!]$. We know that $red\ X \subset [\![red]\!] \cap [\![X]\!]$ because we have the following trees

$$\frac{\overline{\text{All } red\ X \text{ are } red\ X}}{\overline{\text{All } red\ X \text{ are } X}}$$

we know that $X \in [\![red]\!] \cap [\![X]\!]$ and $X \in [\![blue]\!] \cap [\![green]\!] \cap [\![Y]\!]$. Therefore, $\Gamma \vdash$ All $red\ X$ are $blue\ green\ Y$. $\square$

The proof of the completeness of $\mathcal{L}(all, adjectives)$ concludes our introduction to systems of logic their completeness.

$$\frac{\textit{All } n \textit{ are } m \quad \textit{No } m \textit{ are } l}{\textit{No } n \textit{ are } l} \qquad \frac{\textit{No } n \textit{ are } m}{\textit{No } m \textit{ are } n}$$

$$\frac{\textit{No } n \textit{ are } n}{\textit{All } n \textit{ are } m} \qquad \frac{\textit{No } n \textit{ are red } m}{\textit{No } m \textit{ are red } n}$$

Figure 4: The rules logic of $\mathcal{L}(all, adjectives, no)$ when combined with the rules of logic for $\mathcal{L}(all, adjectives)$.

3 All, No, and Adjectives

Next we look at the combination of sentences of the forms *All n are m* and *No n are m*, where n and m are nouns with or without adjectives. The rules for this system can be found in Figure 4, in addition to those rules for $\mathcal{L}(all, adjectives)$. We claim that $\mathcal{L}(all, no, adjectives)$ is also complete.

Theorem 3.1. *The logic of $\mathcal{L}(all, no, adjectives)$ is complete.*

Proof. Let Γ be a set of sentences in $\mathcal{L}(all, no)$ Suppose $\Gamma \models S$. We consider the model M(Γ) where M = set of sets a such that the following are true

if $n \in a$ and $\Gamma \vdash$ All n are m, then $m \in a$
if $n, m \in a$, then $\Gamma \not\vdash$ No V are W
if $n, red\ m \in a$, then $red\ n \in a$

and we set the semantics of any variable and adjective to be

$$[\![Z]\!] = \{a \in \mathcal{M} : Z \in a\}$$

$$[\![red]\!] = \{b \in \mathcal{M} : red\ p \in a, p \text{ some noun} \in \Gamma\}$$

We claim that $\mathcal{M} \models \Gamma$.

Case I: All n are $m \in \Gamma$ It should be noted that in the case where *All X are Y* belongs to Γ, then by the same argument used from $\mathcal{L}(all, no)$ that $[\![X]\!] \subseteq [\![Y]\!]$. Now we look at the general case with finitely many colors. If *All* $c_1 ... c_j X$ *are* $d_1 ... d_k Y \in \Gamma$, we must show that

$$[\![c_1]\!] \cap [\![c_2]\!] \cap ... \cap [\![c_j]\!] \cap [\![X]\!] \subseteq [\![d_1]\!] \cap [\![d_2]\!] \cap ... \cap [\![d_k]\!] \cap [\![Y]\!]$$

Let $a \in [\![c_1]\!] \cap [\![c_2]\!] \cap ... \cap [\![c_j]\!] \cap [\![X]\!]$. Then $c_1 p_1, ..., c_j p_j, X \in a$. By our third condition since $c_j p_j \in a$ and $X \in a$, it follows that $c_j X \in a$. Repeating this arguement $j - 1$ more times, we hae that $c_1 ... c_j X \in a$. Since $\Gamma \vdash$ *All* $c_1 ... c_j X$ *are* $d_1 ... d_k Y$, we know $d_1 ... d_k Y \in a$. And therefore $a \in [\![d_1]\!]$. And we can get the following proof tree

$$\frac{}{\text{All } d_1 ... d_k Y \text{ are } d_2 ... d_k Y}$$

Which tells us that $d_2 ... d_k Y \in a$. Repeating this argument, we find that $a \in [\![d_1]\!] \cap [\![d_2]\!] \cap ... \cap [\![d_k]\!] \cap [\![Y]\!]$. And therefore we have

$$[\![c_1]\!] \cap [\![c_2]\!] \cap ... \cap [\![c_j]\!] \cap [\![X]\!] \subseteq [\![d_1]\!] \cap [\![d_2]\!] \cap ... \cap [\![d_k]\!] \cap [\![Y]\!]$$

Case II: *No n are m* belongs to Γ. Notice that in this case, if *No X are Y* $\in \Gamma$, the proof from $\mathcal{L}(all, no)$ shows that that $[\![X]\!] \cap [\![Y]\!] = \emptyset$. We look at *No* $c_1...c_jX$ *are* $d_1...d_kY \in \Gamma$, we must show that

$$[\![c_1]\!] \cap [\![c_2]\!] \cap ... \cap [\![c_j]\!] \cap [\![X]\!] \cap [\![d_1]\!] \cap [\![d_2]\!] \cap ... \cap [\![d_k]\!] \cap [\![Y]\!] = \emptyset$$

Suppose not, and let $a \in [\![c_1]\!] \cap [\![c_2]\!] \cap ... \cap [\![c_j]\!] \cap [\![X]\!] \cap [\![d_1]\!] \cap [\![d_2]\!] \cap ... \cap [\![d_k]\!] \cap [\![Y]\!]$, then by our previous arguments $c_1...c_jX \in a$, $d_1...d_kY \in a$ and $a \in \mathcal{M}$. But By our third condition, since *blue q* and $X \in a$, *blue* $X \in a$. Similarly since *red p* and *blue* $X \in a$, we have *red blue* $X \in a$. By the same argument, we can conlude that *green yellow* $Y \in a$. But we have that $\Gamma \vdash$ *No* $c_1...c_jX$ *are* $d_1...d_kY$, which contradicts the second condition for $a \in \mathcal{M}$. So we have that $[\![c_1]\!] \cap [\![c_2]\!] \cap ... \cap [\![c_j]\!] \cap [\![X]\!] \cap [\![d_1]\!] \cap [\![d_2]\!] \cap ... \cap [\![d_k]\!] \cap [\![Y]\!] = \emptyset$ and therefore also that $\mathcal{M} \models \Gamma$.

Since we have that $\mathcal{M} \models \Gamma$ and $\Gamma \models S$, we have $\mathcal{M} \models S$. First we consider the case where S is of the form *All n are m*. Here we consider the general case where S is *All* $c_1...c_jX$ *are* $d_1...d_kY$. Let

$$a = \{n : \Gamma \vdash \text{All } c_1...c_jX \text{ are } n\}$$

Case I: $a \notin \mathcal{M}$ Then there must be some $m, l \in a$ such that $\Gamma \vdash$ *No m are l*. Then we get the following proof tree

$$\dfrac{\dfrac{\vdots \qquad \dfrac{\dfrac{\vdots \qquad \vdots}{\text{All } c_1...c_jX \text{ are } m \quad \text{No } m \text{ are } l}}{\text{No } c_1...c_jX \text{ are } l}}{\text{All } c_1...c_jX \text{ are } l \qquad \dfrac{\text{No } c_1...c_jX \text{ are } l}{\text{No } l \text{ are } c_1...c_jX}}}{\dfrac{\text{No } c_1...c_jX \text{ are } c_1...c_jX}{\text{All } c_1...c_jX \text{ are } d_1...d_kY}}$$

So we have that $\Gamma \vdash$ *All* $c_1...c_jX$ *are* $d_1...d_kY$.

Case II: $a \in \mathcal{M}$. Then since $a \in [\![c_1]\!] \cap ... \cap [\![c_j]\!] \cap [\![X]\!]$, we have $a \in [\![d_1]\!] \cap ... \cap [\![d_k]\!] \cap [\![Y]\!]$. Therefore, $d_1p_1, ..., d_kd_k, Y \in a$. Using previous arguments we have that $d_1...d_kY \in a$, so $\Gamma \vdash$ *All* $c_1...c_jX$ *are* $d_1...d_kY$.

Next we consider the case when S is of the form *No n are m*. Specifically we will consider *No* $c_1...c_jX$ *are* $d_1...d_kY$. Here, we let

$$a = \{n : \Gamma \vdash \text{All } c_1...c_jX \text{ are } nor \Gamma \vdash \text{All } d_1...d_kY \text{ are } n\}$$

Notice that $c_1...c_jX, d_1...d_kY \in a$. We claim that $a \notin \mathcal{M}$. To see this notice that if $a \in \mathcal{M}$, we have that $a \in [\![c_1]\!] \cap ... \cap [\![c_j]\!] \cap [\![X]\!] \cap [\![d_1]\!] \cap ... \cap [\![d_k]\!] \cap [\![Y]\!]$. Which implies that $[\![c_1]\!] \cap ... \cap [\![c_j]\!] \cap [\![X]\!] \cap [\![d_1]\!] \cap ... \cap [\![d_k]\!] \cap [\![Y]\!] \neq \emptyset$. So we do in fact have that $a \notin \mathcal{M}$, implying that there are some $n, m \in a$ such that $\Gamma \vdash$ *No n are m*. There are four possible cases, depending on if $\Gamma \vdash$ *All* $c_1...c_jX$ *are* n or $\Gamma \vdash$ *All* $d_1...d_kY$ *are* n and $\Gamma \vdash$ *All* $c_1...c_jX$ *are* m or $\Gamma \vdash$ *All* $d_1...d_kY$ *are*

$$\frac{\textit{All } n \textit{ are } m \quad \textit{Some } n \textit{ are } l}{\textit{Some } m \textit{ are } l} \qquad \frac{\textit{Some } n \textit{ are } m}{\textit{Some } m \textit{ are } n}$$

$$\frac{\textit{Some } n \textit{ are } m}{\textit{Some } n \textit{ are } n} \qquad \frac{\textit{Some } n \textit{ are red } m}{\textit{Some } m \textit{ are red } n}$$

Figure 5: The rules logic of $\mathcal{L}(all, some, adjectives)$ when combined with the rules of logic for $\mathcal{L}(all, adjectives)$.

m. We first explore the case where $\Gamma \vdash$ *All* $d_1 ... d_k Y$ *are* n and $\Gamma \vdash$ *All* $c_1 ... c_j X$ *are* m.

$$\frac{\begin{array}{c}\vdots\\ \text{All } c_1...c_jX \text{ are } m\end{array} \qquad \frac{\frac{\begin{array}{c}\vdots\\ \text{All } d_1...d_kY \text{ are } n\end{array} \quad \begin{array}{c}\vdots\\ \text{No } n \text{ are } m\end{array}}{\text{No } d_1...d_kY \text{ are } m}}{\text{No } m \text{ are } d_1...d_kY}}{\text{No } c_1...c_jX \text{ are } d_1...d_kY}$$

We also have the case where $\Gamma \vdash$ *All blue green* Y *are* m and $\Gamma \vdash$ *All red* X *are* n which has a similar proof tree. Next we explore the case where $\Gamma \vdash$ *All* $c_1 ... c_j X$ *are* n and$\Gamma \vdash$ *All* $c_1 ... c_j X$ *are* m.

$$\frac{\frac{\frac{\begin{array}{c}\vdots\\ \text{All } c_1...c_jX \text{ are } m\end{array} \qquad \frac{\frac{\begin{array}{c}\vdots\\ \text{All } c_1...c_jX \text{ are } n\end{array} \quad \begin{array}{c}\vdots\\ \text{No } n \text{ are } m\end{array}}{\text{No } c_1...c_jX \text{ are } m}}{\text{No } m \text{ are } c_1...c_jX}}{\text{No } c_1...c_jX \text{ are } c_1...c_jX}}{\text{All } c_1...c_jX \text{ are } d_1...d_kY} \qquad \begin{array}{c}\vdots\\ \text{No } c_1...c_jX \text{ are } c_1...c_jX\end{array}}{\text{No } c_1...c_jX \text{ are } d_1...d_kY}$$

Similarly, we have the case where *All* $d_1 ... d_k Y$ *are* n and *All* $d_1 ... d_k Y$ *are* n. Therefore, we know that $\Gamma \vdash S$. □

We concluded with the completeness of the combinations of systems we have thus far discussed.

4 All, Some, Adjectives

As discussed by Moss[1], there is also a language which combines sentences of the form *All* X *are* Y and *Some* X *are* Y. A direction that we began to explore but can be continued is to add adjectives to $\mathcal{L}(all, some)$. We present here what we believe to be the complete rules of logic for $\mathcal{L}(all, some, adjectives)$.

5 Acknowledgments

I'd like to thank Professor Larry Moss for his support and guidance throughout this summer. Addition, many thanks to Professor Kevin Pilgrim for organizing and making this experience possible.

References

[1] L. S. Moss: Completeness Theorems for Syllogistic Fragments.

Average time until fixation of a mutant allele in a given population

KOMI MESSAN

Abstract

One of the important problems in population genetics is how long it takes for a gene to go to fixation (become established). A mutant gene in a given population will eventually be lost or established. The particular interest of this research is to know the mean time for a mutant gene to become fixed in a population, and we will exclude the case when this gene is lost. A diploid population of N individuals will be considered with a forward and backward mutation of u and v respectively per basis. Using a set of nonlinear equations, we will first calculate the genotype frequencies which will allow us to find the equilibrium points for the infinite population. With the diffusion theory, we will approximate the time to fixation for finite populations. We will then proceed with a numerical approximation using C++ to see a close result for the problem.

1 Introduction

The main idea in population genetics is evolution. Evolution is much different from most studies in biology for the fact that its insights are theoretical rather than observational or experimental. Most evolutionary studies concern the frequencies or the fitnesses of genotype in a given population. Evolution is the change in the frequencies of genotype through time, perhaps due to their differences in fitness (Gillespie 2004). Evolution can also be explained by two forces: forces that introduce variation in phenotypic character such as eye colors, height or certain behaviors and forces that make some traits become rare or more common. The main cause of variation is mutation, which changes the sequence of a gene (Strickberger 2000). In other words, mutation is a change in the DNA sequence of Cell's genome.

The forces that make traits to become common or rare are caused by two main processes. One of these processes is natural selection which is a key term used in genetic evolution (Strickberger 2000). Natural selection is the differential reproductive success of a any given organism. Very often, organisms produce more offspring than their environment can support; because of this, not every individual in a population survives in the generation and this can be one of the main cause of natural selection. Over many generations mutation produces random changes in traits, which are then filtered by natural selection and the beneficial traits retained (Gillespie 2004). Another cause of evolution is genetic drift, which is a change in the relative frequencies in which gene variant occurs

in a population due to random sampling and chance. These random changes affect evolution in two important ways. First, a dispersive force that removes genetic variations from population. Let us note that the rate of removal is very weak since it is inversely proportional to the population size. Tahe other is the effect of drift on the probability of survival of a new mutation (Gillespie 2004).

Another evolutionary process we will see later in this paper is genetic recombination. In this process, the DNA or sometimes the RNA molecule breaks and then joins another DNA molecule. Recombination can also have a big impact on the evolutionary processes and this was shown by Paul G. Higgs (Higgs 1997). We will take the same approach to show this but in a higher dimension.

2 Background

We will consider a diploid population consisting of N individuals and having the variance effective number N_e. Let us note that N_e may be different than N and a good explanation of N_e can be found in "KIMURA and CROW 1963". Throughout this paper, we will develop a model that has been introduced by Paul G. Higgs (Higgs 1997) and some other authors (Michalakis and Slatkin, 1996; Phillips, 1996; Stephan, 1996). Most of these authors develop a model in which mutation is irreversible but we will consider a reversible mutation in this paper using the same model.

Our model will involve 2 loci, each with two alleles. The two alleles will be labeled A and a at one locus and B and b at the other. We will therefore have four genotypes: ab, Ab, aB, and AB. The frequencies of ab and AB will respectively be denoted x_0 and x_2. Both the double mutant genotypes have a frequency denoted x_1. We are therefore assuming these double mutant genotype have the same fitness.T genotype ab has fitness 1 and AB has fitness $1 - s_2$. However, the double mutant genotypes have fitness $1 - s_1$. Let us note that $x_0 + 2x_1 + x_2 = 1$.

Throughout this paper, we will assume both u and v are $< 10^{-6}$, and both s_1 and s_2 will be in the range 0.01 to 0.005. For different values of selection, mutations (forward and backward mutations), computer simulation will be used to approximate the time at which the first allele will arrive at genotype AB. Consider it starts from the genotype ab.

3 Equilibrium points of the infinite population

Prior to looking at the changes in the finite population, we will first look at the genotype frequencies in the infinite population. If we call $x_0, x_1, and x_2$ the frequencies at generation t then the frequencies at generation $t+1$ will be denoted $X_0^*, X_1^*, and X_2^*$. Considering all our parameters are different from zero, we obtain the following set of nonlinear equations:

$$(X_0)^* = (1-2u)x_0 + 2vx_1 + 2s_1x_0x_1 - r\left(-x_1^2 + x_0x_2\right) \tag{1}$$

$$(X_1)^* = 2ux_0 + (2-2u-2v)x_1 - 2s_1x_0x_1 + 2vx_2 + 2(-s_1+s_2)x_1x_2 + 2r\left(-x_1^2 + x_0x_2\right) \quad (2)$$

$$(X_2)^* = 2ux_1 + (1-2v)x_2 - 2(-s_1+s_2)x_1x_2 - r\left(-x_1^2 + x_0x_2\right) \quad (3)$$

Because of the complexity of the equations, we will try to simplify the equations by setting recombination to be zero. Doing so, we get these following set of equations:

$$(X_0)^* = (1-2u)x_0 + 2vx_1 + 2s_1x_0x_1 \quad (4)$$

$$(X_1)^* = 2ux_0 + (2-2u-2v)x_1 - 2s_1x_0x_1 + 2vx_2 + 2(-s_1+s_2)x_1x_2 \quad (5)$$

$$(X_2)^* = 2ux_1 + (1-2v)x_2 - 2(-s_1+s_2)x_1x_2 \quad (6)$$

At the fixed position, $X_0^* = x_0, X_1^* = x_1, and X_2^* = x_2$ (Higgs 1997). Hence solving for all the three variables $(x_0, x_1, and x_2)$, we get the following equilibrium frequencies:

$$x_0 = -\frac{v}{s_1} - \frac{2uv}{{s_1}^2} \quad (7)$$

$$x_1 = \frac{1}{2} - \frac{u}{2s_2} + \frac{v}{2s_1} + \frac{uv}{{s_1}^2} + \frac{5{s_1}^4uv}{{s_2}^6} + \frac{-\frac{s_1u}{2} + uv}{{s_2}^2} + \frac{-\frac{{s_1}^2u}{2} + 2s_1uv}{{s_2}^3} + \frac{-\frac{{s_1}^3u}{2} + 3{s_1}^2uv}{{s_2}^4} + \frac{-\frac{{s_1}^4u}{2} + 4{s_1}^3uv}{{s_2}^5} \quad (8)$$

$$x_2 = -\frac{37{s_1}^4}{{s_2}^4} + \frac{7{s_1}^3}{{s_2}^3} - \frac{{s_1}^2}{{s_2}^2} + \frac{s_1}{s_2} \quad (9)$$

Now because we assumed $u^2 = 0$ and $v^2 = 0$ then we can say that $u^2 \approx v^2 \approx uv$. The equations (7), (8), and (9) hence become

$$x_{0=} - \frac{v}{s_1} \quad (10)$$

$$x_1 = \frac{1}{2} + \frac{v}{2s_1} - \frac{us_1^4}{2s_2^5} - \frac{us_1^3}{2s_2^4} - \frac{us_1^2}{2s_2^3} - \frac{us_1}{2s_2^2} - \frac{u}{2s_2} \quad (11)$$

$$x_{2=} - \frac{37{s_1}^4}{{s_2}^4} + \frac{7{s_1}^3}{{s_2}^3} - \frac{{s_1}^2}{{s_2}^2} + \frac{s_1}{s_2} \quad (12)$$

We can clearly see that equations (9) and (12) are the same. This is because the frequency x_2 does not depend on any mutation rate after simplification of the original solutions. In the above solutions, Both u and v are less than s_1 and s_2. These are the simplified version of the original solutions. We also obtain a number of complex solutions, but we are only interested in the real solution as written above. Here we assumed that all mutations (u and v) to the power ≥ 2 are equal to zero since since $u < 10^{-6}$ and $v < 10^{-6}$. Also because the selection coefficients are very low (between 0.01 and 0.005), we assumed $s_1^n = 0$ and $s_2^n = 0$ whenever n ≥ 5.

4 Dynamics of finite populations

Now we will look into the change in the finite population. Let's remember there are four genotype and their frequencies must be equal to 1 ($x_0 + 2x_1 + x_2 = 1$). Hence there are four independent frequency variables, but we are assuming the two single mutants (Ab and aB) are the same and have the same fitnesses of $1 - s_1$. This assumption leads us to work with a three dimensional system. Since we know that the total frequency is 1, it will be easier to work only with two variables and once we get the results, we can find the third variable in term of the others. Our system is therefore reduce to a two dimensional system. Previously Kimura and Ohta (1968) have developed a 1 dimensional model using the diffusion models. Higgs (1997) has also shown that it is possible to solve a 1D system with the diffusion models. We will also use the same model to solve our problem.

Here, because of the mutation and selection forces, we need a drift term which we will call $m(x)$. The $m(x)$ or the infinitesimal mean is the change of frequency in one generation. $m(x)$ can also be called the expected mean change in our variable of interest. A variance and covariance will be needed since they follow a multinomial distribution. Our variance and covariance will respectively be denoted $v(x_i, x_i)$ and $cov(x_i, x_j)$ since we are working in 2 dimensional system. From Lynch's appendix (2008), we see that

$$v(x_i, x_i) = \frac{x_i * (1 - x_i)}{2 * N_e} \tag{13}$$

$$cov(x_i, x_j) = \frac{x_i * x_j}{2 * N_e} \tag{14}$$

We could use the Kolmogorov forward equation (or KFE) as described by Kimura and Ohta (1968), but since our system is two dimensional this diffusion model will not work for us. We will instead use the extended KFE shown by Lynch in his appendix (2008)

$$\frac{\partial[\rho(x,p,t)]}{\partial t} = \frac{1}{2}\sum_{i=1}^{k-1}\frac{\partial^2}{\partial x_i^2}[\rho(x,p,t)\frac{x_i(1-x_i)}{N_e}] - \sum_{i<j}\frac{\partial^2}{\partial x_i \partial x_j}[\rho(x,p,t)\frac{x_i x_j}{N_e}] - \sum_{i=1}^{k-1}\frac{\partial m(x)\rho(x,p,t)}{\partial x_i} \quad (15)$$

where the first part of the equation (15) is the allele-frequency variances, the second part involves the covariances between allele frequencies, and the third involves the change of frequency in generation or the mean. In this equation (15) $\rho(x,p,t)$ denotes the density function with x being the vector of allele frequencies, p the vector of their starting values, and t the time (Lynch appendix 2008). Applying this extended KFE (15) to our specific model, we get

$$\frac{\partial[\rho(x_0,x_2,t)]}{\partial t} = [\frac{1}{2}\frac{\partial^2}{\partial^2 x_0}(\rho(x,p,t)\frac{x_0(1-x_0)}{2N_e}) + \frac{1}{2}\frac{\partial^2}{\partial^2 x_2}(\rho(x,p,t)\frac{x_2(1-x_2)}{2N_e})] - [\frac{\partial^2}{\partial x_0 \partial x_2}(\rho(x,p,t)\frac{x_0 x_2}{2N_e})] - [\frac{\partial\rho(x,p,t)(-2ux_0+2vx_1+2s_1x_0x_1)}{\partial x_0} + \frac{\partial\rho(x,p,t)(2ux_1-2vx_2-2(s_2-s_1)x_1x_2)}{\partial x_2}] \quad (16)$$

As it was shown in equation (15), the first part of (15) involves the alleles-frequency variances, the second part involves the covariances between allele frequencies, and the third part is the mean. We are using $2N_e$ in (16) instead of N_e because we are now considering diploid population. $\rho(x_0, x_2, t)$ is the probability distribution for the random variables x_0 and x_2 at time t. Solving for our probability distribution, we can see the changes in frequencies throughout our generation and for $x_0 = 0$ and $x_2 = 0$, we will be able to see the time to fixation which is the main purpose of this research.

5 Discussion and conclusions

Our study here is an extended version of what Higgs (1997) has done. In his model, Higgs assumes reversible mutation with 2 loci, each with two alleles. The two alleles are labels as in our model but in his study, Higgs assumes both the AB genotype and double mutant ab to have fitness 1, while the two single mutants Ab and aB have a reduced fitness 1-s. Let us remember that we are working in discrete generation for both our model and Higgs' model. Prior to do any modification of Higgs' model, we will first look at this model.

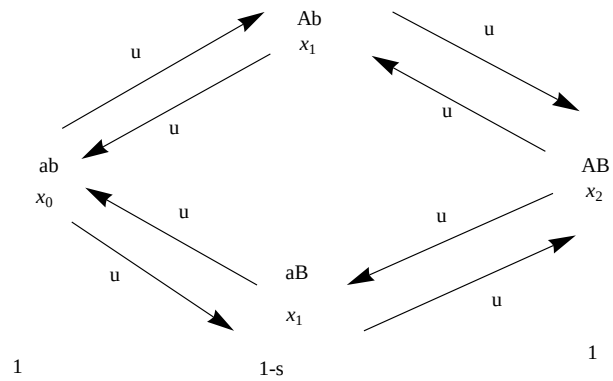


Figure 1: Higgs general model for the infinite population with reversible mutation where u=v and $s_2 = 0$

$$X_0 = (1 - 2u + 2sx_1)x_0 + 2ux_1 - r(x_0x_2 - x_1^2) \tag{17}$$
$$X_1 = (1 - 2u - s + 2sx_1)x_1 + u(x_0 + x_2) + r(x_0x_2 - x_1^2) \tag{18}$$
$$X_2 = (1 - 2u + 2sx_1)x_2 + 2ux_1 - r(x_0x_2 - x_1^2) \tag{19}$$

From this set of equations, and considering we do not have recombination, Higgs got the following set of solutions

$$x_0 = x_2 = \frac{1}{2} - \frac{u}{s} \tag{20}$$
$$x_1 = \frac{u}{s} \tag{21}$$

and the second solutions are

$$x_0 = x_2 = \frac{1}{2} - \frac{2u}{s} \tag{22}$$
$$x_1 = \frac{u}{s} \tag{23}$$

In the equations (17), (18), and (19), we have reversible mutation but the forward and backward mutation rates are equal (u=v). Also the fitnesses of the double mutant genotypes ab and AB are equal (1), and this is the same for the single mutant genotype aB and Ab (1-s). Now we will make the first step assumption by assuming $u \neq v$, but the fitnesses of the double mutant genotypes are still equal. Our model will become

$$X_0 = (1 - 2u + 2sx_1)x_0 + 2vx_1 - r(x_0x_2 - x_1^2) \tag{24}$$

$$X_1 = 2(1 - u - v - sx_0 - sx_2)x_1 + 2ux_0 + 2vx_2 + 2r(x_0x_2 - x_1^2) \tag{25}$$

$$X_2 = (1 - 2v + 2sx_1)x_2 + 2ux_1 - r(x_0x_2 - x_1^2) \tag{26}$$

and from here, we get the following solutions

$$x_0 = x_2 = \frac{-v}{s} \tag{27}$$

$$x_1 = \frac{1}{2} + \frac{u+v}{2s} \tag{28}$$

Now we will go on with our last assumption which is the main solution we are interested in. We will assume $u \neq v$ and also the double mutant genotype with different fitnesses; ab (1) and AB $(1 - s_2)$, but the single mutant genotype still have the same fitnesses $(1 - s_1)$. From here we get the equations (1), (2), and (3) with their respective solutions (10),(11), and (12).

References

[1] . Kimura, T. Ohta, 1968. The average number of generations until fixation of a mutant gene in a finite population. Genetics 61: 763-771.

[2] .W. Strickberger (3rd edition), 2000. Evolution. United States.

[3] .H. Gillespie (2nd edition), 2004. Population genetics: a concise guide. Baltimore, Maryland.

[4] . Iwasa et al., 2005. Population genetics of tumor suppressor genes. Journal of theoretical biology 233:15-23.

[5] . Lynch, A. Abegg, 2010. The rate of establishment of Complex Adaptations. Mol. Bio. Evol. 27(6):1404-1414.

[6] . Bulmer, 1991. The selection-mutation-drift Theory of Synonymous codon usage. Genetica 129:897-907.

[7] . Lynch, 2008, Appendix 1: Diffusion Theory.

[8] . G. Higgs, 1998. Compensatory neutral mutations and the evolution of RNA. Genetica 102/103:91-101.

The Links Between Smale's Mean Value Conjecture and Complex Dynamics

Hayley Miles-Leighton, UC San Diego

Abstract

In this paper we study the links between Smale's Mean Value Conjecture (SMVC) and the convergence of critical points under iteration. We begin by introducing SMVC and discussing what progress has been made towards proving this conjecture. From there we give a brief introduction to a few concepts in Complex Dynamics and Complex Analysis including conjugacy, conjugations by $1/z$, and orbits. We then show that conjugacy conserves SMVC. At this point, we go through the quadratic case to show the patterns that we are looking for as well as to show a basic structure of how this problem is being looked at. From there we look at the cubic case, going through SMVC for the cubic polynomial. We then introduce the Petal Theorem and show how it is used for the cubic case specifically. We conclude with the connections between the SMVC in the cubic case and the convergence of the same critical points explained and discuss future work.

1 Introduction and Background

Complex Dynamics is the study of complex functions under iteration. We are specifically interested in where certain points in the complex plane converge to the origin and to infinity. It is this idea that we kept in mind when we looked at Smale's Mean Value Conjecture (SMVC) which states [7]:

Conjecture 1.1. *Let $f(z) = z + \sum_{i=2}^{n} a_i z^i$ be a complex valued polynomial of degree $n \geq 2$ for which $f(0) = 0$ and $f'(0) = 1$. Then, there exists a critical point c of this polynomial that satisfies the following statement:*

$$|\frac{f(c)}{c}| \leq 1$$

A variation of this conjecture has been announced for all polynomials of degree 10 or less by Sendov and Marinov [6]. This variation replaces 1 with an even better bound of $(d-1)/d$. However, their paper has no proofs to show this result. Several other papers use another variation on this conjecture. Namely, this following proposition:

Proposition 1.2. *Let p be a polynomial of degree $N \geq 2$ over $\mathbb{C}$, and supposed that $x \in \mathbb{C}$ is not a critical point of p. Then there exists a critical point ζ of p such that*

$$|\frac{p(\zeta) - p(x)}{\zeta - x}| \leq 4|p'(x)|. \tag{1}$$

In a paper written by E. Crane [2], he shows that there can be a better bound than the bound of 4 that was proven by S. Smale in 1981 [7] in the proposition given above. He proves that for all polynomials of degree 8 and higher, $K(d) < 4 - \frac{2.263}{\sqrt{d}}$ where d is the degree of the polynomial and $K(d) = sup\{S(p,x) : deg(p) = d, p'(x) \neq 0\}$ $(S(p,x) = min(|\frac{p(\zeta)-p(x)}{(\zeta-x)p'(x)}| : p'(\zeta) = 0))$. For $2 \leq d \leq 7$, Schmeisser [5] proved that we can have $K(d) \leq \frac{2^d-(d+1)}{d(d-1)}$ which is a better bound than what Crane would have for those values of d. In 2003, T.W. Ng proved that this original bound could be further reduced to 2 for all nonlinear odd polynomials with a nonzero linear term [3].

Although there have been many papers published on the possibility of this conjecture (and its variations) being true, none have been able to prove it completely. The goal of this paper is not to prove the conjecture, however. This paper will look into possible connections between SMVC and Complex Dynamics. SMVC says nothing about which critical point or points satisfy the bound. It only states that there is one. On the other hand, complex dynamics says that given a polynomial $f(z) = z + \sum_{i=2}^{n} a_i z^1$, there must exist a critical point tending to the origin under iteration (this fact is presented formally in Proposition 1.8). We hope to combine these two thoughts to see if there is a relation between them. Specifically, we will prove the following main theorem:

Theorem 1.3. *For any cubic of the form $p(z) = z + a_2 z^2 + a_3 z^3$, there is a critical point satisfying the SMVC bound and converging to the origin.*

To begin with, we will be wanting to show that this works in general. It is generally nice to be able to work with one form instead of many so to do this we will use the idea of conjugacy:

Definition 1.4. If there exists a function $\phi(z)$ that is either a linear fractional transformation (such as $1/z$) or a so called afine map (such as $Az + B$), then two functions are conjugate if $\phi \circ f = g \circ \phi$.

Below is an example of this concept:

Example 1.5. Let $f(z) = z^2 - z$ and $\phi(z) = \alpha z + \beta$. We want to find out if this form is conjugate to $g(w) = w^2 + c$ where c is some constant. To find out if these two are conjugate, we will follow this chart:

$$\begin{array}{ccc} Z & \xrightarrow{z^2-z} & \mathbb{C} \\ \phi \downarrow & & \downarrow \phi \\ W & \xrightarrow{w^2+c} & \mathbb{C} \end{array}$$

Following the chart, we see that we get $\alpha(z^2 - z) + \beta = (\alpha z + \beta)^2 + c \Rightarrow \alpha z^2 - \alpha z + \beta = \alpha^2 z^2 + 2\alpha\beta z + \beta^2 + c$. For two polynomials to be equal to one another, all of the coefficients must equal one another. With this in mind, we find that $\alpha = \alpha^2, -\alpha = 2\alpha\beta, \beta = \beta^2 + c \Rightarrow \alpha = 1, \beta = -1/2, c = -3/4$. Since there exists an α and β that makes this true, we know that these two polynomials are conjugate for $c = -3/4$. This means that if we let $\phi(z) = \alpha z + \beta$, we have that $\phi \circ f = g \circ \phi$.

From here, we need to show that we can use the conjugated form to look at SMVC. Below is a lemma that shows this for $\phi(z) = \alpha z$.

Lemma 1.6. *If $p(z) = z + h.o.t.$ satisfies SMVC and if $q = A \circ p \circ A^{-1}$ where $A(z) = \alpha z\,(\alpha \neq 0)$, then q also satisfies SMVC.*

Proof. Suppose that c is a critical point of $p(z)$ and that it satisfies $|p(c)/c| \leq 1$. Since they are conjugate via multiplication by α, the point $x = \alpha c$ is a critical point of q. Since $q = A \circ p \circ A^{-1}$ we can see that $|q(x)/x| = |q(\alpha c)/(\alpha c)| = |\frac{\alpha(p((\alpha c)/\alpha))}{\alpha c}| = |p(c)/c| \leq 1$. Hence, if $p(z)$ satisfies SMVC and if $q = A \circ p \circ A^{-1}$ where $A(z) = \alpha z\,(\alpha \neq 0)$, then $q(z)$ will also satisfy SMVC. $\square$

At this point, I need to introduce the concept of an orbit and a general proposition that uses this concept as well as talk about the conjugacy of $1/z$ which will be used in the cubic case.

Definition 1.7. Suppose X is any set and $f : X \mapsto X$ is any function. Let $x_0 \in X$. Then, the orbit of x_0 is defined as x_0, $f(x_0)$, $f(f(x_0))$, ...

This concept is used in the following proposition:

Proposition 1.8. *There exists a critical point c_i such that the orbit of c_i converges to the origin.*

From here, I will introduce the conjugacy of $1/z$. In general, $1/z$ maps "circles" to "circles." When I say "circles" I mean that they can fit into this general form: $A(x^2 + y^2) + Bx + Cy + D = 0$. At this point, we want to see what x and y are equal to in terms of the new coordinates u and v. After some computations, we see that $x = \frac{u}{u^2+v^2}$ and $y = \frac{-v}{u^2+v^2}$. After we use these to substitute into the general form, we find that in the new plane we get $D(u^2 + v^2) + Bu - Cv + A = 0$. So, lines become circles, circles become circles, and circles can sometimes become lines. If we are mapping a circle that is around the origin, then everything inside goes outside and vice versa. This is because the origin goes to ∞ and ∞ goes to 0. This is the method that I will use in the future.

Now that we have a basic background in Complex Analysis as well as Complex Dynamics, we can look into the quadratic case as an example.

2 The Quadratic Case

To begin with, we will go through the quadratic case. The following Theorem describes what we are planning to prove in this section:

Theorem 2.1. *For each quadratic polynomial $f(z) = z + a_2 z^2$, there is exactly one critical point. It satisfies the bound in SMVC and it converges to the origin*

Proof. The form that we use is a very general form of the quadratic polynomials. However, we want to use a simpler version by conjugating this polynomial by $\phi(z) = \alpha z$ using the process explained in the previous section.

$$\begin{array}{ccc} Z & \xrightarrow{z+a_2z^2} & \mathbb{C}_\infty \\ \phi(z)\downarrow & & \downarrow\phi(z) \\ Z' & \xrightarrow{z-z^2} & \mathbb{C}_\infty \end{array}$$

This means that we need $\alpha(z+a_2z^2) = \alpha z - \alpha^2 z^2 \Rightarrow z + a_2z^2 = z - \alpha z^2 \Rightarrow \alpha = -a_2$. So, we can now do all of our calculations using $f(z) = z - z^2$. First, we need to find its critical point: $f'(z) = 1 - 2z = 0 \Rightarrow z = 1/2$. Referring back to SMVC, we see that $|\frac{f(1/2)}{1/2}| = |\frac{1}{4} \cdot 2| = |\frac{1}{2}| \leq 1$. Hence, SMVC is satisfied by all quadratic polynomials of this form.

Now, we want to look at where this critical point converges to the origin. To get an idea of this, we will refer back to Proposition 1.8. In this case, $i = 1$. So, since there is only one critical point, this point must converge to the origin. Hence, this one critical point both satisfies SMVC and converges to the origin. $\square$

At this point, we want to see if this pattern continues for the cubic case.

3 The Cubic Case

In this section, we plan to prove the main theorem (Theorem 1.3). We will do this both indirectly and directly. The indirect approach will use Proposition 1.8 and handles "areas" or "regions." The direct approach will use Beardon's Petal Theorem which will be introduced later on. To begin with, we will be looking at the following general form of cubic polynomials that satisfy the conditions in SMVC:

$$p(z) = z + az^2 + bz^3 \tag{2}$$

However, we would like this in a more useful form. To do this, we will conjugate this polynomial by $\phi(z) = \alpha z$:

$$\begin{array}{ccc} Z & \xrightarrow{z+az^2+bz^3} & \mathbb{C} \\ \phi\downarrow & & \downarrow\phi \\ Z & \xrightarrow{z-\frac{1}{2}(c+\frac{1}{c})z^2+\frac{1}{3}z^3} & \mathbb{C} \end{array}$$

Following the same process described in example 1.5, we see that we have $z + az^2 + bz^3 = z - \frac{\alpha}{2}(c + \frac{1}{c})z^2 + \frac{\alpha^2}{3}z^3$. For two polynomials to be equal, their coefficients must be equal: $a = -\frac{\alpha}{2}(c + 1/c)$ and $b = \frac{\alpha^2}{3}$. Hence, $\alpha = -\frac{2a}{c+1/c}$. Since there exists an α we know that these two polynomials are conjugate and we can continue with the following cubic:

$$f_c(z) = z - \frac{1}{2}(c + \frac{1}{c})z^2 + \frac{1}{3}z^3 \tag{3}$$

After taking the derivative, it is easy to see the two critical points: $c_1 = c$ and $c_2 = 1/c$. These two critical points are nice to work with since they create

the same map. In other words, $f_c(z) = f_{1/c}(z)$. This means that by looking at one of the critical points, we have a way of looking at the other. For example, if we can show for one critical point that $f_c^{on}(c_1(c)) \to \infty \Leftrightarrow f_{1/c}^{on}(c_1(c)) \to \infty \Leftrightarrow f_{1/c}^{on}(c_2(1/c)) \to \infty$ then we see that in an area that is $1/c$ we have the same thing for the other critical point. So, we will look at what happens to c_1 to get an idea what happens to c_2 as well. At this point, we need to look at where c_1 and c_2 satisfy SMVC:

$$f_c(c) = c - \frac{1}{2}(c + \frac{1}{c})c^2 + \frac{1}{3}c^3 = \frac{3c - c^3}{6} \Rightarrow |\frac{f_c(c)}{c}| = |\frac{3 - c^2}{6}| \leq 1 \Rightarrow |3 - c^2| \leq 6$$

$$f_c(\frac{1}{c}) = \frac{1}{c} - \frac{1}{2}(c + \frac{1}{c})\frac{1}{c^2} + \frac{1}{3c^3} = \frac{3c^2 - 1}{6c^3} \Rightarrow |\frac{f_c(\frac{1}{c})}{\frac{1}{c}}| = |\frac{3c^2 - 1}{6c^2}| \leq 1 \Rightarrow |3 - \frac{1}{c^2}| \leq 6$$

I claim that with these two graphs combined, we see that SMVC is satisfied for any c throughout the parameter plane. Below is a figure that shows this fact.

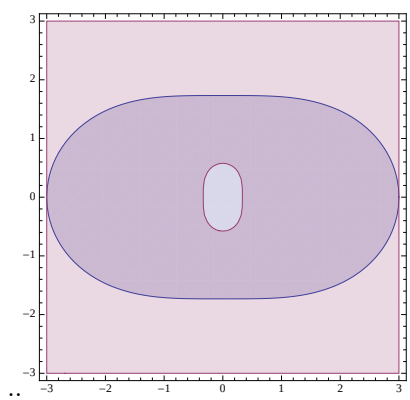


Figure 1: The smallest area is where only c_1 satisfies SMVC, the darkest area shows where both critical points satisfy SMVC, and the area outside the larger curve shows where only c_2 satisfies SMVC. Put together, they cover the entire complex plane.

At this point, we want to look at areas of convergence to infinity. For $c_1 = c$, we will prove the following lemma to show where it converges to infinity:

Lemma 3.1. *Let $\delta = \frac{3}{4}\sqrt{30}$. If $|c| > \delta$ and $|z| > |c|^3/6$ then $|f_c(z)| > |z|$ and, under iteration of f, z converges to infinity.*

Proof. We know that $|f_c(z)| = |z||1 - \frac{1}{2}(c + 1/c)z + \frac{1}{3}z^2|$ and that we want $|z||1 - \frac{1}{2}(c + 1/c)z + \frac{1}{3}z^2| > |z|$. So, proving that $|1 - \frac{1}{2}(c + 1/c)z + \frac{1}{3}z^2| > 1$ will imply the previous inequality. However, if we prove that $|-\frac{1}{2}(c + 1/c)z + \frac{1}{3}z^2| > 2$ then the previous inequality will be true. We can rewrite this as $|z||-\frac{1}{2}(c + 1/c) + \frac{1}{3}z| > 2$, which can be implied by the following: $|-\frac{1}{2}(c + 1/c) + \frac{1}{3}z| > 2/|z| > 2/(|c|^3/6) >$

0.17. Using the triangle inequality we see that $|-\frac{1}{2}(c+1/c)+\frac{1}{3}z| \geq ||\frac{1}{2}(c+1/c)|-\frac{1}{3}|z|| = |\frac{1}{3}|z|-|\frac{1}{2}(c+1/c)|| > |\frac{|c|^3}{18}-|\frac{1}{2}(c+1/c)|| \geq |\frac{|c|^3}{18}-\frac{1}{2}(|c|+|1/c|)|$. This function is monotonically increasing as a function of $|c|$. So, we can re-write it as $\frac{|c|^3}{18}-\frac{1}{2}(|c|+|1/c|) > 1.67$. Since $1.67 > 0.17$ we see that all of the previous inequalities hold. Hence, under iteration of f, z converges to infinity. $\square$

This means that for all values of $|c|$ greater than $\frac{3\sqrt{30}}{4}$, c_1 will converge to infinity. Based on the earlier explanation, we know that c_2 will also converge to infinity in the circle of $|c| > \frac{3}{4}\sqrt{30}$ mapped by $1/c$ in the parameter plane. This will give us the area where c_2 converges to infinity. This circle will be $|c| < \frac{4}{3\sqrt{30}}$. Now that we have two areas where c_1 and c_2 converge to infinity, we want to refer back to Proposition 1.8. In this case, $i = 2$. This means that, for example, in the area where c_1 converges to infinity, we know that c_2 must converge to the origin under iteration. The same happens in the area where c_2 converges to infinity. From here we can combine these circles with Figure 1 to create five distinct areas:

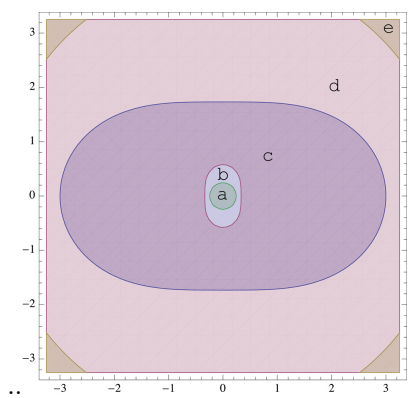


Figure 2: There are five different points representing the five different areas created by four graphs.

We are interested in what happens within each of those areas. Below are five charts that give us an idea of what is going on:

Region a	c_1	c_2
SMVC	y	n
$\to 0$	y	n

Region b	c_1	c_2
SMVC	y	n
$\to 0$	?	?

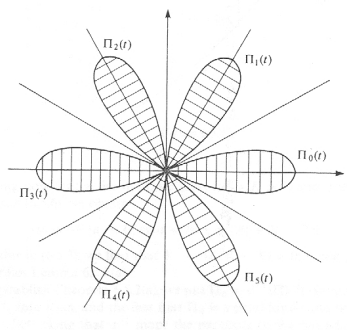
Region c	c_1	c_2
SMVC	y	y
$\to 0$	?	?

Region d	c_1	c_2
SMVC	n	y
$\to 0$	?	?

Region e	c_1	c_2
SMVC	n	y
$\to 0$	n	y

As you can see, only regions a and e have full charts. We can see that each of those have a column of yeses. This is the pattern that we are looking for.

This means that the chart for region c does satisfy this pattern automatically based on Proposition 1.8. Since in this area both c_1 and c_2 satisfy SMVC and we know that one of those critical points will converge to the origin under iteration, we know that we will have a column of yeses. Hence, the two regions that we are interested in are b and d. However, since our critical points have the same map, we can study the area that contains point d and be able to have an understanding of the area that contains point b. To be able to get a good understanding of what is happening in region d, we will need to look at the annulus $1.75 \leq |c| \leq 4.11$ which encompasses the entire region of d. In this annulus, we will be looking at the following theorem:



..

Figure 3: This shows what is meant in the following theorem by "petal." In our specific case, we will only have one petal.

Theorem 3.2. *(The Petal Theorem) Suppose that the analytic map f has a Taylor Expansion*

$$f(z) = z - z^{p+1} + O(z^{2p+1}) \tag{4}$$

at the origin. Then for all sufficiently small t:

1. *f maps each petal $\Pi_k(t)$ into itself;*
2. *$f^n(z) \to 0$ uniformly on each petal as $n \to \infty$;*
3. *$argf^n(z) \to 2k\pi/p$ locally uniformly on $\Pi_k asn \to \infty$;*
4. *$|f(z)| < |z|$ on a neighborhood of the axis of each petal;*
5. *$f : \Pi_k(t) \to \Pi_k(t)$ is conjugate to a translation. [1]*

Since our $f_c(z)$ is not in the form that is needed to be able to use this theorem, we will need to conjugate it to create the form that is needed:

$$\begin{array}{ccc} Z & \xrightarrow{z-\frac{1}{2}(c+\frac{1}{c})z^2+\frac{1}{3}z^3} & \mathbb{C} \\ \phi\downarrow & & \downarrow\phi \\ Z & \xrightarrow{z-z^2+c'z^3} & \mathbb{C} \end{array}$$

where $\phi(z) = \alpha z$. This shows that we need to find what α is as well as c':

$$\alpha(z-\frac{1}{2}(c+\frac{1}{c})z^2+\frac{1}{3}z^3) = \alpha z-\alpha^2z^2+c'\alpha^3z^3 \Rightarrow z-\frac{1}{2}(c+\frac{1}{c})z^2+\frac{1}{3}z^3 = z-\alpha z^2+c''\alpha^2z^3$$

So, for two polynomials to be equal, their coefficients must be equal. This means that $\alpha = \frac{1}{2}(c+\frac{1}{c})$ and that $\frac{1}{3} = c'(\frac{1}{2}(c+\frac{1}{c}))^2$. So, we now know that $\alpha = \frac{1}{2}(c+\frac{1}{c})$ and $c' = \frac{1}{3}(\frac{2c}{c^2+1})^2$. Since we are now working in a new plane, which we will call the u-plane, we need to convert our annulus into the new coordinates. The following lemma shows this:

Lemma 3.3. *Let $R = \{c : 1.75 \leq |c| \leq 0.96\}$ and let $R' = \{c' : 0.07 < |c'| < 0.96\}$. Then $c \in R \Rightarrow c' \in R'$.*

Proof. To show this, we need to reduce c' to a more useful form: $|c'| = |\frac{1}{3}(\frac{c2}{c^2+1})^2| = \frac{4}{3}|\frac{1}{c+1/c}|^2$. To begin with, we want to look at the denominator of the main fraction. Using the triangle inequality, we see that $|c+\frac{1}{c}| \geq ||c| - \frac{1}{|c|}|$. To be able to get a better bound on this, we need to show that $|c| - \frac{1}{|c|}$ is monotonically increasing as a function of $|c|$. To do this, we want to make sure that the derivative is always positive for values between 1.75 and 4.11: $1 + \frac{1}{|c|^2}$ is always positive. This means that the smallest value will happen at the bound. So, we have that $||c| - \frac{1}{|c|} \geq |\frac{7}{4} - \frac{4}{7}| = |\frac{33}{28}|$. This means that $\frac{1}{|c+1/c|} \leq |\frac{28}{33}|$. Combining this with the rest of the function, we get an upper bound: $|c'| \leq \frac{4}{3}|\frac{28}{33}|^2 < 0.96$.

For the lower bound, we will follow a similar process. By the triangle inequality, we see that $|c+\frac{1}{c}| \leq |c| + \frac{1}{|c|}$. Before we can continue, we need to show that this function is monotonically increasing as a function of $|c|$. The derivative is $1 - \frac{1}{|c|^2}$ which is always positive between 1.75 and 4.11. This means that the greatest value will happen at the boundary. So, we have that $|c| + \frac{1}{|c|} \leq 4.11 + \frac{1}{4.11} < 4.35$. Putting this into the rest of the function, we find that the lower bound is $|c'| > 0.07$. □

Now that we have bounds in our new coordinate system, we want to go back to the Petal Theorem. We will show that this theorem works for our case which means that $p = 1$ and $k = 0$ (assuming that we are ignoring the points $c = \pm i$). To do this, we want to prove the following lemma:

Lemma 3.4. *Given z in the dynamical plane of f_c, let $z' = (1/2)(c + 1/c)z$ be the corresponding point in the dynamical plane, and put $w = 1/z'$. If $c \in R$ and $Re(w) > 4.36$ then under iteration of f_c, the point z converges to the origin.*

Proof. To begin with, we want to conjugate $p(z) = z - z^2 + c'z^3$ to the w-plane by the conjugation $\sigma : z \mapsto 1/z$. First, we will look at what $1/p(z)$ is. Using polynomial division, we find:

$$\frac{1}{z - z^2 + c'z^3} = \frac{1}{z} + 1 + Az + \nu(z), \tag{5}$$

where $A = 1 - c'$ and $|\nu(z)| = |\frac{(1-2c')z^2 - c'(1-c')z^3}{1-z+c'z^2}| \leq B|z|^2, B > 0$.

At this point we need to find bounds on both B and A by proving this corollary:

Corollary 3.5. *If $c' \in R'$ and $|z| < 0.23$ then $|A| < 1.96$ and $|\nu(z)| \leq B|z|^2$ where $B = 4.66$.*

Proof. First, we will look at finding a bound on A. We know that $A = 1 - c'$ so $|A| = |1 - c'|$. Using the triangle inequality, we see that $|A| = |1 - c'| \leq 1 + |c'|$. Since $c' \in R'$, we know that $|c'| < 0.96$. So, $1 + |c'| < 1.96$ and hence $|A| < 1.96$.

Now we want to find B and so we need a bound on $\nu(z)$. By definition, $|z| < 0.23$. This choice will be explained later. First, we will look at the denominator of $\nu(z)$. By the triangle inequality, we see that $|1 - z + c'z^2| \geq ||1 - z| - |c'z^2|| = |1 - z| - |c'||z|^2 \geq 1 - |z| - |c'||z|^2 \geq 1 - 0.23 - (0.96)(0.23)^2 \approx 0.719$.

Now that we have a bound on the denominator of $\nu(z)$, we can continue to get a bound on the full function which will become our B. To do this, let's first look at the numerator separately: $|(1 - 2c')z^2 - c'(1 - c')z^3| = |z|^2 \cdot |(1 - 2c') - c'(1 - c')z|$. Since the bound on $|\nu(z)|$ is $B|z|^2$, we can disregard the $|z|^2$ term. Using the triangle inequality and the fact that $|z| < 0.23$ we see that $|(1 - 2c') - c'(1 - c')z| \leq |1 - 2c'| + |c'(1 - c')||z| < |1 - 2c'| + 0.23|c'(1 - c')|$. Earlier we were able to find a bound on the denominator, so if we let $B = \frac{|1-2c'|+0.23|c'(1-c')|}{0.719}$, we can guarantee that $|\nu(z)| \leq B|z|^2$.

At this point,we want to find a bound on B. Since $B = \frac{|1-2c'|+0.23|c'(1-c')|}{0.719}$, we know by the triangle inequality that $\frac{|1-2c'|+0.23|c'(1-c')|}{0.719} \leq \frac{1+2|c'|+0.23|c'|(1-|c'|)}{0.719}$. Using the fact that $|c'| < 0.96$ since $c' \in R'$, we have that $\frac{1+2|c'|+0.23|c'|(1-|c'|)}{0.719} < 4.66$. □

At this point we want to complete the proof by showing the following corollary:

Corollary 3.6. *Given c in R, the critical point $c_2 = 1/c$ converges to the origin.*

Proof. I claim that we can show that c_2 converges by using the second iterate. If we can show that the second iterate converges to the origin, then we can conclude that c_2 converges to the origin as well. Recall that $f_c(c_2) = \frac{3c^2-1}{6c^3}$. Iterating one more time, we find that

$$f_c(f_c(c_2)) = \frac{6c(6c^3)^2(3c^2 - 1) - 3c^2(6c^3)(3c^2 - 1)^2 - 3(6c^3)(3c^2 - 1)^2 + 2c(3c^2 - 1)^3}{6c(6c^3)^3}.$$

We need to convert this into the new coordinate system. To do this, multiply the above function by $\alpha = \frac{c^2+1}{2c}$. After working through the computation, the result is $\frac{5}{48c^2} - \frac{5}{72c^4} + \frac{1}{72c^6} - \frac{1}{1296c^8} - \frac{1}{1296c^{10}} + \frac{3}{16}$. From here, we want an upper bound on $\frac{5}{48c^2} - \frac{5}{72c^4} + \frac{1}{72c^6} - \frac{1}{1296c^8} - \frac{1}{1296c^{10}}$ to create a circle whose radius is that upper bound, centered at $\frac{3}{16}$. To get an upper bound on this, we will use the triangle inequality: $|\frac{5}{48c^2} - \frac{5}{72c^4} + \frac{1}{72c^6} - \frac{1}{1296c^8} - \frac{1}{1296c^{10}}| \leq |\frac{5}{48c^2}| + |\frac{5}{72c^4}| + |\frac{1}{72c^6}| + |\frac{1}{1296c^8}| + |\frac{1}{1296c^{10}}| \leq |\frac{5}{48(7/4)^2}| + |\frac{5}{72(7/4)^4}| + |\frac{1}{72(7/4)^6}| + |\frac{1}{1296(7/4)^8}| + |\frac{1}{1296(7/4)^{10}}| \approx 0.042$.

Figure 4: The graph on the left shows the areas where the denominator of $\nu(z)$ does not vanish and the values of the second image of c_2. The graph on the right are these same graphs but mapped by $1/z$. The half-plane in the right hand graph is the half-plane that we want to show is forward invariant.

Now that we have most of our information, we can continue with the mapping by finding $g(z)$ in the w-plane. To do this, we can take what we have for the $1/p(z)$ function and replace z with $w = 1/z$. This creates $g(w) = w+1+\frac{A}{w}+\theta(w)$ where $|\theta(w)| = |\nu(\sigma^{-1}(w))| \leq B|\sigma^{-1}(w)|^2 \leq \frac{B}{|w|^2}$ and B and A are the same as earlier. Figure 4 shows where all the values of the second iterate of c_2 go under the mapping of $1/z$ as well as where the circle of $|z| < 0.23$ goes under the same mapping. There is also a half-plane at $Re(w) > 4.36$. We want to show that this half-plane is forward invariant. To do this, we need to make sure that when we move the real values of w by $w \mapsto w + 1 + \frac{A}{w} + \theta(w)$, it moves forward. To do this, we need the most negative vales that A and $\theta(w)$ can take.

Recall that $|A| < 1.96$. This means that the most negative A can be is -1.96. Hence, the most negative $\frac{A}{w}$ can be is about -0.45 since $w = 4.36$. Now, we know that $|\theta(w)| \leq \frac{B}{|w|^2}$ which means that $-|\theta(w)| \geq -\frac{B}{|w|^2}$. Earlier, we also showed that $B < 4.66$ hence $-B > -4.66$. This means that since $\frac{1}{|w|} = 0.23$, the most negative $\theta(w)$ can be is -0.25. Combining all of these together, we see that $4.36 \mapsto 4.36 + 1 - 0.45 - 0.25 = 4.66$. This leads us to the following proposition:

Proposition 3.7. *Under iteration of $w \mapsto w + 1 + \frac{A}{w} + \theta(w)$, every point w in the right half-plane $Re(w) > 4.36$ converges to infinity: the real parts of points in its orbit converge to $+\infty$.*

Proof. We will show this using induction. The first step has already been shown above. If we let $w_1 = 4.36$, we will start with the base case $n = 1$. Following the same process as above, we found that $w_1 \mapsto 4.66 = w_2$. Hence, $w_1 < w_2$ and so the base case is true.

Now, we want to assume that n holds. This means that $w_{n-1} \mapsto w_n$ and hence $w_{n-1} < w_n$. So, we need to look at where w_n gets mapped to: $w_n \mapsto w_n+1+A/w_n+\theta(w_n)$. Since we know that $w_n > w_{n-1}$, we can say that $1/w_n < 1/w_{n-1}$. So, we have that $A/w_n < A/w_{n-1}$ which means that the most negative it can be is $-A/w_{n-1}$. Now, we need to look at a bound on $\theta(w_n)$. By definition, $|\theta(w_n)| \leq B/|w_n|^2 < B/|w_{n-1}|^2$ which means that $-|\theta(w_n)| > -B/|w_{n-1}|^2$. So, with these replacements, we see that $w_n \mapsto w_n + 1 + A/w_n + \theta(w_n) > w_n + 1 - A/w_{n-1} - B/|w_{n-1}|^2$. However, since $w_n > w_{n-1}$, we know that $w_{n-1}+1-A/w_{n-1}-B/|w_{n-1}|^2 < w_n+1-A/w_n-B/|w_{n-1}|^2$. We also know that $w_{n-1}+1-A/w_{n-1}-B/|w_{n-1}|^2 = w_n$. So, if we let $w_{n+1} = w_n+1+A/w_n+\theta(w_n) > w_n+1-A/w_{n-1}-B/|w_{n-1}|^2 > w_{n-1}+1-A/w_{n-1}-B/|w_{n-1}|^2 = w_n$, we have that $w_n < w_{n+1}$ which completes the proof. □

What this proposition has told us is that the real values in the half-plane in the w-plane converges to infinity. Since this line is mapped backwards by $1/z$, when this line converges to infinity, we have that the resulting circle converges to the origin. To see this, consider the half-planes given by $x > 4.36$ and $x > 4.66$. Using the method explained in section 1, we can see that these two lines map to these two circles, respectively: $(x - 0.115)^2 + y^2 < 0.013$ and $(x - 0.107)^2 + y^2 < 0.012$. When these two circles are graphed together, we see that the second is mapped inside the first but both are tangent to the origin. Hence, as the real value of the half-plane converges to infinity, these circles converge to the origin. This means that since this area that represents all values of the second iterate of c_2 within our annulus is contained within the half-plane, we can see that the second iterate c_2 must converge to the origin. Hence, c_2 must converge to the origin as well and this completes our proof of Corollary 3.6. □

Combining the proofs of Corollaries 3.5 and 3.6 as well as the proof for Proposition 3.7, we can see that Lemma 3.4 has been proven.

□

Since we were able to show that c_2 converges to the origin within our annulus, we can say something about c_1 in the other region under consideration. Because $f_c(z) = f_{1/c}(z)$, we can say that $f_c(c_2(c)) \Leftrightarrow f_{1/c}(c_2(c)) \Leftrightarrow f_{1/c}(c_1(1/c))$. Hence, within the other area (since it is where c_2 is mapped by $1/c$) we can say that c_1 converges to the origin within region b. With these new pieces of information, we can look back to the charts for regions b and d:

Region b	c_1	c_2
SMVC	y	n
$\to 0$	y	?

Region d	c_1	c_2
SMVC	n	y
$\to 0$	?	y

As you can see, we can now say that c_1 converges to the origin in region b and that c_2 converges to the origin in region d which completes the pattern that we were looking for. Hence, we know that for any cubic, there is a critical point satisfying the SMVC bound and converging to the origin.

4 Concluding Remarks

This paper has shown that Complex Dynamics can help give an idea of which critical points satisfy SMVC for quadratic and cubic polynomials. It began with a brief introduction the problem as well as some basic background information in Complex Dynamics and Complex Analysis. From there, it described the quadratic case as an example for the cubic case. In the cubic case, we showed that the intuition that convergence of a critical point and SMVC do have a relation by using a specific instance of the Petal Theorem. From here, if we can show this intuition is true for higher order polynomials, we may have a different possibility for proving that this form of SMVC is true. There is also the possibility of looking into the following stronger form of SMVC:

Conjecture 4.1. *Let $f(z) = z + \sum_{i=2}^{n} a_i z^i$ of degree $n \geq 2$ be a complex valued polynomial for which $f(0) = 0$ and $f'(0) = 1$. Then, there exists a critical point c of f that satisfies: $|\frac{f(c)}{c}| \leq \frac{d-1}{d}$ where d is the degree of the polynomial.*

This has been shown to be true for polynomials of degree 2, 3, and 4 by algebraic methods in a paper by Q. Rahman and G. Schmeisser [4]. However, the link between this version of SMVC and the convergence of the orbits of the critical points does not seem to have been looked into.

5 Acknowledgments

I would like to thank Professor Kevin Pilgrim for is support and guidance these past 8 weeks. I would also like to thank him for organizing this great opportunity for all of us.

References

[1] Alan Beardon. *Iteration of Rational Functions*. Number 132 in Graduate Texts in Mathematics. Springer-Verlad, New York, 1991.

[2] Edward Crane. A bound for smale's mean value conjecture for complex polynomials. *Bulletin of the London Mathematical Society*, 39(5):781–791, 2007.

[3] T.W. Ng. Smale's mean value conjecture for odd polynomials. *Journal of the Australian Mathematical Society*, 75(3):409–411, 2003.

[4] Q. Rahman and G. Schmeisser. Analytic theory of polynomials. *London Mathematical Society Monographs*, 2002.

[5] G. Schmeisser. The conjectures of sendov and smale. pages 353–369, 2002.

[6] Blagovest Sendov and Pencho Marinov. Verification of smale's mean value conjecture for $n \leq 10$. *Bulgarian Academy of Sciences*, 60(11):1151–1156, 2007.

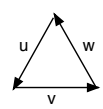
[7] Steve Smale. The fundamental theorem of algebra and complexity theory. *Bulletin of the American Mathematical Society*, 4(1):1–33, 1981.

On the Combinatorics of Schubert Calculus

Linh Truong

1 Introduction

We construct a lattice by choosing three unit vectors u, v, w in the plane such that $u + v + w = 0$.



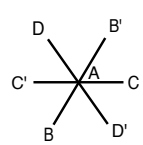
The points $iu + jv$ with i, j integers will be called *lattice points*, and a segment joining two nearest lattice points will be called a *small edge*. We consider positive measures m which are supported by a union of small edges, that satisfy the following properties:

(1) The restriction of m to each small edge is a multiple of a linear measure. This multiple is called the *density* of m on the small edge.

(2) m satisfies the balance condition

$$m(AB) - m(AB') = m(AC) - m(AC') = m(AD) - m(AD')$$

whenever A is a lattice point and the neighboring lattice points B, C', D, B', C, D' are in cyclic order around A.



The density of a measure m will be considered to be zero on segments outside its support. A lattice point incident to at least three small edges in the support of m is called a *branch point* of the measure m. We only consider measures with at least one branch point.

Fix an integer $r \geq 1$, and denote by Δ_r the (closed) triangle with vertices $0, ru, ru + rv = -rw$. We use the notation

$$A_j = ju, \quad B_j = ru + jv, \quad C_j = (r - j)w,$$

for $j = 0, 1, \ldots, r$, for the lattice points on the boundary of Δ_r. We also set

$$X_j = A_j + w, \quad Y_j = B_j + u, \quad Z_j = C_j + v$$

for $j = 0, 1, 2, \ldots, r$.

We denote by $\mathcal{M}_r$ the collection of all measures m satisfying conditions (1) and (2), whose branch points are contained in Δ_r, and such that

$$m(A_j X_{j+1}) = m(B_j Y_{j+1}) = m(C_j Z_{j+1}) = 0, \quad j = 0, 1, \ldots, r.$$

The numbers

$$\alpha_j = m(A_j X_j), \quad \beta_j = m(B_j Y_j), \quad \gamma_j = m(C_j Z_j),$$

where $j = 0, 1, \ldots, r$, will be called the *exit densities* of m. A measure $m \in \mathcal{M}_r$ is said to be rigid if there is no other measure $m' \in \mathcal{M}_r$ with the same exit points and exit densities as m. In other words, a rigid measure is entirely determined by its exit densities.

Given a measure $m \in \mathcal{M}_r$, we define its weight $w(m) \in \mathbb{R}_+$ to be

$$w(m) = \sum_{j=0}^{r} m(A_j X_j) = \sum_{j=0}^{r} m(B_j Y_j) = \sum_{j=0}^{r} m(C_j Z_j).$$

The equality of the three sums giving $w(m)$ is an easy consequence of the balance condition.

The remainder of the paper is organized as follows. In Section 2 we formulate the Littlewood-Richardson Rule in terms of measures. In Section 3 we focus our discussion of measures on a special kind - tree measures. This leads us to our main results in Section 4, where we develop a set of rules for constructing rigid tree measures. We conclude the paper with possible directions for future research in Section 5.

2 The Littlewood-Richardson Rule

We can describe the Littlewood-Richardson rule in terms of measures, and this turns out to be a very useful way to study intersections of Schubert varieties. Given integers $n, 1 \leq k \leq n-1$, the Grassmanian manifold $Gr(n,k)$ is defined to be

$$Gr(n,k) = \{k\text{-dimensional linear vector subspaces of } \mathbb{C}^n\}$$

For every flag

$$\mathcal{E} = \{\{0\} = E_0 \subset E_1 \subset E_2 \subset \cdots \subset E_n = \mathbb{C}^n\},$$

where E_j is a subspace of dimension j, $Gr(n,k)$ can be written as a union of Schubert varieties described as follows. For each set

$$I = \{i_1 < i_2 < \cdots < i_r\} \subset \{1, 2, \ldots, n\},$$

one defines the Schubert variety

$$S(\mathcal{E}, I) = \{M \in Gr(n,k) : \dim(M \cap E_{i_x}) \geq x, x = 1, 2, \ldots, k\}.$$

Given sets $I, J, K \subset \{1, 2, \ldots, n\}$ of cardinality k such that

$$\sum_{l=1}^{k} (i_l + j_l + k_l - 3l) = 2k(n-k)$$

the Littlewood-Richardson rule provides a non-negative integer c_{IJK} with the property that the set

$$S(\mathcal{E}, I) \cap S(\mathcal{F}, J) \cap S(\mathcal{G}, K)$$

has a finite intersection, equal to c_{IJK}. The integer c_{IJK} is called the Littlewood-Richardson coefficient, and c_{IJK} can be defined in terms of measures.

Assume that $m \in \mathcal{M}_r$ assigns integer densities to all small edges. Let $\alpha_n, \beta_n, \gamma_n$ be the exit densities of m. We can then define an integer

$$n = r + w(m),$$

and sets $I, J, K \subset \{1, 2, \ldots, n\}$ of cardinality r by setting $I = \{i_1, i_2, \ldots, i_r\}$, where

$$i_l = l + \sum_{n=0}^{l-1} \alpha_n, \quad l = 1, 2, \ldots, r, \tag{1}$$

$$j_l = l + \sum_{n=0}^{l-1} \beta_n, \quad l = 1, 2, \ldots, r, \tag{2}$$

$$k_l = l + \sum_{n=0}^{l-1} \gamma_n, \quad l = 1, 2, \ldots, r. \tag{3}$$

These are precisely the triples of sets (I, J, K) which satisfy the Littlewood-Richardson rule. The *Littlewood-Richardson coefficient* c_{IJK} equals the number of measures $m \in \mathcal{M}_r$ with integer densities which satisfy (1), (2), and (3), i.e. which have the same exit densities as m.

Given a measure $m \in \mathcal{M}_r$, we formulate the associated Schubert intersection problem. A measure m determines sets I, J, K as above. The problem is to compute explicit elements in the intersection of three Schubert varieties,

$$S(\mathcal{E}, I) \cap S(\mathcal{F}, J) \cap S(\mathcal{G}, K),$$

where $\mathcal{E}, \mathcal{F}, \mathcal{G}$ are generic flag varieties. An explicit solution of the Schubert intersection problem associated with a measure can be produced in the rigid case, and the method is described in [2].

Example. In Figure 1, the measure m assigns density 1 to the thickened edges and density 0 to the other edges in the triangle.

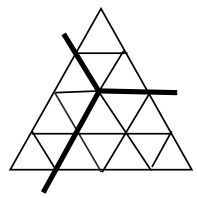


Figure 1: This measure has one branch point.

The sets I, J, K are determined by m as follows:

$$I = \{1, 3, 4, 5\}$$
$$J = \{1, 3, 4, 5\}$$
$$K = \{1, 2, 4, 5\}$$

The subspace M of $\mathbb{C}^5$ in the intersection of three Schubert varieties

$$S(\mathcal{E}, I) \cap S(\mathcal{F}, J) \cap S(\mathcal{G}, K)$$

is exactly $M = E_1 + F_1 + G_2$, where E_1, F_1, G_2 are elements of the flags $\mathcal{E}, \mathcal{F}, \mathcal{G}$, respectively.

3 Trees and Measures

Some measures $m \in \mathcal{M}_r$ have an underyling tree structure which we describe next. We consider trees embedded in the plane $\mathbb{R}^2$ such that

(T.1) each edge of the tree is a straight line segment of unit length,

(T.2) each vertex has order 2 or 3, and

(T.3) there are only finitely many vertices of order 3.

These conditions imply that a tree is infinite, but has a finite number of *ends*. Ends are sequences of vertices of the form $V_0 V_1 \ldots$, where V_0 is a branch point, V_j has order 2 for $j \geq 1$, and $V_j V_{j+1}$ is an edge for each $j \geq 0$. We will require one more condition on our trees.

(T.4) The shortest path joining two different ends contains an odd number of vertices of order 3.

All trees discussed in the sequel satisfy the above four properties.

An *immersion* of a tree $T \subset \mathbb{R}^2$ is a continuous map $f : T \to \mathbb{R}^2$ which satisfies the following properties:

(1) f is isometric on each edge.

(2) If VA and VB are two edges meeting at a vertex of order 2, then

$$2f(V) = f(A) + f(B).$$

(3) If VA, VB, VC are three edges meeting at a vertex of order 3, then

$$3f(V) = f(A) + f(B) + f(C),$$

and the restriction of f to $VA \cup VB \cup VC$ preserves the orientation of the tree.

A tree is endowed with an arclength measure. Given an immersion f of T, we consider the push-forward m_f of this measure. That is, if $f(T)$ is contained in the small edges of a lattice determined by the vectors u, v, w, then m_f assigns to each edge a density equal to the number of its preimages in T. The resulting measure satisfies the balance condition at all vertices. Since T has a finite number of ends, we can arrange f so that m_f belongs to $\mathcal{M}_r$ for sufficiently large r.

A measure m will be called a *tree measure* if $m = m_f$ for some immersion f of a tree. In the next section, we will construct a set of rules for which an immersion produces a rigid tree measure.

4 Results

There is a certain class of loops which indicate non-rigidity if they exist in the support of a measure. Let $A_1A_2 \dots A_kA_1$ be a loop consisting of small edges A_jA_{j+1} contained in the support of a measure $m \in \mathcal{M}_r$. We will say that this loop is *evil* if each three consecutive points $A_{j-1}A_jA_{j+1} = ABC$ forms an *evil turn*, i.e. one of the following situations occurs:

(E.1) $C = A$, and the small edges BX, BY, BZ which are $120°, 180°$, and $240°$ clockwise from AB are in the support of m.

(E.2) BC is $120°$ clockwise from AB.

(E.3) $C \neq A$ and A, B, C are collinear.

(E.4) BC is $120°$ counterclockwise from AB and the edge BX which is $120°$ clockwise from AB is in the support of m.

(E.5) BC is $60°$ counterclockwise from AB and the edges BX, BY which are $120°$ and $180°$ clockwise from AB are in the support of m.

The existence of an evil loop in the support of a measure implies non-rigidity. This was proven in [1]:

Theorem 4.1. *A measure $m \in \mathcal{M}_r$ is rigid if and only if its support contains none of the following configurations:*

(1) Six edges meeting at one lattice point.

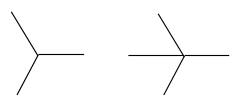
(2) An evil loop.

Our main result is to prove the following proposition:

Proposition 4.2. *Let T be a tree that satisfies properties (T.1) through (T.4), and f an immersion $f : T \to \mathbb{R}^2$. Suppose f satisfies the following conditions:*

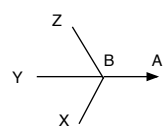
(1) There is a vertex $A \in T$ (called the root of T) such that $f^{-1}(f(A)) = \{A\}$.

(2) The only branch points of $f(T)$ are of the following forms (up to rotation):



(3) f has consistent orientation. That is, suppose X_1X_2 and Y_1Y_2 are edges of T such that X_1 and Y_1 are closer to the root of T than X_2 and Y_2, respectively, and $f(X_1X_2) = f(Y_1Y_2)$. Then $f(X_1) = f(Y_1)$ (and consequently, $f(X_2) = f(Y_2)$).

(4) If four edges meet at a lattice point B, then the orientation of one of the edges is determined as follows. Let AB be the small edge such that the other small edges BX, BY, BZ are located $120°, 180°, 240°$ clockwise from AB. Then the orientation of AB must point from B to A.



Then m_f is a rigid tree measure.

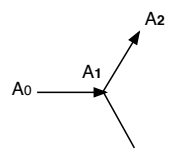
Proof. Suppose $f : T \to \mathbb{R}^2$ satisfies the conditions of the hypothesis. By Rule (2), a branch point in $f(T)$ cannot have six surrounding edges all belonging to $f(T)$. By Theorem 4.1, it is enough to show that $f(T)$ contains no evil loops. We show that if an evil loop exists in $f(T)$, then the loop lifts to a single branch of T. This will lead to a contradiction of property T.3.

By the conditions imposed on the immersion f, the only evil turns that can arise in $f(T)$ are E.1, E.2, E.3 and E.4. The evil turn E.5 cannot arise because of Rule (2). The turns E.2 and E.4 are reversible under our conditions, in the sense that they are evil either way we traverse them.

Let $A_0A_1A_2 \cdots A_n$ be an evil loop in $f(T)$ (where $A_0 = A_n$). Since an evil loop must have a turn of the form E.2, E.4, or E.5 (because collinear "turns" alone cannot form a loop), we may choose A_0 so that the evil turn $A_0A_1A_2$ is of the form E.2 or E.4. Since the turns E.2 and E.4 are reversible, we may assume that A_0A_1 assumes the orientation of the tree (if not, we traverse the evil loop

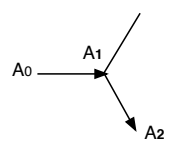
in the reverse order). That is, we can assume that there is a lift B_0B_1 to the tree of the edge A_0A_1. We show that A_1A_2 can also be lifted by showing that the orientation of A_1A_2 matches the tree orientation. The possible configurations of $A_0A_1A_2$ are:

(E.2) A_1A_2 is 120° clockwise from A_0A_1.
The edge A_1X which is 240° clockwise from A_0A_1 must also be in the support of $f(T)$.



The edge A_1A_2 must assume the orientation of the tree by consistency of orientation of the immersion f.

(E.4) A_1A_2 is 120° counterclockwise from A_0A_1 and the edge A_1X which is 120° clockwise from A_0A_1 is in the support of $f(T)$.



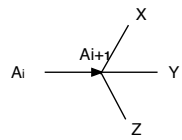
The edge A_1A_2 must assume the orientation of the tree by consistency of orientation of the immersion f.

Thus, given a lift B_0B_1 of A_0A_1, there is a vertex B_2 such that B_1B_2 lifts A_1A_2. So the evil turn $A_0A_1A_2$ lifts to a simple path $B_0B_1B_2$ on T, consistent with the orientation of T.

We inductively show that each edge in the evil loop $A_0 \dots A_n$ lifts to an edge of the tree. For simplicity, we let $A_{k+mn} = A_k$ for all $m \in \mathbb{N}$, $0 \leq k < n$.

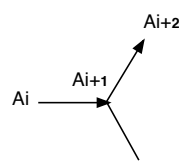
Suppose there is a lift B_iB_{i+1} of the edge A_iA_{i+1} so that A_iA_{i+1} assumes the orientation of the tree. We show that there is also a lift $B_{i+1}B_{i+2}$ of the edge $A_{i+1}A_{i+2}$. The possible configurations of the evil turn $A_iA_{i+1}A_{i+2}$ are:

(E.1) $A_{i+2} = A_i$, and the small edges $A_{i+1}X, A_{i+1}Y, A_{i+1}Z$ which are $120°, 180°$, and $240°$ clockwise from A_iA_{i+1} are in the support of $f(T)$.



The orientation of the edge A_iA_{i+1} in this evil turn contradicts Rule (4) of the immersion f, so this evil turn does not arise in $f(T)$.

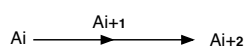
(E.2) $A_{i+1}A_{i_2}$ is $120°$ clockwise from A_iA_{i+1}.
The edge $A_{i+1}X$ which is $240°$ clockwise from A_iA_{i+1} must also be in the support of $f(T)$.



The edge $A_{i+1}A_{i+2}$ must assume the orientation of the tree by consistency of orientation of the immersion f.

(E.3) $A_{i+2} \neq A_i$ and A_i, A_{i+1}, A_{i+2} are collinear. There are two cases:

(a) The lattice point A_{i+1} has only two surrounding edges, A_iA_{i+1} and $A_{i+1}A_{i+2}$, in the support of $f(T)$.



Then the edge $A_{i+1}A_{i+2}$ must assume the orientation of the tree by consistency of orientation of the immersion f.

(b) The lattice point A_{i+1} has four surrounding edges in the support of $f(T)$.
By Rule (4), the edge $A_{i+1}A_{i+2}$ must assume the orientation of the tree.

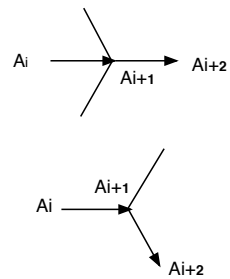
(E.4) $A_{i+1}A_{i+2}$ is $120°$ counterclockwise from A_iA_{i+1} and the edge $A_{i+1}X$ which is $120°$ clockwise from A_iA_{i+1} is in the support of $f(T)$.

The edge $A_{i+1}A_{i+2}$ must assume the orientation of the tree by consistency of orientation of the immersion f.

Thus, given a lift B_iB_{i+1} of A_iA_{i+1}, there is a vertex B_{i+2} such that $B_{i+1}B_{i+2}$ lifts $A_{i+1}A_{i+2}$, completing the induction.

We have constructed inductively an infinite path

$$B_0B_1B_2 \cdots B_nB_{n+1} \cdots$$



of the tree T with the following properties:

(1) $f(B_j) = A_j$

(2) $B_j B_{j+1}$ is oriented away from the root.

We claim that infinitely many of the vertices B_j have order 3.

By construction, the vertex A_1 is a branch point of m_f. Thus lifted vertex $B_1 \in T$ with $f(B_1) = A_1$ is a vertex of order 3, because the immersion f cannot map a vertex of order 2 to a turn of the form E.2 or E.4.

Moreover, each vertex B_{1+mn}, $m \in \mathbb{N}$, has order 3 because

$$f(B_{1+mn}) = A_{1+mn} = A_1.$$

Thus, infinitely many vertices B_j have order 3. But this contradicts property T.3 of the tree. Thus, we've shown that en evil loop cannot exist in $f(T)$, and therefore m_f is a rigid tree measure. □

The above proposition guarantees that if we construct a tree measure following the four stated rules, then the resulting measure will be rigid.

5 Future Research

The finiteness of the number of vertices of order 3 (i.e. property T.3) was crucial in the proof of Proposition 4.2. Our next question is whether Proposition 4.2 still applies to trees with an infinite number of vertices of order 3. That is, if a tree satisfies properties T.1, T.2, and T.4, and if an immersion $f : T \to \mathbb{R}^2$ satisfies the four rules of Proposition 4.2, then will the resulting measure $f(T)$ be rigid?

Another question we investigated is the following: Given a triangle of size r, what is the maximum weight that a rigid tree measure $m \in \mathcal{M}_r$ could have? We have a lower bound on this number: for large r, the maximum weight of a measure in $\mathcal{M}_r$ is at least $2^{\lfloor r/3 \rfloor}$. The lower bound can be easily explained with the help of a couple of figures.

Given a measure $m \in \mathcal{M}_r$ of weight w, Figures 2 and 3 demonstrate how to produce a measure $m' \in \mathcal{M}_{r+3}$ of weight $2w$. By branching the exit densities on one side of Δ_r in the illustrated way, one exit branch of density w is added

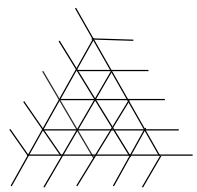


Figure 2: A measure $m \in \mathcal{M}_4$ of weight w

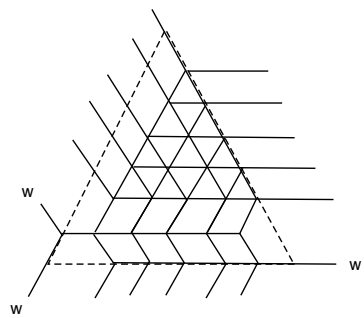


Figure 3: A measure $m' \in \mathcal{M}_7$ of weight $2w$

to each side of the triangle, thereby increasing the total weight by a factor of 2. In this process, we also increase the size of the triangle by 3. Thus, for large r, the maximum weight of $m \in \mathcal{M}_r$ is $\geq 2^{\lfloor r/3 \rfloor}$. We are interested in finding a least upper bound on the maximum weight of a measure.

A tree measure of weight w has $3w - 2$ branch points. This relation allows us to investigate the maximal weight problem using a different approach. The problem becomes that of finding the maximum number of branch points of a rigid tree measure in $\mathcal{M}_r$.

Thus, we attempted to characterize trees with rigid immersions. That is, we are interested in the types of trees that can be immersed onto the plane to produce rigid measures. Does every tree have a rigid immersion? Are there some tree configurations that absolutely prohibit a rigid immersion? By characterizing trees with rigid immersions, we might be able to find an upper bound on the number of branch points that a rigid tree measure $m \in \mathcal{M}_r$ can have, and ultimately find the maximum weight that such a measure could have.

6 Acknowledgements

I would like to thank my faculty advisor Hari Bercovici for his guidance throughout the project and his insight into the problems. I thank him for teaching me a great deal of mathematics, in topics related to my project and beyond.

I want to thank Wing Suet Li for helping with my understanding of the ideas of my project. I also thank Kevin Pilgrim for running the Indiana University

REU in mathematics, and the National Science Foundation for funding the REU program. Finally, I thank my fellow REU students for an enjoyable and memorable summer.

References

[1] H. Bercovici, B. Collins, K. Dykema, W. S. Li and D. Timotin, *Intersections of Schubert varieties and eigenvalue inequalities in an arbitrary finite factor*, Journal of Functional Analysis 258 (2010) 1579-1627.

[2] H. Bercovici, W. S. Li, and D. Timotin, *A family of reductions for Schubert intersection problems.* ArXiv 0909.0908 (2009): http://adsabs.harvard.edu/abs/2009arXiv0909.0908B.

Moduli spaces of square-tiling deformations and single-tile real line tilings

Garrett Proffitt

Moduli spaces of square-tiling deformations and single-tile real line tilings

Garrett Proffitt

October 27, 2010

Moduli Spaces of Square Tiling Deformations and Single-Tile Real Line Tilings

Garrett Proffitt
mentored by
Dr. Matthias Weber

October 26, 2010

Abstract

Classifying Single-Tile Periodic Tilings of the Real Line and Realizing the Deformation Spaces of Two and Three Square Periodic Tilings of the Plane through Combinatorial Structure.

1 Introduction

A tiling can be thought of as an assortment of objects that covers a space cleanly and without any gaps. A periodic tiling can be thought of as a base object that is a subset of the space (what we will call a fundamental domain) being translated around the space by a consistent method of instructions (what we will deem translation by the group of translational vectors). Generally, in $\mathbb{R}^n$, we can think of a fundamental domain as an n-dimensional object (or subset of $\mathbb{R}^n$) with a set of n translational vectors. We will be looking at periodic tilings in $\mathbb{R}$ and $\mathbb{R}^2$, so we will only need one or two translational vectors.

Figure 1: Example of a periodic tiling in $\mathbb{R}$

More specifically, in $\mathbb{R}^2$, we will be looking at periodic tilings by squares, most notably tilings by one, two, and three squares.

As for periodic tilings of $\mathbb{R}$, we will be looking at tiles consisting of disjoint intervals. Many examples will be given for both. (Note we cannot show all of $\mathbb{R}$ or $\mathbb{R}^2$ in our tiling, so we will only show a subset of these regions being tiled.)

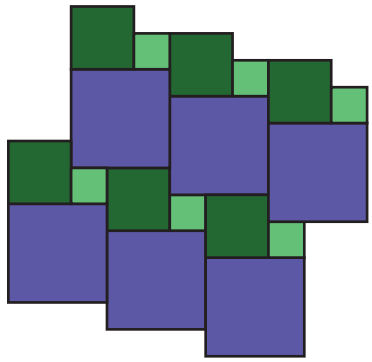


Figure 2: Example of a tiling in $\mathbb{R}^2$

2 Tilings of the Real Line

2.1 Defining the Tiling

Given any length interval in $\mathbb{R}$, call length L, we can subdivide the interval into n parts, say, of lengths $a_1, ..., a_n$ such that the sum of the a_i's equals L. Note the a_i's do not have to be in any specific order.

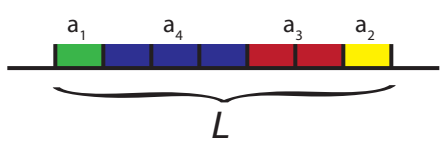


Figure 3: A divided interval of length L

We can then deform the intervals in a modularly arithmetic specific manner (breaking the parts up and putting them apart from each other, called a **lift** mapping) to create a union of disjoint intervals. We call this an **n-prong tile**.

Definition 2.1. Let $a_1, ..., a_n$ represent interval lengths in $\mathbb{R}$. An **n-prong tile** is a union of disjoint intervals $a_1, ..., a_n$ (a_i's in increasing order) with $a_1 + ... + a_n = L$ such that the tile modulo L is a single connected tile of length L.

Figure 4: Creating an n-prong tile ($n = 4$) from a single interval of length L

Let the a_1 length interval always be the first disjoint interval (we know we can do this by simply moving any interval in front of a_1 to the back (e.g. $a_5a_{10}a_1a_n...a_{n-25}$ is equivalent to $a_1a_n...a_{n-25}a_5a_{10}$ by modular arithmetic mod

L). Then this n-prong tile will tile $\mathbb{R}$ by translation with period L. We would like to say this is the only way to tile $\mathbb{R}$ with translates of one tile, but it will not be proven in this paper (see the results of Lagarias and Wang in [1] for showing how all tilings of the real line by a single tile will eventually be periodic).

2.2 Examining the Tiling

Note that with n disjoint intervals, (we will refer to them as prongs, numbered in increasing order: "1, 2, . . . , n"), we have one prong that is fixed (namely prong 1), which we must have in order to define our n-prong tile.

Figure 5: Because of the numbering of the prongs (1-2-3, 2-3-1, and 3-1-2), the above three tiles are equivalent; they are just starting at different prongs. We fix prong 1 as the starting prong to make sure we don't repeat the same tilings.

This means the other $n-1$ prongs can be patterned in any desired order over the period-length interval (i.e. when the tile is taken modulo L), which allows for $(n-1)!$ **classes** of countably infinite many different tilings. We control the different tilings by modulating the gap sizes, or if the **prong pattern** is in consecutive order (e.g. 1, 2), we can set the gap between the two equal to zero and make an $n-1$ prong tile, or set the gap equal to L, and by modulating this length we get $0+L+L+...+L$ (m times), or simply mL. (For modulating other gaps, we take the shortest length between the consecutive prongs (say $a_1 + a_3$ or something) then go around the interval again, which will of course add on L to the shortest gap length, and we can do this as many times as we please.)

Figure 6: Modulating the second gap in a 3-prong tile

Also it should be noted that gaps can be modulated independently. If we call gaps b and call the gap after each a_i, (for $i = 1, ..., n$) b_i (so that b_n is the gap after the end prong a_n), we can associate an m_i in $\mathbb{N}$ to each b_i to indicate how modulated a gap is, and m_i does not necessarily have to equal m_j. This ensures there are countably infinite different tilings for each class of tiles.

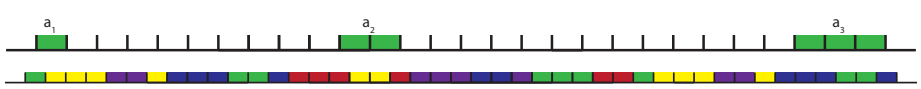


Figure 7: We use the same tile as in the previous example. This time, the first gap is modulated once and the second gap is modulated twice. We also show it tiles.

Here are the definitions of some of the terms emphasized above:

Definition 2.2. The **prong pattern** of an n-prong tiling is the numerical ordering of the prongs when the tile is taken modulo L. The prong pattern determines which tiling **class** we are in, which is useful in determining if two tilings are **equivalent**.

Definition 2.3. Two tilings are in the same **class** if their prong patterns are the same. If two tilings are in the same class, they are said to be **equivalent**.

2.3 Building up the Classes

Let us look at the classes for each n-prong tile, starting with $n = 1$:

For a 1-prong tile, all we do to tile is shift by the period (or interval) length L. We can call this **Class 0** or "class-less" since the prong pattern is always "1, 1, 1, 1, . . .", and no other tilings have the same prong number adjacent to one another.

Figure 8: Example of a Class 0 tiling

For a 2-prong tile, we have prong number 1 fixed in the first spot, so prong number 2 must come next, yielding $1! = 1$ class(es), so there is only one class of countably infinite tilings. The prong pattern will always be "1, 2, 1, 2, . . .". Let us examine further:

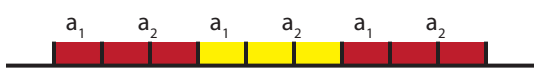


Figure 9: A two prong tile with a gap length of zero

We can have gap $b_1 = 0$, but that reduces to a class-less 1-prong tile. (Note the obvious fact that combining any two prongs by setting $b_i = 0$ will reduce an n-prong tile to an $(n-1)$-prong tile, drastically reducing the number of classes for decent-sized n (say $n \geq 4$).) To make things interesting, we take $b_1 = a_2 + a_1$ and get:

Figure 10: A two prong tile, tiling by translation by $a_1 + a_2$

As previously stated, we can take $b_1 = m_1(a_2 + a_1)$ to get the countably infinite many tilings in this class. Note the increasing modular pattern "1, 2, 1, 2, . . ." where the prong number is increasing until all the prongs are named

(e.g. "1, 2, 3, 1, 2, 3, . . ." for 3-prongs and "1, 2, 3, 4, 1, 2, 3, 4, . . ." for 4-prongs). This pattern shows up as a class for all n-prong tilings, $n \geq 2$, so we call this **Class 1**.

Definition 2.4. For any n-prong tile, $n \geq 2$, a tiling with prong pattern "1, 2, 3, 4, . . . , n" is defined to be a **Class 1** tiling.

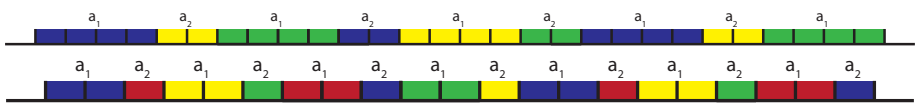


Figure 11: Examples of 2-prong Class 1 tilings; note the second example has its gap modulated

For a 3-prong tile, we have prong number 1 fixed in the first spot, leaving prongs 2 and 3 free, yielding $2! = 1$ classes with prong patterns "1, 2, 3, 1, 2, 3, . . ." or "1, 3, 2, 1, 3, 2, . . .". We have already looked at the first and named it Class 1 (and it is important to note Class 1 tilings have the longest unmodulized gaps, for all the gaps are always equal to at least L). We now must look at this second interesting prong pattern that is decreasing modularly (like "3, 2, 1, 3, 2, 1, . . ." with 1 moved to the front by convention):

Figure 12: Showing the lift mapping for a Class 2, 3-prong tile

Completely opposite of the Class 1 tilings, these **Class 2** tilings have the shortest unmodulized gaps (in the case of the 3-prong tiles). As with Class 1, Class 2 tilings appear for every n-prong tiling (for $n \geq 3$ this time).

Definition 2.5. For any n-prong tile, $n \geq 3$, a tiling with prong pattern "1, n, $n-1$, $n-2$, . . . , 3, 2" is defined to be a **Class 2** tiling.

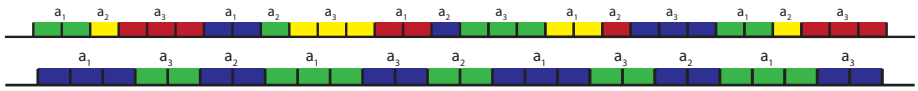


Figure 13: Examples of 3-prong tilings; the first is a Class 1 tiling and the second is a Class 2 tiling

For a 4-prong tile, we have prong number 1 fixed in the first spot, leaving three free prongs, yielding $3! = 6$ classes, and this is where the classes start to become overwhelming. We only name Class 1 and Class 2 and call the other classes by the prong pattern they have. For the 4-prong tiles, the six classes are:

Class 1 (1,2,3,4)
Class 2 (1,4,3,2)
Class 1-2-4-3
Class 1-3-2-4
Class 1-3-4-2
Class 1-4-2-3

As always, the gaps can be modulized, and the period L is the sum of the prong lengths.

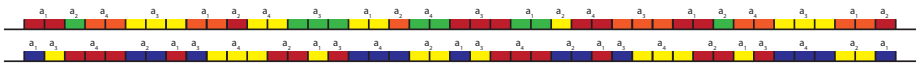


Figure 14: Examples of 4-prong tilings; the first is a Class 1-2-4-3 tiling and the second is a Class 1-3-4-2 tiling

We can keep going, to five, six, and seven prong tiles, but, in general, for n-prong tiles, we fix the first prong, leaving $n-1$ free prongs, which yield $(n-1)!$ classes. Class 1 is always included for $n \geq 2$, and Class 2 is always included for $n \geq 3$. Gaps can always be modulized independently, and the tiles are always translated by L, the sum of the prong lengths.

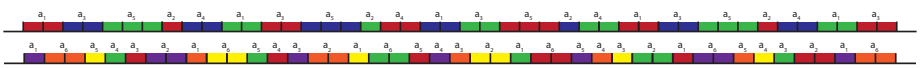


Figure 15: Examples of n-prong tilings; the first is a Class 1-3-5-2-4, 5-prong tiling, and the second is a Class 2, 6-prong tiling

2.4 Number Classification by Tilings Theorem

With the above information we have gathered empirically, we can now form a theorem relating these tilings of the real line to number theory, namely, that we can produce countably infinite real line tilings (ignoring the obvious uncountably infinitely many choices for each prong's length) for any finite ordered list of natural numbers. For ease, we will require the list of numbers to always start with 1 to avoid any repetition.

Theorem 2.6. *Given a finite ordered list N of n natural numbers beginning with 1, there exist countably-infinite many n-prong tiles (ignoring the choice of prong length) whose prong patterns are the same as the ordered list N.*

Proof. The proof simply follows from what we have already found. If we have our ordered list N that begins with 1, we can relate this list to a prong pattern of some n-prong tile. To each number i in N, we can associate a real value a_i in order to construct this tile. As before, let $L = a_1 + ... + a_n$ and then create an L-length interval broken up into these n parts with a prong pattern equivalent

to the list N. Then we can lift the tile from its modulo L state into a tile with n prongs such that we tile $\mathbb{R}$ periodically. The "countably-infinite many" part comes from the fact that we can modulate the size of each gap to whatever we please. $\square$

This theorem is useful for giving us a geometric representation of a finite pattern of natural numbers, and it is interesting to note that for *any* ordered list of natural numbers, we can find an n-prong tiling (and not just one, but *always* countably infinitely many).

3 Tilings by Squares

3.1 Definitions and Examples

Recall in the introduction we mentioned the criteria for a tiling. For tilings by squares, the space we are looking to tile is $\mathbb{R}^2$ and our **fundamental domain** is a square or a union of several squares (without any gaps). The **translational vectors** are two linearly independent vectors in $\mathbb{R}^2$, inherently with an x component and a y component. Below is an example of a tiling by one square:

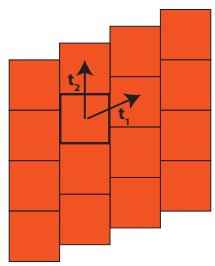


Figure 16: A tiling by one square with translational vectors $\vec{t_1}$ and $\vec{t_2}$

Note the fundamental domain in this case is just one square, but the situation changes with two squares:

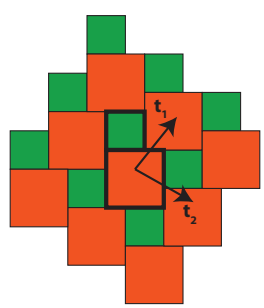


Figure 17: A tiling by two squares with translational vectors $\vec{t_1}$ and $\vec{t_2}$

It would be wise to define what we mean by fundamental domain and translational vectors in this case:

Definition 3.1. Given a topological space X and a group G acting on it, we can find that the images of a single point under the group action form an orbit, and a **fundamental domain** is a subset of X that contains exactly one point from each orbit.

This definition can be simplified for our purpose of periodically tiling $\mathbb{R}^2$ by squares since we are looking at $\mathbb{Z}^2$ acting on $\mathbb{R}^2$ by translation, which will always form a lattice.

Definition 3.2. For $\mathbb{Z}^2$ acting on $\mathbb{R}^2$ by translation (the case for our square tilings), a **fundamental domain** of a tiling is the most irreducible and irredundant piece of the tiling such that the whole tiling can be determined (in accordance with translations by a constant set of two translational vectors). We define irreducible to mean that no subsets of the fundamental domain will be able to tile the plane by itself, and we define irredundant to mean the intersection of any two fundamental domains is the empty set. In our case with a tiling by n squares, this means each of the n squares can only be represented once in the fundamental domain.

Definition 3.3. In general, for a vector space V of dimension n, the set of **translational vectors** is the consistent set of instructions that determines the n-linearly independent directions in which the fundamental domain will move, and this set of instructions is repeated indefinitely until the entire space is tiled. For $\mathbb{Z}^2$ acting on $\mathbb{R}^2$ by translation, the set of translational vectors are two linearly independent vectors $\vec{t_1}$ and $\vec{t_2}$ in $\mathbb{R}^2$.

Looking back at our first two-square example, this becomes obvious:

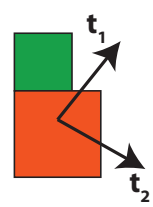


Figure 18: A single tile of two squares that tiles $\mathbb{R}^2$

The "L"-shaped fundamental domain is being translated by the vectors $\vec{t_1}$ and $\vec{t_2}$ throughout all of $\mathbb{R}^2$, and we can see this will tile the plane periodically. We can prove this by looking at the fundamental domain and seeing where another copy of it would have to go to tile periodically. In this case, we look at the corner where the "L"-shape bends and look at what can fill that corner. The only possible fit is the bottom left corner of the fundamental domain, so we attach another copy there and continue:

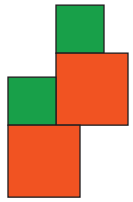


Figure 19: The two tiles must fit together in this manner

Note that this automatically determines one of the vectors, since we must do this for all copies and so we will eventually form an infinite band.

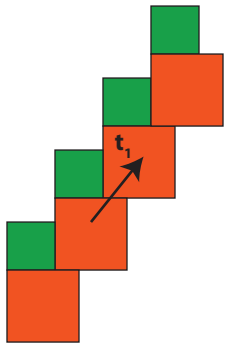


Figure 20: One translational vector is determined

To find the other vector, we must note how the long bands can fit together. Since there is one band, there must be another. We look at the zig-zagging pattern on the bands, and note how the two bands fit together.

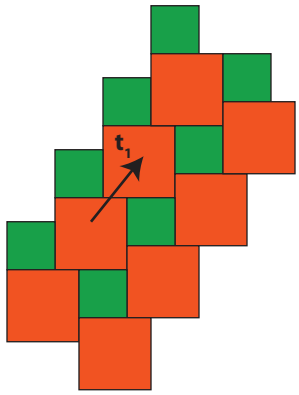


Figure 21: Two bands can only fit together in this manner

Since these two bands fit together, we can connect all bands together and get a periodic tiling of the entire plane.

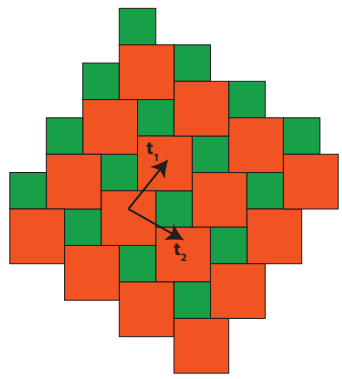


Figure 22: The second vector is determined, and we tile $\mathbb{R}^2$

We can construct a useful proposition of the process we just went through for all "L"-shaped tiles:

Definition 3.4. A fundamental domain (or tile) is "L"-shaped if it consists of six vertices, five of which form outward-oriented corners and one of which forms an inward-oriented corner (as one would expected for any block "L"). A **perfect** "L"-shaped fundamental domain is one that can take no other forms when tiled, i.e. its tiling cannot have an alternate fundamental domain that is a rectangle or a "Z"-shape (we will not discuss "Z"-shaped tiles however).

Proposition 3.5. *Given a perfect "L"-shaped fundamental domain in* $\mathbb{R}^2$*, there exists only one unique tiling of* $\mathbb{R}^2$

Proof. To tile, all gaps must be filled. That means the one inward corner of our "L" shape must be filled. However, the only possible translate that will fill that corner is diagonal, so that the corner opposite the inward one meets it cleanly (a figure above represents this step). If we don't do this, we will have a gap or an overlap near the inward corner. We can translate in this direction infinitely long to get an infinite band, and now we must find how the band interlocks.

Since we plan on tiling the entire plane, we must have another one of these infinitely long bands, and the two must cleanly match-up (otherwise we will have gaps or overlaps). Note the zig-zag patterns on either side of the bands. We can cleanly match-up the bands in multiple ways (just by shifting the bands against each other), but each of these match-ups produces the same tiling (see one of the above figures for a visualization). Continuing to add these bands on in this manner will tile the plane.

We tried all possibilities to create this tiling (looking at options for the inward corner and the zig-zag), but we can see there is only one tiling in which the "L"-shaped tiles will fit together. □

The process discussed above can be used to determined if a tile will actually tile the plane if it is not "L"-shaped, but as one can see the process can be quite lengthy and fairly inefficient. Because of this we introduce the new idea a **deformation space**.

Definition 3.6. Given a periodic tiling of the plane by squares whose sides are parallel to the axes, a **deformation space** is the set of all possible periodic tilings that can be obtained by a single translation or a single scaling of a subset of squares in the fundamental domain. Both the scaling and shifting are done in a continuous manner.

For ease, we will disregard any deformation that uniformly scales all parts of the fundamental domain, and we will disregard any rotations that are used to deform the tiling. Since this can be done with any tiling, it is best to mention it once but then ignore it due to triviality.

We can see what this looks like in our "L"-shaped example. We can deform this tile by scaling the smaller square on top until the two squares are of equal size.

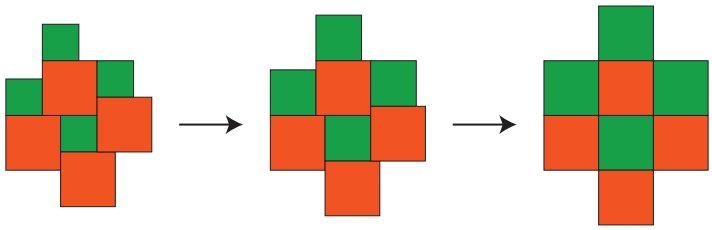


Figure 23: Deforming the "L"-shaped two square tiling by growing the small square

Note we can continue this deformation by now shrinking the square to the other side:

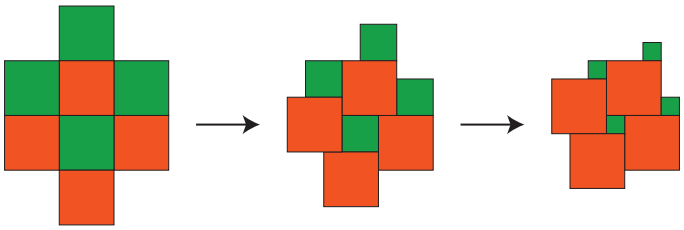


Figure 24: Deforming the equal-size two square tiling by shrinking the green square to the right side

If we continue this deformation, we will end up shrinking one square to a point, leaving us with one square. We can also start with our original tile and

shrink the top square to a point, again leaving us with one square, the same one as in the first case. This completes a deformation space with two squares of different sizes. Note the importance of color in the example as well; a tiling by a single orange square and a tiling by a single green square are not the same.

To make sure a fundamental domain tiles, we still must go through the steps outlined above. Allowing for the deformation spaces gives us a better intuition on what may tile and what may not, and we are better able to understand the underlying structure of a **tiling system** with their help.

Definition 3.7. A **tiling system** for a periodic tiling by n squares is the complete classification of all possible tilings by n squares. That is, it is the union of all existing deformation spaces.

The tiling system is usually made evident by drawing the entire moduli space, which we will now define.

3.2 Moduli Spaces

After the insight deformation spaces give us about the underlying structure of a tiling system, one begs the question, "are there multiple deformation spaces?" and "how are they related?". We introduce the concept of **moduli space** to answer these questions and give us a full view of the structure of the tiling system.

Definition 3.8. A **moduli space** is a geometric object whose points uniquely represent another mathematical object (in this case different periodic tilings by squares) that are connected by isomorphism classes (in this case they are of the same deformation space). Furthermore, note different moduli spaces can intersect in different ways.

In our case, we use moduli spaces to represent deformation spaces, where each point in the moduli space represents a unique tiling. It is important to note that two different points cannot represent the same tiling. A given deformation space maps bijectively to a corresponding moduli space; if two deformation spaces share the same tiling, these two deformation spaces must intersect at a point.

In the end of the previous section, we mapped out the deformation space of the "L"-shaped two-square tile. We can represent this as a moduli space in the form of a loop. Since the translational vectors are uniquely determined by the shape of the fundamental domain (i.e. it only admits one specific tiling), we only have a one-parameter family (save for scaling and trivial rotation) since we can only choose the size of the second square, hence the one-dimensional moduli space (the loop).

It follows that since all our deformation spaces will be "cyclic" in the one-, two-, and three-square cases, all our moduli spaces must be "cyclic", or closed into a loop for each dimension the deformation space admits. (Here, "cyclic" means that continuing the deformation will return us to the tiling with which we

started, same as what we did in the previous section. There is a slight exception in the case where we reduce our tiling to a fewer number of squares, and we will address that issue in this section.)

For dimension greater than or equal to two, it would seem that a sphere may fit this description. However, in a sphere the "loop size" will vary depending on where you are on the sphere. If we look at all loops in the x-direction in say the two-sphere, we notice the loop is full at the equator but shrinks to a point as we increase or decrease our value for y. This cannot be; it doesn't follow the structure of the deformation space we want. A torus, however, does keep a constant-sized loop for each dimension, indicating that continuing a deformation will eventually return us to the tiling with which we started. (Note this also makes more sense topologically. The loop in the one-dimensional case is simply S^1, and for an n-dimensional moduli space, we want all n-dimensions to have the independent property of following a loop. We can then take the Cartesian product of all these S^1 objects to maintain the "independence" of each dimension, giving us $S^1 \times S^1 \times ... \times S^1$, which is homeomorphic to the n-torus.) Therefore in our cases of one-, two-, and three-square tilings, the moduli spaces will either be loops or tori, and we will see that the interest lies in how they are all connected.

3.2.1 Moduli Space for One Square

Let us first look at the one-square moduli space. In determining how to begin looking at the moduli space, it is often best to start with the most basic tiling one can think of. In this case, we want to look at the "grid" tiling where all the squares line up (in the next section we will understand combinatorially why this is the case).

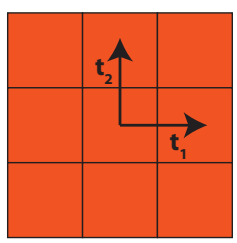


Figure 25: The basic "grid" tiling of one square

Notice that the fundamental domain in the one-square case is simply one square, so any scaling done to part of the fundamental domain will inherently tile the entire fundamental domain. Therefore we can eliminate the "independent scaling" deforming tactic mentioned previously and just look at shifts/translates. Notice that two shifts can occur by moving vertical bands up or down or moving horizontal bands left or right. Let us first look at moving the vertical bands and look at how the translational vectors are affected:

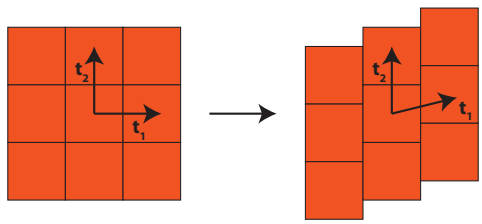


Figure 26: Deforming the one square "grid" tiling by a vertical shift

Right away note that the vertical vector stays the same, but the previously horizontal vector now has a y-component. Also note that there are no free parameters for this tiling. The new tiling created by the vertical shift still has two well-defined translational vectors coming from each fundamental domain (i.e. each square), so the tiling is determined uniquely.

We can continue this deformation until we reach the point where the tiling becomes a grid again, but note how the translational vectors have changed:

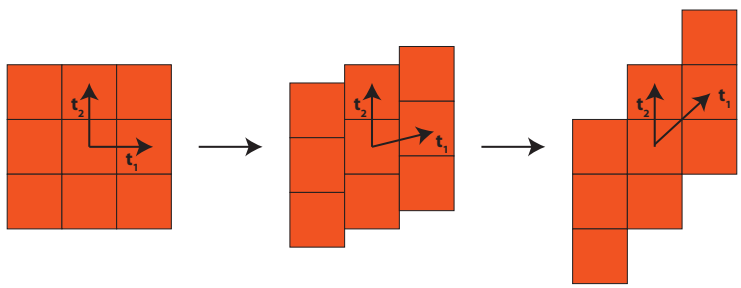


Figure 27: Continuing the vertical shift deformation

Fortunately, the first grid and the last grid are isomorphic by the group action by $SL_2(\mathbb{Z})$, so we have

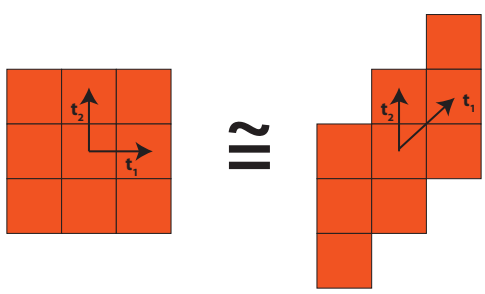


Figure 28: Isomorphic by group action by $SL_2(\mathbb{Z})$

and this shows that we have completed the loop for this moduli space.

Similarly, we can follow the same process for the horizontal band shift. There will be no free parameters, so we will again have a one-parameter family (another loop). Continuing the deformation, we will reach the original grid as before with different translational vectors, but again these tilings are isomorphic by the group action by $SL_2(\mathbb{Z})$. That closes the loop for the horizontal shift moduli space, and since there are no more deformations we can do (save for the trivial ones of scaling and rotating), we have a complete picture of all tilings by one square, and it turns out to be a nice figure eight:

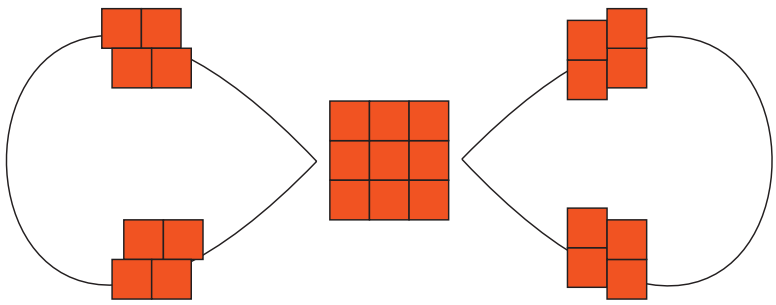


Figure 29: The entire one-square moduli space

3.2.2 Moduli Space for Two Squares

Things get more interesting with two-square tilings. As before, we want to start with the most basic tiling we can think of, which would be where all the squares are of the same size and line up in a grid. There are two colors this time instead of one, so we must think what the most basic and homogeneous way is to arrange the tiles. If we have long horizontal or vertical bands of the same color, we eliminate the homogeneous aspect we were going for, so we will instead arrange the initial tiling as a checkerboard pattern.

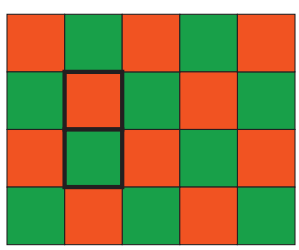


Figure 30: The checkerboard pattern two square tiling

Note there are four fundamental domains we can take from this tiling: a tall rectangle and a wide rectangle, both have a case where orange is the top-most or left-most color, and both have a case where green is the top-most or left-most color.

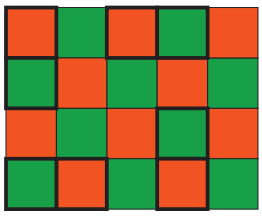


Figure 31: The checkerboard pattern with four possible fundamental domains highlighted

In this case (as we will see), the same tilings and deformation spaces are created with the four different fundamental domains, but we cannot always assume this, so it is important to make note of which fundamental domains we have available.

Let us first take the case where we scale one square to be smaller, yielding an "L"-shape tiling that we discussed in the previous section. We will deform our checkerboard tiling by shrinking the orange square relative to the green square. Note that we can shrink it two ways:

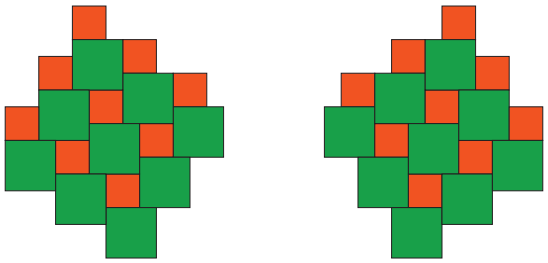


Figure 32: Both orientations of the shrinking orange square deformation

Also make note of the various fundamental domains for each tiling. Each of the tilings admits four fundamental domains, highlighted below:

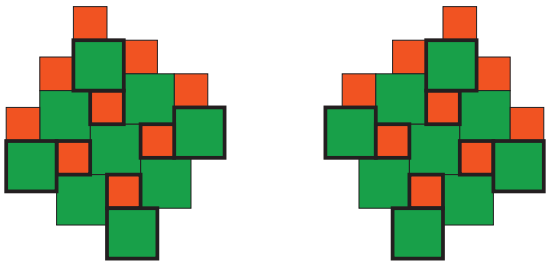


Figure 33: Both orientations of this deformation with fundamental domains highlighted

We can continue shrinking the orange square in both tilings until it becomes a point, giving us two green squares with an orange point on a corner (the corner on which the point lies depends on which fundamental domain we choose).

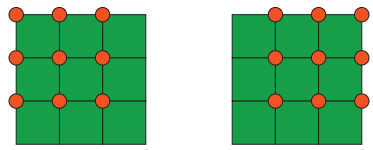


Figure 34: Both orientations in which the orange square shrinks to a point

However, it is obvious that these two new tilings are the exact same. If we look at the tiling that the green square and the orange point admits, we find that we have a grid of green squares with an orange point at every single vertex.

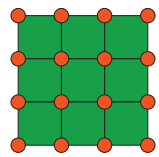


Figure 35: The single tiling at which the orange square has shrunk to a point

It now becomes unclear which green square/orange point orientation we started with, so we know that all orientations admit the same tiling. This holds true for every square with an accompanying point. Therefore we can essentially disregard the point, and this leaves us with a tiling by one square, whose "figure eight" moduli space we already know about. Then we have a complete loop with what we will call a **limit point** where we attach the moduli space for one square.

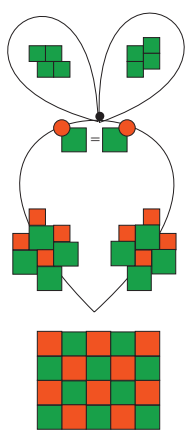


Figure 36: The moduli space in which we shrink the orange square

Definition 3.9. A **limit point** is a point in a moduli space for periodic square tilings by n squares ($n \leq 3$, in our case) where one or more of the squares is deformed into a point. At this point, we can attach the moduli space for square tilings by m squares (where m is the number of remaining squares that haven't been deformed to points). We perform this process to understand how all the different moduli spaces are connected.

We can repeat what we just did by shrinking the green square relative to the orange square, getting two different tilings depending on which side the green square was shrunk, and then getting a tiling by a single orange square where the green square is deformed to a point at the limit point in the loop. We then attach the moduli space for one square as stated in the above definition.

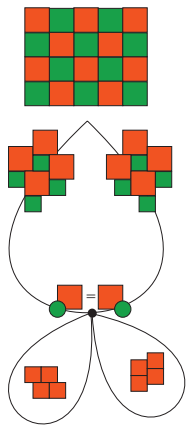


Figure 37: The moduli space in which we shrink the green square

Now that we have done all the independent scaling we could do, we must look for the possible shifts that can be done. Notice right away that a vertical band shift and a horizontal band shift is possible as in the one-square case. Let us first look at the horizontal band shift:

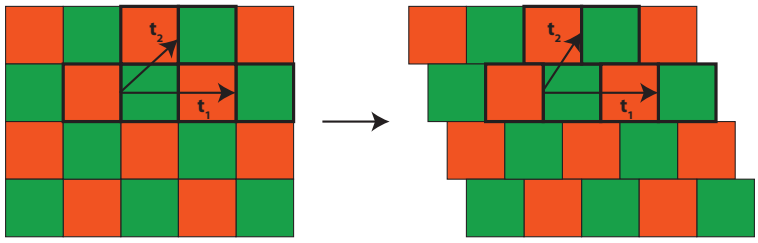


Figure 38: The horizontal shift deformation from the original checkerboard tiling

Note that as with all the previous deformation cases, the translational vectors are determined by the tiling. We have one horizontal vector and one diagonal

vector. However, let us continue this horizontal shift. Eventually notice that the green and orange squares will be aligned in vertical bands. At this point, we could continue with our horizontal shift, eventually returning back to the checkerboard tiling and staying in the same moduli space, or we could vertically shift the same-color bands, entering a whole new moduli space. Since we already have a firm grasp on the first case, we will enter the new moduli space.

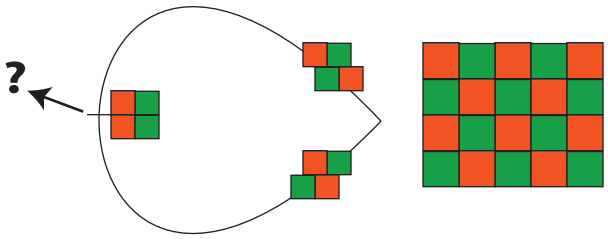


Figure 39: The moduli space for the horizontal shift deformation

The new way to deform this band tiling is to shift the vertical bands up or down. Let us do that and look at the translational vectors.

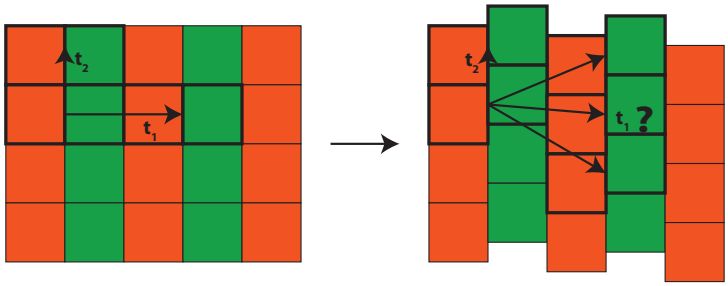


Figure 40: Shifting the vertical bands; note the one free vector

One of the vectors is defined, but the other can go anywhere past the adjacent green band. This gives us our first free translational vector in our moduli spaces, and since we already can shift the green band freely, we have two variables in this deformation space. This means instead of a loop we get a two-dimensional torus as a moduli space (the reasoning for this was explained earlier in the section).

We can repeat this whole process from the beginning at the checkerboard tiling, this time starting with a vertical shift deformation rather than a horizontal shift, and we will end up with horizontal bands of the same color, which when shifted against each other will also admit a two-dimensional torus as a moduli space.

These are all the deformations we can do with a two-square tiling, so we can combine all these moduli spaces into a single picture, giving us the entire system of periodic tilings by two squares.

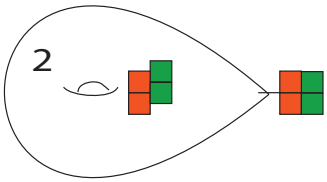


Figure 41: The toroidal moduli space for shifting the vertical bands

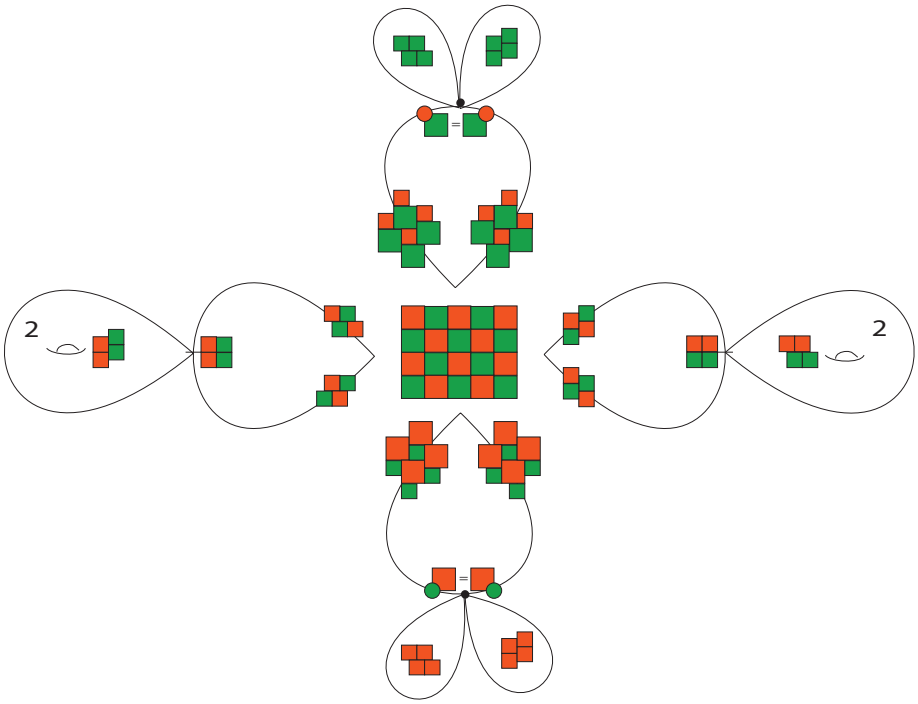


Figure 42: The entire two-square moduli space

It is important to notice that we can deform any tiling by two squares into another tiling by two squares without ever touching a limit point (the same applies for tilings by one square). This means the entire space of tilings by two squares is connected, which is a critical observation. Also, at the beginning of the discussion about the two-square space, we mentioned that the fundamental domain we look at can be any one of the four possible ones on the base tilings, and that any deformations we produce from the four will produce equivalent tilings. Instead of explaining with words, look now at the four pictures of the moduli spaces where we observe a different starting fundamental domain:

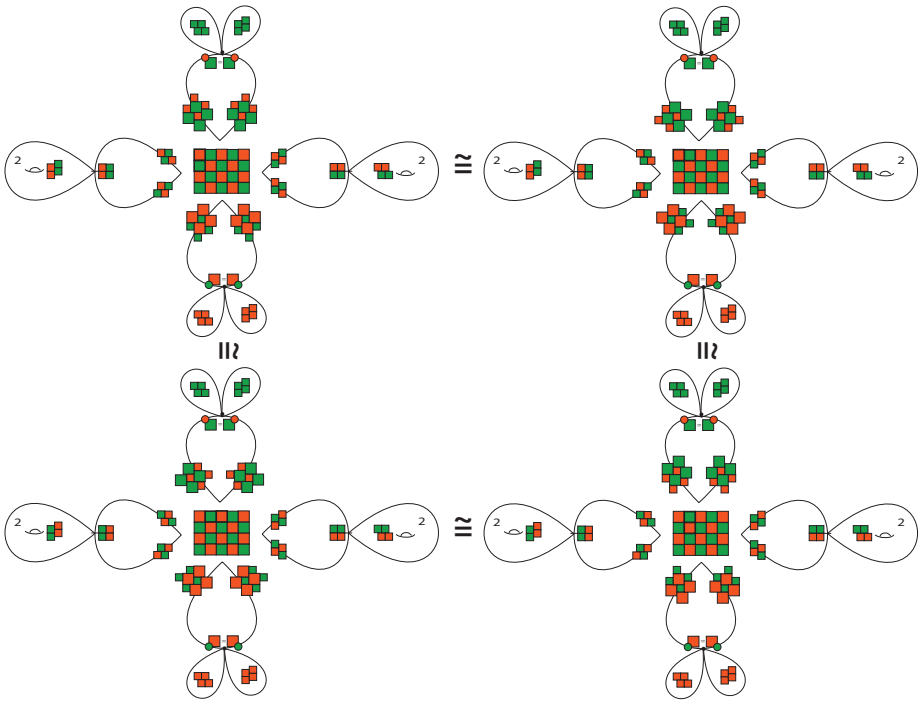


Figure 43: The four equivalent variations of the two-square moduli space, implying the choice of a fundamental domain does not matter for two squares

Note that for each deformation space, the tilings are the same, though different fundamental domains exist. The translational vectors for each case are different, but since the tilings are equivalent it does not matter. (Since each point in the moduli space uniquely represents a tiling, it does not matter if the translational vectors are different; it only matters that we have the same tiling.) There is always an "L"-shape with one color bigger than the other, there is always a checkerboard base tiling, there are limits to one square of each color, and there are the bands of color; each has multiple fundamental domains, but the tilings are equivalent. This brings us to an important proposition for two-square tilings:

Proposition 3.10. *Given a periodic tiling of the plane by two squares of colors a and b, we can take any fundamental domain we like and deform however we like and the deformation space (and thereby the moduli space) will be the same. That is, the choice of a fundamental domain does not affect the moduli space for the two-square tilings.*

Proof. The four illustrations above clearly depict the equivalence of the moduli (and deformation) spaces. Since the above moduli space represents all possible periodic tilings by two squares, the proposition holds true. □

This proposition becomes useful when studying the three-square tiling, since there are multiple places where a tiling at a limit point will have a different fundamental domain, but will produce the same tiling. Since the tilings are what we care about, we merely make these the same point to adhere to the moduli space's definition of unique representation.

3.2.3 Moduli Space for Three Squares

With tilings by three squares, everything breaks loose. We no longer have a simple, homogeneous, basic tiling with which we can start. Instead, we have four. As with the previous two cases, we find our beginning tilings by taking the most basic and homogeneous tilings, i.e. having all squares with the same size and having no two squares of the same color share an edge. Furthermore, it helps if we fix one of the square colors as the "lead color" in order to eliminate any redundant reordering of fundamental domains (e.g. an "orange-green-blue" horizontal rectangle as a fundamental domain will clearly admit the same tiling as a "green-blue-orange" and a "blue-orange-green" horizontal rectangle). However, reorderings are not the same as permutations (e.g. an "orange-green-blue" horizontal rectangle as a fundamental domain is *not* the same as an "orange-blue-green"). This is an important proposition.

Proposition 3.11. *Consider a periodic tiling by two or three squares in which the squares are all the same size and the fundamental domain can be represented as a rectangle that is the union of two or three squares (i.e. two or three squares tile a rectangle). The tilings will be equivalent if the colors in the rectangle can be reordered modularly and* not *permuted. That is, given colors a and b, the fundamental domain a-b is equivalent to the fundamental domain b-a (where the fundamental domain is a rectangle (union of two squares), horizontal or vertical). Similarly for three squares, given colors a, b, and c, the fundamental domain a-b-c is equivalent to b-c-a is equivalent to c-a-b, and the fundamental domain a-c-b is equivalent to c-b-a is equivalent to b-a-c. We would refer to color a as the "lead color" in this case.*

Proof. If this observation isn't obvious at first, consider the band that a rectangular fundamental domain would make. If we look at the coloring encoded on this band, we will see *a-b-a-b-*. . . for two squares, and either *a-b-c-a-b-c-*. . . or *a-c-b-a-c-b-*. . . for three squares. But notice: the *a-b-a-b-*. . . is the same as a *b-a-b-a-*. . . pattern, the *a-b-c-a-b-c-*. . . is the same as the *b-c-a-b-c-a-*. .

. and the c-a-b-c-a-b-. . . pattern, and the a-c-b-a-c-b-. . . is the same as the c-b-a-c-b-a-. . . and the b-a-c-b-a-c-. . . pattern; it simply depends on where we begin our encoding. However, the encoding is infinitely long since we are in $\mathbb{R}^2$, so for every pattern there is another continuance of that pattern before and after it. This makes the choosing of the beginning color arbitrary, for the same pattern will yield the same tiling. Therefore we can just deem color a the "lead color" in order to encompass the various possible reorderings. □

This eliminates the possibility of redundant tilings in a moduli space, and it allows us to connect different points of the moduli space that we might have overlooked. This becomes evident for the three-square moduli space.

With this proposition, we can now look at the four base tilings for the three-square case. Remember we will not include any redundant fundamental domains.

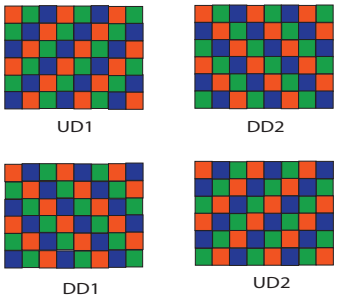


Figure 44: The four base tilings for tilings by three squares, named UD1, DD1, UD2, and DD2

Note we have labeled the base cases "UD1", "DD1", "UD2", "DD2" for ease of identification. Right away, we can see that we can deform DD1 to UD1 by a downward shift by the tall rectangular fundamental domain, and we can deform DD2 to UD2 by a downward shift by the tall rectangular fundamental domain.

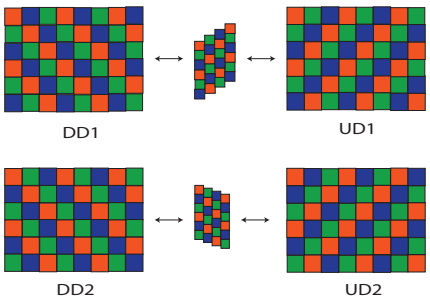


Figure 45: Deforming the base tilings by a downward shift of the tall rectangular fundamental domain

Note we also can deform DD2 to UD1 by a right shift by the wide rectangular fundamental domain, and we can deform DD1 to UD2 by a right shift by the wide rectangular fundamental domain.

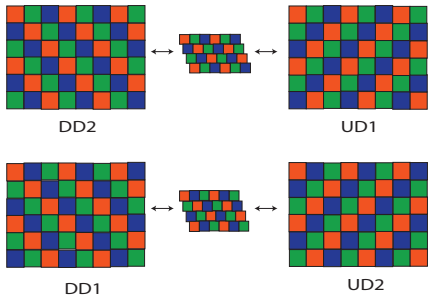


Figure 46: Deforming the base tilings by a right shift of the wide rectangular fundamental domain

But now note we can connect these four base tilings into a nice square figure, forming a more interesting pattern than the simple point we had as our base case in the one and two-square cases.

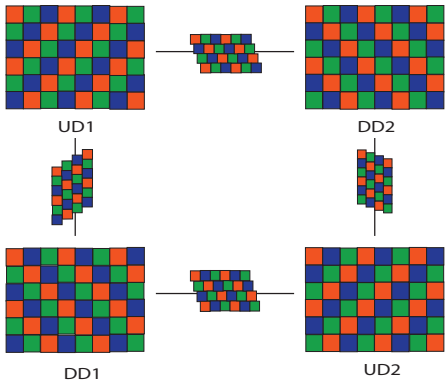


Figure 47: The four base tilings and the deformations that connect them

There is more shifting we can do, however. Let us return to the first case of deforming DD1 to UD1 by a downward shift by the tall rectangular fundamental domain. We can continue the deformation after reaching UD1, still shifting downward, until we reach a new line-up. In this tiling, which we will name "V1", we have a horizontal band of orange above a horizontal band of green above a horizontal band of blue above a horizontal band of orange again, and the pattern continues. We could continue the downward shift deformation we were originally doing, and this would complete the loop and bring us back to DD1, or we could also stay on V1 and shift the bands horizontally against each other like we did in the two-square case.

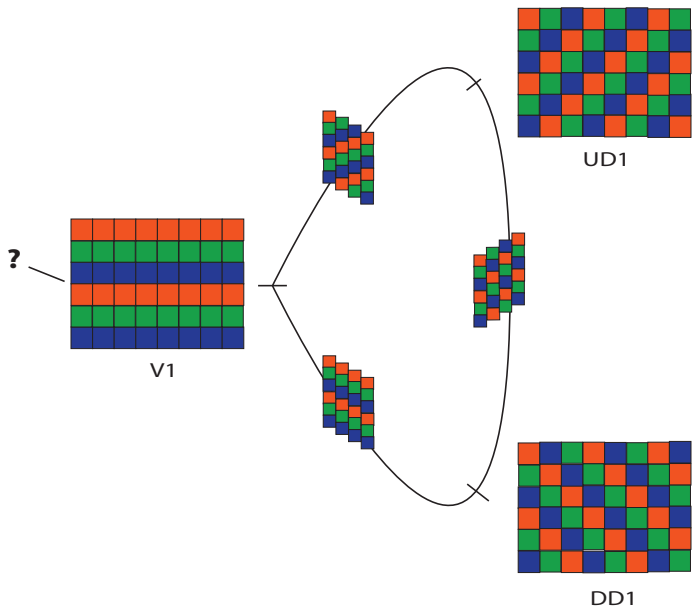


Figure 48: Continuing the downward shift deformation to get to the new tiling V1 and to complete a moduli space

As one may have guessed, we have free variables with these horizontal band shifts as in the two-square case, but here we have one more free variable than in the two-square case. We can fix the orange band (thereby fixing one of the translational vectors), then we can shift the green band independently, we can shift the blue band independently, and we can shift the next orange band independently, making the second translational vector free. This gives us three variables for the deformation space of the V1 tiling, which means we can represent this moduli space with a three-dimensional torus, *not* a two-dimensional torus or a loop.

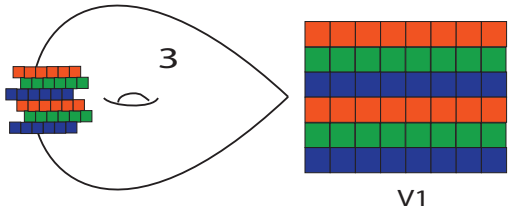


Figure 49: The 3-D toroidal moduli space on which V1 lies

This adds a new type of moduli space to our tiling system, and as one may guess it does not happen one time, but four times. We can continue the defor-

mation of DD2 to UD2 to get the tiling "V2", we can continue the deformation of DD2 to UD1 to get "H1", and we can continue the deformation of DD1 to UD2 to get the tiling "H2". With each of these, we can continue on the original deformation path to complete the loop, or we can branch off from V2, H1, and H2 into three-dimensional tori. We then get a very interesting and structured moduli space combining all of these spaces.

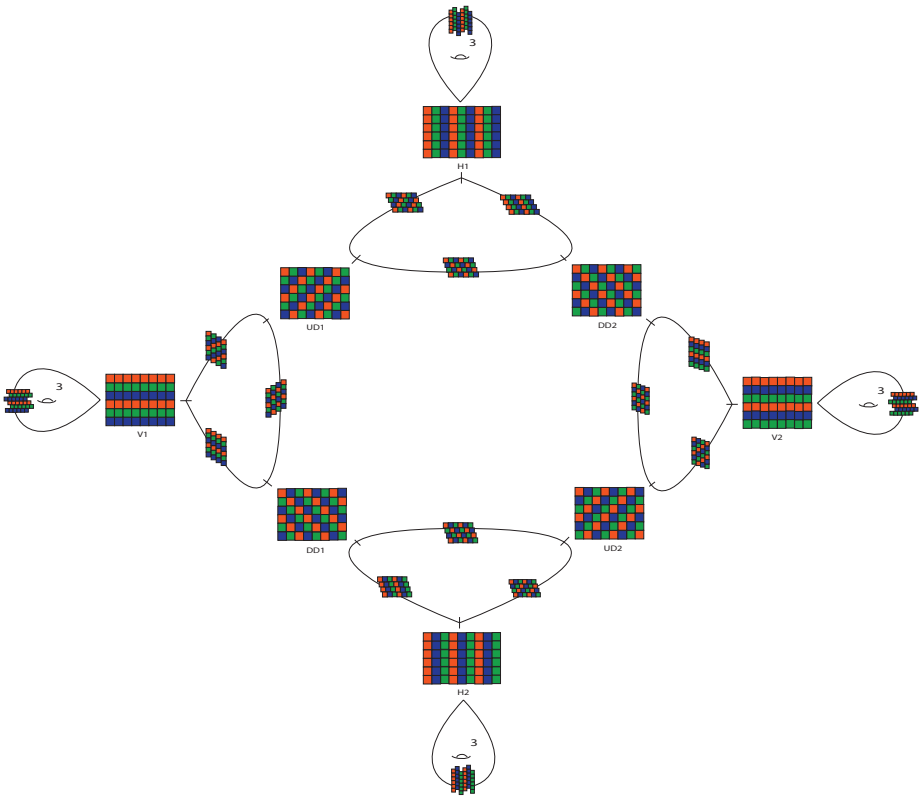


Figure 50: The main structure of the three square tiling system, which includes the four base tilings and the four tilings with bands (V1, H1, V2, H2)

This is only the beginning of the three-square tiling system, however. We still have yet to deform by independent scaling, which we will now do. As before with the shifts, we will have to look at the two main fundamental domains for each of the four base tilings, the vertical rectangle and the horizontal rectangle.

To find the new deformation spaces, we want to look at scaling one or two of the squares in the fundamental domain; that brings up the possibilities of growing one square, shrinking one square, growing two squares, and shrinking two squares. (The two-square scaling must be done uniformly with both of the squares, otherwise we will get a fundamental domain that doesn't tile. The same goes for three-square scaling; they must all be scaled uniformly. But that

is the trivial uniform scaling deformation that we cannot use, making the above four options the only four.) There is a way to reduce these four cases to two cases, however, making our hunt for new deformation spaces much easier; we just have to look at the cases relativistically.

We want to say that growing one square is the same deformation as shrinking two squares and that growing two squares is the same as shrinking one. We can show this is true; it is merely a difference of perspective. We will do an example that will be universal for all of the possible fundamental domains.

Let us grow the orange square, but look at the other two squares from the point of view of the orange square. The orange square may not realize he is growing, so to him, it looks like the blue and the green squares are shrinking, i.e. we are deforming by shrinking two squares at the same time we are deforming by growing one square. Since we do not account for uniform scaling, we can conclude that the two deformations are equivalent. Similarly, if we shrink the orange square and observe from his point of view, it looks like the green and the blue squares are getting bigger. That is, we are deforming by growing two squares at the same time we are deforming by shrinking one. Again, since we do not account for uniform scaling, we can conclude that the two deformations are equivalent. (Note that another way to make this evident is to fix the area of the three squares to add up to 1.)

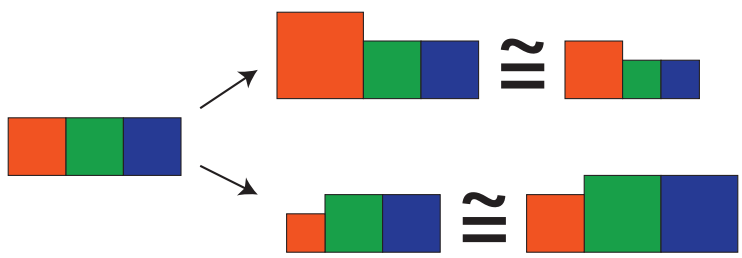


Figure 51: A diagram of the "relativistic" proof detailed above; note the cases are equivalent modulo scaling

This simplification means that we only have to worry about growing or shrinking a single square in the four base tilings. Unfortunately, this allows for twenty-four possible new deformation spaces, but, luckily, we can reduce this number.

The first and simpler deformation is shrinking one of the squares. This will lead us to a limit point and into the two-square moduli space. But first, recall Proposition 3.10: that color reordering yields the same tiling for colors a and b. From this, we can determine the equivalence of multiple tilings. As long as we have a "gridlock" tiling (like in the base cases, where everything is lined up), rectangular fundamental domains of the same orientation (either vertical or horizontal), and the same colors (e.g. blue followed by green and green followed by blue), then the tilings will be the same. The picture below gives an example:

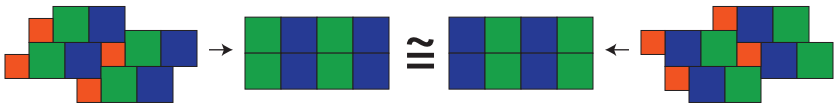


Figure 52: Note how different tilings can reach the same limit point

It would be helpful to define the term **gridlock** for ease of notation.

Definition 3.12. A periodic tiling by squares is called **gridlocked** if all the squares are lined up such that each vertex is touched by four squares. Necessarily, the squares are all of the same size.

Examples of gridlock tilings we have seen include the checkerboard tiling, the vertical band tiling, and the horizontal band tiling from the two-square case. We also have UD1, UD2, DD1, DD2, H1, H2, V1, and V2 from the three-square case. We know there are only those three gridlock tilings for the two-square case (since we know the entire moduli space for two squares), and we will see later that those eight tilings for the three-square case are the only gridlock tilings as well.

Note that at the limit point, the tiling will always be gridlocked (this will be explained combinatorially in the next section). Therefore the tiling will either be the checkerboard type, the vertical band type, or the horizontal band type. However, the previous example is indicative of the type of deformation to two squares we will always have, so the vertical or horizontal band tilings will always come from deforming the four base cases into two squares. Always remember that each point in the moduli space is unique and that the same tilings will have to occur at the same point, so we will have many instances where these limit points connect deformation spaces. We will see soon that this moduli space becomes too complicated and entangling to be accurately represented in a two-dimensional drawing, but we will try to make the diagrams as meaningful and readable as possible.

Let us first take UD1 with its vertical rectangular fundamental domain and look at shrinking the blue square, say to the left side to start. Notice that right away we get a perfect "L"-shaped fundamental domain. But notice DD1 has the same vertical rectangular fundamental domain, and that if we shrink the blue square there to the left side we get the same perfect "L"-shaped fundamental domain. Yet we know by Proposition 3.5 that if we have the same perfect "L"-shaped fundamental domain, we have the same tiling; therefore these two deformations must be represented by a single point in the moduli space. We call this a **shared deformation space** or **shared space** for short.

Definition 3.13. In a tiling system, a **shared deformation space** (or **shared space** for short), is a deformation space which is formed from similarly deforming two different tilings that share a common fundamental domain. The deformations of these two base tilings create the exact same deformed tiling, and the two base cases essentially "share" this deformation space.

This means for our moduli space, UD1 and DD1 must be infinitesimally close but not touching, yet they must be close enough to be connected by a point. In fact, they must be close enough to be connected by three points that do not touch (one for shrinking the blue square, one for shrinking the green square, and one for shrinking the orange square). Then from each of these points we can form a one-dimensional deformation space like we did with the "L"-shaped tiles for two squares. All three loops will reach a limit point into two squares, and the tilings will be horizontal band tilings.

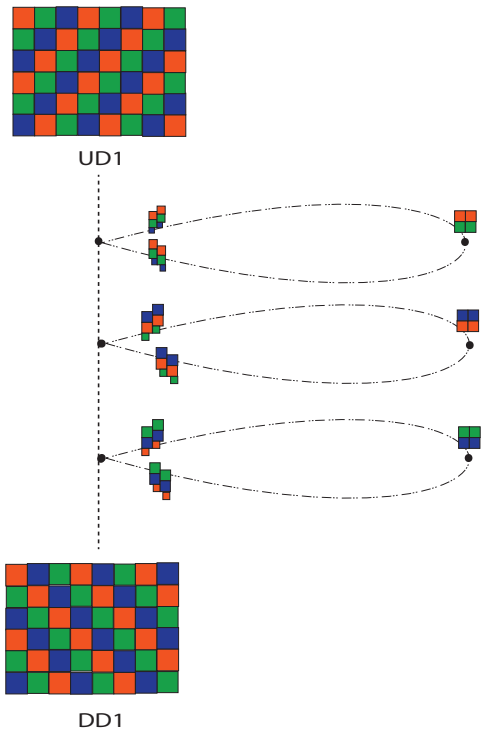


Figure 53: The limit points resulting from shrinking one square in the UD1 and DD1 tilings; note the short dashed lines to indicate a shared space

We can also do the same for the vertical fundamental domains of UD2 and DD2 since they are the same. Shrinking a square in UD2 yields the same tiling as shrinking the same color square in DD2, so, as in the above case, UD2 and DD2 must be infinitesimally close but not touching, and there must be three points connecting them. We can then branch off and form one-dimensional deformation spaces like before, and each loop will reach a limit point into two squares that yield horizontal band tilings.

But we just said a little while ago that horizontal or vertical band tilings composed of the same colors were identical. That means each pair of loops that has eliminated the same color must touch at the limit point, so we have six

loops but only three limit points.

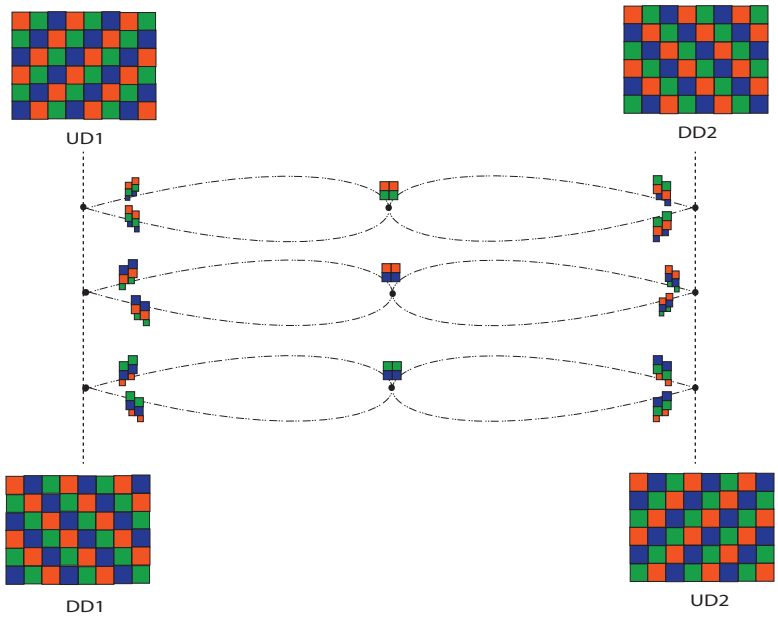


Figure 54: The equivalent limit points between the two pairs of base tilings sharing the same vertical fundamental domain

Now we can do the same process for the horizontal fundamental domains of the four base tilings. Note that UD1 and DD2 share the same horizontal fundamental domain, and DD1 and UD2 share the same one. This means, like above, UD1 and DD2 must be infinitesimally close and be connected by three points that have one-dimensional deformation spaces coming from them, and DD1 and UD2 must be infinitesimally close and be connected by three points that have one-dimensional deformation spaces coming from them. We get vertical band tilings at each limit point in this case, so like before, each pair of loops that has eliminated the same color must touch at the limit point, giving six loops but only three limit points.

The moduli space now gets very hard to draw effectively on a flat, two-dimensional surface, so we will use a different style line for the loops resulting from the vertical fundamental domains than for the loops resulting from the horizontal fundamental domains. Below is a preview of the drawing so that it will make sense later when put into the entire moduli space.

Recall that in the two-square moduli space, horizontal and vertical band tilings of the same color would eventually deform into each other. This means that out of our six limit points, the three pairs that have the same colors will connect in the two-square moduli space. We won't draw any lower square moduli spaces since the drawing is already complicated enough, but it is important to notice this fact.

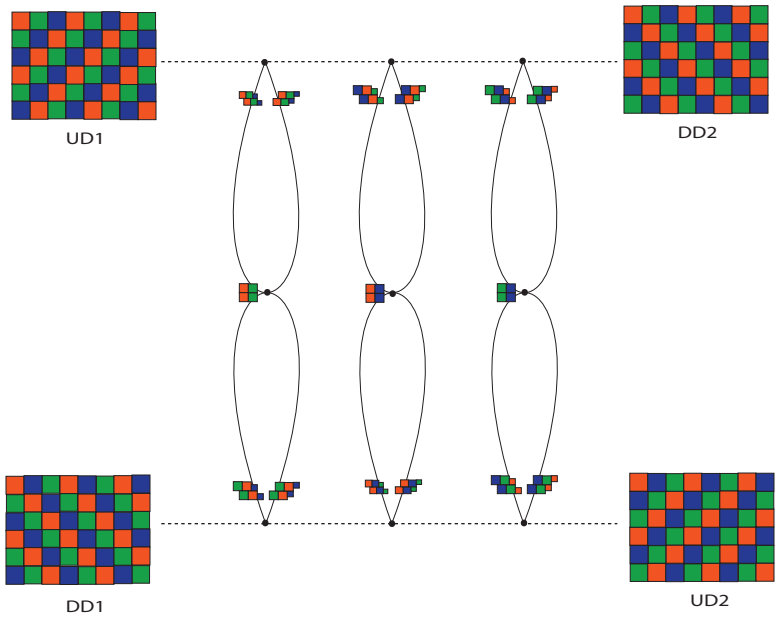


Figure 55: The equivalent limit points between the two pairs of base tilings sharing the same horizontal fundamental domain

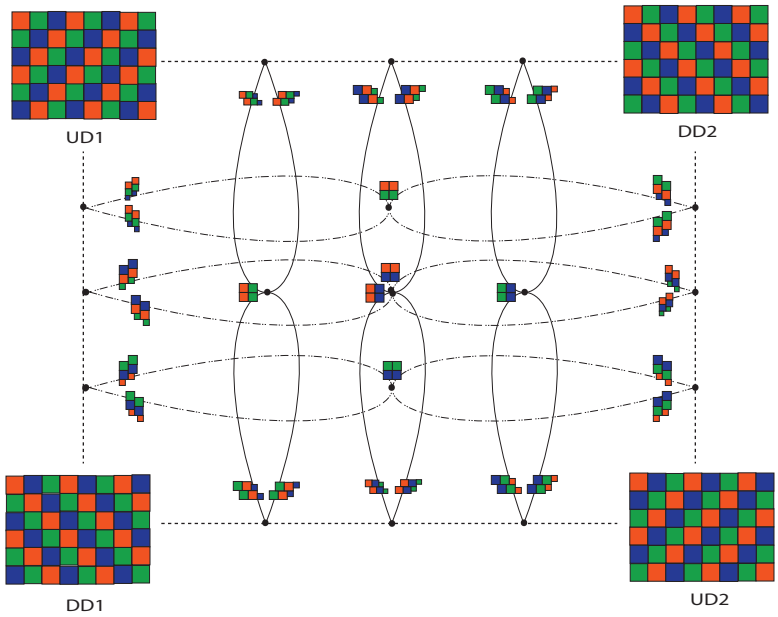


Figure 56: All one square shrinking deformations resulting from the four base cases

Now that we have the shrinking square case settled, we can look at what happens when we grow a square. First, let us look at growing the vertical fundamental domain of UD1 and DD1. Since they share the same vertical fundamental domain and any deformation by growing a square will turn the fundamental domain into a perfect "L"-shape, we again know by Proposition 3.5 that the two deformations produce the same tilings for each color square that is being grown. Therefore we have another three points connecting UD1 and DD1, and we can look at the deformation spaces for each.

Let us look at growing the orange square. Growing the other cases produce similar results, so this case will be indicative of the other two. First, let the orange square grow to the right, such that its left edge still matches up with the left edges of the green and blue squares. There will come a point in this growth where the orange square's side length is equal to the combined length of the side lengths of the blue and green square (which are of equal size). Note that this "L"-shaped tiling is one that is not perfect, for there exists another fundamental domain for this tiling (namely the rectangular one).

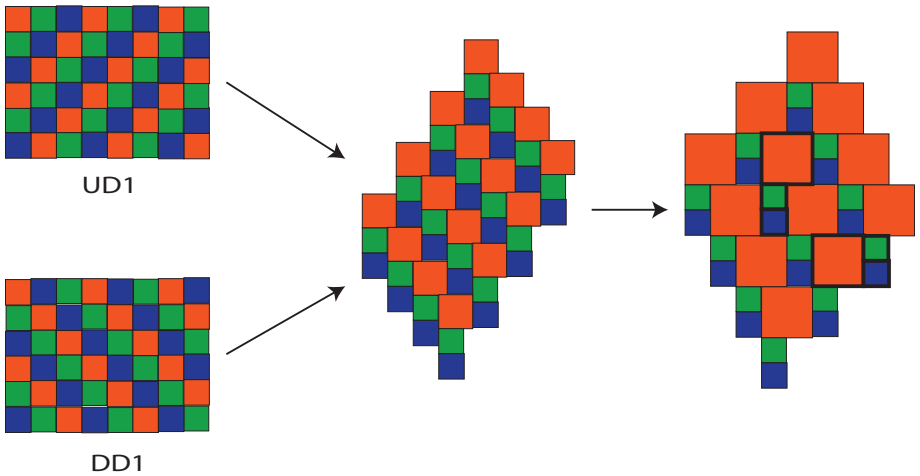


Figure 57: Growing the orange square from UD1 and DD1, making note of the particular tiling in which the three squares form a rectangular fundamental domain

There are two deformations we could take from this point. One involves shifting and the other involves a continuation of the growing we have done. We will come back to the shift and will continue on with the growing.

Recall growing one square is the same as shrinking two, so if we continue growing the orange one, we are constantly shrinking the green and the blue one (which are the same size). This means that this deformation will end in a limit point to one square, which is our first three-to-one square limit point.

Now we can grow the other two squares back from the other side of the orange one to complete the deformation space. We will go back through the

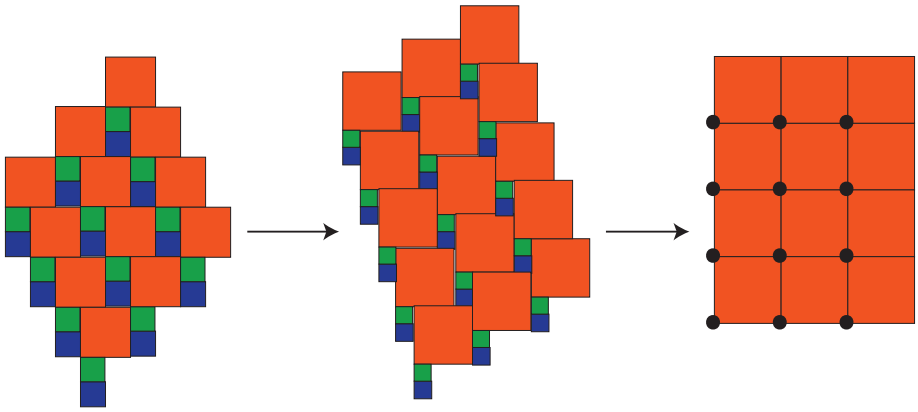


Figure 58: Growing the orange square and reaching a three-to-one square limit point

tiling with the non-perfect "L"-shaped fundamental domain and return to our starting place at UD1 and DD1. Note that the non-perfect "L"-shaped tiling is slightly different this time. As one may guess, the two non-perfect "L"-shaped tilings will be connected.

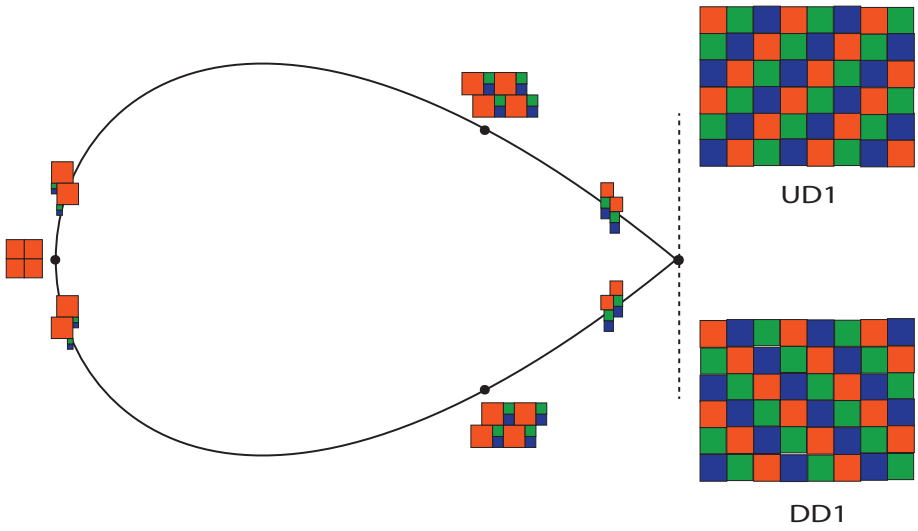


Figure 59: The entire single loop moduli space of growing one square

Let us return to the non-perfect "L"-shapes. Recall that we can use the shift as a deformation by shifting the horizontal bands of rectangles. If we look at the first of these tilings that we encountered, we can shift to the right or left (both will be the same but the involve moving around the moduli space in a

different direction). Let us shift to the right.

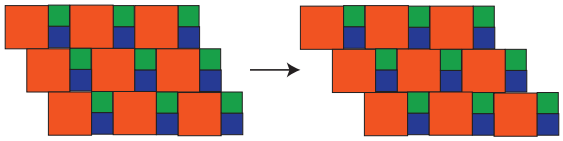


Figure 60: Shifting the rectangular fundamental domain to the right

There will come a point where the rectangular tiles are all aligned; the orange squares line up in vertical bands, and the green and blue squares line up in an alternating vertical band. This is a special case that we will discuss after we complete the shifting deformation which we are now on.

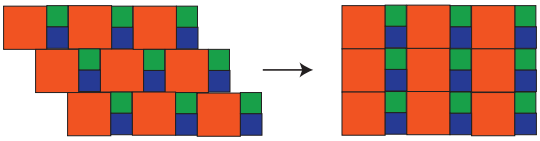


Figure 61: Further shifting the rectangular fundamental domain to the right, reaching a unique point in which we have vertical bands of squares

If we continue with the right shift, we will eventually reach the second non-perfect "L"-shape tiling we discovered in the above square-growth deformation. This is the connection between the two that was mentioned earlier. Now we can simply continue this right shift, and we will return back to the tiling at which we started.

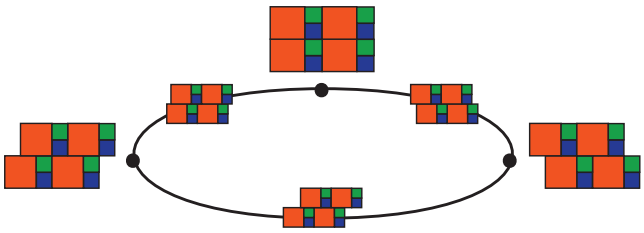


Figure 62: The moduli space of shifting the rectangular fundamental domain

Now we can look at the special rectangular tiling we just mentioned, where the orange squares are lined up in vertical bands and the green and blue squares line up in an alternating vertical band. This is the final barrier that stands between us and completing the three-square tiling system. Notice right away that it is possible to shift the two vertical bands independently, almost exactly like we did in the two-square case with vertical bands. As before, this will form a two-dimensional deformation space that is represented by a 2-D torus.

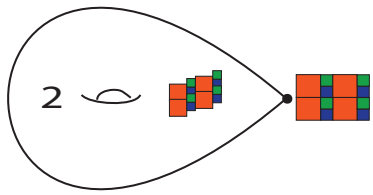


Figure 63: The special rectangular fundamental domain with the 2-D torus to which it is attached; note the example of a tiling that lies inside this torus

There are some special cases that arise in this torus, however. Notice if we only shift the alternating green and blue vertical band one square length up, we have a different rectangular fundamental domain than we started with. We could then perform a horizontal rectangle shift to the non-perfect "L"-shape, then shrink the orange square, and we would end up in the UD2 and DD2 base cases instead of UD1 and DD1. This tells us that the 2-D torus overlaps with the opposite one, connecting the two opposing sides in a new way. We will illustrate this important observation soon, but there are still two more special cases on the torus that we must address. Both of the cases involve breaking the horizontal bands and lining the vertical ones up in a certain way. Below are the two cases; we will call the first Case 1 and the second Case 2:

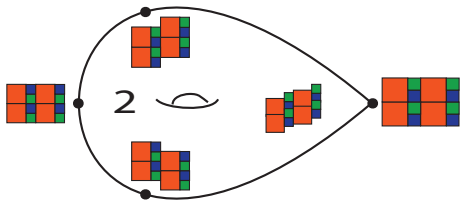


Figure 64: The two special cases on the 2-D torus; Case 1 is located near the top of the figure, while Case 2 is located near the bottom

Notice how in either case, two different rectangular fundamental domains arise depending on if you look at the orange square being on the left or right of the other two.

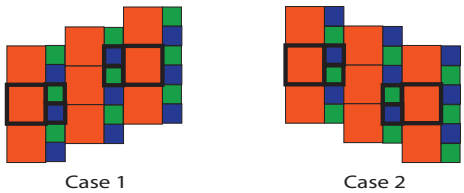


Figure 65: The Case 1 tiling and the Case 2 tiling; note the two different fundamental domains in each one

These two special alignments allow us to perform one final deformation that will let us complete the tiling system for three squares. There exists a tiling by three squares in which all squares are different sizes, but there are a few restrictions. The first is that the sum of the side lengths of the two smaller squares must be the same as the entire side length of the larger square. The next (which essentially follows from the first) is that the tile must be a perfect "L"-shape.

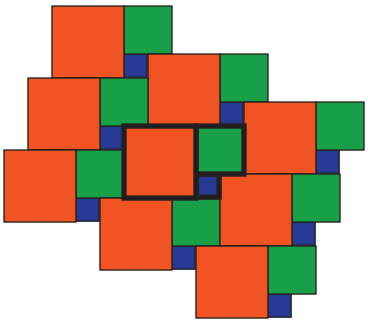


Figure 66: An example of a tiling by three squares in which all three squares are different sizes

We will see in the next section how the first restriction arises mathematically, but for now we will just perform the deformation. Let us take Case 1. First we will grow the green square while shrinking the blue one, and we will take the fundamental domain that places the orange square on the left of the other two.

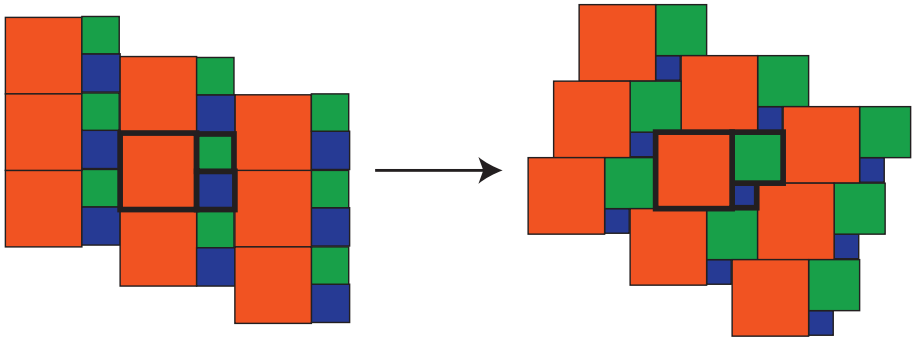


Figure 67: The beginning of the deformation starting from Case 1 in which we grow the green square while shrinking the blue square

If we continue this deformation of growing green while shrinking blue, we will eventually reach a limit point with the blue square disappearing to a point, leaving the orange and the green square. However, notice the pattern through which we reach the limit point:

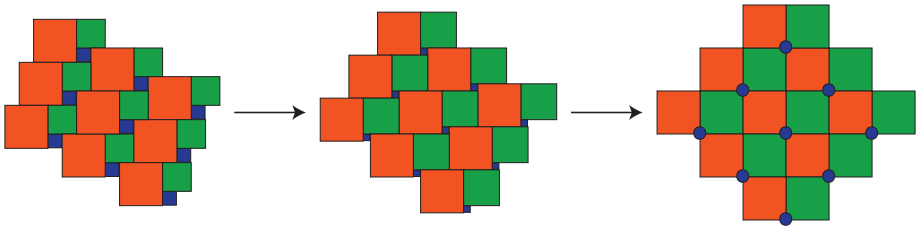


Figure 68: Continuing the above deformation, this time shrinking the blue square to a point

Notice that the squares do not line up in vertical bands like they did when we shrunk one square. Instead, they form the checkerboard pattern that is at the center of the two-square tiling system. This means that we have three points in the three-square tiling system that reduce to three different cases in the two-square system. We can think of it as an "underground network" for the two-square moduli spaces; they act as shortcuts to different parts of the three-square moduli space, creating a much more interesting link between the three-square and the two-square system than we had with the two-square and the one-square system.

Now we must continue the deformation. Since we already shrunk the blue square to a point, we must bring it back on the other side like we have done in the past. As before, we do this by taking a vertical mirror image of what we just did. Continuing to grow the blue square and shrink the green, we end up at the Case 2 tiling, not back at our original starting point.

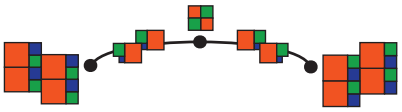


Figure 69: The half loop that we get from the deformation in which the green square is larger than the blue square

We can do the same thing again for the Case 1 tiling if we grow the blue square while shrinking the green. This time at the limit point we will have an orange and blue square checkerboard pattern. Continuing the deformation, we again end up at the Case 2 tiling, not at our original starting point.

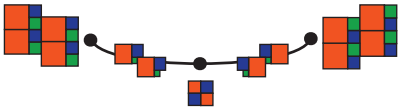


Figure 70: The half loop that we get from the deformation in which the blue square is larger than the green square

But now note that we have already taken care of the deformations that come from the Case 2 tiling. This allows us to complete the internal structure of the 2-D toroidal moduli space.

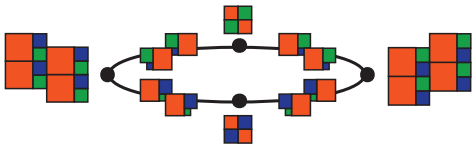


Figure 71: The internal loop of the 2-D torus

Notice the two checkerboard tilings at the two reduction points. We will see that these exact same tilings come up multiple times throughout the entire tiling system. For the ease of illustrating the moduli spaces on a two-dimensional page, we will not connect them but instead mark each point by a two-color star. This two-color star will stand for the checkerboard pattern with two squares that have the star's colors. Any stars that are the same are the same point. This will help us in decreasing the overwhelming crowdedness and complexity of the three-square system, and it will better allow us to see how the spaces connect. We will do the same for the one-square case. Instead of connecting all one-square limit points of the same color, we will just mark those points with a solid-colored star of the square's color.

Figure 72: The various stars we will use in the three-square moduli space; the two-color stars stand for checkerboard tilings while the solid-colored stars stand for gridlock tilings by the same-colored square

Now we can combine this information to get a clear picture of what is going on in the 2-D torus that we have looked at so deeply:

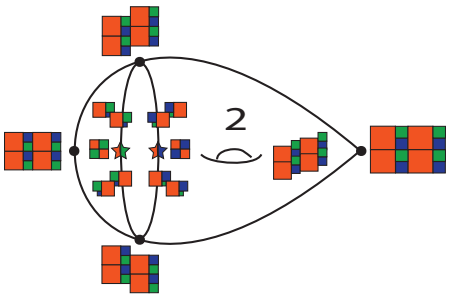


Figure 73: The 2-D torus moduli space along with its internal structure

It is important to remember that the 2-D torus is shared between opposite sides (e.g. the UD1/DD1 side and the UD2/DD2 side). When illustrating the entire moduli space, we will not connect the two since it would make our picture impossible to read. Instead, we will indicate a copy of the 2-D torus with a dashed line. We will still mark the two tilings that join the opposing sides, but it will be made clear that there is only one of that type of 2-D torus.

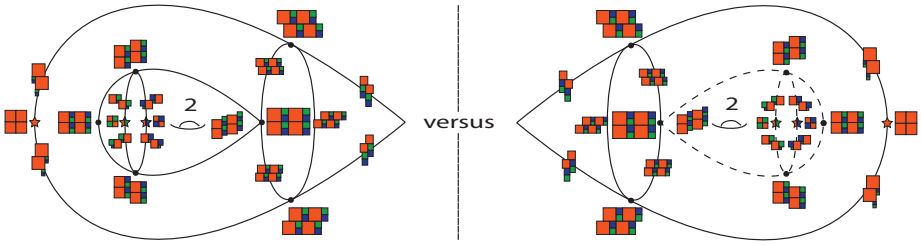


Figure 74: An example of the 2-D torus (and the surrounding moduli space) on both sides of the entire three-square moduli space; note the dashed lines on the torus to the right to indicate that it is repeated

Now that we have the final piece figured out, we can repeat the deformation for each color in the fundamental domain. Notice that this repetition will create many more two-color stars, indicating the same two-square checkerboard patterns at limit points. For the UD1 and DD1 starting case, we still must grow the green square, and we must grow the blue square. Completing those two tasks will yield the picture depicted below.

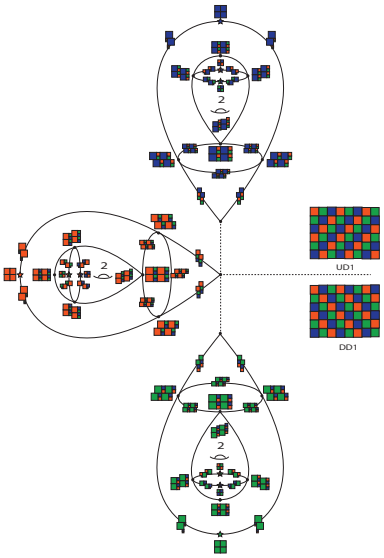


Figure 75: The complete one-square growth for the UD1 and DD1 starting case

We can also complete the opposing UD2/DD2 starting case since the two sides share tori. They also share three-to-one square limit points, so we can now make use of the single colored stars. (As one may surmise, the single colored stars will occur on each side.) They share three-to-two limit points as well (and so do all of the other sides).

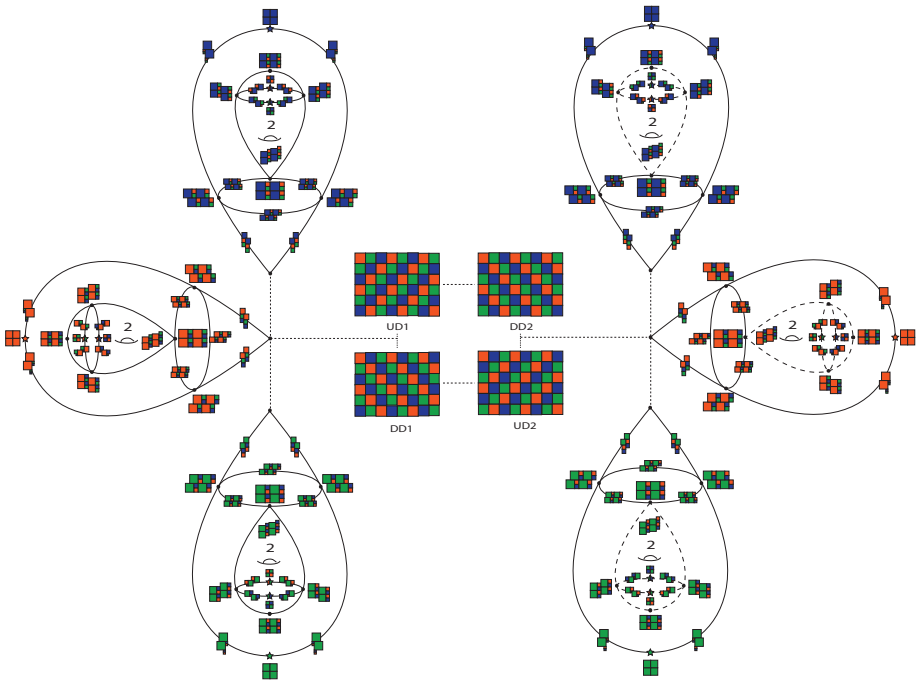


Figure 76: The complete one-square growth for the UD1 and DD1 starting case along with the UD2 and DD2 starting case; note the dashed lines for the 2-D tori on the right side of the figure to indicate they are repeated

Now we only have the horizontal fundamental domains from the four base cases to worry about. However, it is all the same idea, only with rectangles that are tall instead of wide. We apply everything we have discussed with this "one growing square" deformation, and we finally lay eyes upon our long sought after prize: the complete moduli space for the three-square tiling system.

The end result is a beautiful topological object. Though it is hard to effectively represent on a two-dimensional surface, the moduli spaces and the connections are evident. The symmetry that this tiling system creates is unbelievable. After examining/admiring the three-square system for a little while, we now must focus in on the most important question: why do these tiling systems look the way they do? What is the mathematics behind it all?

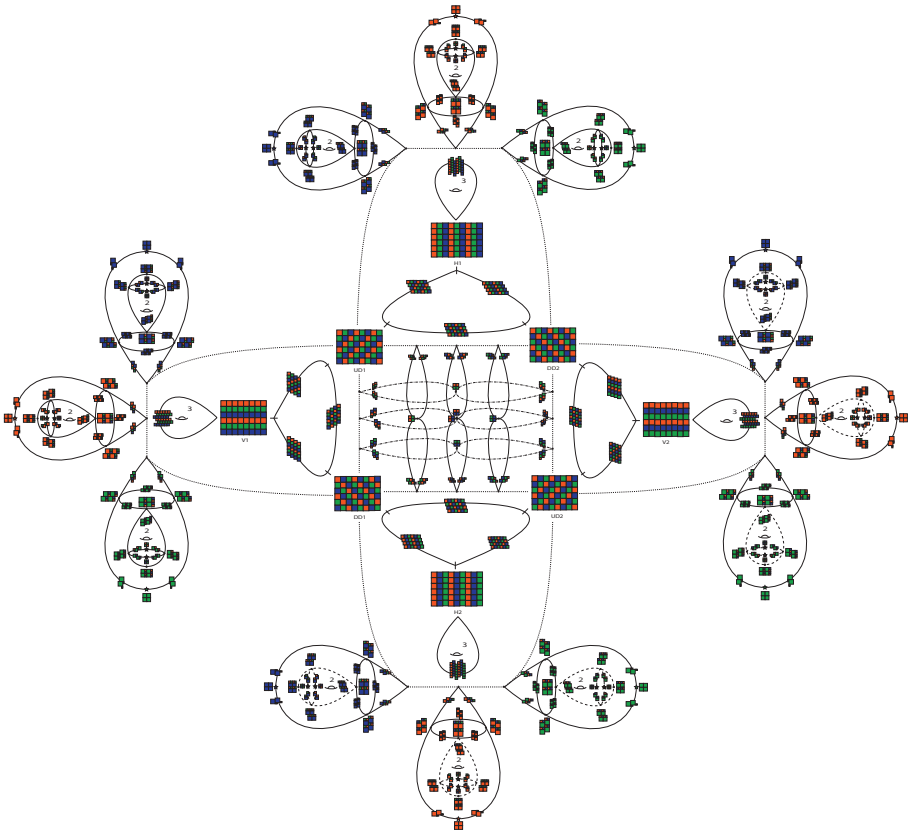


Figure 77: The entire three-square moduli space

3.3 Combinatorial Structure

We now know what each tiling by one, two, and three squares looks like, and we know how their deformation spaces connect, but it is important to understand *why* they look the way they do and deform the way they do. To do this, we use what is called the **dual graph** of a tiling.

Definition 3.14. The **dual graph** of a tiling is a graph we place over the entire tiling to look at how each individual square interacts with its neighboring squares. First, we draw a vertex in the middle of each square and label the vertex with the size of the square (say a, b, or c, since we are only concerned with three different squares in our case). Then, given a square, we look at what is touching its top and bottom edges and what is touching its left and right edges. For each square touching our square's top or bottom edges, we draw a colored line (say red) to that neighboring square to indicate a vertical edge, and for each square touching our square's left or right edges, we draw a different colored line (say yellow) to that neighboring square to indicate a horizontal edge. Note that squares that touch at a vertex do not qualify as neighbors. Also note that the vertical and horizontal edges naturally make the graph a directed graph. We repeat this process for all squares in the tiling to obtain the dual graph.

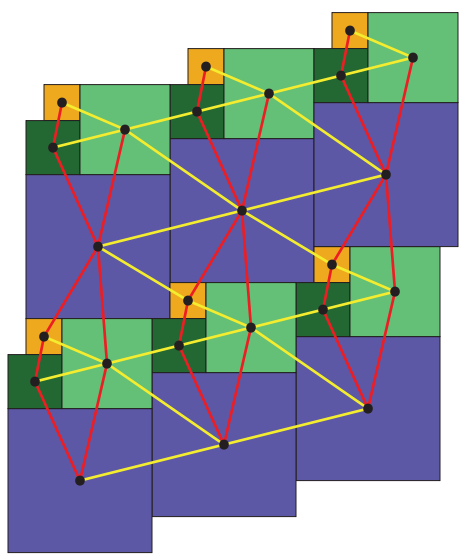


Figure 78: An example of a dual graph

So why introduce this dual graph? One of the things that the dual graph can tell us is how we can deform a tiling and how many different deformation spaces intersect the tiling. Let us first look at deforming the tiling, then we will discuss the deformation space intersections.

What we want to do now is attach data to the dual graph. We have already labeled the vertices by square size, so the next thing we do is "reduce" the dual graph to avoid any redundant edges. Let us think of these edges as vectors, directing us from one vertex to another. The only constraint we will put on the vectors is that they all point in the positive x-direction. If this does not determine the y-direction, then we will require the vector to point in the positive y-direction as well. Now for the reduction: we need to look at a single fundamental domain and the pieces around it (dual graph still intact). Now we must label each edge as a vector according to our above constraints. For example, we can have a vector going from square a to square a and one from square a to square b (where a and b simultaneously denote the square's size and name). Note that if we have a square a to square a vector, all other vectors that direct square a to a translate of itself will be equivalent, i.e. there is only one path from square a to square a. The a to b vector however, will not be the same as the b to a vector as long as we follow our vector constraints above. We continue on like this, labeling the dual graph edges as vectors and making sure we do not count the same vector twice (modulo translation). The process is harder to describe in words than to show, so below is a generic example.

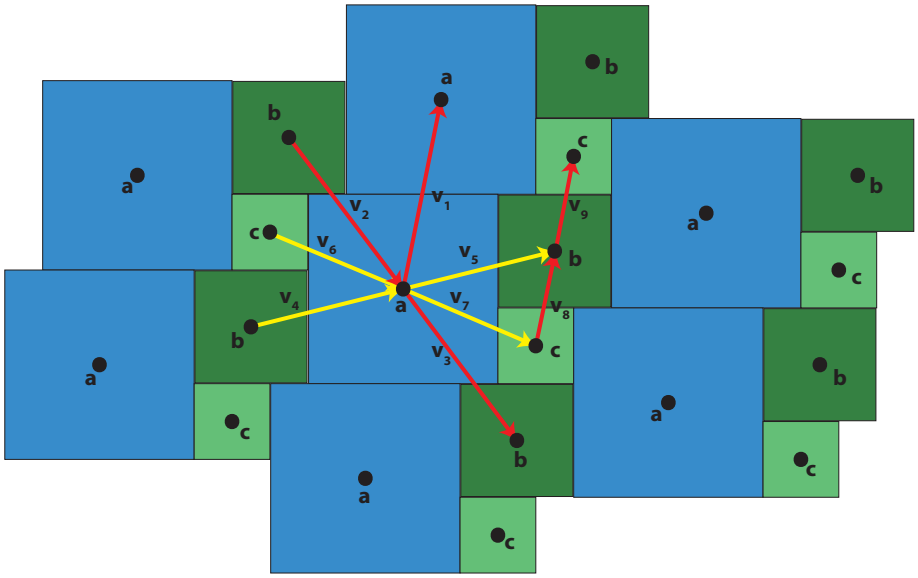


Figure 79: An example of a tiling by three squares with a "reduced" dual graph, in which all three squares are labeled and all nine vectors are labeled

The next step is to explicitly define each of the vectors we have just labeled. Let us continue with the above example, as it helps us understand the process better. Looking at v_1, we will notice that in the vertical direction, it moves from the vertex of square a to another vertex of square a, so it travels $2a$ vertically. (We will say the size a, b, or c is half of the square's side length to avoid

unnecessary fractions in our vectors.) We can see that the vector moves some unidentified amount in the horizontal direction, so we call this value x_1 and get $v_1 = (x_1, 2a)$.

For v_2, note that it moves by a variable amount x_2 horizontally and moves downward $a + b$, so we get $v_2 = (x_2, -(a + b)) = (x_2, -a - b)$

One could easily identify the remaining vectors, calling unknown horizontal movements x_i and unknown vertical movements y_i for vector v_i, and we find the following values for each vector: $v_3 = (x_3, -a - b)$, $v_4 = (a + b, y_4)$, $v_5 = (a+b, y_5)$, $v_6 = (a+c, -y_6)$, $v_7 = (a+c, -y_7)$, $v_8 = (x_8, b+c)$, and $v_9 = (x_9, b+c)$.

Now, how can we find these x_i and y_i values? The answer lies in the loops composed of three (or sometimes four) vectors that the dual graph holds. With each of these loops, we can create a vector equation that equals zero, and with all of these equations, we have a system that allows us to solve for the unknown values. But, as before with the vectors, we cannot count the same loop twice. That is, loops composed of the same vectors located at the same places in the loop are equivalent (modulo translation). We will label the loops with Roman numerals, and we will continue on with our above example.

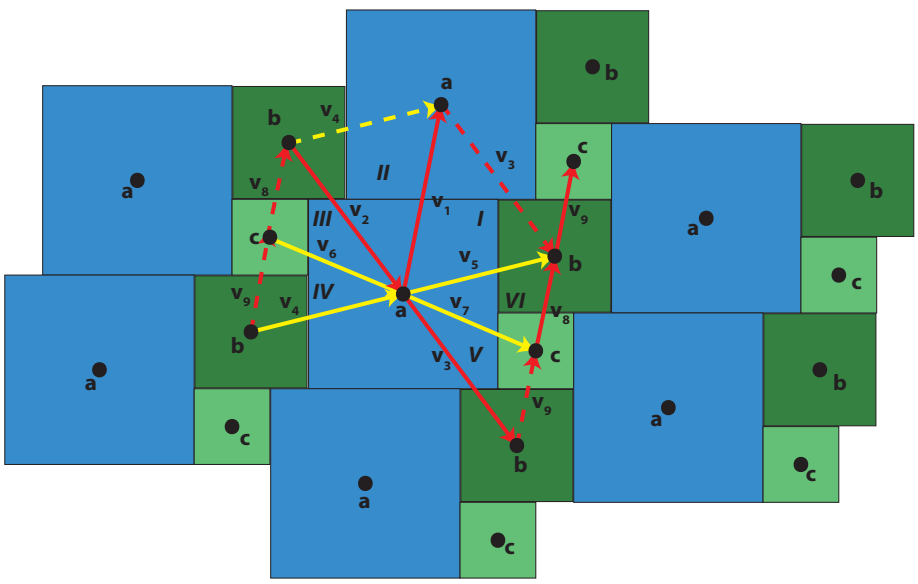


Figure 80: All of the necessary data attached to the tiling; note the dashed vectors indicate a vector that is copied by translation in order to complete a loop

In this case we find six loops, which we label with Roman numerals. The

six equations are:

$$
\begin{aligned}
I &: v_5 - v_3 - v_1 = 0 \\
II &: v_1 - v_4 + v_2 = 0 \\
III &: v_6 - v_2 - v_8 = 0 \\
IV &: v_4 - v_6 - v_9 = 0 \\
V &: v_9 - v_7 + v_3 = 0 \\
VI &: v_8 - v_5 + v_7 = 0
\end{aligned}
$$

Note that we can write any one equation in terms of the other five (e.g. $VI = -I - II - III - IV - V$), so the equations are linearly dependent. (In general, the equations are dependent because together they sum to zero; this means we can always write one equation in terms of the others.) This allows us to eliminate one equation and lets us not use any redundant information. We will toss out equation VI, but in cases with more squares, it would help to carefully choose which equation to throw out to make the linear algebra much easier to do.

For each equation, we get two separate equations: one for the horizontal movement, and one for the vertical movement. We will use the subscript 1 for horizontal movement and the subscript 2 for vertical movement for ease of notation. So now we have ten equations, and they are as follows:

$$
\begin{aligned}
I_1 &: a + b - x_3 - x_1 = 0 \\
I_2 &: y_5 + a + b - 2a = 0 \\
II_1 &: x_1 - a - b + x_2 = 0 \\
II_2 &: 2a - y_4 - a - b = 0 \\
III_1 &: a + c - x_2 - x_8 = 0 \\
III_2 &: -y_6 + a + b - b - c = 0 \\
IV_1 &: a + b - a - c - x_9 = 0 \\
IV_2 &: y_4 + y_6 - b - c = 0 \\
V_1 &: x_9 - a - c + x_3 = 0 \\
V_2 &: b + c + y_7 - a - b = 0
\end{aligned}
$$

Now all we do is solve each equation. What we want to do in this case is get each x_i and y_i in terms of a, b, and c, and then check for the constraints on a, b, and c. First we want to look at the equations where we only have a single x_i or a single y_i. This includes equations I_2, II_2, III_2, IV_1, and V_2, and we find that $y_5 = a - b$, $y_4 = a - b$, $y_6 = a - c$, $x_9 = b - c$, and $y_7 = a - c$, respectively, from each of these equations.

Next we want to find an equation where we know one of the x_i or y_i quantities in terms of a, b, and c, but don't know the other. Equation V_1 is the case for which we are looking, for we know the quantity x_9 and now can find the quantity x_3. We get $x_3 = a - b + 2c$.

Now we simply continue this process, looking for the equations with one known variable and one unknown variable. We can now find x_1 from I_1, then we can find x_2 from II_1, and finally x_8 from III_1. We find $x_1 = 2b - 2c$, $x_2 = a - b + 2c$, and $x_8 = b - c$.

Notice how we never used equation IV_2. This equation tells us the constraints on the square size variables a, b, and c. If we plug in y_4 and y_6 into IV_2, we find $2a - 2b - 2c = 0$, or $2a = 2b + 2c$. Recall that we used a, b, and c as half of the side length for each square to avoid fractions in our equations. This means $2a$, $2b$, and $2c$ indicate the entire side length for each square. This final constraint tells us that the smaller squares b and c must have their side lengths always sum to the larger square a's side length.

Now, how does this relate to the deformation space? The last constraint tells us most of the information we need to know. Since we don't count scaling, the size a will be given, and we can only choose sizes b or c. But if we choose one, the other one will automatically be given by the size constraint above, so we only get to freely choose one of the sizes. Furthermore, since all other variables are in terms of a, b, and c, this free choice is the only free choice we get with this space of tilings. That means we only have one free variable with this set up, which means we have a one-dimensional deformation space and therefore a loop for the moduli space, all of which we predicted in the last section. The use of the dual graph and the named vectors and loops gives us the information we need on how we can deform a tiling, though in some cases this deformation space may be easier to recognize on some tilings than on others (the above case was pretty simple).

Before moving on, note how there were three squares, six loops, and nine vectors. This is no coincidence; it turns out this is directly related to the Euler characteristic for *tori*. (It makes sense that we look at tori since our fundamental domain repeats itself on all sides.)

Recall that the Euler characteristic $\chi = V - E + F$, where V denotes vertices, E denotes edges, and F denotes faces. For tori, we know that $\chi = 0$, always (see Matveev's entry in the Encyclopedia of Mathematics [2]). This means $E = V + F$. We already know V, the vertices, since each squares' center represents a vertex in the dual graph, so we have three of them. Since edges simply connect vertices, it makes sense that our nine vectors would be the edges that we want. Going one step further, faces connect edges, so our loop equations represent the faces in this case. (Normally faces would be two-dimensional objects, but in our case we created loops to represent each of these two-dimensional objects to allow us to use algebra to solve for the unknowns. Though the dimensions are not the same, the numbers are, and that is what we are concerned about.)

We can generalize this relationship, writing E and F in terms of V. Given a set of vertices V, we will have $3V$ edges connecting each of the vertices, and we will have $2V$ faces (or loops) connecting each of the edges. That means the original equation $E = V + F$ is transformed into $3V = V + 2V$, which is still true. This is the relationship that will always hold in the dual graph analysis of the tiling, so if one applies the analysis and finds that the squares, vectors, and loops don't make this formula true, there is a mistake. This formula turns

out to be very useful to ensure we are on the right track with trying to find the dimension of the deformation space.

Below are more examples in which we find the dimension of a deformation space using the dual graph. Note we can analyze cases in which there are more than three squares, and the procedure still works.

Example 3.15. We begin by putting a "reduced" dual graph on the tiling and labeling all squares, vectors, and loops.

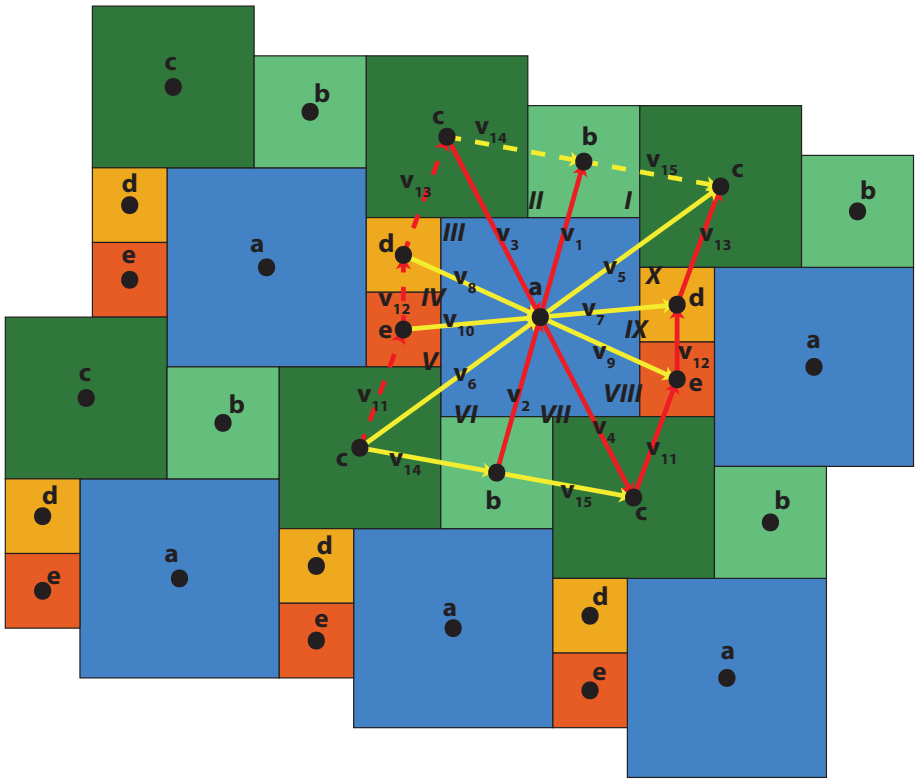


Figure 81: A tiling by five squares with all squares, vectors, and loops labeled

We next define each of the vectors in the picture.

$$v_1 = \begin{pmatrix} x_1 \\ a+b \end{pmatrix} \quad v_2 = \begin{pmatrix} x_2 \\ a+b \end{pmatrix} \quad v_3 = \begin{pmatrix} x_3 \\ -a-c \end{pmatrix} \quad v_4 = \begin{pmatrix} x_4 \\ -a-c \end{pmatrix}$$

$$v_5 = \begin{pmatrix} a+c \\ y_5 \end{pmatrix} \quad v_6 = \begin{pmatrix} a+c \\ y_6 \end{pmatrix} \quad v_7 = \begin{pmatrix} a+d \\ y_7 \end{pmatrix} \quad v_8 = \begin{pmatrix} a+d \\ y_8 \end{pmatrix}$$

$$v_9 = \begin{pmatrix} a+e \\ y_9 \end{pmatrix} \quad v_{10} = \begin{pmatrix} a+e \\ y_{10} \end{pmatrix} \quad v_{11} = \begin{pmatrix} x_{11} \\ c+e \end{pmatrix} \quad v_{12} = \begin{pmatrix} x_{12} \\ d+e \end{pmatrix}$$

$$v_{13} = \begin{pmatrix} x_{13} \\ c+d \end{pmatrix} \quad v_{14} = \begin{pmatrix} b+c \\ y_{14} \end{pmatrix} \quad v_{15} = \begin{pmatrix} b+c \\ y_{15} \end{pmatrix}$$

Now we write down the equation for each loop. Note we have ten equations in this example.

$$\begin{aligned}
&I : v_5 - v_{15} - v_1 = 0 \\
&II : v_1 - v_{14} + v_3 = 0 \\
&III : v_8 - v_3 - v_{13} = 0 \\
&IV : v_{10} - v_8 - v_{12} = 0 \\
&V : v_6 - v_{10} - v_{11} = 0 \\
&VI : v_{14} + v_2 - v_6 = 0 \\
&VII : v_{15} - v_4 - v_2 = 0 \\
&VIII : v_{11} - v_9 + v_4 = 0 \\
&IX : v_{12} - v_7 + v_9 = 0 \\
&X : v_{13} - v_5 + v_7 = 0
\end{aligned}$$

Since the equations are linearly dependent, we can toss one out (say equation X). Then we can split up the remaining equations into their x and y components, leaving us with eighteen equations in total.

$$\begin{aligned}
&I_1 : a + c - b - c - x_1 = 0 \\
&I_2 : y_5 - y_{15} - a - b = 0 \\
&II_1 : x_1 - b - c + x_3 = 0 \\
&II_2 : a + b - y_{14} - a - c = 0 \\
&III_1 : a + d - x_3 - x_{13} = 0 \\
&III_2 : y_8 + a + c - c - d = 0 \\
&IV_1 : a + e - a - d - x_{12} = 0 \\
&IV_2 : y_{10} - y_8 - d - e = 0 \\
&V_1 : a + c - a - e - x_{11} = 0 \\
&V_2 : y_6 - y_{10} - c - e = 0 \\
&VI_1 : b + c + x_2 - a - c = 0 \\
&VI_2 : y_{14} + a + b - y_6 = 0 \\
&VII_1 : b + c - x_4 - x_2 = 0 \\
&VII_2 : y_{15} + a + c - a - b = 0 \\
&VIII_1 : x_{11} - a - e + x_4 = 0 \\
&VIII_2 : c + e - y_9 - a - c = 0 \\
&IX_1 : x_{12} - a - d + a + e = 0 \\
&IX_2 : d + e - y_7 + y_9 = 0
\end{aligned}$$

In this example we want to find all x_i and y_i in terms of a, b, c, d, and e like we did in the previous example. The first set of equations we can do this for is I_1, II_2, III_2, IV_1, V_1, VI_1, VII_2, VIII_2, and IX_1. These equations reduce to the following:

$$\begin{aligned}
&I_1 \rightsquigarrow x_1 = a - b \\
&II_2 \rightsquigarrow y_{14} = b - c \\
&III_2 \rightsquigarrow y_8 = d - a \\
&IV_1 \rightsquigarrow x_{12} = e - d \\
&V_1 \rightsquigarrow x_{11} = c - d \\
&VI_1 \rightsquigarrow x_2 = a - b \\
&VII_2 \rightsquigarrow y_{15} = b - c \\
&VIII_2 \rightsquigarrow y_9 = d - a \\
&IX_1 \rightsquigarrow x_{12} = d - e
\end{aligned}$$

Note right away we can combine equations IV_1 and IX_1 to find:

$$x_{12} = d - e = e - d \Rightarrow \{x_{12} = 0 \text{ and } d = e\}.$$

Now we can combine what we found in the above nine equations with I_2, II_1, IV_2, VI_2, VII_1, and IX_2 to find six more x_i and y_i values.

$$\begin{aligned}
&I_2 \rightsquigarrow y_5 - b + c - a - b = 0 \Leftrightarrow y_5 = a + 2b - c \\
&II_1 \rightsquigarrow a - b - b - c + x_3 = 0 \Leftrightarrow x_3 = 2b + c - a \\
&IV_2 \rightsquigarrow y_{10} - d + a - d - d = 0 \Leftrightarrow y_{10} = 3d - a \\
&VI_2 \rightsquigarrow b - c + a + b - y_6 = 0 \Leftrightarrow y_6 = a + 2b - c \\
&VII_1 \rightsquigarrow b + c - x_4 - a + b = 0 \Leftrightarrow x_4 = 2b + c - a \\
&IX_2 \rightsquigarrow d + d - y_7 + d - a = 0 \Leftrightarrow y_7 = 3d - a
\end{aligned}$$

With the values obtained from these equations, we can write equations V_2 and VIII_1 using only a's, b's, c's, and d's.

$$\begin{aligned}
V_2 &\rightsquigarrow a + 2b - c - 3d + a - c - d = 0 \Leftrightarrow 2a + 2b = 2c + 4d \Leftrightarrow a + b = c + 2d \\
VIII_1 &\rightsquigarrow c - d - a - d + 2b + c - a = 0 \Leftrightarrow 2a - 2b = 2c - 2d \Leftrightarrow a - b = c - d
\end{aligned}$$

We can now solve for x_{13} located in III_1 using the x_3 value we found in II_1 and the constraints on a, b, c, and d we found in VIII_2.

$$\begin{aligned}
III_1 &\rightsquigarrow a + d - 2b - c + a - x_{13} = 0 \Leftrightarrow 2a - 2b - c + d = x_{13} \\
&\Leftrightarrow 2c - 2d - c + d = x_{13} \Leftrightarrow x_{13} = c - d
\end{aligned}$$

Now let us look at combining V_2 and VIII_1 in order to get more explicit constraints on the square sizes. We already know $d = e$, so we take $V_2 - VIII_1$ to get:

$$V_2 - VIII_1 : (a + b = c + 2d) - (a - b = c - d) \rightsquigarrow 2b = 3d \Leftrightarrow b = \frac{3}{2}d$$

Now we take $V_2 + VIII_1$ to get:

$$V_2 + VIII_1 : (a + b = c + 2d) + (a - b = c - d) \rightsquigarrow 2a = 2c + d \Leftrightarrow c = a - \frac{d}{2}$$

With these new constraints, we claim that a and d are free variables, and we will show this by explicitly writing each variable in terms of a and d.

$$\begin{aligned}
&a = a \\
&b = \frac{3}{2}d \\
&c = a - \frac{d}{2} \\
&d = d \\
&x_1 = x_2 = a - \frac{3}{2}d \\
&x_3 = x_4 = 2(\frac{3}{2}d) + (a - \frac{d}{2}) - a = 3d - \frac{d}{2} = \frac{5}{2}d \\
&y_5 = y_6 = a + 2(\frac{3}{2}d) - (a - \frac{d}{2}) = \frac{7}{2}d \\
&y_7 = y_{10} = 3d - a \\
&y_8 = y_9 = d - a \\
&x_{11} = x_{13} = (a - \frac{d}{2}) - d = a - \frac{3}{2}d = x_1 = x_2 \\
&x_{12} = 0 \\
&y_{14} = y_{15} = (\frac{3}{2}d) - (a - \frac{d}{2}) = 2d - a
\end{aligned}$$

We must choose a to define our tiling, leaving d as our only free parameter, indicating that we are on a one-dimensional deformation space. □

Example 3.16. As before, we begin by putting a "reduced" dual graph on the tiling and labeling all squares, vectors, and loops. Note that we have three squares, nine vectors, and six loops, as expected.

We next define each of the vectors in the picture.

$$v_1 = \begin{pmatrix} a+b \\ y_1 \end{pmatrix} \quad v_2 = \begin{pmatrix} a+c \\ y_2 \end{pmatrix} \quad v_3 = \begin{pmatrix} a+b \\ y_3 \end{pmatrix}$$

$$v_4 = \begin{pmatrix} x_4 \\ 2a \end{pmatrix} \quad v_5 = \begin{pmatrix} a+c \\ y_5 \end{pmatrix} \quad v_6 = \begin{pmatrix} a+b \\ y_6 \end{pmatrix}$$

$$v_7 = \begin{pmatrix} a+c \\ y_7 \end{pmatrix} \quad v_8 = \begin{pmatrix} x_8 \\ b+c \end{pmatrix} \quad v_9 = \begin{pmatrix} x_9 \\ b+c \end{pmatrix}$$

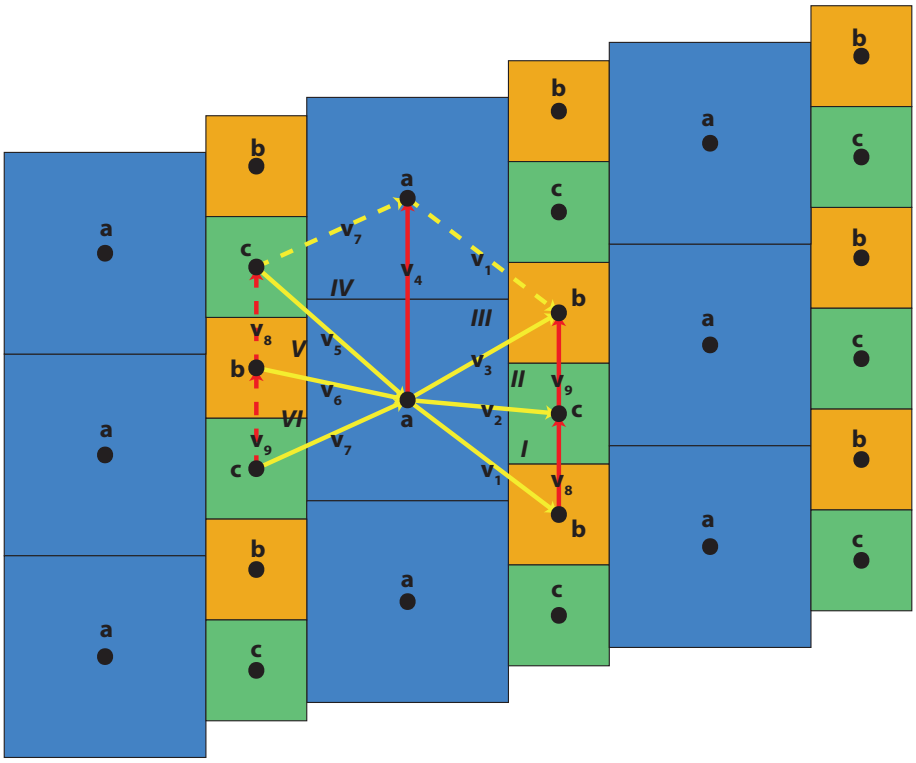


Figure 82: A tiling by three squares with all squares, vectors, and loops labeled

Now we write down the equation for each loop. Note there are six equations in this example.

$$
\begin{aligned}
&I : v_1 + v_8 - v_2 = 0 \\
&II : v_2 + v_9 - v_3 = 0 \\
&III : v_3 - v_1 - v_4 = 0 \\
&IV : v_4 - v_7 + v_5 = 0 \\
&V : v_6 - v_5 - v_8 = 0 \\
&VI : v_7 - v_6 - v_9 = 0
\end{aligned}
$$

Because of linear dependency, we will toss out one equation (say VI). Now we break up each vector equation into its x and y components, yielding ten equations.

$$
\begin{aligned}
&I_1 : a + b + x_8 - a - c = 0 \\
&I_2 : y_1 + b + c - y_2 = 0 \\
&II_1 : a + c + x_9 - a - b = 0 \\
&II_2 : y_2 + b + c - y_3 = 0 \\
&III_1 : a + b - a - b - x_4 = 0 \\
&III_2 : y_3 - y_1 - 2a = 0 \\
&IV_1 : x_4 - a - c + a + c = 0 \\
&IV_2 : 2a - y_7 + y_5 = 0 \\
&V_1 : a + b - a - c - x_8 = 0 \\
&V_2 : y_6 - y_5 - b - c = 0
\end{aligned}
$$

Next we reduce and rearrange each equation, writing the x_i or y_i terms on one side of the equality and the a, b, and c terms on the other side.

$$
\begin{aligned}
&I_1 \rightsquigarrow x_8 = c - b \\
&I_2 \rightsquigarrow y_2 - y_1 = b + c \\
&II_1 \rightsquigarrow x_9 = b - c \\
&II_2 \rightsquigarrow y_3 - y_2 = b + c \\
&III_1 \rightsquigarrow x_4 = 0 \\
&III_2 \rightsquigarrow y_3 - y_1 = 2a \\
&IV_1 \rightsquigarrow x_4 = 0 \\
&IV_2 \rightsquigarrow y_7 - y_5 = 2a \\
&V_1 \rightsquigarrow x_8 = b - c \\
&V_2 \rightsquigarrow y_6 - y_5 = b + c
\end{aligned}
$$

Right away note that I_1 and V_1 give:

$$x_8 = c - b = b - c \Rightarrow \{x_8 = 0 \text{ and } b = c\}.$$

Also note that $I_2 + II_2 = III_2$, which gives an important relation:

$$\begin{aligned}
&I_2 + II_2 = y_2 - y_1 + y_3 - y_2 = y_3 - y_1 = III_2 \\
\Leftrightarrow &I_2 + II_2 = (b+b) + (b+b) = 4b = 2a = III_2 \\
\Rightarrow &a = 2b
\end{aligned}$$

With these two relations, we can now write our ten equations in an even more reduced form:

$$\begin{aligned}
&I_1 \rightsquigarrow x_8 = 0 \\
&I_2 \rightsquigarrow y_2 - y_1 = 2b \Leftrightarrow y_2 - y_1 = a \\
&II_1 \rightsquigarrow x_9 = 0 \\
&II_2 \rightsquigarrow y_3 - y_2 = 2b \Leftrightarrow y_3 - y_2 = a \\
&III_1 \rightsquigarrow x_4 = 0 \\
&III_2 \rightsquigarrow y_3 - y_1 = 2a \\
&IV_1 \rightsquigarrow x_4 = 0 \\
&IV_2 \rightsquigarrow y_7 - y_5 = 2a \\
&V_1 \rightsquigarrow x_8 = 0 \\
&V_2 \rightsquigarrow y_6 - y_5 = 2b \Leftrightarrow y_6 - y_5 = a
\end{aligned}$$

We claim that a, y_2, and y_5 are free variables, and we will show this by explicitly writing each variable in terms of a, y_2, and y_5.

$$\begin{aligned}
&a = a \\
&b = \frac{a}{2} \\
&c = \frac{a}{2} \\
&y_1 = y_2 - a \\
&y_2 = y_2 \\
&y_3 = y_2 + a \\
&x_4 = 0 \\
&y_5 = y_5 \\
&y_6 = y_5 + a \\
&y_7 = y_5 + 2a \\
&x_8 = 0 \\
&x_9 = 0
\end{aligned}$$

We must choose a to define our tiling, so we are left with two free parameters, indicating that we are on a two-dimensional deformation space. □

Notice how the deformation spaces from these examples all have a dimension of at least 1. Now let us shift gears and look at how the deformation spaces intersect and why these examples have such dimensions.

Usually (as one has seen from the moduli spaces above), a tiling doesn't intersect any additional deformation spaces other than the one on which it sits. This is the relatively "boring" case, but we can make it interesting by looking at what happen at each vertex of each square in the fundamental domain.

Notice that in a complete tiling, the area around each vertex must be covered. This means every part of the 360° degrees around each vertex must be covered. The vertex itself has a square coming from it, so this covers a fourth of the space, or 90°. Now we must find a way to cover the other 270°. Clearly we can have three more vertices all touch at this one point, each adding 90° to cover the space, so the area around the vertex will indeed be covered. Then at the vertex we have a "+" alignment. But we can also have an edge touch a vertex, and this will cover half of the space, or 180°. Then all we have left to cover is 90°, for which we can use another vertex. This creates a "T" alignment. We claim these are the only two alignments, and this follows directly from our progression above. The only two parts that can touch at a vertex are an edge and other vertices, so this means we can only add discrete values of 90° and 180° to get to 360°. Adding 180° to 180° does give 360°, but we are missing the important part of looking at what surrounds a vertex, so this case cannot happen. If we have an edge, the only way to get to 360° is to add on two more vertices, giving us the "T" alignment. If we do not have an edge, the only way to get to 360° is to add all vertices, giving us the "+" alignment.

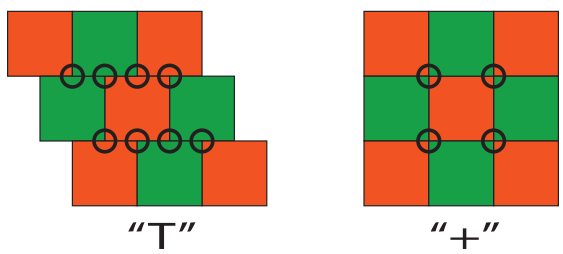


Figure 83: Examples of "T" alignments and "+" alignments

Proposition 3.17. *In a periodic tiling by squares, the only two possible alignments at each and every vertex are the "T" alignment and the "+" alignment.*

Proof. Our argument above suffices as a proof to this proposition; we showed how to get to the "T" and "+" alignments, and we showed that they were the only possible ones. □

So how does this relate to intersecting deformation spaces? Well, first we must look at only the alignments in and around the fundamental domain and disregard any redundant ones (the same thing we did when reducing the dual graph, its vectors, and its loops modulo translation). Since there are copies of the fundamental domain all around, we must make sure we don't count the same alignment twice: only once on one fundamental domain and not again in a

different spot on a different fundamental domain. Drawing a piece of the tiling helps with this dilemma.

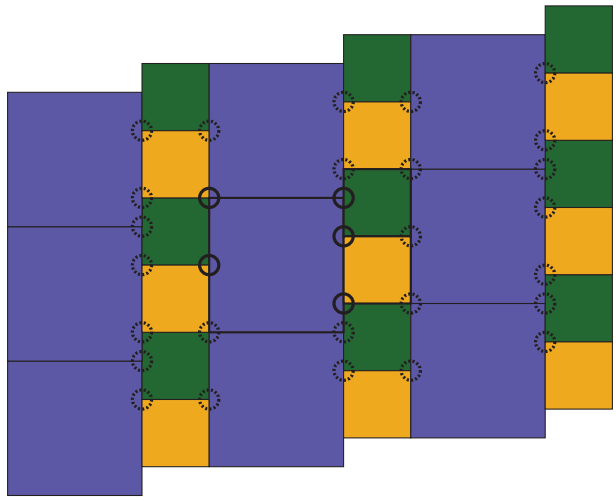


Figure 84: An example detailing how to accurately determine the number of alignments; note the dashed circles to indicate a repeated alignment (i.e. an alignment that is equivalent modulo translation to a previous one)

We claim that if our tiling only has "T" alignments, then we will stay on the same deformation space, yet if we "+" alignments, there exist different deformations to take us to different deformation spaces. If we have one "+" alignment, we have an intersection of two deformation spaces. If we have two "+" alignments, we have four different deformation spaces intersecting. With three "+"'s we have eight. There seems to exist a pattern with the amount of "+" alignments and the number of intersecting deformation spaces at that point. This does not hold true at limit points (obviously, since we are no longer dealing with the same number of squares). We have not proven this yet, but we will make a conjecture:

Remark 3.18. ***Deformation Space Intersection Conjecture (DSI Conjecture)***: For n number of "+" alignments in a given tiling ($n \geq 0$), there exist 2^n one-dimensional deformation spaces that intersect at that tiling.

We have an idea to start the proof, and it involves looking at the nearby moduli space. At a tiling with "+" alignments, the space nearby looks like the union of open subsets of finite dimensional vector spaces corresponding to the moduli spaces of tilings where each "+" alignment has been deformed into a "T" alignment. For n "+" alignments, there can be at most 2^n such spaces (since a single "+" alignment admits two "T" alignment deformations). The tricky part comes when we try to show this number is always reached. We can

show this on a case-by-case basis, but we have not been able to prove this in general.

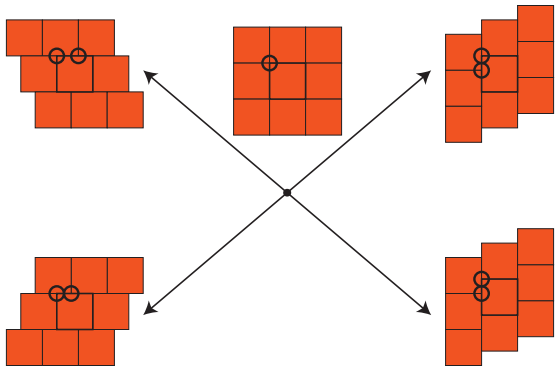


Figure 85: The intersecting moduli spaces in the one-square case; note there is one "+" alignment in the gridlock tiling and two one-dimensional moduli spaces that intersect at that tiling

We have to "maneuver" through the conjecture some to make it work for "+" alignments for tilings in a torus (and pay great attention to the one-dimensional detail). For instance, we can have the following case appear in the two-dimensional torus moduli space for two squares:

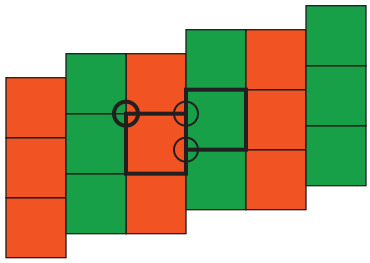


Figure 86: An example of a tiling located in the 2-D torus moduli space in the two-square tiling system; note there is one "+" alignment

In this case we have one "+" alignment, but we are already located deep inside a deformation space and are not near any others with which to intersect. How can this be? Well, recall that the two-dimensional torus is homeomorphic to $S^1 \times S^1$, which means we actually can think of this torus as two loops working together in this space, so anytime in this deformation space that we have a "+" alignment, we can think of the two S^1 loops crossing each other. This topological detail allows us to fit in the "+" alignments that we find while in toroidal moduli spaces. Note that in the three-square tilings V1, H1, V2, and H2, we have a three-dimensional torus as a moduli space, so we are working with three S^1

loops, which can cross each other in different ways, accounting for the various "+" alignments we may find.

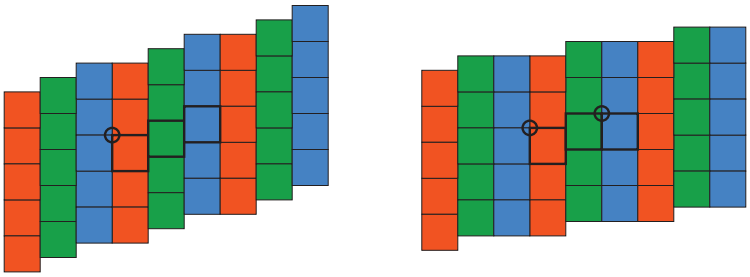


Figure 87: Examples of tilings in the 3-D torus attached to V1; note the tiling on the left has one "+" alignment and the tiling on the right has two "+" alignments

Another issue arises when we have the shared spaces that we brought up in the last section. Clearly, in the three-square tiling system, shared spaces are very prominent, with six shared spaces associated to every adjacent pair of the base tilings.

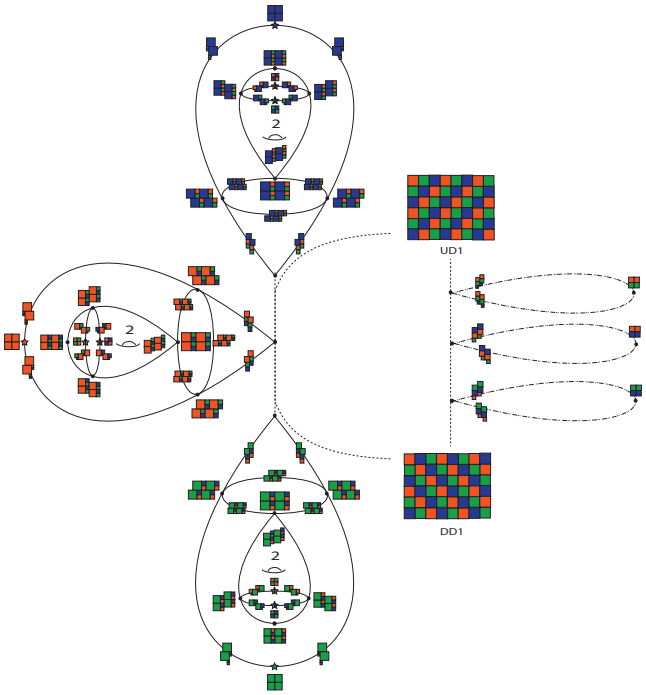


Figure 88: UD1 and DD1 and their six shared spaces

But if we count all the deformation spaces that surround a base tiling (let us take UD1 as an example), we get a massive number of fourteen deformation spaces: the horizontal-sliding space, the vertical-sliding space, the six shared spaces with DD1, and the six shared spaces with DD2.

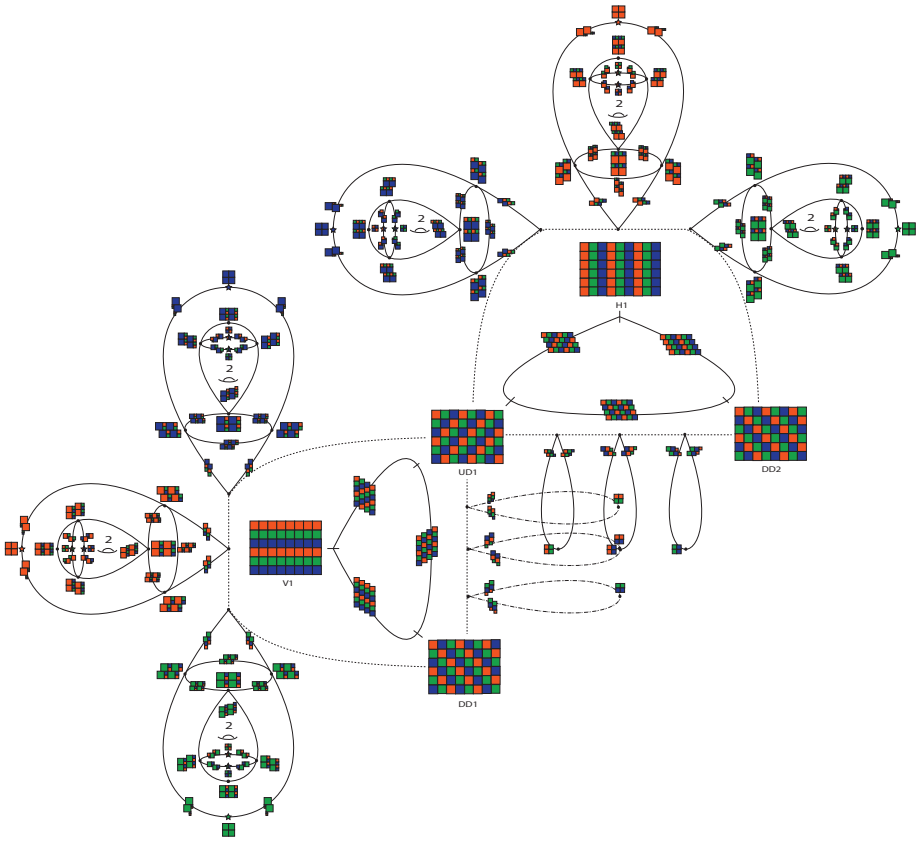


Figure 89: UD1 and all the deformation spaces that surround it

Since all four base cases have three "+" alignments, we expect $2^3 = 8$ intersecting deformation spaces. So what is the problem here? It comes from overcounting the shared spaces.

We should be counting each shared space once, but if we get fourteen for each of the four base cases, we have counted them all twice. For instance in the UD1 example, we included all the shared spaces with DD1. But if we look at DD1 and include all of the shared spaces with UD1, we have just become redundant. There are two ways around this. One is to include only one set of shared spaces. That is, for UD1 only including the shared spaces with DD2, for DD2 only including the shared spaces with UD2, for UD2 only including the shared spaces with DD1, and for DD1 only including the shared spaces with UD1. This way each shared space in counted only once, and each of the four

base cases have eight intersecting deformation spaces.

This procedure is alright, but we would like a better way that incorporates both sets of shared spaces without overcounting. For this, we count the number of shared spaces between two tilings, divide by two, and add that number to the number of intersecting deformation spaces. It is the same idea in principal as the first, but in a more mathematical and property-preserving manner. That means for UD1, the shared spaces with DD1 account for $6 \div 2 = 3$ of the intersecting deformation spaces, and the shared spaces with DD2 account for $6 \div 2 = 3$ of the intersecting spaces as well. Adding this onto the two sliding deformation spaces, we get eight total intersecting spaces. This works for all four of the base cases, and the problem is settled.

One may ask what the difference is between overcounting the shared spaces and overcounting other spaces. Shouldn't it be that if two tilings are on the same deformation space that we add $\frac{1}{2}$ to each of their values for intersecting spaces? Clearly not, but it may not be obvious at first glance. In the case of shared spaces, we have two different tilings that act as the same point on a deformation space. This is why we divide by two in that case. For any other case like the UD1-DD1-V1 loop, UD1 and DD1 are on the same deformation space, but they are different points. This is why we do not divide by two (as we had been doing up to this point).

One final issue comes up with the intersection of a torus with another moduli space, namely a loop. Let us look at the intersection point of the 2-D torus with vertical band alignments for the two-square case, and then the intersection point of the 3-D torus with the vertical band alignments (the tiling V1) for the three-square case.

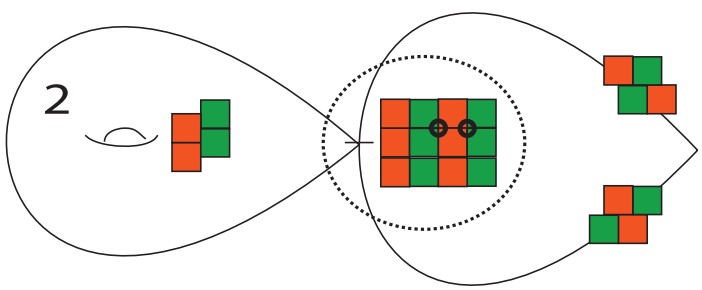


Figure 90: The 2-D torus, its neighboring loop, and the gridlock tiling at their intersection; note the gridlock tiling has two "+" alignments

Right away note that there are two "+" alignments, but we only have two intersecting moduli spaces, and our conjecture predicts there should be four. The issue arises with the underlying S^1 loops in the torus again. Instead of looking at the loop and the 2-D torus intersecting at a point, we must look at the loop intersecting with two other loops at a point, essentially quantizing the intersection of the loop and the 2-D torus. Yet these other two loops form the torus, so they behave differently. Let us call the torus loops α and β, and call the

other moduli space loop M_0. Now we need four deformation spaces to intersect at this point. Clearly, we will have M_0 intersect. The trick comes when looking at the α- and β-loops. We will have the α-loop by itself intersect, and we will have the β-loop by itself intersect. We will also have a combination of the loops intersect: an α-β-loop so to speak. We include this combination to account for the direct product in $S^1 \times S^1$. Though we cannot realize these additional intersections when simply looking at the 2-D torus and the loop intersecting at a point, we can quantize the torus into its loops and see that mathematically, there are indeed four one-dimensional deformation spaces that intersect this point.

Now we can look at the case with three squares and the tiling V1:

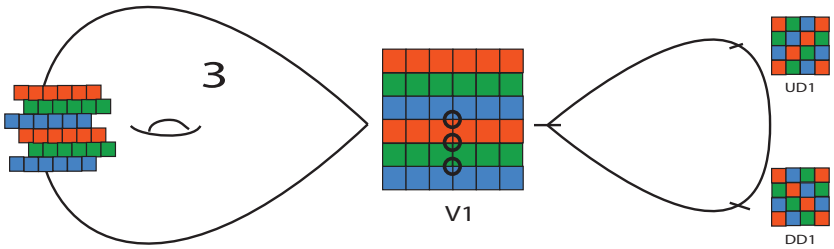


Figure 91: The 3-D torus, its neighboring loop, and the gridlock tiling (V1) at their intersection; note V1 has three "+" alignments

In this case, we have three "+" alignments, but again we only see two intersecting moduli spaces. This time our conjecture predicts there should be eight. Now, as before, we must break up the 3-D torus and look at its three underlying S^1 loops. Let us call them α, β, and γ, and let us call the other moduli loop M_0. Clearly we will have M_0 intersect, and clearly we will have the single α-loop, the single β-loop, and the single γ-loop intersect at this point as well. Now we must account for the direct product in $S^1 \times S^1 \times S^1$. With this, we find we also have the α-β-loop intersect, the α-γ-loop intersect, the β-γ-loop intersect, and the α-β-γ-loop intersect, all at this point V1. That gives us the eight intersecting one-dimensional deformation spaces we need, and our conjecture still holds.

The single internal loop inside of the 2-D tori in the three-square tiling system follows this precedence as well. Either side of internal loop accounts for one of the four possible intersecting moduli spaces, while the other three come from the combinations of loops in the 2-D torus.

We can also do this with a tiling by four squares that has vertical bands (similar to the V1 case). In this case we have a 4-D torus and a neighboring loop as our moduli spaces, and our tiling has four "+" alignments, which means we need sixteen intersecting one-dimensional deformation spaces.

We can label each of the four toroidal loops as before (and the original M_0 one as well) and look at all their different combinations as before. At the intersection, we can have the four loops by themselves, six different combinations with two loops, four different combinations with three loops, and one combina-

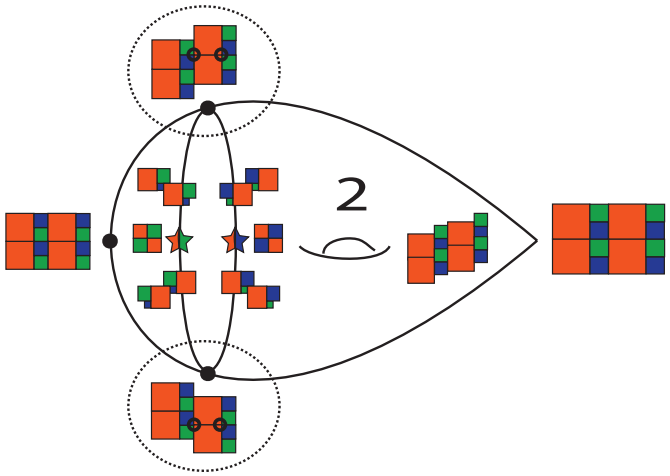


Figure 92: The internal loop inside of the 2-D torus in the three-square tiling system; note that the tiling at either intersection of the loop and the torus has two "+" alignments

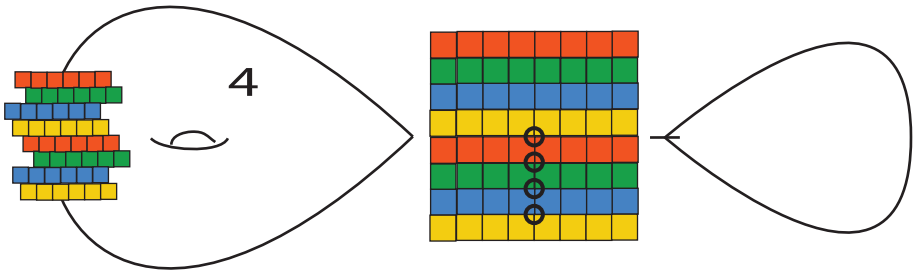


Figure 93: The 4-D torus, its neighboring loop, and the gridlock tiling at their intersection; note the gridlock tiling has four "+" alignments

tion with all four loops. Those fifteen intersecting loops plus the original M_0 loop account for all sixteen one-dimensional deformation spaces that intersect at that point.

Note that for this final issue of tori intersecting loops, the 2^n follows from the combinatorial formula involving the binomial coefficient:

$$2^n = \sum_{k=0}^{n} \binom{n}{k}$$

where k is the number of loops in consideration and n in the number of "+" alignments in the tiling (also the dimension of the torus in consideration). For any torus, $k = 0$ corresponds to M_0 intersecting, $k = 1$ corresponds to all single loops intersecting, $k = 2$ corresponds to all combinations of two loops, $k = 3$ corresponds to all combinations of three loops, and so on. It becomes clear that our conjecture will hold for all cases.

3.4 Conclusions

With the DSI Conjecture and the dual graph analysis, we can theoretically determine all parts of a tiling system for a finite number of squares. As we saw in the moduli space section, we were able to find the entire tiling systems for two and three squares without the use of the DSI Conjecture and the dual graph analysis, but when we think about four or five squares, the tiling system become so complicated that it would be extremely difficult to manually complete them. There are also so many more available tilings in four and five square cases (for instance, instead of just the "L"-shaped tiles and the rectangle tiles, "Z"-shaped tiles begin to appear) that it would be hard enough to merely think of all the different possible cases, let alone classify them and see how they interact with other deformation spaces.

With these two tools we described in the last section, we can look at a given tiling, use the dual graph analysis to determine the dimension of its deformation space and how to deform the tiling, and then use the DSI Conjecture to see how many other deformation spaces intersect the point we are looking at. Usually, we will have no "+" alignments and will have no intersections with other deformation spaces. But there will be tilings with these interesting "+" alignments after some deforming, and we will find beautiful mathematical patterns and topological objects through the various intersections of these deformation spaces as we further explore tilings by squares.

References

[1] J.C. Lagarias, Y. Wang: Tiling the line with translates of one tile. Inventiones Mathematicae. **124**(1-3), 341-365 (1996).

[2] S.V. Matveev: Euler characteristic. Encyclopedia of Mathematics. Facts on File (2005).