Research Experiences for Undergraduates

Student Reports

INDIANA UNIVERSITY

Summer 2008

Bloomington, Indiana

The Summer 2008 REU Program At Indiana University

During the summer of 2008 ten students participated in the Research Experiences for Undergraduates program in Mathematics at Indiana University. The program ran for eight weeks, from June 16 through August 8. Eight faculty served as research advisers. Two faculty members oversaw pairs 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. Students also received orientations to the mathematics library and to our computing facilities. We were saddened during week two at the departure of participant Ben Schweinhart, who returned home for medical reasons. 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. During week six, we hosted the Indiana Mathematics Undergraduate Research conference, which featured 22 lectures by 33 students from Rose-Hulman Institute of Technology, Goshen College, Wabash College, and Indiana University. The program concluded with the students giving formal, hourlong presentations to the REU students and faculty, and the turning in 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.

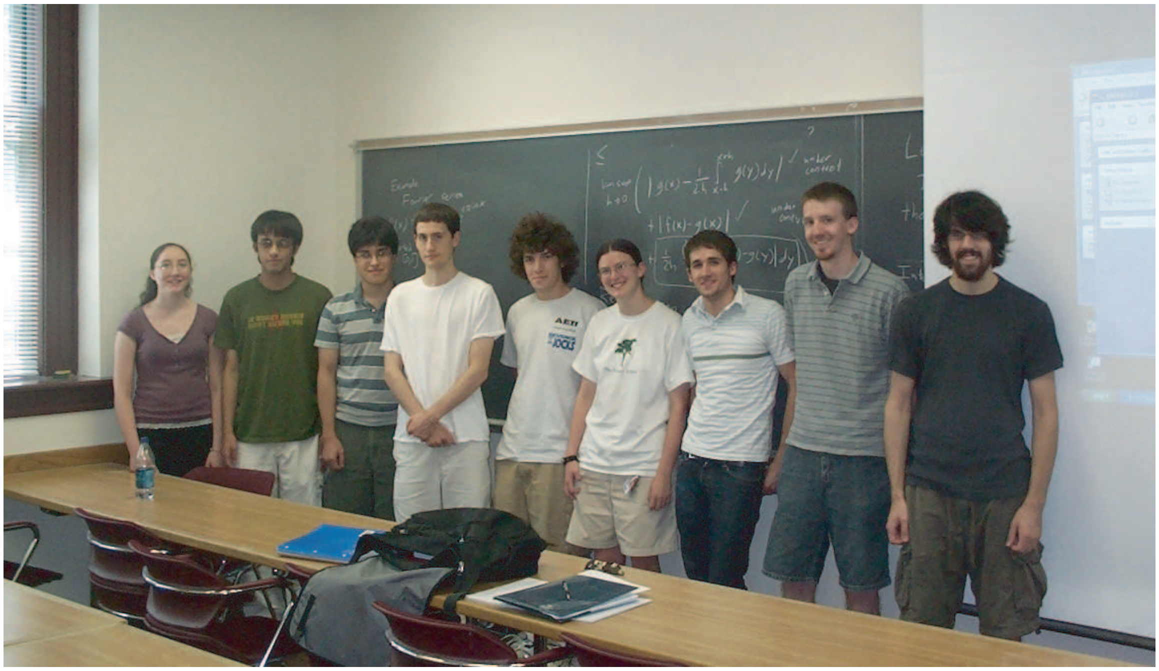
We thank the National Science Foundation for major financial support through the REU program. We also thank the College of Arts and Sciences for crucial additional funding. We thank Indiana University for the use of facilities, including library, computers, and recreational facilities. We thank the staff of the Department of Mathematics for support, especially Mandie McCarty for coordinating the complex logistical arrangments (housing, paychecks, information packets, meal plans, etc.) 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.

We thank Professors Eric Bedford, Hari Bercovici, Jiri Dadok, James Davis, David Hoff, Paul Kirk, Larry Moss, Bruce Solomon, and Matthias Weber for volunteering to give lectures during the program. We also thank Professor Michael Larsen for his plenary lecture at the Indiana Mathematics Undergraduate Research conference.

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: Hari Bercovici, Chris Connell, Allan Edmonds, Michael Jolly, Chris Judge, Larry Moss, William Orrick, and Kevin Pilgrim.

Chris Judge
Kevin Pilgrim

REU Participants Summer 2008



From left: Sharon Ulery
Theodore Spaide
Carlo Angiuli
John Ullman
Samuel Ziegler
Anne Carter
Michael Grilli
Michael Meaden
Adam Vollrath

-6

REU Participants
Summer 2008

Carlo Angiuli	Indiana University
Annie Carter	Swarthmore College
Michael J. Grilli	Providence College
Michael Meaden	Elmhurst College
Ted Spaide	Harvey Mudd College
Sharon Ulery	Bowdoin College
John Ullman	Indiana University
Adam Vollrath	Northwestern University

Faculty Advisors

Hari Bercovici
Chris Connell
Allan Edmonds
Michael Jolly
Chris Judge
Larry Moss
William Orrick
Kevin Pilgrim

Contents

On the combinatorics of honeycombs

Carlo Angiuli
Indiana University

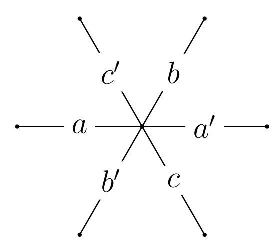
Indiana University REU Summer 2008
Advisor: ADVISOR

A.1 Introduction

Consider three unit vectors u, v, and w where $u+v+w=0$. These vectors, which form an equilateral triangle in the plane, generate a set of *lattice points* $\{iu+jv : i,j \in \mathbb{Z}\}$. Define a *small edge* as a unit segment joining two nearest lattice points. Our research is concerned with *measures* as defined in [1]—unions of small edges assigned positive densities, which satisfy the "balance condition" that

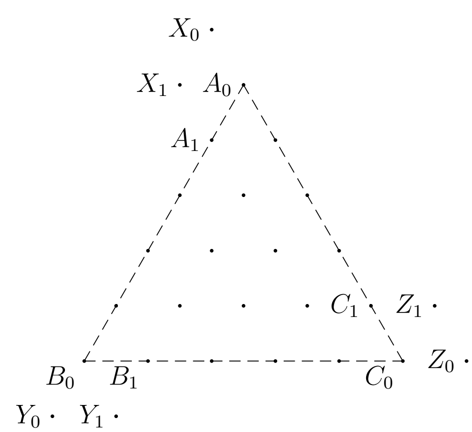
$$m(a) - m(a') = m(b) - m(b') = m(c) - m(c')$$

whenever the six edges a, c', b, a', c, and b' are located in cyclic order around a single lattice point, and where $m(e)$ denotes the density assigned to a small edge e.



Define the support of a measure m as the set of small edges in the measure, $\{e \in m : m(e) > 0\}$. A *branch point* is any lattice point incident to at least three edges in the support of a measure.

Consider in particular a closed triangle, denoted $\triangle_r$ for fixed integral $r \geq 1$, with vertices at 0, ru, and $ru+rv$. Name the lattice points on its borders $A_j = ju$, $B_j = ru + jv$, and $C_j = rw - jw$. The lattice points immediately outside its borders will be denoted $X_j = A_j + w$, $Y_j = B_j + u$, and $Z_j = C_j + v$. A few of these points are depicted below on $\triangle_5$.



Define $\mathcal{M}_r$ as the set of measures with all branch points contained in $\triangle_r$ where

$$m(A_jX_{j+1}) = m(B_jY_{j+1}) = m(C_jY_{j+1}) = 0, \quad j \in \{0,1,\ldots,r\}.$$

Similarly, define $\mathcal{M}_r^*$ as the set of measures with all branch points contained in $\triangle_r$ where

$$m(A_jX_j) = m(B_jY_j) = m(C_jY_j) = 0, \quad j \in \{0,1,\ldots,r\}.$$

$\mathcal{M}_r^*$ can be considered a reflection of $\mathcal{M}_r$ across any angle bisector of $\triangle_r$.

Note that all measures in $\mathcal{M}_r$ are determined entirely by their restrictions on $\triangle_r$. Indeed, because no branch points are permitted outside of $\triangle_r$, all densities outside the triangle must propagate as half-lines.

Define an *attachment point* of $m \in \mathcal{M}_r$ as any non-corner point on the border of $\triangle_r$ incident to an exterior edge in the support of m. More precisely, an attachment point is any A_j (or B_j, C_j) for $j \in \{1, 2, \ldots, r-1\}$ where $m(A_j X_j) > 0$ (or $B_j Y_j$, $C_j Z_j$). Attachment points are defined analogously for $\mathcal{M}_r^*$. The number of attachment points in a measure m will be denoted $ap(m)$.

For a measure $m \in \mathcal{M}_r$, we define its weight $\omega(m)$ as

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

The balance condition necessitates that these sums be equal. The weight of a measure in $\mathcal{M}_r^*$ can be defined analogously, using the edges $A_j X_{j+1}$, etc.

Define the *boundary* ∂m of $m \in \mathcal{M}_r$ as a triple of r-tuples, $(\alpha, \beta, \gamma) \in (\mathbb{R}^r)^3$ where, for $i \in \{1, 2, \ldots, r\}$,

$$\alpha_i = \sum_{j=0}^{i-1} m(A_j X_j), \beta_i = \sum_{j=0}^{i-1} m(B_j Y_j), \gamma_i = \sum_{j=0}^{i-1} m(C_j Z_j).$$

We are concerned with *rigid* measures, measures determined entirely by their weight and boundary; for the rest of this paper, we will only consider rigid measures, and indeed, many of the following results do not hold for non-rigid measures.

A.2 Puzzles and Duality

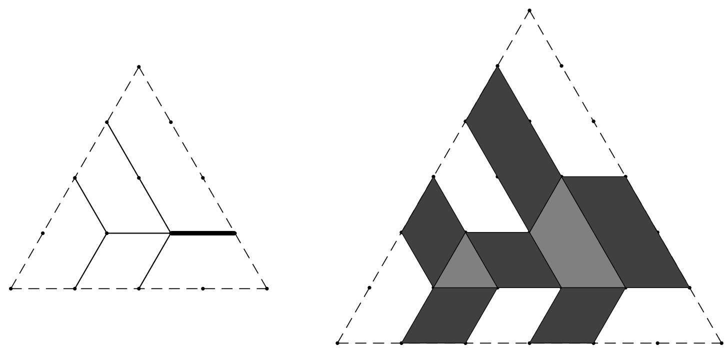
In this section, we define the notions of inflating a measure m into an object termed a "puzzle," for reasons to be made clear, and subsequently the notion of *deflating a puzzle into a dual measure m^*.

Define the *inflation* of a measure $m \in \mathcal{M}_r$ as the following procedure. First, cut $\triangle_r$ along the edges of the support of m, forming a collection of *white* puzzle pieces corresponding to the resulting shapes.

Translate each small edge e in m along segments 60° clockwise from e, with length $m(e)$. Together, the four resulting segments form a parallelogram with two edges of e's original length, parallel to e, and two edges of length $m(e)$, 60° clockwise from e. This parallelogram is termed the *inflation of* e, and is illustrated as a *dark gray* puzzle piece.

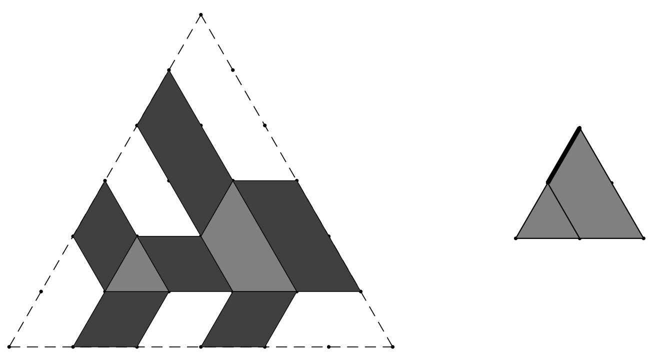
Translate the white puzzle pieces away from each other, fitting the newly-created parallelograms in place of their corresponding small edges. These pieces all fit together, leaving spaces corresponding to each branch point in m, which are filled in as *light gray* puzzle pieces.

The resulting *puzzle* is a triangle of size $r + \omega(m)$, consisting of three kinds of pieces, corresponding to shapes carved out of $\triangle_r$ by m, small edges of m, and branch points of m.



In the example above, the thinner lines in the support have density one, and the thicker one density two. The inflation is drawn to the right.

This puzzle can now be **deflated* to yield a measure in $\mathcal{M}_r^*$, as follows. Remove the white pieces, and deflate the dark gray pieces in the opposite direction as they were inflated—shrink the original sides to points, while maintaining the edges of length $m(e)$. The light gray pieces remain, separated by edges corresponding to the dark gray pieces.



In the above figure, the puzzle is deflated to yield a dual measure. The thicker edge, again, has density two; the borders of the triangle, as well as an inner edge, have density one.

This *deflation results in a triangle with sides of $\omega(m)$, and a measure in $\mathcal{M}_r^*$ denoted m^*, the *dual measure* of m. Each original edge is rotated $60°$ clockwise from its original location, and its length and density have swapped places.

The original measure m can be generated from m^* by applying *inflation followed by deflation, in the opposite directions as inflation and *deflation. This replaces the original length and density of each small edge.

A.3 Descendance and Skeletons

We now define a partial order on small edges. For incident small edges e and f, we say $e \to f$ if either:

1. they are $120°$ apart, and the edge opposite e has zero density;
2. they are opposite, and an edge $120°$ from e has zero density.

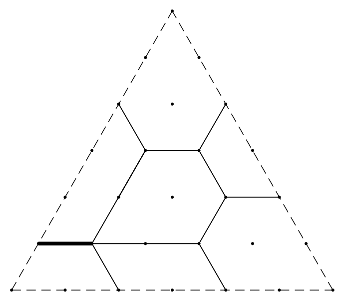
$e \to f$ implies that $m(e) \leq m(f)$, as a simple consequence of the balance condition. $m(e) = m(f)$ only if $f \to e$ as well.

Furthermore, we say $e \Rightarrow f$ if there exists a sequence of small edges $e_1, e_2, \ldots, e_n$ such that

$$e = e_1 \to e_2 \to \cdots \to e_n = f.$$

f is said to be a *descendant* of e, its *ancestor*. If $e \Rightarrow f$ and $f \Rightarrow e$, then $e \Leftrightarrow f$ and the edges are *equivalent*. Edges that are minimal with respect to descendance are called *root edges*.

In the skeleton below, all the thinner edges have density one and are equivalent root edges; the thicker edge has density two and is not a root edge.



The sequence $e_1, e_2, \ldots, e_n$ is called the *descendance path from e to f*, and induces a natural orientation on f, the direction that f is traversed in its descendance path. As shown in [1], if f is not a root edge, then this *orientation of f away from the root edges* is the same for any descendance path from any root edge that is an ancestor of f.

A support $s \subset m = \{e \in m : r \Rightarrow e\}$ containing all descendants of a root edge r is called a *skeleton*. A measure whose support is a skeleton is called an *extremal measure*. It is a property of skeletons that they contain no proper subset which can support a measure, which in fact serves as a definition in [1]. As a result, an extremal measure is entirely determined by its density on any small edge.

By the nature of the partial ordering $\Rightarrow$, all members of each equivalence class of root edges generate the same skeleton. Clearly, every edge in m is a descendant of at least one equivalence class of root edges. Thus, m can be treated simply as a sum of the skeletons generated by a maximal collection of inequivalent root edges. This decomposition of m into a sum of skeletons is unique, and the number of skeletons in this decomposition—alternatively, the number of equivalence classes of root edges—will be denoted $sk(m)$.

Finally, we define an order relation on skeletons. For skeletons S_1 and S_2, we say $S_1 \prec_0 S_2$ if S_1 contains collinear small edges a and b and S_2 contains collinear small edges c and d, such that a, b, c, and d are incident at a single point, and a is $60°$ clockwise from c. It is shown in [1] that $\prec_0$ is well-defined for skeletons contained in a rigid measure.

A.4 Results

Measures are a representation of the Horn inequalities from linear algebra, as discussed in [1], [2], and [3]. This paper is concerned, however, with the combinatorial aspects of measures, in particular, Bercovici's conjecture that

$$sk(m) + sk(m^*) = ap(m) + 1$$

for all rigid m. The proof for the case $sk(m) = sk(m^*) = 1$ is located in [1]; here we use a much more general approach to prove that

$$sk(m) + sk(m^*) \leq ap(m) + 1.$$

Two measures are *homologous* if there exists a bijection between edges in m and m' that sends all edges to parallel edges, and keeps all concurrent edges concurrent. The measures m and m' are homologous if and only if m^* and m'^* are homologous. Two puzzles are said to be homologous if their deflations are homologous.

Consider a measure

$$m = \alpha_1 m_1 + \alpha_2 m_2 + \cdots + \alpha_{sk(m)} m_{sk(m)}$$

and its dual,

$$m^* = \beta_1 \mu_1 + \beta_2 \mu_2 + \cdots + \beta_{sk(m^*)} \mu_{sk(m^*)}$$

where $\alpha_i, \beta_i \in \mathbb{R}^+$ and $m_1, \ldots, m_{sk(m)}, \mu_1, \ldots, \mu_{sk(m^*)}$ are extremal measures.

Altering any α_i changes the densities of small edges in m, but this clearly results in a measure m' homologous to m, as long as α_i remains positive.

Altering any β_i changes the densities of small edges in m^*, affecting the *lengths* of edges in m, since duality swaps lengths and densities. However, because the resulting $m^{*\prime}$ is homologous to the original m^*, m' is also homologous to m (again, as long as β_i remains positive).

Adjusting any number of α, β is the only way to generate a measure m' homologous to m—it generates all possible densities for the support of m, and all possible lengths that maintain homology with the original measure.

Define $\mathcal{P}_m$ as the set of all puzzles homologous to the inflation of a rigid measure m.

Lemma A.4.1. *There exists a bijection between* $\mathcal{P}_m$ *and* $(\mathbb{R}^+)^{sk(m)+sk(m^*)}$.

Proof. We have just produced a bijection between the set of all m' homologous to m and the set of tuples

$$(\alpha_1, \ldots, \alpha_{sk(m)}, \beta_1, \ldots, \beta_{sk(m^*)}), \quad \alpha_i, \beta_i \in \mathbb{R}^+.$$

Additionally, inflation is a bijection between all m' and all puzzles in $\mathcal{P}_m$. These can be composed to form a bijection between $\mathcal{P}_m$ and $(\mathbb{R}^+)^{sk(m)+sk(m^*)}$. □

The bijection described in Lemma A.4.1 is one way to distinguish puzzles in $\mathcal{P}_m$. We will need to introduce one more set describing such puzzles.

Every attachment point corresponds to the edge of a puzzle piece on the border of its puzzle. Consider the location of an attachment point on a puzzle to be the clockwise vertex of that edge, unless otherwise stated, and number the attachment points counterclockwise starting at A_1.

Let ℓ denote the length of the side of a puzzle, and let $\epsilon_i \in (0, 1]$ denote the relative position of the ith attachment point in the puzzle, as a fraction of the side length. Define $\mathcal{Q}_m$ as the set of tuples

$$(\ell, \epsilon_1, \epsilon_2, \ldots, \epsilon_{ap(m)}).$$

Clearly, $\dim \mathcal{Q}_m \leq ap(m) + 1$.

We now construct a map from $\mathcal{Q}_m$ to $\mathcal{P}_m$.

We first use a construction on two extremal measures $\mu_1, \mu_2 \in \mathcal{M}_r$ introduced in [1], called the *stretch of* μ_1 *to the puzzle of* μ_2. Essentially, μ_2 is inflated to yield a puzzle of size $r + \omega(\mu_2)$, on which is placed a "stretched" version of μ_1 homologous to μ_1. This new measure $\mu_1' \in \mathcal{M}_{r+\omega(\mu_2)}$ is then inflated by itself, yielding a puzzle whose pieces correspond to the locations of edges in μ_1 in the inflation of $\mu_1 + \mu_2$.

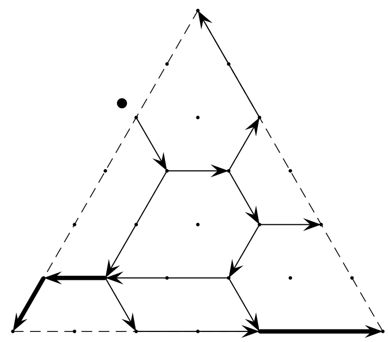
We will use this construction because it yields a puzzle that corresponds only to edges in μ_1, while maintaining the positions of μ_1's inflated edges in the inflation of $\mu_1 + \mu_2$. Specifically, attachment points have the same location in the inflation of μ_1' as in the inflation of $\mu_1 + \mu_2$.

Note, however, that the procedure outlined in [1] is only possible if $S_2 \not\succ_0 S_1$, where S_1, S_2 are the supports of μ_1, μ_2.

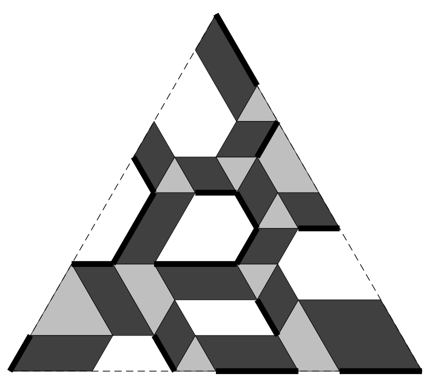
Define the set of S_i as the skeletons supporting the extremal measures μ_i which comprise m. Order these skeletons in a non-increasing order, such that $S_1 \not\succ_0 S_2 \not\succ_0 \cdots \not\succ_0 S_{sk(m)}$. Then, consider the stretch of each μ_i to the puzzle of $\mu_{i+1} + \cdots + \mu_{sk(m)}$. Call this new measure m_i.

Now we construct a certain extremal measure m_i' homologous to m_i, located on the inflation of m_i. Choose an attachment point whose exterior edge is a root edge.

Call that root edge r, and orient all other edges in m_i according to their descendance paths from r. In the example below, the location of the exterior root edge is marked by a dot.



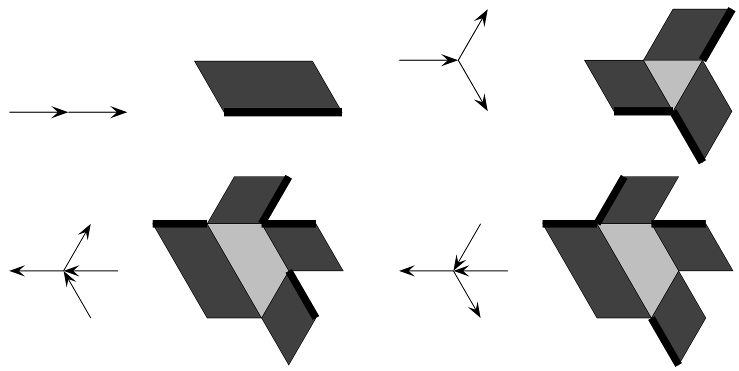
Attach each edge to the *right* side, according to this orientation, of its inflation in the puzzle. This results in a collection of disconnected segments on the puzzle, as shown below.



These segments will now be connected using the process outlined in Lemma A.4.2.

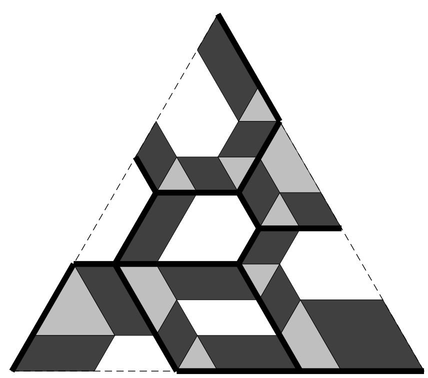
Lemma A.4.2. *Assume e, f are incident small edges in m_i and have been placed on a puzzle as described. Then the edges can be extended to intersect at a vertex, without entering the interior of a puzzle piece, as long as $e \to f$.*

Proof. Using the definition of $e \to f$ and the balance condition, we can easily enumerate all possible oriented supports of m_i around the lattice point to which e and f are incident, up to rotation. As shown below, in all four cases, the edges can be extended to meet at a point without entering a puzzle piece.



□

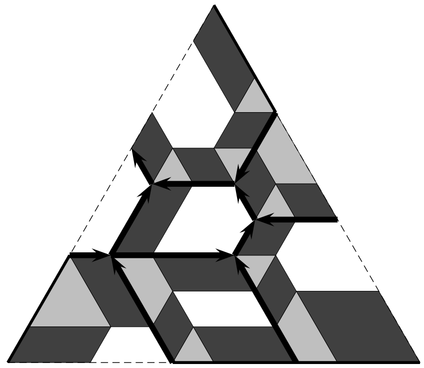
Because m_i is extremal, every edge has a descendance path from r. Clearly, we can follow these descendance paths, connecting every two consecutive edges as prescribed by Lemma A.4.2, until every edge is connected. It is easy to see that this procedure results in a new measure m_i' homologous to m_i.



By this construction, all the attachment points of m_i' except our initial point are reached from interior edges oriented outward, and thus these edges in m_i' always intersect the clockwise end of each attachment point. Similarly, m_i' always intersects the original attachment point at its counterclockwise end.

Now suppose we perform this construction in reverse, given a tuple in $\mathcal{Q}_m$. Since their locations are already known, we start at the clockwise end of each attachment point except one which has an exterior root edge. By following the descendance paths in reverse, we can obtain the location of the counterclockwise end of the one attachment point we omitted at the beginning.

Below is an example of how the reverse process would proceed. Interior vertices can be obtained by intersecting lines extended from the attachment points. The arrows indicate the order in which those vertices are determined.



Notice also that the final arrow determines the counterclockwise end of the initial attachment point.

By taking the difference of the location of that counterclockwise end and the location of its corresponding clockwise end, which as given by the tuple in $\mathcal{Q}_m$, we obtain the width of that attachment point on the puzzle, which is equal to the density of its exterior root edge. Becaues m_i is extremal, knowing the density of this one edge allows us to determine the densities of all its edges.

Thus, from the tuple in $\mathcal{Q}_m$, we have determined the edge densities of m_i, which correspond directly to the densities of μ_i. Since we have fully determined μ_i, we now perform the same operation on μ_{i+1}.

Once this process is finished for every skeleton, the resulting measures $\mu_1, \ldots, \mu_{sk(m)}$ can be summed to determine m, which in turn determines the specific member of $\mathcal{P}_m$ corresponding to the given tuple in $\mathcal{Q}_m$.

Therefore, this entire process determines a map from $\mathcal{Q}_m$ to $\mathcal{P}_m$. Since every puzzle in $\mathcal{P}_m$ corresponds to a tuple in $\mathcal{Q}_m$, this map must be surjective; indeed, this process must work for every tuple corresponding to a puzzle in $\mathcal{P}_m$. We can now use this fact to prove the upper bound asserted earlier.

Theorem A.4.3. *$sk(m) + sk(m^*) \leq ap(m) + 1$ for any rigid m.*

Proof. We have just shown there exists a surjective map from $\mathcal{Q}_m$ to $\mathcal{P}_m$, implying that $\dim \mathcal{Q}_m \geq \dim \mathcal{P}_m$. But since $\dim \mathcal{P}_m = sk(m) + sk(m^*)$ and $\dim \mathcal{Q}_m \leq ap(m) + 1$, this implies that $sk(m) + sk(m^*) \leq ap(m) + 1$. □

A.5 Future Research

Because the above proof was only recently discovered, it lacks some details which merit further investigation. Most notably, we must prove that every extremal measure must have an external root edge.

We believe it will be fairly straightforward to prove that $\dim \mathcal{Q}_m = ap(m) + 1$, and that the map demonstrated above is a bijection. The full conjecture should follow directly from these additional facts. We also hope to further study this map and related maps, in an effort to discover further properties of measures.

A.6 Acknowledgements

I would like to thank Hari Bercovici and Wing Suet Li for their great expertise and insight into this problem. I would also like to thank the National Science Foundation for funding this REU program, and Indiana University for hosting it.

Bibliography

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, arXiv:math.OA/0805.4817v1.

2. A. Knutson and T. Tao, The honeycomb model of $\mathrm{GL}_n(\mathbb{C})$ tensor products. I. Proof of the saturation conjecture, *J. Amer. Math. Soc.* **12**(1999), no. 4, 1055–1090.

3. A. Knutson, T. Tao, and C. Woodward, The honeycomb model of $\mathrm{GL}_n(\mathbb{C})$ tensor products. II. Puzzles determine facets of the Littlewood-Richardson cone, *J. Amer. Math. Soc.* **17**(2004), no. 1, 19–48.

Pure braid homomorphisms and complex dynamics

Annie Carter
Swarthmore College

Indiana University REU Summer 2008
Advisor: ADVISOR

B.1 Generalities

The braid group may be thought of as a group of equivalence classes of sets of descending, non-intersecting paths (strands) which begin at a given (finite) set of points in the plane and end at the same set of points in a copy of the plane which has been translated downwards. The equivalence relation is homotopy relative to the endpoints; that is, two braids are equivalent if one can be deformed into the other, leaving the endpoints fixed, without passing strands through each other. Two braids are multiplied by concatenation, joining the bottom of the first braid to the top of the second. The pure braid group is the subset of the braid group for which each strand returns to its original position in the new plane.

We can give the set of points a standard labeling which allows us to identify sets of paths with particular braid elements. The pure braid group on n strands PB_n admits the following presentation [1][3]: it is generated by elements A_{ij} (see figure B.1[1]), where $1 \leq i < j \leq n$, which correspond to twisting the ith

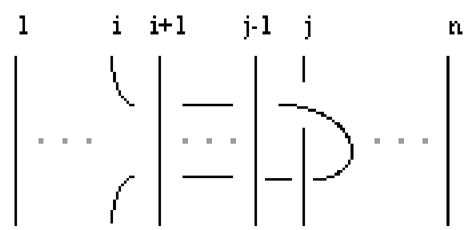


Figure B.1: The generator A_{ij}

strand around the jth strand, subject to the following relations:

$$A_{rs}^{-1} A_{ij} A_{rs} = \begin{cases} A_{ij} & \text{if } r < s < i < j \\ & \text{or } i < r < s < j \\ A_{rj} A_{ij} A_{rj}^{-1} & \text{if } r < i = s < j \\ A_{ij} A_{sj} A_{ij} A_{sj}^{-1} A_{ij}^{-1} & \text{if } r = i < s < j \\ A_{rj} A_{sj} A_{rj}^{-1} A_{sj}^{-1} A_{ij} A_{sj} A_{rj} A_{sj}^{-1} A_{rj}^{-1} & \text{if } r < i < s < j. \end{cases}$$

Call these relations $R_1, \ldots, R_4$. For illustrations of the first two Artin relations, see figure B.2. Note that the first relation tells us that two pure braids commute if the strands involved in each are completely disjoint (far commutativity)—compare figures B.2(a) and B.2(d)—or if the strands indexed by one lie strictly between the strands indexed by the other—compare figures B.2(b) and B.2(e). For the second relation, imagine pulling the first strand in front (see figure B.2(c)) or behind (see figure B.2(f)) the others.

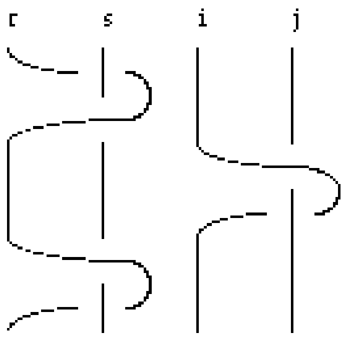
We can project the paths onto one of the planes (see figure B.3). In this case an oriented curve or curves in the plane traces the path along which each strand travels. This will be our standard projection.

The center of the pure braid group is generated by the full twist D^2 [2], which is given by

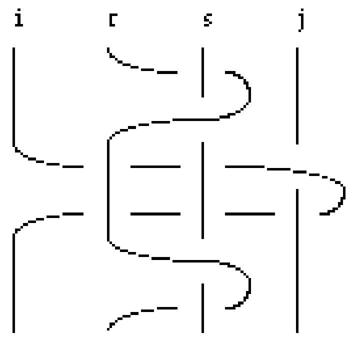
$$D^2 = (A_{12} A_{13} A_{14} \cdots A_{1n})(A_{23} \cdots A_{2n})(A_{34} \cdots A_{3n}) \cdots (A_{(n-1)n})$$

for the braid on n strands. Note that each parenthesized factor in this product corresponds to one loop moving around all those numbered above it (see figure B.4). Since no loop interacts with any of the others, each loop and thus each factor commutes with the others.

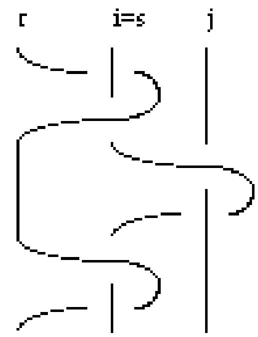
[1] We have used the positive orientation of the plane, while Artin uses the negative orientation. However, the relations are indifferent to the choice of orientation.



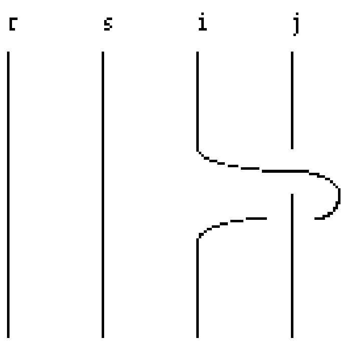
(a) $A_{rs}^{-1}A_{ij}A_{rs}$, where $r < s < i < j$



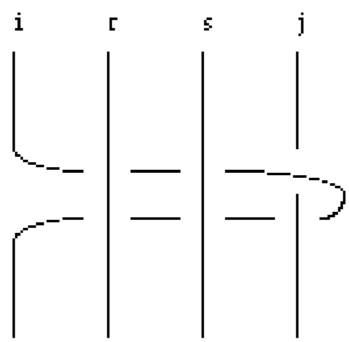
(b) $A_{rs}^{-1}A_{ij}A_{rs}$, where $i < r < s < j$



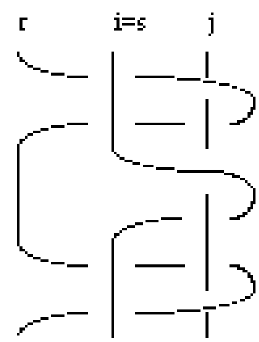
(c) $A_{rs}^{-1}A_{ij}A_{rs}$, where $r < i = s < j$



(d) A_{ij}, where $r < s < i < j$



(e) A_{ij}, where $i < r < s < j$



(f) $A_{rj}A_{ij}A_{rj}^{-1}$, where $r < i = s < j$

Figure B.2: Illustrations of the first two Artin relations for pure braids

Figure B.3: The generator A_{ij}

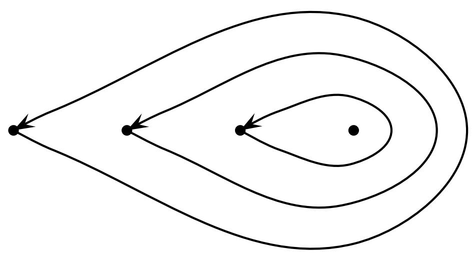


Figure B.4: The full twist D^2 on four strands

Let $f : \mathbb{C} \to \mathbb{C}$ be the function $f(z) = z^2 + c$ for some $c \in \mathbb{C}$ such that the point 0 is periodic or pre-periodic. We will primarily be concerned with functions for which the forward orbit of 0 consists of three or four points. To find examples of such functions we may simply solve equations of the form

$$f^m(c) = f^n(c)$$

for c, choosing appropriate values of m and n. For example, to find a value of c for which the orbit of 0 enters a 2-cycle after two iterations, we set

$$c^2 + c = ((c^2 + c)^2 + c)^2 + c.$$

Putting all the terms on one side and factoring gives us

$$c^3(c + 1)^2(c + 2)(c^2 + 1) = 0.$$

The point 0 is a fixed point for $c = 0$, is periodic of period 2 for $c = -1$, and enters a 1-cycle after two iterations for $c = -2$. So the two values of c for which 0 enters a 2-cycle after two iterations are i and $-i$.

Let $X = \{f^n(0) \mid n \in \mathbb{N}\}$ and let $X' = X \cup (-1)X$. Note that $f(X') = f(X) \subseteq X$. Let $G = PB_X$ denote the group of pure braids based at the points in X. Say $|X| = n$, so that $G \simeq PB_n$. Call the strand based at the point c the nth strand, and label the remaining strands $1, \ldots, n-1$. Consider an oriented closed curve based at a point in $X - \{c\}$ and not passing through any other points in X. This corresponds to a braid in G (we may parametrize the curve so that it gives a path, and treat the other strands as moving straight down). When we take its inverse image under f, we get either two oriented closed curves based at points in X' or one oriented closed curve passing through two points of X' and mapping by degree two. These curves induce a permutation of the points of X' in the natural way. This gives rise to a homomorphism $\rho : G \to S_{X'}$, where $S_{X'}$ denotes the symmetric group on X', called the monodromy homomorphism.

In general, pure braid generators A_{ij} which do not have strands running around the singularity c will be sent under ρ to the identity, while generators A_{in} will be sent to transpositions of the form $(z\ -z)$, where $f(z)$ is the base point of the ith strand. See figures B.5, B.8, B.9, and B.10.[2]

[2]These figures were created using the OTIS applet at http://www.math.nagoya-u.ac.jp/ kawahira/programs/otis.html.

We can give the image of ρ explicitly. Since each generator of the form A_{ij} where $j \neq n$ is sent to the identity, we need only consider the image of generators of the form A_{in}. Suppose that $\rho(A_{i_1 n}) = (z_1 \ -z_1)$ and $\rho(A_{i_2 n}) = (z_2 \ -z_2)$. If $z_1 = \pm z_2$, then $f(z_1) = f(z_2)$ is the base point of both the (i_1)th and (i_2)th strands, and thus $A_{i_1 n} = A_{i_2 n}$. So each A_{in} permutes a pair of elements of X' which are not affected by any other generators. There are $n-1$ generators of the form A_{in}, so we have $\mathrm{im}(\rho) \simeq \mathbb{Z}_2^{n-1}$.

We are interested in pure braids which are taken under f^{-1} to other pure braids; that is, elements in the kernel of ρ. We have $A_{ij} \in \ker(\rho)$ for $i < j < n$, and $A_{in}^2 \in \ker(\rho)$ for $i < n$; since $\ker(\rho) \triangleleft G$, all conjugates of these are also in $\ker(\rho)$. We can give the generators for $\ker(\rho)$ explicitly.

Theorem B.1.1. *Let $f : \mathbb{C} \to \mathbb{C}$ be the quadratic function $f(z) = z^2 + c$, where c is some complex constant such that the orbit of 0 is finite. Let X be this orbit, let $n = |X|$, let $X' = X \cup (-1)X$, and let $G = PB_X \simeq PB_n$ be the group of pure braids based at X. Call the strand based at the point c the nth strand, and label the remaining strands $1, \ldots, n-1$. Let $\rho : G \to S_{X'}$ be the monodromy homomorphism which sends the pure braid A to the permutation induced by $f^{-1}(A)$, where we consider our standard projection of A. Then $H = \ker(\rho)$ is generated by the set of elements of the following form:*

$$\begin{cases} A_{ij} & \text{where } i < j < n \\ A_{in}^2 & \text{where } i < n \\ A_{in} A_{ij} A_{in}^{-1} & \text{where } i < j < n. \end{cases}$$

Proof. Let S be this set of elements and let K be the subgroup generated by S. The proof is in two parts. The first shows that all conjugates of these elements are in K, that is, that K is a normal subgroup. We certainly have $K \subseteq H$, so the second shows that all elements of H are elements of K, that is, that $H \subseteq K$ and thus that $H = K$.

First, we show that K is a normal subgroup. It suffices to show that conjugating each element of S by each generator of G again gives an element of K, since if $s \in S$, $xsx^{-1} = s_1 \cdots s_k$ where each $s_i \in S$ or $s_i^{-1} \in S$, and $g \in G$, then $gxsx^{-1}g^{-1} = gs_1 \cdots s_k g^{-1} = gs_1 g^{-1} g \cdots g^{-1} g s_k g^{-1}$. Each $gs_i g^{-1}$ can then be reduced once again to a product of elements of S or their inverses.

We divide the argument into cases. Most of the cases depend on strings of calculations using the relations $R_1, \ldots, R_4$, which I have omitted.

- **Conjugates of A_{ij}, where $j \neq n$**
  We want to show that elements $A_{rs} A_{ij} A_{rs}^{-1}$ are products of elements of S. We divide this further into subcases depending on the values of r and s.
  - $r, s \in [1, n)$
    The element $A_{rs} A_{ij} A_{rs}^{-1}$ is trivially in K, since $A_{rs}, A_{ij} \in S$.
  - $r \in [1, i) \cup (j, n), s = n$
    By R_1, the elements A_{rs} and A_{ij} commute, so $A_{rs} A_{ij} A_{rs}^{-1} = A_{ij} \in K$.
  - $r = i, s = n$
    The element $A_{rs} A_{ij} A_{rs}^{-1}$ is in K by construction.
  - $r \in (i, j), s = n$
    By a long series of calculations, we have $A_{rs} A_{ij} A_{rs}^{-1} =$
    $(A_{in}^{-2})(A_{in} A_{ij} A_{in}^{-1})(A_{in}^2)(A_{rj}^{-1})(A_{rn}^{-2})(A_{rn} A_{rj} A_{rn}^{-1})(A_{rj}^2)(A_{in}^{-2})$
    $(A_{in} A_{ij}^{-1} A_{in}^{-1})(A_{ij})(A_{in} A_{ir}^{-1} A_{in}^{-1})(A_{rj}^{-1})(A_{in} A_{ij} A_{in}^{-1})(A_{in}^2)(A_{ir})$
    $(A_{in}^{-2})(A_{in} A_{ij}^{-1} A_{in}^{-1})(A_{in}^2)(A_{ij})(A_{ir}^{-1})(A_{rn}^{-2})(A_{in}^{-2})(A_{in} A_{ir} A_{in}^{-1})$

$(A_{in}^{-2})(A_{rn}A_{rj}^{-1}A_{rn}^{-1})(A_{in}^{-2})(A_{in}A_{ir}^{-1}A_{in}^{-1})(A_{in}^{2})(A_{rn}^{-2})(A_{ir})(A_{rj})$
$(A_{ir}^{-1})(A_{in}^{-2})(A_{in}A_{ir}A_{in}^{-1})(A_{in}^{2})$, a product of elements of S and their inverses and thus in K.

– $r = j, s = n$
By a shorter set of calculations we have $A_{rs}A_{ij}A_{rs}^{-1} = (A_{in}^{-2})(A_{in}A_{ij}A_{in}^{-1})(A_{in}^{2})$, a product of elements of S and their inverses and thus in K.

- **Conjugates of** A_{in}^2
We want to show that elements $A_{rs}A_{in}^2A_{rs}^{-1}$ are products of elements of S. We divide this into subcases depending on the values of r and s.

– $r, s \in [1, n)$
The element $A_{rs}A_{in}^2A_{rs}^{-1}$ is trivially in K, since $A_{rs}, A_{in}^2 \in S$.

– $r \in [1, i), s = n$
By another brief set of calculations, we have $A_{rs}A_{in}^2A_{rs}^{-1} = (A_{ri}^{-1})(A_{in}^{2})(A_{ri})$, a product of elements of S and thus in K.

– $r = i, s = n$
In this case we have $A_{rs}A_{in}^2A_{rs}^{-1} = A_{in}^2 \in K$.

– $r \in (i, n), s = n$
By another brief set of calculations we have $A_{rs}A_{in}^2A_{rs}^{-1} = (A_{in}^{-2})(A_{in}A_{ir}^{-1}A_{in}^{-1})(A_{in}^{2})(A_{in}A_{ir}A_{in}^{-1})(A_{in}^{2})$, a product of elements of S and their inverses and thus in K.

- **Conjugates of** $A_{in}A_{ij}A_{in}^{-1}$**, where** $j \neq n$
We want to show that the elements $A_{rs}A_{in}A_{ij}A_{in}^{-1}A_{rs}^{-1}$ are products of elements of S. Once again we divide into subcases depending on the values of r and s.

– $r, s \in [1, n)$
As before, in this case the element $A_{rs}A_{in}A_{ij}A_{in}^{-1}A_{rs}^{-1}$ is trivially in K, since $A_{rs}, A_{in}A_{ij}A_{in}^{-1} \in S$.

– $r \in [1, i), s = n$
By another calculation we have $A_{rs}A_{in}A_{ij}A_{in}^{-1}A_{rs}^{-1} = (A_{ri}^{-1})\,(A_{rn}^{-1}A_{ri}A_{rn})(A_{in}A_{ij}A_{in}^{-1})(A_{rn}^{-1}A_{ri}^{-1}A_{rn})(A_{ri})$, a product of elements of S and their inverses and thus in K.

– $r = i, s = n$
We have $A_{rs}A_{in}A_{ij}A_{in}^{-1}A_{rs}^{-1} = (A_{in}^{2})(A_{ij})(A_{in}^{-2}) \in K$.

– $r \in (i, j), s = n$
By another hideous calculation we have $A_{rs}A_{in}A_{ij}A_{in}^{-1}A_{rs}^{-1} =$
$(A_{in}^{-2})(A_{in}A_{ir}^{-1}A_{in}^{-1})(A_{in}^{2})(A_{ir})(A_{ij})(A_{rj}^{-1})(A_{ir}^{-1})(A_{rn}^{-2})(A_{ir})$
$(A_{in}A_{rn}A_{in}^{-1})(A_{rj})(A_{in}A_{rn}^{-1}A_{in}^{-1})(A_{rj}^{2})(A_{ij}^{-1})(A_{in}^{-2})(A_{in}A_{ij}A_{in}^{-1})$
$(A_{in}^{2})(A_{ir}^{-1})(A_{in}^{-2})(A_{rj}^{-1})(A_{in}^{2})(A_{ij})(A_{in}A_{ir}A_{in}^{-1})(A_{in}^{-2})(A_{ij}^{-1})$
$(A_{in}A_{ij}A_{in}^{-1})(A_{in}A_{ir}^{-1}A_{in}^{-1})(A_{ir}^{-1})(A_{rn}^{-2})(A_{ir})(A_{ir})(A_{in}^{-4})$
$(A_{in}A_{rn}A_{in}^{-1})(A_{rj}^{-1})(A_{in}A_{rn}^{-1})(A_{in}^{-1})(A_{in}^{-2})(A_{in}^{2})(A_{ir}^{-2})(A_{rn}^{-2})(A_{ir})$
$(A_{in}A_{ir}A_{in}^{-1})(A_{rj})(A_{in}A_{ir}^{-1}A_{in}^{-1})(A_{in}^{-2})(A_{in}A_{ir}A_{in}^{-1})(A_{in}^{2})$, a product of elements of S and their inverses and thus in K.

- $r = j, s = n$
  By yet another calculation we have $A_{rs}A_{in}A_{ij}A_{in}^{-1}A_{rs}^{-1} =$
  $(A_{in}^{-2})(A_{in}A_{ij}^{-1}A_{in}^{-1})(A_{in}^{2})(A_{ij})(A_{in}^{-2})(A_{in}A_{ij}A_{in}^{-1})(A_{in}^{2})$, a product of elements of S and their inverses and thus in K.
- $r \in (j, n), s = n$
  Finally, we have $A_{rs}A_{in}A_{ij}A_{in}^{-1}A_{rs}^{-1} = (A_{in}^{-2})(A_{in}A_{ir}^{-1}A_{in}^{-1})(A_{in}^{2})$
  $(A_{ir})(A_{in}A_{ij}A_{in}^{-1})(A_{ir}^{-1})(A_{in}^{-2})(A_{in}A_{ir}A_{in}^{-1})(A_{in}^{2})$, a product of elements of elements of S and their inverses and thus an element of K.

So every element of S, when conjugated by any generator and thus any element of G, remains in K; so $K \lhd G$.

Now we show that $H \subseteq K$. First, recall that the image of ρ is isomorphic to $\mathbb{Z}_2^{n-1}$, and that every generator of the form A_{in} is mapped under ρ to a transposition. We have $G/H \simeq \mathrm{im}(\rho)$, so the cosets of H correspond exactly to ordered $(n-1)$-tuples of 0s and 1s. There is a 1 in the ith position for the coset xH whenever the generator A_{in} appears an odd number of times as a factor of the coset representative x. (Note that the inverses of A_{in} may be either added or subtracted, since we are counting modulo 2.) This also tells us that any product of the generators A_{ij} is in H if each generator A_{in} appears an even number of times and each generator A_{ij} where $j \neq n$ appears any number of times.

We will use an inductive argument on the length of an element in the generators A_{ij}. Let $|g|$ denote this length, so that if $g = A_{i_1j_1}\cdots A_{i_kj_k}$, then $|g| = k$, and say that $|1| = 0$. Suppose $g \in H$. We want to show that $g \in K$.

There are four base cases to consider. If $|g| = 0$, then we have $g = 1 \in K$. If $|g| = 1$, then g must be of the form $g = A_{ij}$ where $j \neq n$, and thus $g \in K$. If $|g| = 2$, then g is either of the form $g = A_{i_1j_1}A_{i_2j_2}$, where $j_1, j_2 \neq n$, or of the form $g = A_{in}^2$, both in K. If $|g| = 3$, then g may be of any form in which each generator A_{in} appears an even number of times; for example, we may have $g = A_{in}A_{jk}A_{in} = (A_{in}A_{jk}A_{in}^{-1})(A_{in}^2)$. The first part of the proof gives us that $(A_{in}A_{jk}A_{in}^{-1}) \in K$, and so $g \in K$. The other cases are similar.

Now suppose that for any $h \in H$ with $|h| \leq m-1$ we have $h \in K$. Let $g \in H$ be a word of length m. We wish to show that $g \in K$.

Write $g = A_{i_1j_1}^{\pm 1}\cdots A_{i_mj_m}^{\pm 1}$. Let x be the first factor $A_{i_1j_1}^{\pm 1}$ and let y be the second factor $A_{i_2j_2}^{\pm 1}$. If $j_1 \neq n$, then we are done, for $g, x \in H$ and so $A_{i_2j_2}^{\pm 1}\cdots A_{i_mj_m}^{\pm 1}$, a word of length $m-1$, must also be in H and thus in K. So consider the case when $j_1 = n$. Similarly, if $j_2 \neq n$, then $g = xyx^{-1}xA_{i_3j_3}^{\pm 1}\cdots A_{i_mj_m}^{\pm 1}$, we have $g, xyx^{-1} \in H$, and thus $xA_{i_3j_3}^{\pm 1}\cdots A_{i_mj_m}^{\pm 1}$, a word of length $m-1$, is also in H and thus in K. So consider the case when $j_2 = n$.

Because the element g is in H, the factors x and y must appear an even number of times and hence at least once more. We will consider the two cases when x appears before y and when x appears after y separately.

Suppose x appears as the factor $A_{i_kj_k}^{\pm 1}$ and y appears as the factor $A_{i_lj_l}^{\pm 1}$. Let $c = A_{i_3j_3}^{\pm 1}\cdots A_{i_{k-1}j_{k-1}}^{\pm 1}$, $d = A_{i_{k+1}j_{k+1}}^{\pm 1}\cdots A_{i_{l-1}j_{l-1}}^{\pm 1}$, and $e = A_{i_{l+1}j_{l+1}}^{\pm 1}\cdots A_{i_mj_m}^{\pm 1}$; that is, we have $g = xycx^{\pm 1}dy^{\pm 1}e$. If we remove the factors x and y, the remaining element cde must still be in H, since each factor A_{in} still occurs an even number of times. So the elements cd and e must be in the same coset uH. Recall that if the element A_{in} appears as a factor of the coset representative u, then it must appear an odd number of times, and thus at least once, in each of the elements cd and e. Thus, we must have $|u| \leq |cd|$ and $|u| \leq |e|$. Write $e = ue'$, where $e' \in H$.

If $cd \in uH$, then $cx^{\pm 1}d \in xuH$; but left and right cosets of normal subgroups are equal, so we can write $cx^{\pm 1}d \in xHu$. Let $cx^{\pm 1}d = xku$, where $k \in H$.

Now we can write

$$\begin{aligned} g &= xyxkuy^{\pm1}ue' \\ &= (xyxy)(y^{-1}ky)(y^{-1}uy^{\pm1}u)(e'). \end{aligned} \tag{B.1}$$

Each of the parenthesized factors is in H. We have

$$\begin{aligned} xyxy &= A^{\pm1}_{i_k n}A^{\pm1}_{i_l n}A^{\pm1}_{i_k n}A^{\pm1}_{i_l n} \\ &= A^{\pm1}_{i_k n}A^{\pm1}_{i_l n}A^{\mp1}_{i_k n}A^{\pm2}_{i_k n}A^{\pm1}_{i_l n} \\ &= \begin{cases} A^{-1}_{i_k i_l}A^{\pm1}_{i_l n}A_{i_k i_l}A^{2}_{i_k n}A^{\pm1}_{i_l n}, \text{ by } R_2 & \text{if } x = A_{i_k n} \text{ and } i_k < i_l \\ A^{-2}_{i_k n}A^{-1}_{i_k i_l}A^{\pm1}_{i_l n}A_{i_k i_l}A^{2}_{i_k n}A^{-2}_{i_k n}A^{\pm1}_{i_l n}, \text{ by } R_2 & \text{if } x = A^{-1}_{i_k n} \text{ and } i_k < i_l \\ A^{-1}_{i_l n}A^{-1}_{i_l i_k}A^{\pm1}_{i_l n}A_{i_l i_k}A_{i_l n}A^{2}_{i_k n}A^{\pm1}_{i_l n}, \text{ by } R_3 & \text{if } x = A_{i_k n} \text{ and } i_l < i_k \\ A^{-2}_{i_k n}A^{-1}_{i_l n}A^{-1}_{i_l i_k}A^{\pm1}_{i_l n}A_{i_l i_k}A_{i_l n}A^{2}_{i_k n}A^{-2}_{i_k n}A^{\pm1}_{i_l n}, \text{ by } R_3 & \text{if } x = A^{-1}_{i_k n} \text{ and } i_l < i_k \end{cases} \\ &= \begin{cases} (A^{-1}_{i_k i_l})(A^{\pm1}_{i_l n}A_{i_k i_l}A^{\mp1}_{i_l n})(A^{\pm1}_{i_l n}A^{2}_{i_k n}A^{\pm1}_{i_l n}) & \text{if } x = A_{i_k n} \text{ and } i_k < i_l \\ (A^{-2}_{i_k n})(A^{-1}_{i_k i_l})(A^{\pm1}_{i_l n}A_{i_k i_l}A^{\mp1}_{i_l n})(A^{\pm1}_{i_l n}A^{2}_{i_k n}A^{-2}_{i_k n}A^{\pm1}_{i_l n}) & \text{if } x = A^{-1}_{i_k n} \text{ and } i_k < i_l \\ (A^{-2}_{i_l n})(A_{i_l n}A^{-1}_{i_l i_k}A^{-1}_{i_l n})(A^{1\pm1}_{i_l n})(A_{i_l i_k})(A_{i_l n}A^{2}_{i_k n}A^{-1}_{i_l n})(A^{1\pm1}_{i_l n}) & \text{if } x = A_{i_k n} \text{ and } i_l < i_k \\ (A^{-2}_{i_k n})(A^{-2}_{i_l n})(A_{i_l n}A^{-1}_{i_l i_k}A^{-1}_{i_l n})(A^{1\pm1}_{i_l n})(A_{i_l i_k})(A^{1\pm1}_{i_l n}) & \text{if } x = A^{-1}_{i_k n} \text{ and } i_l < i_k \end{cases} \end{aligned}$$

Each parenthesized factor is a conjugate of an element of K and hence in K. We have $k \in H$ and

$$|k| \le |x^{-1}cx^{\pm1}du| \le |x| + |c| + |x| + |d| + |u| \le |x| + |c| + |x| + |d| + |e| = m - 2,$$

so $k \in K$ and thus $y^{-1}ky \in K$. We have $y^{-1}uy^{\pm1}u \in H$ and

$$|y^{-1}uy^{\pm1}u| \le |y| + |u| + |y| + |u| \le |y| + |cd| + |y| + |e| = m - 2,$$

so $y^{-1}uy^{\pm1}u \in K$. And we have $e' \in H$ and

$$|e'| \le |u^{-1}e| \le |u| + |e| \le |cd| + |e| = m - 4,$$

so $e' \in K$. Thus all of the parenthesized factors in (B.1) are in K, so we have $g \in K$.

The other case is similar. Suppose y appears as the factor $A^{\pm1}_{i_k j_k}$ and x appears as the factor $A^{\pm1}_{i_l j_l}$. Let $c = A^{\pm1}_{i_3 j_3} \cdots A^{\pm1}_{i_{k-1} j_{k-1}}$, $d = A^{\pm1}_{i_{k+1} j_{k+1}} \cdots A^{\pm1}_{i_{l-1} j_{l-1}}$, and $e = A^{\pm1}_{i_{l+1} j_{l+1}} \cdots A^{\pm1}_{i_m j_m}$; that is, we have $g = xycy^{\pm1}dx^{\pm1}e$. Once again we must have cd and e in the same coset uH, so $|u| \le |cd|$ and $|u| \le |e|$. Moreover we have $ycy^{\pm1}d \in uH = Hu$, and we write $ycy^{\pm1}d = ku$, where $k \in H$, and $e = ue'$, where $e' \in H$.

So we have

$$\begin{aligned} g &= xkux^{\pm1}ue' \\ &= (xkx^{-1})(xux^{\pm1}u)(e'). \end{aligned} \tag{B.2}$$

We have $k \in H$ and

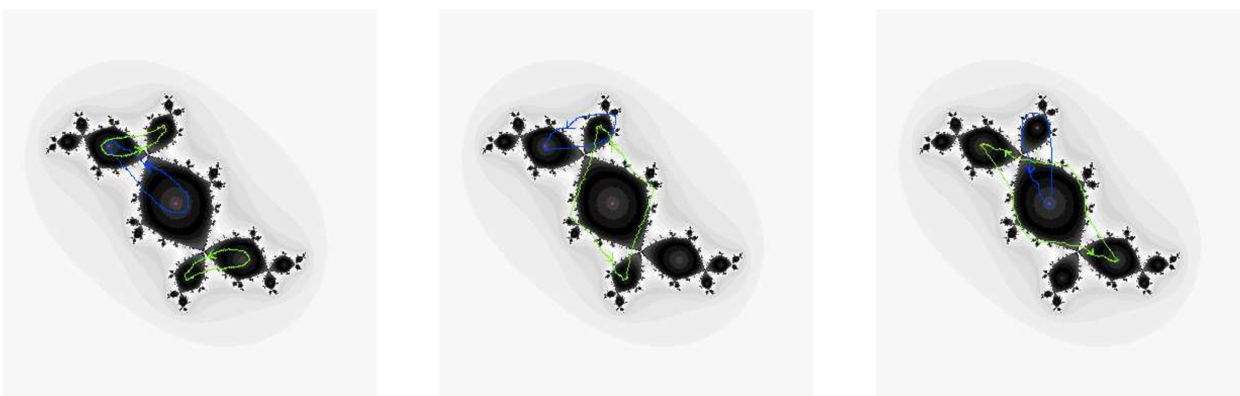
$$|k| \le |ycy^{\pm1}du^{-1}| \le |y| + |c| + |y| + |d| + |u| \le |y| + |c| + |y| + |d| + |e| = m - 2,$$

so k and thus xkx^{-1} are in K. The case for $xux^{\pm1}u$ is almost identical to the case for $y^{-1}uy^{\pm1}u$ above, and the case for e' is the same as before. Thus, each of the parenthesized factors in (B.2) is in K, so $g \in K$. Therefore we have $H \subseteq K$, and thus $H = K$. $\square$

We now turn to the task of constructing homomorphisms between pure braid groups. We know that the inverse image under f of a pure braid in H is a pure braid in $PB_{X'}$. Let $\tilde{\phi}_f : H \to PB_{X'}$ be the homomorphism that carries an element of H to its inverse image under f. Let $\iota_* : PB_{X'} \to PB_X$ be the "forgetful map" which loses track of the strands based at points in $X' - X$. This map is also a homomorphism. Recalling that $G = PB_X$, let $\phi_f : H \to G$ be the homomorphism $\phi_f = \iota_* \circ \tilde{\phi}_f$. A homomorphism is defined by its action on the generators of its domain, so we move to our examples to give ϕ_f explicitly.

B.2 The Rabbit; or, the 0-3 Case

For the braids on three strands, let $a = A_{12}, b = A_{13}, c = A_{23}$. Let $f(z) = z^2 + c_R$, where $c_R \approx -.122561 + .744862i$. Under this function, the point 0 is periodic of period 3. Figure B.5 shows the action of f^{-1} on the pure braid group generators. To illustrate how we evaluate ϕ_f on the generators of H, consider the

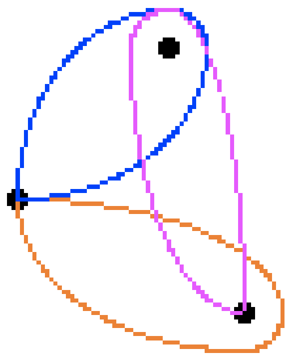


(a) A_{12} and $f^{-1}(A_{12})$ (b) A_{13} and $f^{-1}(A_{13})$ (c) A_{23} and $f^{-1}(A_{23})$

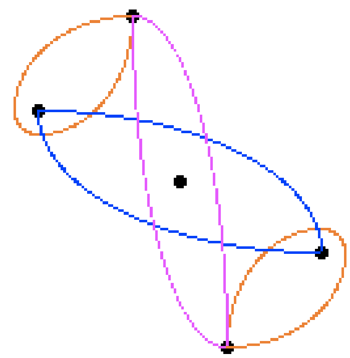
Figure B.5: The pure braid group generators and their inverse images under f for $f(z) = z^2 + c_R$. The blue curves are the pure braid group generators, and the green curves are the inverse images of these generators. The first, second, and third strands are, respectively, those based at the points $c_R^2 + c_R, (c_R^2 + c_R)^2 + c_R, c_R$.

progression shown in figure B.6 for the element b^2 and in figure B.7 for the element bab^{-1}. That is, we evaluate $\tilde{\phi}_f$ on elements of H by tracing the paths along their inverse images. We then use ι_* to ignore the strands based at $-c_R$ and $-c_R^2 - c_R$. Proceeding in this way for all of the generators of H, we find that

$$\begin{aligned}
\phi_f(a) &= b \\
\phi_f(b^2) &= c \\
\phi_f(c^2) &= a \\
\phi_f(bab^{-1}) &= 1
\end{aligned}$$

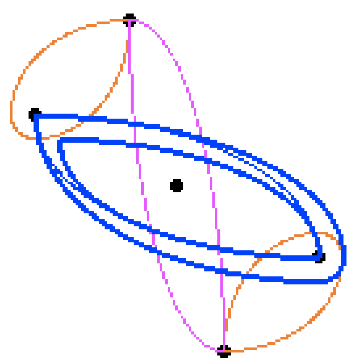


(a) As a reminder, the pure braid group generators...

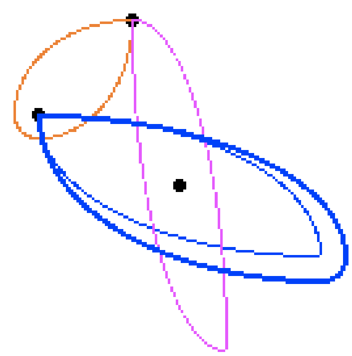


(b) ...and their inverse images under f

(c) We start by tracing the curve for b twice...

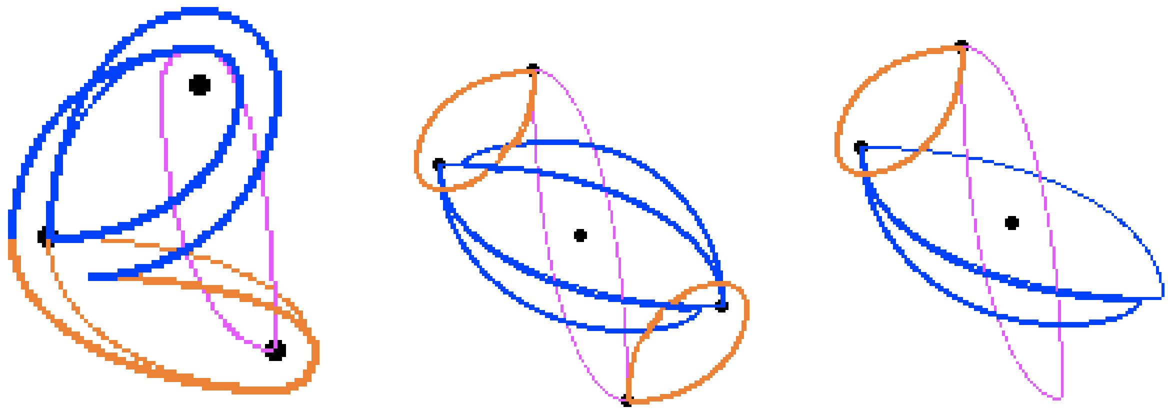


(d) ...then apply $\tilde{\phi}_f$, tracing the inverse image of b twice...



(e) ...then apply ι_*, forgetting the strands from extraneous points to themselves

Figure B.6: An illustration of the evaluation of ϕ_f on the generator b. The orange curve is a, the blue curve is b, and the purple curve is c. All curves are oriented counter-clockwise.



(a) We start by tracing the curves for bab^{-1}...

(b) ...then apply $\tilde{\phi}_f$, tracing the inverse image of these curves...

(c) ...then apply ι_*, forgetting the strands from extraneous points to themselves. These strands can be pulled straight—so $\phi_f(bab^{-1}) = 1$.

Figure B.7: An illustration of the evaluation of ϕ_f on the generator bab^{-1}. The orange curve is a, the blue curve is b, and the purple curve is c. All curves are oriented counter-clockwise.

We note that ϕ_f is surjective and not injective. We also examine where ϕ_f sends the center of G. In general, the element D^2 is not in H, but the element D^4 is in H. On three strands we have

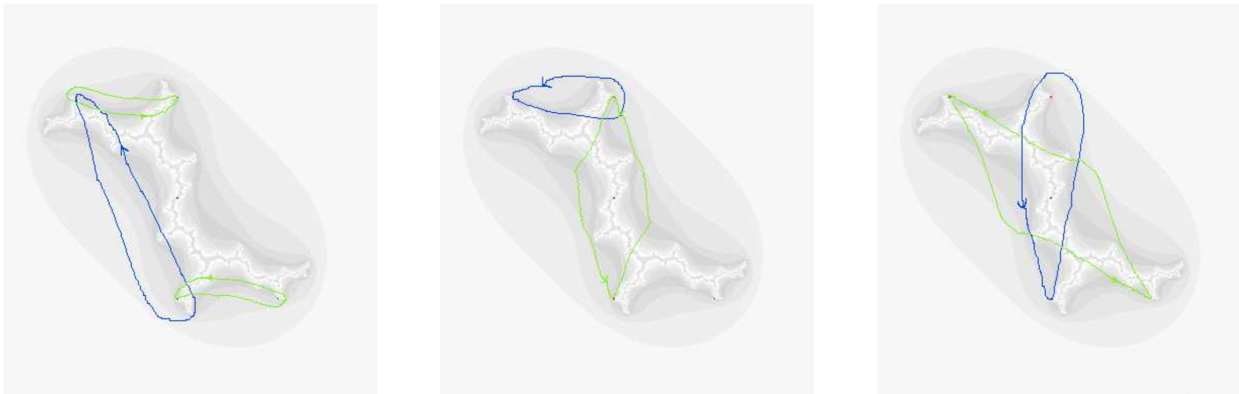
$$\begin{aligned} D^4 &= (abc)^2 \\ &= abcbca \\ &= aa^{-1}babc^2a \\ &= b^2b^{-1}abc^2a \end{aligned}$$

so that $\phi_f(D^4) = D^2$ for this case.

B.3 i; or, the 2-2 Case

Let $f(z) = z^2 + i$. Under this function, the point 0 is enters a 2-cycle after two iterations. Figure B.8 shows the action of f^{-1} on the pure braid group generators. For this case we will replace the generator bab^{-1} with the generator $b^{-1}ab$. As with the rabbit, we evaluate ϕ_f on the generators of H by tracing inverse images and discarding the strands based at points outside of the forward orbit of 0. In this case we find that

$$\begin{aligned} \phi_f(a) &= b \\ \phi_f(b^2) &= c \\ \phi_f(c^2) &= 1 \\ \phi_f(b^{-1}ab) &= a \end{aligned}$$



(a) A_{12} and $f^{-1}(A_{12})$ (b) A_{13} and $f^{-1}(A_{13})$ (c) A_{23} and $f^{-1}(A_{23})$

Figure B.8: The pure braid group generators and their inverse images under f for $f(z) = z^2 + i$. The blue curves are the pure braid group generators, and the green curves are the inverse images of these generators. The first, second, and third strands are, respectively, those based at the points $i - 1, -i, i$.

Once again, we have ϕ_f surjective and not injective, and $\phi_f(D^4) = \phi_f(b^2b^{-1}abc^2a) = cab = D^2$.

B.4 The 3-1 Case

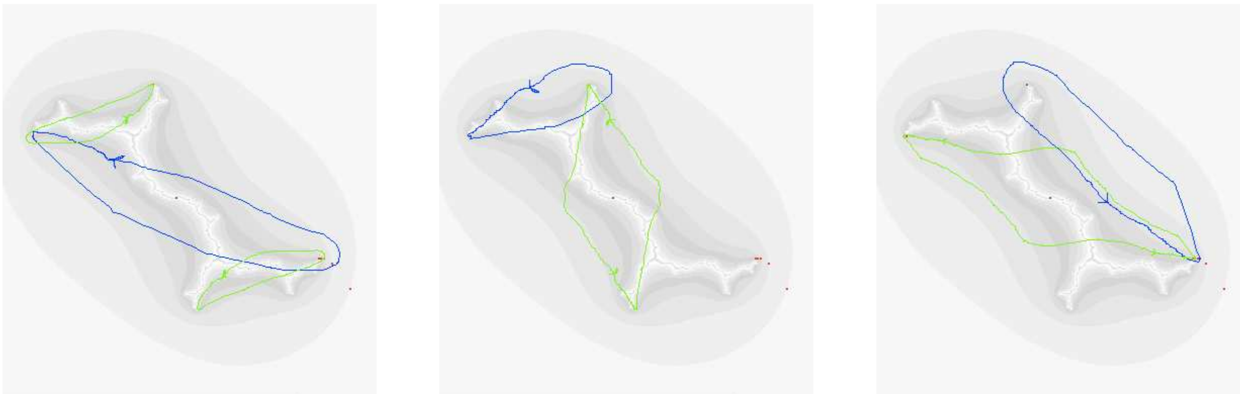
Let $f(z) = z^2 + c_3$, where $c_3 \approx -.228155 + 1.11514i$. In this case the point 0 enters a 1-cycle after three iterations of f. Figure B.9 shows the action of f^{-1} on the pure braid group generators. Evaluating ϕ_f on the generators of H in the same way gives

$$\begin{aligned} \phi_f(a) &= b \\ \phi_f(b^2) &= 1 \\ \phi_f(c^2) &= a \\ \phi_f(bab^{-1}) &= c \end{aligned}$$

Once again: ϕ_f is surjective and not injective, and $\phi_f(D^4) = \phi_f(b^2b^{-1}abc^2a) = cab = D^2$.

B.5 The 0-4 Case

Let $f(z) = z^2 + c_4$, where $c_4 \approx -.15652 + 1.03225i$. In this case the point 0 is periodic of period 4. Figure B.10 shows the action of f^{-1} on the pure braid group generators. In this case we have



(a) A_{12} and $f^{-1}(A_{12})$ (b) A_{13} and $f^{-1}(A_{13})$ (c) A_{23} and $f^{-1}(A_{23})$

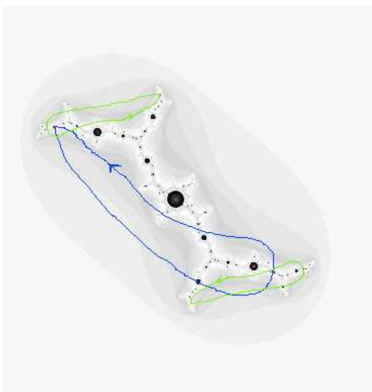
Figure B.9: The pure braid group generators and their inverse images under f for $f(z) = z^2 + c_3$. The blue curves are the pure braid group generators, and the green curves are the inverse images of these generators. The first, second, and third strands are, respectively, those based at the points $c_3^2 + c_3, (c_3^2 + c_3)^2 + c_3, c_3$.

$$
\begin{aligned}
\phi_f(A_{12}) &= A_{14} \\
\phi_f(A_{13}) &= 1 \\
\phi_f(A_{23}) &= 1 \\
\phi_f(A_{14}^2) &= A_{34} \\
\phi_f(A_{24}^2) &= A_{13} \\
\phi_f(A_{34}^2) &= A_{23} \\
\phi_f(A_{14}A_{12}A_{14}^{-1}) &= 1 \\
\phi_f(A_{14}A_{13}A_{14}^{-1}) &= A_{24} \\
\phi_f(A_{24}A_{23}A_{24}^{-1}) &= A_{12}
\end{aligned}
$$

Once again, ϕ_f is surjective and not injective.

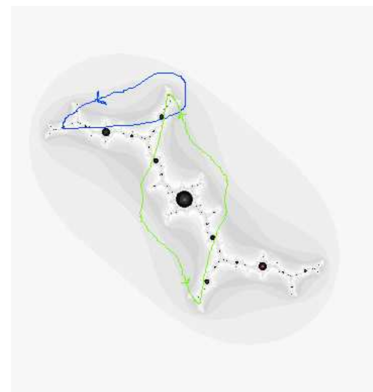
B.6 Extensions

We are interested in using our braid homomorphism to tell us more about the long-term behavior of simple closed curves under repeated iteration of f^{-1}. One principle we would like to verify is that of contraction of word length; that is, given a word of a certain length in the generators of G, we would like to show that after a certain number of iterations of ϕ_f the length of the resulting word has decreased by some definite factor. The action of ϕ_f on the generators of H certainly suggests this; the generator A_{12} is the only one in each case we have examined whose length is not immediately decreased by application of ϕ_f. However, verifying this property is not as simple as verifying it on the generators, since, as was suggested above in proving the generating set for H, writing an element which is in H as a product of generators of H has the potential to dramatically increase word length.

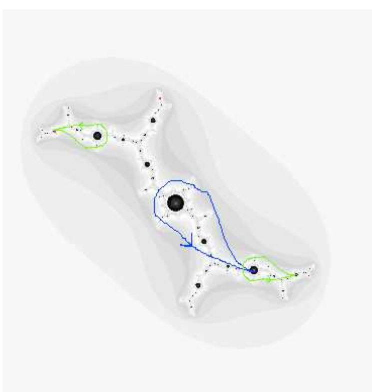


(a) A_{12} and $f^{-1}(A_{12})$

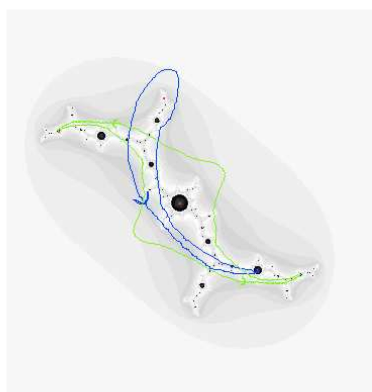
(b) A_{13} and $f^{-1}(A_{13})$



(c) A_{14} and $f^{-1}(A_{14})$



(d) A_{23} and $f^{-1}(A_{23})$



(e) A_{24} and $f^{-1}(A_{24})$

(f) A_{34} and $f^{-1}(A_{34})$

Figure B.10: The pure braid group generators and their inverse images under f for $f(z) = z^2 + c_4$. The blue curves are the pure braid group generators, and the green curves are the inverse images of these generators. The first, second, third, and fourth strands are, respectively, those based at the points $c_4^2 + c_4, (c_4^2 + c_4)^2 + c_4, ((c_4^2 + c_4)^2 + c_4)^2 + c_4, c_4$.

Thus we introduce two extensions which have the potential for use in verifying the contraction property. The first is an extension of ϕ_f. Consider the case where $n = 3$. The domain of ϕ_f is a subgroup of index 4 of G. We extend the homomorphism ϕ_f to a map $\overline{\phi_f} : G \to G$ as follows:

$$\overline{\phi_f}(w) = \begin{cases} \phi_f(w) & \text{if } w \in H \\ \phi_f(b^{-1}w) & \text{if } w \in bH \\ \phi_f(c^{-1}w) & \text{if } w \in cH \\ \phi_f(c^{-1}b^{-1}w) & \text{if } w \in bcH. \end{cases}$$

Note that $\overline{\phi_f}$ is not a homomorphism.

The second extension is an expanded form of the elements of G. Label the cosets H, bH, cH, bcH 1 through 4, respectively. Let $\rho' : G \to S_4$ give the permutation of the cosets induced by an element of G. Then the expanded form is

$$g = \rho'(g)(\overline{\phi_f}(g), \overline{\phi_f}(gb), \overline{\phi_f}(gc), \overline{\phi_f}(gbc)).$$

Consider a set containing the generators of G, their inverses, and any product of two of these. For each element s in this set, examine its expanded form. If any of the elements $\overline{\phi_f}(s), \overline{\phi_f}(sb), \overline{\phi_f}(sc)\overline{\phi_f}(sbc)$ are not in the set, add them and repeat. If eventually the set closes off, then we can show the contraction property.

We give the expanded forms of the generators for our three examples where $n = 3$.

- For $f(z) = z^2 + c_R$, we have

$$\begin{aligned} a &= (b, 1, 1, b) \\ b &= (1\ 2)(3\ 4)(1, c, 1, a^{-1}ca) \\ c &= (1\ 3)(2\ 4)(1, b, a, ab^{-1}) \\ a^{-1} &= (b^{-1}, 1, 1, b^{-1}) \\ b^{-1} &= (1\ 2)(3\ 4)(c^{-1}, 1, a^{-1}c^{-1}a, 1) \\ c^{-1} &= (1\ 3)(2\ 4)(a^{-1}, ba^{-1}, 1, b^{-1}) \end{aligned}$$

- For $f(z) = z^2 + i$, we have

$$\begin{aligned} a &= (b, a, a, b) \\ b &= (1\ 2)(3\ 4)(1, c, 1, a^{-1}ca) \\ c &= (1\ 3)(2\ 4)(1, ba^{-1}, 1, ab^{-1}) \\ a^{-1} &= (b^{-1}, a, a^{-1}, b^{-1}) \\ b^{-1} &= (1\ 2)(3\ 4)(c^{-1}, 1, a^{-1}c^{-1}a, 1) \\ c^{-1} &= (1\ 3)(2\ 4)(1, ba^{-1}, 1, ab^{-1}) \end{aligned}$$

- For $f(z) = z^2 + c_3$, we have

$$\begin{aligned} a &= (b, c, a^{-1}ca, b) \\ b &= (1\ 2)(3\ 4)(1, 1, 1, 1) \\ c &= (1\ 3)(2\ 4)(1, bc^{-1}, a, cab^{-1}) \\ a^{-1} &= (b^{-1}, c, a^{-1}c^{-1}a, b^{-1}) \\ b^{-1} &= (1\ 2)(3\ 4)(1, 1, 1, 1) \\ c^{-1} &= (1\ 3)(2\ 4)(a^{-1}, ba^{-1}, 1, cb^{-1}) \end{aligned}$$

We hope that further work in this direction will give us the desired property.

Bibliography

1. E. Artin, Theory of braids, *Ann. of Math. (2)* **48** (1947), 101–126

2. Joan S. Birman, *Braids, Links, and Mapping Class Groups*, Annals of Mathematical Studies, no. 82, Princeton University Press, Princeton, NJ, 1974, ix+228

3. Dan Margalit and Jon McCammond, Geometric presentations for the pure braid group, to appear, *J. Knot Theory Ramifications*. Available at http://arxiv.org/abs/math/0603204

16×16 **Hadamard matrices and their codes**

Michael J. Grilli
Providence College

Indiana University REU Summer 2008
Advisor: ADVISOR

C.1 Preliminary Definitions in Coding Theory

Let $\mathbb{F}_2$ denote the finite field having two elements, namely $\{0,1\}$. The vector space $\mathbb{F}_2^n$ consists of all length n vectors with entries either 0 or 1. We define a **binary linear code**, $\mathcal{C}[n,k]$, to be a k dimensional subspace of $\mathbb{F}_2^n$, in which the vectors in $\mathcal{C}[n,k]$ are called **codewords**. We define the **weight** of a codeword to be the number of places in which the codeword has a nonzero entry. To see this, let's consider an example in $\mathbb{F}_2^8$; let $u = (1,0,0,1,0,0,1,1) \in \mathbb{F}_2^8$. Clearly, the weight of u, denoted $w(u)$, is 4. A code in which every codeword has weight divisible by four is said to be **doubly even**.

For any given code $\mathcal{C}[n,k]$ we define its **dual**, $\mathcal{C}^{\perp}[n,n-k]$, to be the set of vectors $v \in \mathbb{F}_2^n$ that are orthogonal to every codeword in $\mathcal{C}[n,k]$. The concept of orthogonality is the same as always, except since we are in $\mathbb{F}_2$, operations are performed modulo 2. A code is said to be **self dual** if $\mathcal{C}[n,k] = \mathcal{C}^{\perp}[n,n-k]$. It follows that a self dual code in $\mathbb{F}_2^n$ has dimension $n/2$. On the other hand, a code is said to be **self orthogonal** if $\mathcal{C}[n,k] \subseteq \mathcal{C}^{\perp}[n,n-k]$, thus a self orthogonal code will have dimension less than or equal to $n/2$.

In the next section, we will define and discuss Hadamard matrices and their connection to codes. To see this connection more clearly, we need to define what the *generator matrix* for a code is.

Definition C.1. Let $\mathcal{C}[n,k]$ be a k dimensional code whose codewords have length n. Define a *generator matrix* for $\mathcal{C}$ to be a $k \times n$ matrix whose rows are basis vectors for $\mathcal{C}$.

Definition C.2. Let $\mathcal{C}_1[n,k]$ and $\mathcal{C}_2[n,k]$ be two k dimensional codes. We say $\mathcal{C}_1$ is *equivalent* to $\mathcal{C}_2$ if we can permute the columns of their generator matrices so that they have the same basis.

C.2 Introduction to Hadamard Matrices

Let H be an $n \times n$ matrix with entries ± 1 that satisfies $HH^T = nI_n$. Such a matrix is called a **Hadamard matrix** and will exist and have maximal determinant $n^{n/2}$ only if $n = 1, 2$, or $n \equiv 0 \pmod 4$ [5]. Hadamard matrices are non-singular and (by negating and permuting rows and columns) will contain a first row and column in which every entry is 1. Doing this is called *normalizing* the matrix. The resulting matrix will be equivalent to the *non-normalized* matrix, where equivalent means up to row and column permutation and negation.

For any Hadamard matrix, there is an equivalence class of binarly linear codes associated with it. To obtain this code, we first normalize the matrix, then replace the -1's with 0's to obtain a $\{0,1\}$ matrix. The code is the linear span of these rows, and equivalent matrices give rise to equivalent codes.

For this paper, we will find the codes of certain Hadamard matrices and their subcodes in hopes that they too will be the codes of a Hadamard matrix. The codes in question will be doubly even, self orthogonal or self dual codes, because when $n \equiv 0 \pmod 8$, where n is the size of the Hadamard matrix, the resulting code will be doubly even and self orthogonal [1]. Because of the large number of possibilities of subcodes for certain codes, we will make a definition to narrow our search.

Definition C.3. A code $\mathcal{C}$ is said to be *suitable* if it contains the all 1 vector and its generator matrix does not contain duplicate columns.

Clearly, the code of a Hadamard matrix must contain the all 1 vector, since for any Hadamard matrix, we can make the first row have all 1's. Also, from a result found in [8], the code for an $n \times n$ Hadamard matrix must have dimension $d \geq \log_2 n + 1$, therefore we only need to consider subcodes down to this dimension. The generator matrix for the code may not have duplicate columns because, if it did, every codeword in the code would have duplicate columns. This would translate to the associated Hadamard matrix having duplicate columns, which we know is not true since Hadamard matrices are non-singular.

As stated above, this paper will deal solely with binary linear codes, more specifically doubly even self dual and self orthogonal codes. Although we are not limiting ourselves to codes only generated by weight four codewords, special mention of them must be made. Any code that is generated by weight four codewords is the direct sum of the following codes: the duadic codes (d_{2n}), e_7, and e_8 [2]. Therefore we will focus our discussion on these codes to get a better picture of where the codes generated by weight four codewords come from.

C.3 The Duadic Codes (d_{2n}), e_7, and e_8

We'll begin our discussion of codes generated by weight four codewords with the **duadic codes** d_{2n}. These codes are of dimension $n-1$ and length $2n$, whose basis is formed as such:

$$\begin{aligned} d_{2n} = &\langle 11110000\cdots 00, 11001100\cdots 00, \\ &11000011\cdots 00, \cdots, 11000000\cdots 11\rangle \end{aligned}$$

where each codeword is of length 2n. As an example, the basis for d_8 is $\langle 11110000, 11001100, 11000011\rangle$, and contains the codewords

11110000
11001100
11000011
00111100
00001111
00110011
11111111
00000000

These codes respect **duads**, or have a **duadic** structure. This means that we are free to permute columns 1 and 2, 3 and 4, 5 and 6, and so on. This is clear by how the basis is formed; column 1 is the same as column 2, column 3 is the same as column 4, and so on. It is also important to note that d_4 respects **tetrads**, meaning we are free to permute columns 1 through 4 and columns 5 through 8. Both notions of a code respecting duads and tetrads will be used later.

The other codes generated only by weight four codewords are e_7 and e_8. To construct e_7, begin with the basis for d_6, adding a zero to the end of each codeword to make it length 7, and add the **self glue vector** 1010101 to the basis. Adding a self glue vector (a vector not in the original code) to the basis of a code increases its dimension by one and thus increases the amount of words in the code by a factor of two. Note that e_7 has length 7, though for the purposes of this paper, we will add zeros to the end of each codeword to make its length a multiple of 8 (this process is called "padding"). For example: to make every codeword of e_7 have length 8, we will pad each codeword with a single zero, making $e_7 = \langle 11110000, 11001100, 10101010\rangle$. We construct e_8 in a similar way, glueing 10101010 to d_8, thus e_8 is $\langle 11110000, 11001100, 11000011, 10101010\rangle$.

Note that adding a self glue vector to a code is different than adding a glue vector, a notion that will be used when considering codes that have a direct sum structure.

Definition C.4. Let $\mathcal{C}_1 \oplus \mathcal{C}_2 \oplus \cdots \oplus \mathcal{C}_k$ be a code formed by the direct sum of k codes. Define a *glue vector* x of the form $(x_1, x_2, \cdots x_k)$, where $x_j \notin \mathcal{C}_j$ for $1 \leq j \leq k$, to be a vector that adds a dimension to the code, increases the number of words in the code by a factor of two, and breaks the direct sum structure of the code.

Consider the example $\langle d_8 \oplus d_8, (10101010, 10101010)\rangle$, which has the following basis

$$\begin{array}{c} 11110000\ 00000000 \\ 11001100\ 00000000 \\ 11000011\ 00000000 \\ 00000000\ 11110000 \\ 00000000\ 11001100 \\ 00000000\ 11000011 \\ 10101010\ 10101010 \end{array}$$

Notice that without the glue vector $(10101010, 10101010)$, any codeword of the code will retain the direct sum structure. However, adding this glue vector essentially "glues" the two codes d_8 and d_8 together. Therefore, with the addition of this glue vector, not every codeword in the code will retain this direct sum structure. The following lemma is helpful in showing the uniqueness of glue vectors within codes.

Lemma C.3.1. *Let $\mathcal{C} = \mathcal{C}_1 \oplus \mathcal{C}_2 \oplus \cdots \oplus \mathcal{C}_k$ be a code made up of the direct sum of k codes. Let u and v be codewods such that $v = u + c$, where $c \in \mathcal{C}$. Then the code formed by gluing u to $\mathcal{C}$ is equivalent to the code formed by gluing v to $\mathcal{C}$.*

The proof is trivial and not shown here, but to see its application, let's consider an example. Consider the above example $\langle d_8 \oplus d_8, (10101010, 10101010)\rangle$. If we glue

$$(10101010, 10101010) + (11110000, 00000000) = (01011010, 10101010)$$

to $d_8 \oplus d_8$ instead of gluing $(10101010, 10101010)$ to $d_8 \oplus d_8$, we will clearly get a code equivalent to $\langle d_8 \oplus d_8, (10101010, 10101010)\rangle$. This follows because we can just add $(11110000, 00000000)$ back to the glue vector displayed above to make the last basis vector for the code be $(10101010, 10101010)$, thus the codes are equivalent.

Because of the frequency with which certain glue components appear, we will define the glue components $a = 10101010$ and $q = 11000011$. Note that, with this definition for q, the code d_8 is $\langle d_6, q\rangle$.

C.4 The Deletion Process

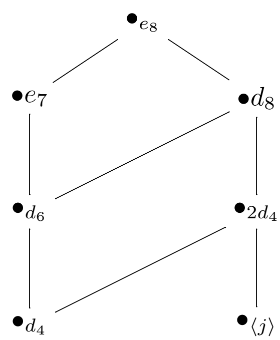
Given a dimension k, doubly even, self orthogonal code, how can we find all of the dimension $k-1$ subcodes of this code? When working in $\mathbb{R}^n$, it is natural to find the $k-1$ dimensional subcodes of a k dimensional code by finding the orthogonal complements of all the 1 dimensional subspaces of $\mathbb{R}^n$. However, since we are working in $\mathbb{F}_2^n$, applying this process would yield the original code back again, since every codeword is orthogonal to both itself and every other codeword.

A process for finding the $k-1$ dimensional subcodes of a doubly even, self orthogonal dimension k code is as follows: let $\mathcal{C}[n,k]$ be a doubly even, self orthogonal code. Choose the basis for $\mathcal{C}[n,k]$ to be $\langle C_1, C_2, \cdots C_k\rangle$. Let $v \in \mathbb{F}_2^k$. Choose $k-1$ linearly independant vectors to form $\langle v\rangle^\perp$. Each of the vectors in $\langle v\rangle^\perp$ will correspond to a vector in $\mathcal{C}[n,k]$. For example; $(1,0,0,\cdots 0)$ will correspond to C_1, $(0,1,0,\cdots 0)$ will correspond to C_2, $(1,1,0,\cdots 0)$ will correspond to $C_1 + C_2$, and so on. So, the $k-1$ dimensional space $\langle v\rangle^\perp$ will correspond to a $k-1$ dimensional subcode of $\mathcal{C}[n,k]$.

To see this, let's consider the example of finding all the 3 dimensional subcodes of the doubly even, self dual code e_8. Since e_8 is of dimension 4, we will choose v from $\mathbb{F}_2^4$. Recall that e_8 is

$$\langle 11110000, 11001100, 11000011, 10101010 \rangle$$

Choosing $v = (1, 0, 0, 0)$ corresponds to the deletion of 11110000 from the basis for e_8, as $\langle v \rangle^{\perp}$ yields $\langle (0, 1, 0, 0), (0, 0, 1, 0), (0, 0, 0, 1) \rangle$. This will be $\langle 11001100, 11000011, 10101010 \rangle$ in the code, which is equivalent to e_7. Thus the resulting 3 dimensional subcode of e_8 is e_7. As it turns out, there is only one other dimension 3 subcode of e_8, namely d_8, which can easily be shown by considering all $2^4 - 1 = 15$ choices of v (we omit the all zero vector). The complete breakdown of dimension 3, 2, and 1 subcodes of e_8 is shown below and will be used later.



Clearly, none of the above codes are suitable except for e_8. However, they will be important in our search as we will need to know all possible subcodes of e_8, d_8, e_7, and so on.

For this paper we will be using the deletion process starting with self dual codes and essentially working our way down in dimension. There is a complete list of the self dual codes of length 8, 16, 24, and 32 given in [2]. We also know that for any $n \equiv 0 \pmod 8$, a self dual code of dimension n will exist, and that every doubly even self orthogonal code is contained in some doubly even self dual code [6]. We'll also be working with sizes $n \equiv 0 \pmod{16}$, since the codes for a Hadamard matrix can be either self dual or self orthogonal, whereas in the $8 \pmod{16}$ case, they are restricted only to self dual codes.

C.5 The Complete Breakdown of the Self Dual, Doubly Even Code $e_8 \oplus e_8$

Our goal here is to provide a complete list of suitable subcodes for the length 16, doubly even self dual code $e_8 \oplus e_8$, similar to that done for e_8. This will be done by using the deletion process, but since $v \in \mathbb{F}_2^8$, a brute force method will take quite some time as there are 255 possible choices for v. So, we will use the following results to provide the complete list.

For this first result, recall from linear algebra that for some space S^n, we can write S^n as the sum of $S^l \oplus S^{n-l}$, where $1 \leq l \leq n - 1$.

Definition C.5. Let Let $v \in \mathbb{F}_2^n \cong \mathbb{F}_2^l \oplus \mathbb{F}_2^{n-l}$. Define *field one* to be the first l positions of the vector v, and define *field two* to be the remaining $n - l$ positions of v.

Before stating the lemma, let's first consider an example to clearly show the meaning of field one and field two. Let $v = (0, 1, 1, 0, 0, 0, 0, 0, 0)$ and let $u = (0, 0, 0, 0, 1, 1, 0, 0, 1)$. Both v and u are in $\mathbb{F}_2^9 \cong \mathbb{F}_2^4 \oplus \mathbb{F}_2^5$. Thus field one refers to the first 4 positions of v and u and field two refers to the last 5 positions of v and u. Clearly, v has ones only in field one, and u has ones only in field two.

Lemma C.5.1. *Let $v \in \mathbb{F}_2^n$. Then the space $\langle v \rangle^\perp$ can be written as the span of $n-1$ vectors of weight 1 or 2, where at most one of which has a 1 in field one and field two.*

Proof. To begin the proof, first consider the two trival cases where $w(v) = w = 1$ or 2. When $w = 1$, we can form $\langle v \rangle^\perp$ with only weight one vectors, and when $w = 2$, we can form $\langle v \rangle^\perp$ with $n-2$ weight one vectors and need only one vector of weight two (namely the vector v itself), which may or may not have a one in field one and field two.

For $w \geq 3$, consider the process for forming $\langle v \rangle^\perp$. Begin by including the weight one vectors that have a one in a position where v has a zero, and zeros everywhere else. Clearly, these vectors are linearly independent and orthogonal to v. Note that there are $n-w$ of these, since v has weight w, leaving $n-w$ positions where v has a zero entry. Since all operations are performed modulo 2, the remaining vectors in $\langle v \rangle^\perp$ need only be of weight two to be orthogonal to v.

Let v_1 be the number of ones in field one of v and let v_2 be the number of ones in field two of v. Clearly, $v_1 + v_2 = w$. From field one alone, there are $v_1 - 1$ linearly independent vectors orthogonal to v. Similarly, from field two alone, there are $v_2 - 1$ linearly independent vectors orthogonal to v. We now have $(n-w) + (v_1 - 1) + (v_2 - 1)$ linearly independent vectors to form $\langle v \rangle^\perp$, none of which have a one in field one and field two. Since $v_1 + v_2 = w$, we have

$$\begin{aligned}(n-w) &+ (v_1 - 1) + (v_2 - 1) \\ &= n - w + (v_1 + v_2) - 2 \\ &= n - 2\end{aligned}$$

linearly independent vectors to form $\langle v \rangle^\perp$, none of which have a one in field one and field two. We are now free to choose the final weight two basis vector for $\langle v \rangle^\perp$ to have a one in field one and a one in field two. $\square$

It is important to note that we can break $\mathbb{F}_2^n$ into a direct sum of two vector spaces in a number of different ways. In the above example, we could have said $\mathbb{F}_2^9 = \mathbb{F}_2^3 \oplus \mathbb{F}_2^6$ and defined field one to be the first three positions and field two to be the last six positions. The above result will still hold no matter how we break $\mathbb{F}_2^n$ into a direct sum of two vector spaces, however it is important to choose the correct vector spaces in order for the lemma to be useful. How to choose the two vector spaces correctly can be seen in the following result:

Proposition C.5.2. *Let $\mathcal{C}_1$ be a code of dimension l, and let $\mathcal{C}_2$ be a code of dimension $n-l$. The n dimensional code formed by their direct sum $\mathcal{C}_1 \oplus \mathcal{C}_2$ will have $n-1$ dimensional subcodes only of the following forms:*

(a) $\bar{\mathcal{C}_1} \oplus \mathcal{C}_2$*, where* $\bar{\mathcal{C}_1}$ *is any* $l-1$ *dimensional subcode of* $\mathcal{C}_1$

(b) $\mathcal{C}_1 \oplus \bar{\mathcal{C}_2}$*, where* $\bar{\mathcal{C}_2}$ *is any* $n-l-1$ *dimensional subcode of* $\mathcal{C}_2$

(c) $\langle \bar{\mathcal{C}_1} \oplus \bar{\mathcal{C}_2}, (x, y) \rangle$*, where* $x \in \mathcal{C}_1$ *but* $x \notin \bar{\mathcal{C}_1}$*, and* $y \in \mathcal{C}_2$ *but* $y \notin \bar{\mathcal{C}_2}$

Furthermore, the resulting code is independent of the choice of x and y.

Proof. For this proof, we will use the "deletion process" described above, where $v \in \mathbb{F}_2^n$. By Lemma C.5.1, $\langle v \rangle^\perp$ will either have no basis vectors with a 1 in field one and a 1 in field two, or at most one basis vector with a 1 in field one and a 1 in field two. In this case, field one is defined as the first l positions of v, since $\mathcal{C}_1$ is of dimension l, and field two will be the last $n-l$ positoins of v, since $\mathcal{C}_2$ is of dimension $n-l$. Since

$\mathcal{C}_1 \oplus \mathcal{C}_2$ is already broken up into a direct sum, the case where $\langle v\rangle^\perp$ has no basis vectors with a 1 in field one and a 1 in field two will yield either (a) or (b).

If $\langle v\rangle^\perp$ does have a basis vector with a 1 in field one and a 1 in field two, this particular vector will correspond to a glue vector in a subcode of $\mathcal{C}_1 \oplus \mathcal{C}_2$. Since there will be at most one of such vectors, the remaining basis vectors of an $n-1$ dimensional subcode of $\mathcal{C}_1 \oplus \mathcal{C}_2$ will be determined by a direct sum of an $l-1$ dimensional subcode of $\mathcal{C}_1$ and an $n-l-1$ dimensional subcode of $\mathcal{C}_2$. It is clear from the definition of a glue vector that the glue vector will be of the form (x, y), where $x \in \mathcal{C}_1$ but $x \notin \bar{\mathcal{C}_1}$ and $y \in \mathcal{C}_2$ but $y \notin \bar{\mathcal{C}_2}$. It is also clear from Lemma C.3.1 that the code is independent of the choice of x and y. □

We now have sufficient results to find all the dimension 7 subcodes of $e_8 \oplus e_8$, and they are

$$\begin{gathered}\langle d_8 \oplus d_8, (a,a)\rangle \\ \langle e_7 \oplus d_8, (q,a)\rangle \\ \langle e_7 \oplus e_7, (q,q)\rangle \\ e_7 \oplus e_8 \\ d_8 \oplus e_8\end{gathered}$$

Because we are only concerned with suitable codes, we need only continue finding subcodes of $\langle d_8 \oplus d_8, (a,a)\rangle$, since it is the only code that does not have duplicate columns and contains the all 1 vector.

When finding the dimension 6 subcodes of $\langle d_8 \oplus d_8, (a,a)\rangle$, we can essentially ignore the glue vector (a,a) and only consider the dimension 5 subcodes of $d_8 \oplus d_8$. We are free to do this because adding (a,a) to any combination of vectors in $d_8 \oplus d_8$ will yield a vector equivalent to (a,a). This is because we are free to swap columns 1 with 2, 3 with 4, 5 with 6, and so on without changing any of the vectors in $d_8 \oplus d_8$. Since adding (a,a) to any combination of vectors in $d_8 \oplus d_8$ will only change the vector with respect to those pairs of columns, we will get a vector equivalent to (a,a). By using the deletion process, one of three things can happen; we will have a vector v without a 1 in position 7, which will correspond to including (a,a) in the basis for the subcode. If v is of weight one and has a 1 in position 7, this will result in the deletion of (a,a) from the code, resulting in the sub code $d_8 \oplus d_8$. If v is of weight ≥ 2, and has a 1 in position 7, this will correspond to a basis vector equivalent to (a,a), since we are free to permute columns 1 and 2, 3 and 4, 5 and 6, and so on. Therefore, unless $v = (0,0,0,0,0,0,1)$, the deletion process will always yield a code with (a,a) in the basis. So, to find all the dimension 6 subcodes of $\langle d_8 \oplus d_8, (a,a)\rangle$, we need only consider the dimension 5 subcdoes of $d_8 \oplus d_8$, since (a,a) will be a glue vector for any subcode of $\langle d_8 \oplus d_8, (a,a)\rangle$. Thus the dimension 6 subcodes of $\langle d_8 \oplus d_8, (a,a)\rangle$ are

$$\begin{gathered}d_8 \oplus d_8 \\ \langle d_6 \oplus d_8, (a,a)\rangle \\ \langle 2d_4 \oplus d_8, (a,a)\rangle \\ \langle d_6 \oplus d_6, (q,q), (a,a)\rangle \\ \langle d_6 \oplus 2d_4, (q,q), (a,a)\rangle \\ \langle 4d_4, (q,q), (a,a)\rangle\end{gathered}$$

Again, since we only want to consider the suitable codes, we will only need to find the subcodes of $\langle 4d_4, (q,q), (a,a)\rangle$, since it is the only code that does not have duplicate columns and that contains the all 1 vector.

By a similar argument used to show when (a,a) will be in the basis for a dimension 6 subcode of $\langle d_8 \oplus d_8, (a,a)\rangle$, we can show both (a,a) and (q,q) will be in the basis for any dimension 5 subcode of $\langle 4d_4, (q,q), (a,a)\rangle$. This will be the case when $v \neq (0,0,0,0,1,0)$ and $v \neq (0,0,0,0,0,1)$, since of course this just corresponds to deleting (a, a) or (q, q) from the code. Therefore, we need only consider the 3 dimensional subcodes of $4d_4$. So, the dimension 5 subcodes of $\langle 4d_4, (q,q), (a,a)\rangle$ are

$$\begin{gathered}
\langle 4d_4, (a,a)\rangle \\
\langle 4d_4, (q,q)\rangle \\
\langle 3d_4, (q,q), (a,a)\rangle \\
\langle j \oplus 2d_4, (q,q), (a,a)\rangle \\
\langle j \oplus j, (p,p), (q,q), (a,a)\rangle \\
\langle j \oplus d_4, (p,w), (q,q), (a,a)\rangle \\
\langle d_4 \oplus d_4, (w,w), (q,q), (a,a)\rangle
\end{gathered}$$

where j is the length 8, dimension 1 code $\langle 11111111\rangle$, p is the length 8 vector 11110000, and w is the length 8 vector 00001111. Again, we only want the suitable codes, which in this case is $\langle j \oplus j, (p,p), (q,q), (a,a)\rangle$. In fact, this code is equivalent to the Reed Muller Code, which can be found in [4].

We now have our complete list of the suitable dimension 7, 6, and 5 subcodes of the doubly even, self dual code $e_8 \oplus e_8$, and they are

$$\begin{gathered}
\langle d_8 \oplus d_8, (a,a)\rangle \\
\langle 4d_4, (q,q), (a,a)\rangle \\
\langle j \oplus j, (p,p), (q,q), (a,a)\rangle
\end{gathered}$$

C.6 The Other Doubly Even, Self Dual Code of Length 16: $\langle d_{16}, (a,a)\rangle$

By applying the deletion process again to $\langle d_{16}, (a,a)\rangle$, we can find the suitable subcodes of dimension 7, 6, and 5. To show this, it is important to know all the possible dimension $n-2$ subcodes of any duadic code d_{2n}.

Lemma C.6.1. *Let d_{2n} be the dimension $n-1$, length $2n$ duadic code. Then the $n-2$ dimensional subcodes of d_{2n} are:*

$$\begin{gathered}
d_{2n-2}, d_{2n-4} \oplus d_4, d_{2n-6} \oplus d_6, \cdots d_n \oplus d_n \textit{ if } n \textit{ is even} \\
\textit{or} \\
d_{2n-2}, d_{2n-4} \oplus d_4, d_{2n-6} \oplus d_6, \cdots d_{n+1} \oplus d_{n-1} \textit{ if } n \textit{ is odd}
\end{gathered}$$

Proof. Recall that the basis for d_{2n} is

$$\begin{gathered}
1111000000 \cdots 00 \\
1100110000 \cdots 00 \\
1100001100 \cdots 00 \\
1100000011 \cdots 00 \\
\vdots \\
1100000000 \cdots 11
\end{gathered}$$

Note that there are $n-1$ basis vectors for d_{2n}, and that we are free to use row operations to change the basis and still get an equivalent code. If we keep the first and last vector the same, and change the remaining $n-3$ vectors to themselves plus the last basis vector $(11000000 \cdots 11)$, we will get a basis that resembles a

$d_4 \oplus d_{2n-4}$ with a glue vector $(11000000\cdots 11)$. But, since this is still $n-1$ dimensional, we can remove the last vector from the basis to get the $n-2$ dimensional subcode $d_4 \oplus d_{2n-4}$ of d_{2n}. In a similar way, if we keep the first two vectors and the last vector, and change the remaining $n-4$ vectors as we did above, we will get the subcode $d_6 \oplus d_{2n-6}$. In general, if we keep the first k basis vectors as they are, where $1 \leq k \leq \lfloor \frac{n-1}{2} \rfloor$, and we perform the procedure as described above, we will get the $n-2$ dimensional subcode $d_{2+2k} \oplus d_{2(n-1-k)}$. Clearly, we ge t all the subcodes listed above except d_{2n-2}, which can be attained simply by keeping the original basis and removing the last vector from it.

We now need to show that this list is in fact complete. Assume then that this list were not complete. There would then exist a code $\mathcal{C}$ generated by weight 4 codewords that is not in the above list. We can say this code must be generated by weight 4 code words by the way we chose $\langle v \rangle^\perp$ in the deletion process. Recall that we chose $\langle v \rangle^\perp$ to only have weight 1 and weight 2 vectors. These will correspond to weight 4 codewords since adding any two codewords in d_{2n} will result in a weight 4 codeword. Because every code generated by weight 4 codewords is a direct sum of d_{2n}, e_7, or e_8 [2], $\mathcal{C}$ would have to contain e_7 or e_8. This is because the above lists all possible ways to break d_{2n} into a direct sum, while still retaining the desired dimension. However, since d_{2n} respects duads, meaning we are free to swap columns 1 and 2, 3 and 4, 5 and 6, and so on without changing the c ode, any subcode if it would also have to respect duads. So if the list were incomplete, $\mathcal{C}$ would have to contain e_7 or e_8, which do not respect duads, thus contradicting the fact that any subcode of a code that respects duads will also respect duads, therefore the above list is complete. $\square$

Lemma C.6.2. *Let d_{2n} be the $n-1$ dimensional duadic code of length $2n$. If any codeword or combination of codewords is added to the glue vector a_{2n}, the resulting codeword is equivalent to a_{2n} by column permutations.*

Proof. First note that a_{2n} refers to the length $2n$ glue vector $a = (\underbrace{10101010\cdots 10}_{2n})$. Since all vectors in d_{2n} respect duads, we are free to swap columns 1 and 2, 3 and 4, 5 and 6, and so on of the generator matrix without changing the code. Adding a_{2n} to any vector in d_{2n} will only swap entries of a_{2n} within duads, therefore we are free to permute the affected columns of this "new" vector to produce a_{2n} back again, therefore this "new" vector is equivalent to a_{2n}. $\square$

We now have sufficient results to find all the dimension 7 subcodes of the other length 16 self dual code $\langle d_{16}, (a, a)\rangle$.

Proposition C.6.3. *All possible dimension 7 subcodes of $\langle d_{16}, (a, a)\rangle$ are*

$$\begin{gathered} d_{16} \\ \langle d_{14}, (a, a)\rangle \\ \langle d_4 \oplus d_{12}, (a, a)\rangle \\ \langle d_6 \oplus d_{10}, (a, a)\rangle \\ \langle d_8 \oplus d_8, (a, a)\rangle \end{gathered}$$

Proof. By using the deletion process, v will be of the following forms: it will not have a 1 in position 8, meaning we include (a, a) in the basis for a subcode; v is of weight 1 with its 1 in position 8, meaning we delete (a, a) from the code resulting in the subcode d_{16}; or v is of weight ≥ 2 with a 1 in position 8, which by C.6.2 will result in a subcode having a glue vector equivalent to (a, a) in the basis. Therefore, aside from the case where v is of weight 1 with its 1 in position 8, any subcode of $\langle d_{16}, (a, a)\rangle$ will have (a, a) as a basis vector. Therefore we need only consider the subcodes of d_{16} to find all possible subcodes of $\langle d_{16}, (a, a)\rangle$. Applying C.6.1 will now get the above complete list of subcodes of $\langle d_{16}, (a, a)\rangle$. $\square$

Again, since we only want suitable codes, this will narrow our findings down to $\langle d_8 \oplus d_8, (a,a)\rangle$, since the other codes either do not contain the all 1 vector or have a generator matrix with duplicate columns. Obviously, we have already found all the subcodes of $\langle d_8 \oplus d_8, (a,a)\rangle$, thus we have our complete list of suitable subcodes of $e_8 \oplus e_8$ and $\langle d_{16}, (a,a)\rangle$. We can now move to a discussion of their relevance to Hadamard Matrices.

C.7 Hadamard Matrices of Order 16

Hall proved in [3] that there are five equivalence classes of Hadamard Matrices of order 16. This means that, given a 16×16 Hadamard matrix, this matrix is equivalent to one of the five matrices given in [3] up to row and column permutation and negation.

Sloane's directory of Hadamard matrices gives these five matrices, labeled 16.0, 16.1, 16.2, 16.3, and 16.4 [7]. The codes associated with these matrices are our suitable codes found above, where 16.0 is $\langle j \oplus j, (p,p), (q,q), (a,a)\rangle$, 16.1 is $\langle 4d_4, (q,q), (a,a)\rangle$, 16.2 is $\langle d_8 \oplus d_8, (a,a)\rangle$, 16.3 is $\langle d_{16}, (a,a)\rangle$, and 16.4 is $e_8 \oplus e_8$.

Bibliography

1. Assmus & Key *Designs and their Codes*, Cambridge University Press, New York, 1992

2. Conway & Pless *O*n the Enumeration of Self-Dual Codes, *Journal of Combinatorial Theory* Series A **28** (1980) 26-53

3. M. Hall Jr. Hadamard Matrices of Order 16, *J.P.L Research Summary No. 36-10* **1** (1961) 21-26

4. Huffman & Pless *Fundamentals of Error-Correcting Codes*, Cambridge University Press, New York, 2003

5. J. H. van Lint & R. M. Wilson, *A Course in Combinatorics*, Cambridge University Press, New York, 2001

6. F. J. MacWilliams, N. J. A. Sloane, & J. G. Thompson Good Self Dual Codes Exist, *J. of Descrete Math.* **3** (1972) 153-162

7. N. J. A. Sloane *A Library of Hadamard Matrices*, available online at `http://www.research.att.com/~njas/hadamard/` (2008).

8. W. D. Wallis Integral Equivalence of Hadamard Matrices, *Israel J. Math.* **10** (1971) 349-358

Growing equilibrium configurations of point vortices

Michael Meaden
Elmhurst College

Indiana University REU Summer 2008
Advisor: ADVISOR

D.1 Introduction

For a given vortex, the core is considered to contain a vortex line perpendicular to the plane the vortex is in, around which particles in the fluid rotate at some angular frequency. We consider a point vortex to be a point singularity with some angular frequency placed at the point the vortex line crosses the plane. Vortices are of interest in fluid dynamics, and are often studied when analyzing and modeling weather and atmospheric patterns, ocean flows, and atomic interactions, among other topics.

Though there has been a huge variety of projects researching various aspects of systems of vortices, there is still much left unknown. The possible vortex configurations that can develop in a system have been studied for nearly 150 years, dating back to 1878-79 when the physicist Alfred Mayer performed experiments utilizing magnets on floating corks in a magnetic field in an effort to demonstrate atomic interactions [1]. Certain steady states observed during this experiment proceeded to launch many other research experiments involving vortices in an effort to model and observe many different phenomena. In 1978 and 1979, Campbell and Ziff published papers describing such equilibrium configurations of vortices in the plane, primarily focusing on stable configurations [3]. Their well known report, often referred to as the "Los Alamos catalog", revealed many nested ring equilibria configurations [2]. While we have unfortunately been unable to obtain a copy of the catalog itself, it is repeatedly referenced in several other papers on point vortex research.

Equilibrium configurations of point vortices are defined as systems of vortices that move such that the size and shape of the configuration do not change, so that the distances between vortices is always the same [1]. More recently, in 1998, Aref and Vainchtein from the University of Illinois at Urbana-Champaign, published a paper outlining a method of "growing" equilibrium configurations of point vortices of equal strength, through which they found many configurations not previously found in the Los Alamos Catalog [2]. The method consists of two main steps. Assuming a valid equilibrium of N vortices to be given, the method first calls for all co-rotating points for that system to be found. A co-rotating point is considered to have the same angular frequency as the vortices given, but a strength of zero. In other words, it is essentially an additional vortex, but with strength zero so that it does not affect the system. The second step is then to take the strength of a given co-rotating point and increase it incrementally from zero to one, adjusting all vortices in the system. As all vortices in the system are considered to have strengths of one, the configuration found when the strength of the co-rotating point is equal to one is found to be the new equilibrium configuration of $N + 1$ vortices. This method is described in full in this paper.

As the method outlined above has clearly been utilized in past research projects, our goal is to first verify results found in past research so as to ensure the method is indeed working, and then shift our focus into other areas of the problem that have not been so extensively researched to this point. While we able to successfully verify past results and analyze the process of growing new configurations fairly extensively, time constraints have unfortunately not allowed our research to proceed much farther. However, the future of point vortex research continues to look very promising, and the stage is set to continue on to other aspects of the problem.

D.2 Vortex Equations of Motion

An equilibrium configuration of N vortices of strength Γ satisfies the differential equation [2]

$$\frac{dz_k^*}{dt} = \frac{\Gamma}{2\pi i} \sum_{j=1}^{N} \frac{1}{z_k - z_j}$$

where z_j and z_k are complex variables, $k = 1, ..., N$, $j \neq k$, and the asterisk denotes complex conjugation.

If we consider such a configuration to rotate uniformly with angular frequency ω - that is $z_k(t) = z_k(0)e^{i\omega t}$ - then we see that this equation easily simplifies to an algebraic system.

$$\frac{dz_k^*}{dt} = -i\omega z_k^*(0)e^{-i\omega t} = \frac{\Gamma}{2\pi i}\sum_{j=1}^{N}\frac{1}{(z_k(0) - z_j(0))e^{i\omega t}}$$
$$z_k^*(0) = \frac{\Gamma}{2\pi\omega}\sum_{j=1}^{N}\frac{1}{z_k(0) - z_j(0)}$$

We must multiply the complex conjugate by $e^{-i\omega t}$ when we consider the rotation of the system, as this is the conjugate of $e^{i\omega t}$. This term then cancels from our equation, and we scale $\frac{\Gamma}{2\pi\omega} = 1$, yielding the greatly simplified algebraic system

$$z_k^* = \sum_{j=1}^{N}\frac{1}{z_k - z_j}$$

Because of the cancellation of the $e^{i\omega t}$ term in the equation, it is evident that a valid equilibrium solution is not dependent on the rotation of the system at a time t and we therefore can choose t arbitrarily. Therefore, given a valid equilibrium configuration of vortices $z_1, ..., z_N$, we see that any rotation of that configuration is also a valid configuration. Therefore, while we take the vortex equations of motion to be satisfied by $z_k(t) = z_k(0)e^{i\omega(t)}$, the algebraic system is satisfied by $z_k(0)$. We take $t = 0$ to be implied in the algebraic system. Because it is already apparent that any rotation of a valid configuration of vortices is also a valid configuration, we consider any rotation of a valid equilibrium configuration to, in fact, be the *same* configuration. This is important to consider when growing new configurations.

D.3 Solving for Co-Rotating Points

In order to grow new configurations of $N + 1$ vortices, we first need a system of vortices $z_1, ..., z_N$ of equal strength that satisfy the system

$$z_k^* = \sum_{j=1}^{N}\frac{1}{z_k - z_j} \tag{D.1}$$

where $k = 1, ..., N$ and $j \neq k$.

Given such a system that solves equation D.1, a valid co-rotating point z_{N+1} is found that satisfies the system [2]

$$z_{N+1}^* = \sum_{j=1}^{N}\frac{1}{z_{N+1} - z_j} \tag{D.2}$$

For any initial configuration, it is expected that several co-rotating points will be found, though the exact number depends on the initial configuration. In order to solve equation D.2, we must first convert the system to be in terms of real variables. Because $z = a + ib$, we find the system it terms of real variables as shown:

$$
\begin{aligned}
z^*_{N+1} &= \sum_{j=1}^{N} \frac{1}{z_{N+1} - z_j} \\
a_{N+1} - ib_{N+1} &= \sum_{j=1}^{N} \frac{1}{(a_{N+1} - a_j) + i(b_{N+1} - b_j)} \cdot \left(\frac{(a_{N+1} - a_j) - i(b_{N+1} - b_j)}{(a_{N+1} - a_j) - i(b_{N+1} - b_j)} \right) \\
a_{N+1} &= \sum_{j=1}^{N} \frac{a_{N+1} - a_j}{(a_{N+1} - a_j)^2 + (b_{N+1} - b_j)^2} \\
b_{N+1} &= \sum_{j=1}^{N} \frac{b_{N+1} - b_j}{(a_{N+1} - a_j)^2 + (b_{N+1} - b_j)^2} \\
f_1 &= a_{N+1} - \sum_{j=1}^{N} \frac{a_{N+1} - a_j}{(a_{N+1} - a_j)^2 + (b_{N+1} - b_j)^2} \\
f_2 &= b_{N+1} - \sum_{j=1}^{N} \frac{b_{N+1} - b_j}{(a_{N+1} - a_j)^2 + (b_{N+1} - b_j)^2}
\end{aligned}
$$

Notice that the $-i$ term cancels in the equation involving b_{N+1}. We then solve f_1 and f_2 for a_{N+1} and b_{N+1} using Newton's method.

D.3.1 Newton's Method

Newton's method involves choosing an initial guess point, extending the tangent line at that point to where it crosses the x axis, and taking that x coordinate as the new guess point. The function and derivative are then calculated at that point, and the process is repeated. To find the residual r that the x coordinate is modified by at every iteration of the algorithm, we examine the Taylor series expansion. Given an initial guess point $x_k \in \mathbb{R}$ and a function $f : \mathbb{R} \to \mathbb{R}$, we then find the Taylor series to be

$$
f(x_k + r) = f(x_k) + f'(x_k)r + \frac{f''(x_k)}{2}r^2 + \cdots
$$

Because we assume r to be small, we take all nonlinear terms to be insignificant and therefore we are able to approximate, setting the function equal to zero

$$
0 \approx f(x_k) + f'(x_k)r
$$

Solving for r, we find $r = \frac{-f(x_k)}{f'(x_k)}$. For the next iteration of Newton's method, the x coordinate is taken to be $x_{k+1} = x_k + r$ and the method is repeated. This easily generalizes to multiple dimensions, giving for a function $\mathbf{f} : \mathbb{R}^{\mathbf{n}} \to \mathbb{R}^{\mathbf{n}}$, $\mathbf{r} = \frac{-\mathbf{f}(\mathbf{x_k})}{\mathbf{Df}}$, where $D\mathbf{f} = \left[\frac{\partial f_i}{\partial x_j}\right]$ is the Jacobian matrix of $\mathbf{f}$.

While Newton's method is a very powerful root approximation algorithm, it is not without its shortcomings. Though the method has a rapid rate of convergence, it is not guaranteed to converge, as this depends upon the initial x coordinate and the derivative at that point. If x happens to fall on or near an extremum, then the first derivative will be found to be 0 or near 0, and the tangent line can be sent off to infinity, resulting in a failure of convergence. It is also possible that the residual will modify the x coordinate such that $x_k = x_{k+2}$ and the algorithm will continually adjust between the same two points, and never converge.

D.3.2 Applying the Method

For the purposes of finding the co-rotating points for a given system of vortices, we have a function $\mathbf{f}$ defined from $\mathbb{R}^2 \to \mathbb{R}^2$. This is because we take the initial configuration to be fixed in the complex plane, and we are only allowing the coordinates of our guess for the co-rotating point to be adjusted through Newton's method. Our Jacobian matrix is then

$$D\mathbf{f} = \begin{bmatrix} \frac{\partial f_1}{\partial a_{N+1}} & \frac{\partial f_1}{\partial b_{N+1}} \\ \\ \frac{\partial f_2}{\partial a_{N+1}} & \frac{\partial f_2}{\partial b_{N+1}} \end{bmatrix}.$$

The partial derivatives are then found as follows:

$$\begin{aligned} \frac{\partial f_1}{\partial a_{N+1}} &= \frac{\partial}{\partial a_{N+1}} \left(a_{N+1} - \sum_{j=1}^{N} \frac{a_{N+1} - a_j}{(a_{N+1} - a_j)^2 + (b_{N+1} - b_j)^2} \right) \\ &= 1 - \sum_{j=1}^{N} \frac{(b_{N+1} - b_j)^2 - (a_{N+1} - a_j)^2}{[(a_{N+1} - a_j)^2 + (b_{N+1} - b_j)^2]^2} \end{aligned}$$

$$\begin{aligned} \frac{\partial f_1}{\partial b_{N+1}} &= \frac{\partial}{\partial b_{N+1}} \left(a_{N+1} - \sum_{j=1}^{N} \frac{a_{N+1} - a_j}{(a_{N+1} - a_j)^2 + (b_{N+1} - b_j)^2} \right) \\ &= \sum_{j=1}^{N} \frac{2(a_{N+1} - a_j)(b_{N+1} - b_j)}{[(a_{N+1} - a_j)^2 + (b_{N+1} - b_j)^2]^2} \end{aligned}$$

$$\begin{aligned} \frac{\partial f_2}{\partial a_{N+1}} &= \frac{\partial}{\partial a_{N+1}} \left(b_{N+1} - \sum_{j=1}^{N} \frac{b_{N+1} - b_j}{(a_{N+1} - a_j)^2 + (b_{N+1} - b_j)^2} \right) \\ &= \sum_{j=1}^{N} \frac{2(a_{N+1} - a_j)(b_{N+1} - b_j)}{[(a_{N+1} - a_j)^2 + (b_{N+1} - b_j)^2]^2} \end{aligned}$$

$$\frac{\partial f_2}{\partial b_{N+1}} = \frac{\partial}{\partial b_{N+1}} \left(b_{N+1} - \sum_{j=1}^{N} \frac{b_{N+1} - b_j}{(a_{N+1} - a_j)^2 + (b_{N+1} - b_j)^2} \right)$$
$$= 1 - \sum_{j=1}^{N} \frac{(a_{N+1} - a_j)^2 - (b_{N+1} - b_j)^2}{[(a_{N+1} - a_j)^2 + (b_{N+1} - b_j)^2]^2}$$

To be confident that all co-rotating points will be found, we wrote a program including methods similar to selected subroutines from [5]. We developed the program such that an initial guess point is placed every 0.02 units along both the real and imaginary axes, starting from (-100, -100) and extending to (100, 100). Many of these points yield the same result for the coordinates of the co-rotating point and many diverge. When all initial points have been run, however, we are left with a list of coordinates for several unique co-rotating points, which we can then plot. Some examples are given in section D.5 for several initial configurations.

D.4 Growing New Configurations of $N + 1$ Vortices

Once we have successfully found all co-rotating points for a valid equilibrium configuration, we then follow a similar method for "growing" the new configuration of $N+1$ vortices. First picking one of the co-rotating points as our initial coordinate for Newton's method, we then solve the system [2]

$$z_k^* = \sum_{j=1}^{N} \frac{1}{z_k - z_j} + \frac{p}{z_k - z_{N+1}} \quad \text{(D.3)}$$

where $k = 1, ..., N$, $j \neq k$, and p is increased from 0 to 1, as well as

$$z_{N+1}^* = \sum_{j=1}^{N} \frac{1}{z_{N+1} - z_j} \quad \text{(D.4)}$$

When $p = 0$, equation D.3 is simply equivalent to equation D.1. As the strength p of the co-rotating point is increased, however, all points in the system (including the initial vortices) are adjusted. We therefore have a function $\mathbf{f} : \mathbb{R}^{2N+2} \to \mathbb{R}^{2N+2}$. Because we are interested in equilibrium configurations where all vortices are of equal strength, and we take the strength of all vortices in the initial system to be 1, we take the new configuration of $N+1$ vortices to be the coordinates found when $p = 1$. To solve, we must find these systems in terms of real variables. Taking odd terms (i.e. x_{2k-1}) to be real components and even terms (i.e. x_{2k}) to be imaginary components, we find:

$$f_{2k-1} = x_{2k-1} - \sum_{j=1}^{N} \frac{(x_{2k-1} - x_{2j-1})}{h_j} - p\frac{(x_{2k-1} - x_{2N+1})}{d_k}$$

$$f_{2k} = x_{2k} - \sum_{j=1}^{N} \frac{(x_{2k} - x_{2j})}{h_j} - p\frac{(x_{2k} - x_{2N+2})}{d_k}$$

$$f_{2N+1} = x_{2N+1} - \sum_{j=1}^{N} \frac{x_{2N+1} - x_{2j-1}}{\tilde{h}_j}$$

$$f_{2N+2} = x_{2N+2} - \sum_{j=1}^{N} \frac{x_{2N+2} - x_{2j}}{\tilde{h}_j}$$

where $h_j = (x_{2k-1} - x_{2j-1})^2 + (x_{2k} - x_{2j})^2$, $d_k = (x_{2k-1} - x_{2N+1})^2 + (x_{2k} - x_{2N+2})^2$, and $\tilde{h}_j = (x_{2N+1} - x_{2j-1})^2 + (x_{2N+2} - x_{2j})^2$,$k = 1, ..., N$, and $j \neq k$. The complete work for this conversion can be found in section D.7.

While we again use Newton's method to solve these functions, we do this through utilizing a software package called AUTO. AUTO is an arc length continuation and bifurcation software package that incrementally increases the parameter p from 0 to 1. Fortunately, AUTO has a program to numerically find the Jacobian matrix for our functions, though the necessary partial derivatives can also be user supplied if the need arises.

The software reads the initial coordinates for the vortices and the co-rotating point of interest from a file, and then outputs data into three separate files. One file (fort.7) contains the data relevant to the continuation curve up through six components of the points being adjusted. This file also includes data for the parameter p and the Euclidean norm. Another file (fort.8) gives complete information for points along the continuation curve that AUTO singles out and assigns a type. These points are of special interest, and include limit (turning) points, endpoints, user defined points (in this case, when $p = 1$), points at a certain frequency, and points at which it is apparent that the method will not converge. Finally, fort.9 includes convergence information. While we have not had a need to examine the convergence data, we are certainly interested in the data given in fort.7 and fort.8, as this is the data used to create the bifurcation diagrams, as well as give the actual coordinates of the vortices for the new equilibrium configuration. Using this data, we are then able to plot the new configurations of $N + 1$ vortices.

D.5 Results and Analysis

D.5.1 The Trivial Case: $N = 3$

The logical place to start when growing new equilibrium configurations of vortices is the simplest configuration that can easily be verified by hand. In our case, we began with three vortices placed equidistantly from one another on the unit circle. As this is clearly a trivial case, we certainly did not expect to find particularly unique results. However, not only is this the most practical place to begin, but it also proved to provide a great introduction to using AUTO, what we can expect in terms of the bifurcation diagrams found and plotting the new equilibrium configurations, as well as an excellent test to assure us that the programs used are indeed working.

Solving for all co-rotating points, we find ten - four within the unit circle and six around the initial configuration. These are shown in figure D.1. Notice the symmetry across the real axis.

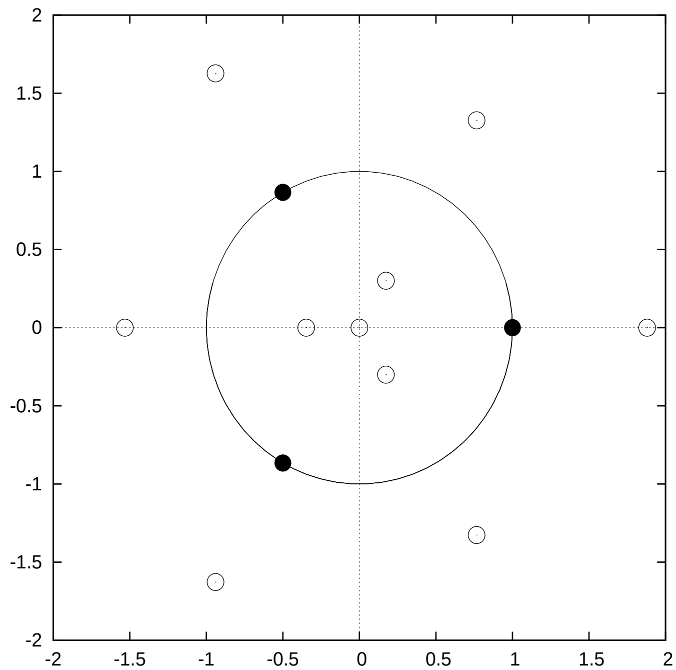


Figure D.1: Co-rotating points for an initial configuration of three vortices placed equidistantly on the unit circle.

Once we have found the coordinates for the co-rotating points, we then take each co-rotating point and utilize AUTO to determine the new configuration that results from increasing the strength of that point from 0 to 1. The resulting bifurcation diagrams of the correlation between p and a_1 for several co-rotating points are given below in figures D.2 through D.5. It may at first seem that we have many configurations resulting even from one co-rotating point, as we often see a range of values for a_1 when $p = 1$. However, upon plotting the new configurations associated with these points on the diagram, we find that many are identical configurations, and another vortex has simply switched places with the vortex that a_1 describes. Though we have often found that not all configurations given from one co-rotating point are exactly identical in terms of the exact coordinates of the vortices, we do see that they are all rotations of one another, and we therefore take them to be the same equilibrium configuration (see section D.2). Up to this point, we have not seen a case where one co-rotating point results in more than one unique new equilibrium configuration. In fact, we have found that several unique co-rotating points often result in growing the same configuration.

Even from this initial trivial case, it is apparent that the continuation curves can exhibit very unique properties. While we would typically expect some variation of figure D.2, where some branches continue through $p = 1$, some turn around, and some diverge, clearly there are many other possibilities. While we cannot immediately observe what configurations will result from which points on the bifurcation diagram, we can gain a better understanding of how the configuration develops as p increases. The clearest example is in figure D.5, when we see that all branches converge to 0 at $p = 1$. This shows us that the point of interest represented in the diagram is located at the center point of $(0, 0)$. We could similarly analyze the continuation curves given for other components to gain a more complete picture of exactly how all of the vortices are adjusted. However, as our primary interest is the resulting configurations, we will not delve into much detail in this report. Other unique bifurcation diagrams are shown in figures D.3 and D.4. One shows a continuation curve with nearly every branch having a limit point at $p = 1$, whereas the other shows a continuation curve with no limit points at all and yields a different configuration than what we had previously found. While we cannot say what is particularly special about the points that would cause such behavior, it is nevertheless intriguing and may be an area of interest in future work.

Two unique configurations are found from this trivial case: three vortices equidistant from one another on a circle of radius $\sqrt{2}$ around a center point at $(0, 0)$ (figures D.2, D.3, and D.5), and four vortices equidistant from one another on a circle of radius $\sqrt{\frac{3}{2}}$ (figure D.4). Intriguingly, all continuation curves that yield the latter configuration have similar properties to those seen in figure D.4 - namely, they all seem to have no limit points. Not surprisingly, we see an increase in the radius of the circular configurations as more vortices are introduced. To see exactly how the radius changes with each added vortex, however, we must continue growing configurations, taking the two configurations of four vortices as our initial configurations and applying the same method.

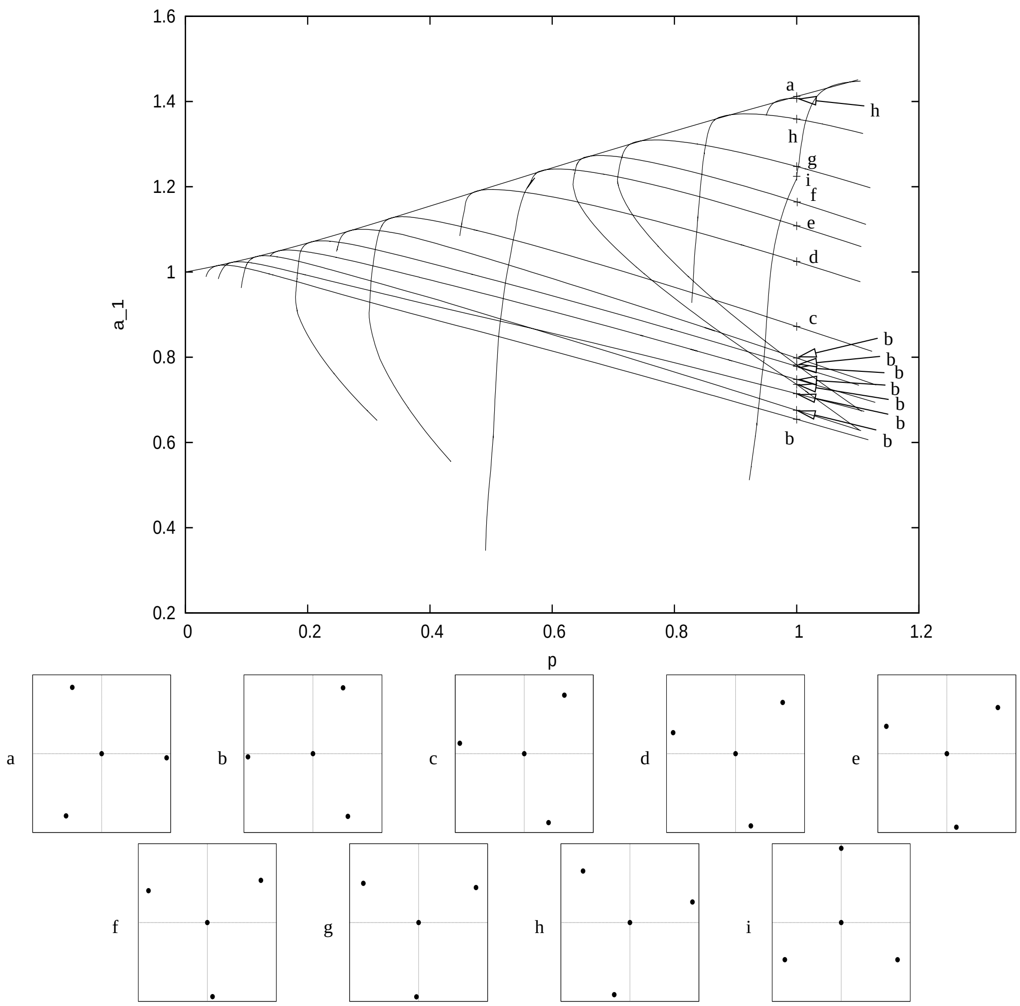


Figure D.2: Typical bifurcation diagram of a_1 vs. p. Each point where a branch crosses p=1 corresponds to a valid new equilibrium configuration.

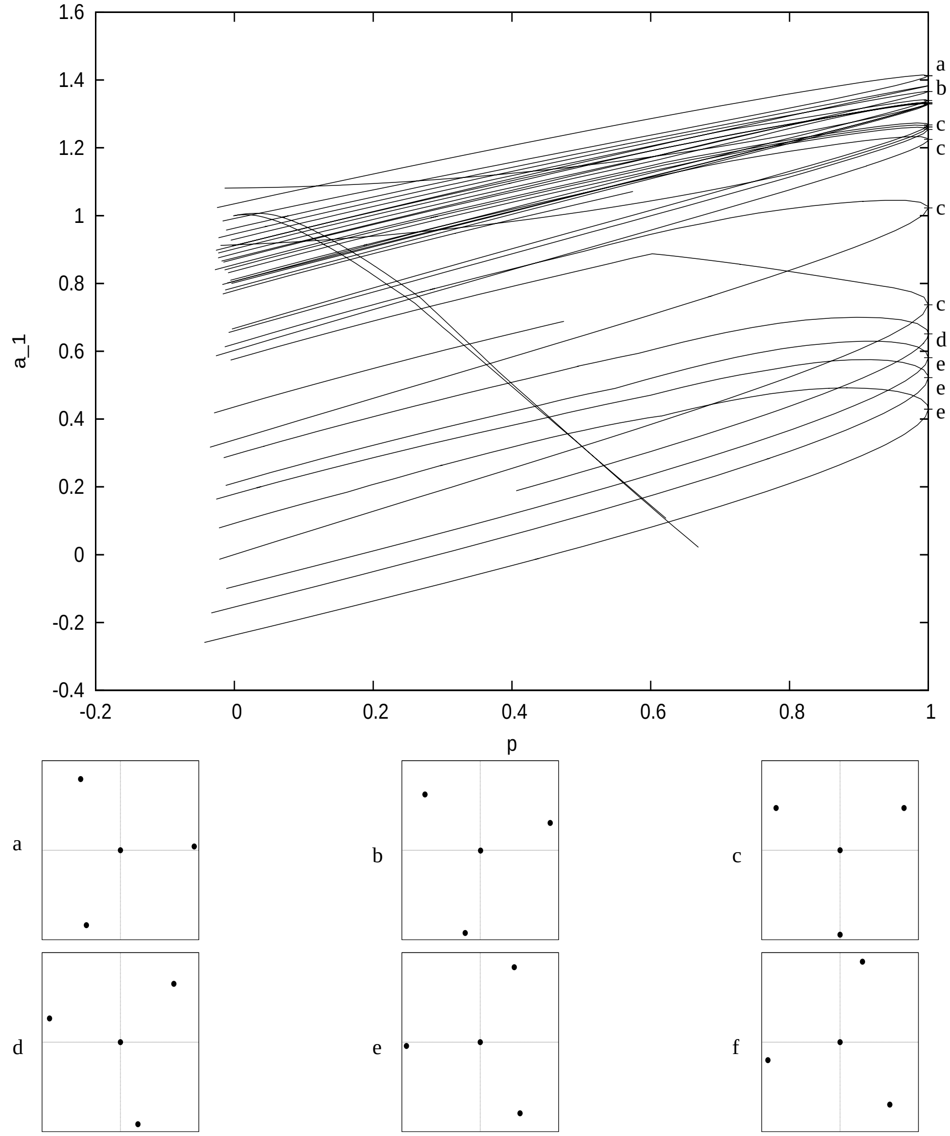


Figure D.3: Bifurcation diagram with nearly all branches having limit points at $p = 1$. Despite this uniqueness, we find the same configuration previously found.

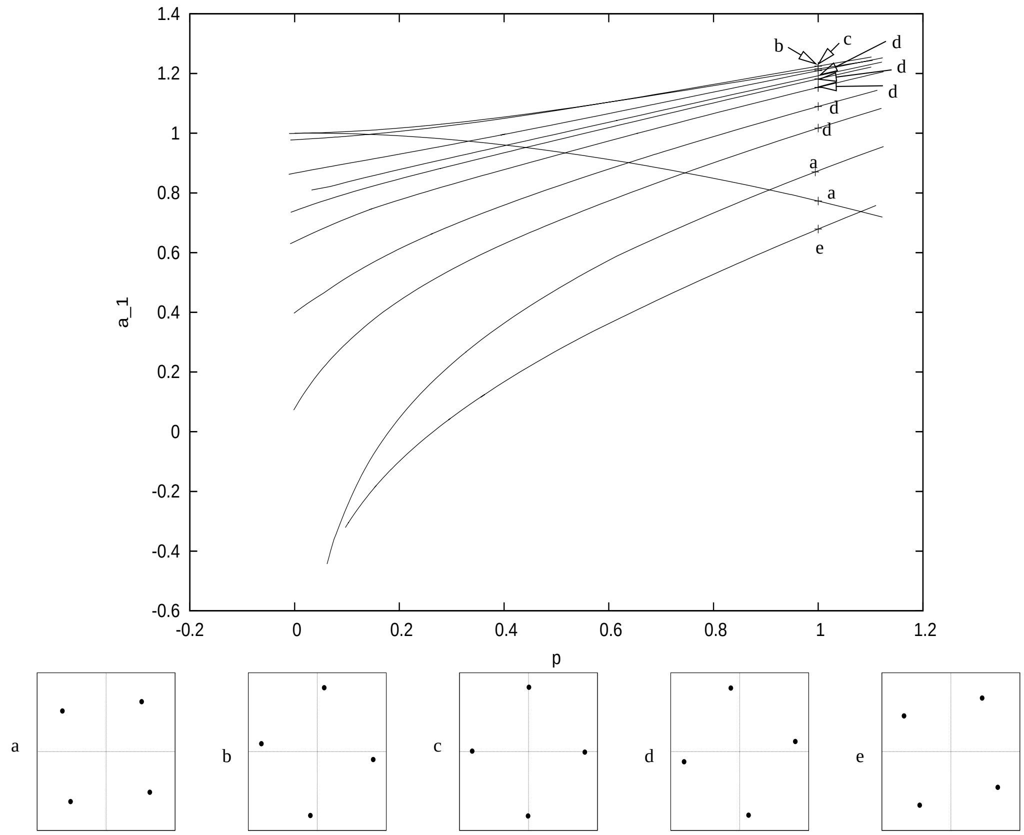


Figure D.4: Bifurcation diagram with no limit points. This co-rotating point also yields a new configuration, with four vortices placed equidistantly from one another on a circle of radius $\sqrt{\frac{3}{2}}$.

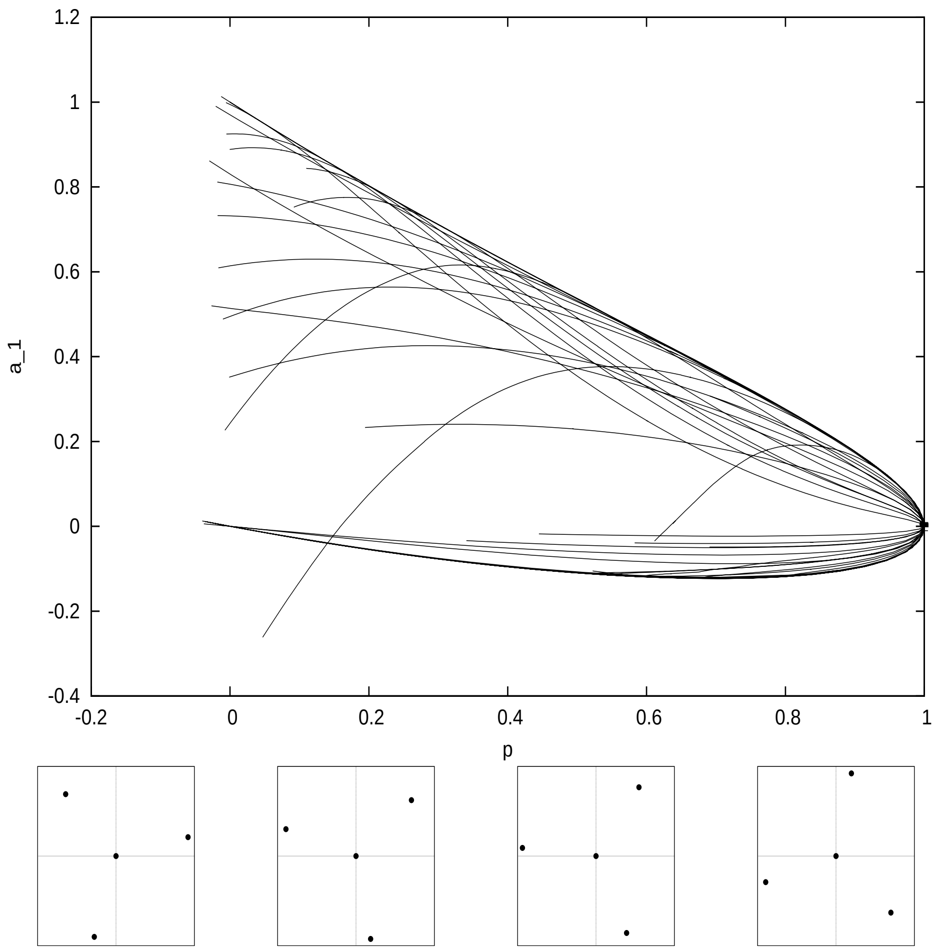


Figure D.5: Bifurcation diagram showing all branches converging to (and having limit points at) $p = 1$. For the configuration found, we see that a_1 correlates to the real component of the center point.

D.5.2 Higher Values of N

As new configurations are continually found and taken to be the new initial configuration in growing the next set of systems, it is very easy to become inundated with very large amounts of vortex configurations, for each of which we must test several co-rotating points. Several of these initial systems are shown with their co-rotating points in figure D.6.

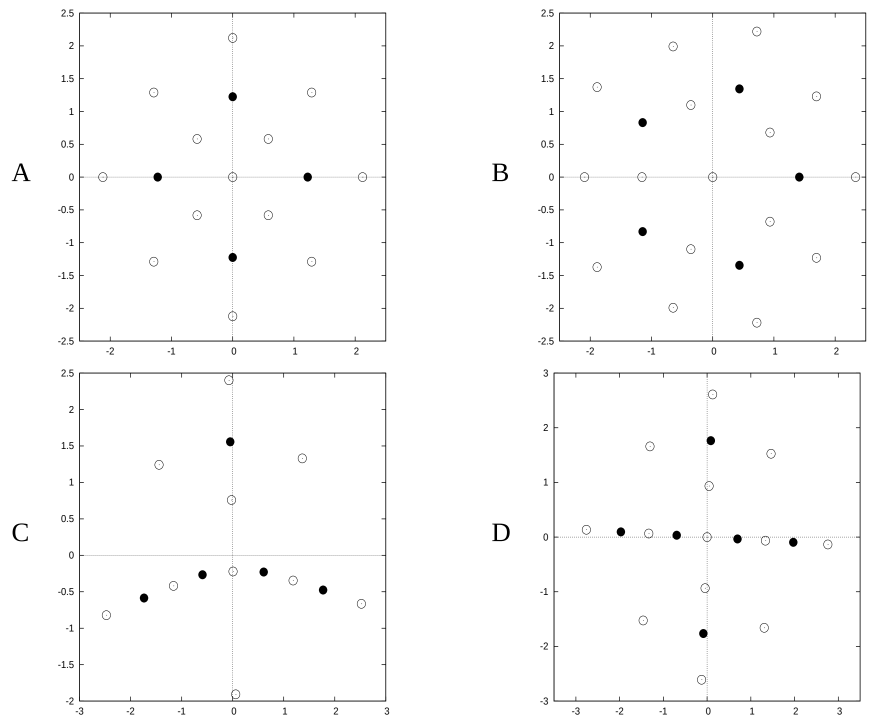


Figure D.6: Examples of co-rotating points for several initial systems of vortices. Filled circles represent initial vortices while open circles are co-rotating points.

Through a systematic and exhaustive search for equilibrium configurations of vortices for $N = 3, ..., 6$, we feel confident that we have found all equilibrium configurations for these N values. These configurations are shown in figure D.7.

Upon measuring the radii of configurations for N vortices on a circle and $N-1$ vortices rotating around a center vortex, we find that a pattern appears to emerge. We find that, for N vortices placed equidistantly on a circle the radius is $\sqrt{\frac{N-1}{2}}$ and that, for $N-1$ vortices on a circle around a center vortex, the radius is $\sqrt{\frac{N}{2}}$. Similar relations may emerge from other families, though the systems' complexities make it considerably more difficult to determine.

While previous research has shown configurations that lack both rotational and reflectional symmetry for $N \geq 8$, we see configurations that lack only rotational symmetry for $N \geq 5$ [2] [1]. We also see "families" of configurations emerge - systems that display similar patterns across different N values. While we cannot say with certainty that these families will be apparent for all N values greater than that at which they initially

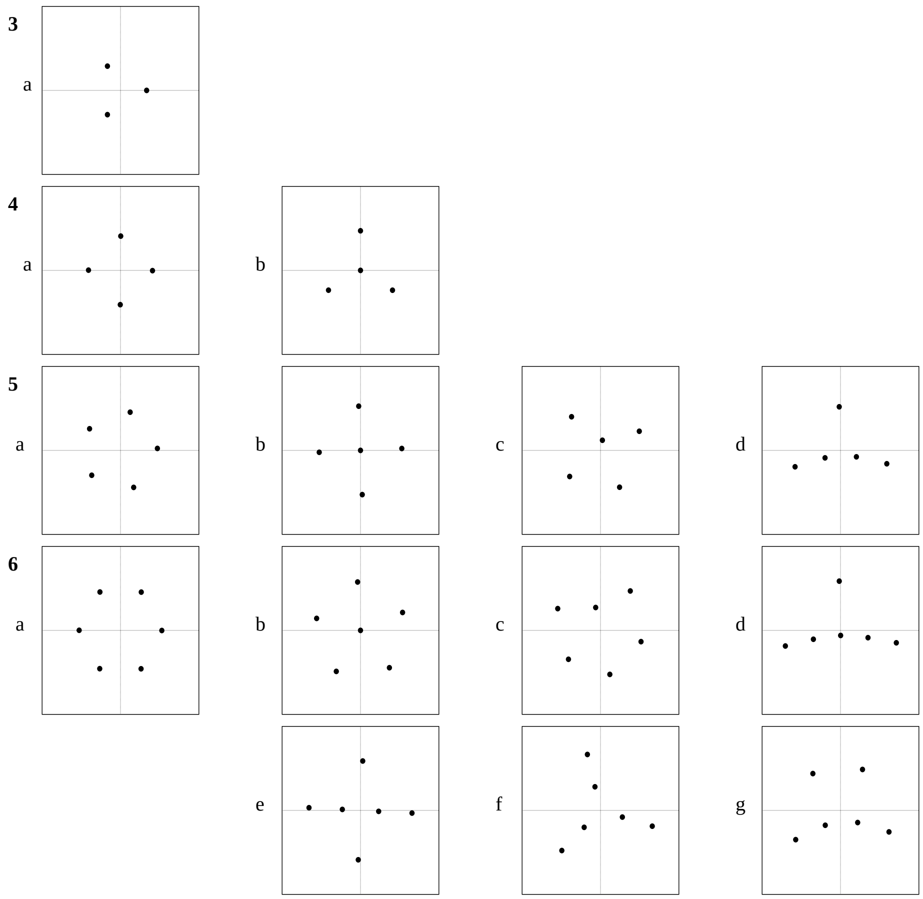


Figure D.7: Configurations grown for $N = 3, ..., 6$. Plots shown are from $(-3, -3)$ to $(3, 3)$. As N increases, we see new "families" of configurations - patterns that emerge and appear to be sustained even for different (increasing) N values.

appear, they seem to translate consistently to higher N values in our narrow scope of configurations. As figure D.7 clearly shows, it seems that every increase in N yields many more valid configurations and many more families of configurations, making it difficult at best to continue exhaustively finding all configurations for increasing N values. Fortunately, realizing that growing configurations from a given co-rotating point will yield identical configurations to those found from that co-rotating point's symmetric counterpart, we can considerably reduce the number of co-rotating points that we need to use as starting points in growing our configurations. Unfortunately, as we will see in the next section, there are cases when this does not necessarily hold true.

D.5.3 Asymmetric Configurations

As aforementioned, previous research has shown that asymmetric configurations can be found for $N \geq 8$. We will briefly analyze one of these cases.

Consider the initial configuration of seven points given in figure D.8 with all co-rotating points. The points labeled A and B are our primary points of interest.

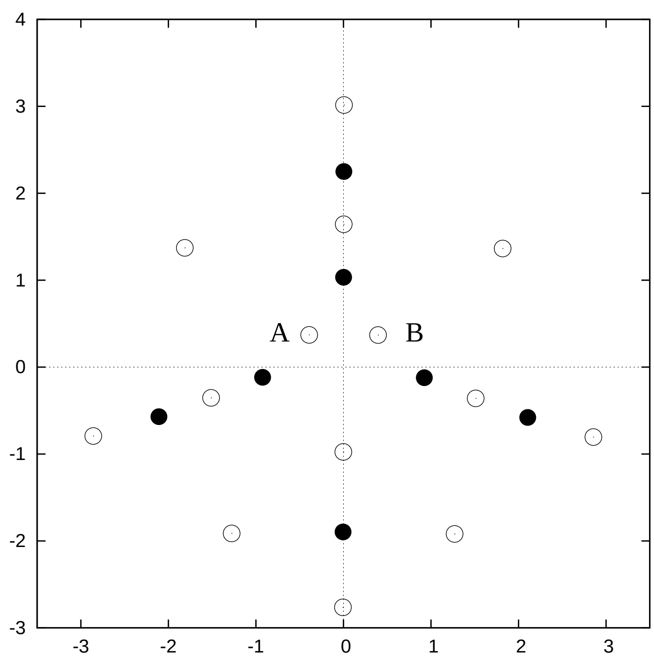


Figure D.8: Initial configuration of 7 vortices shown with all co-rotating points. Our points of interest are the co-rotating points labeled A and B.

Taking point A as our initial co-rotating point, we find the asymmetric configuration of eight points given in figure D.9

Clearly asymmetric configurations such as this are intriguing. Though we see no rotational or reflectional symmetry, there are still vortices that share a common radius. However, more importantly, we see that, if we take co-rotating point B (the symmetric counterpart of A), and choose to grow a configuration, we do

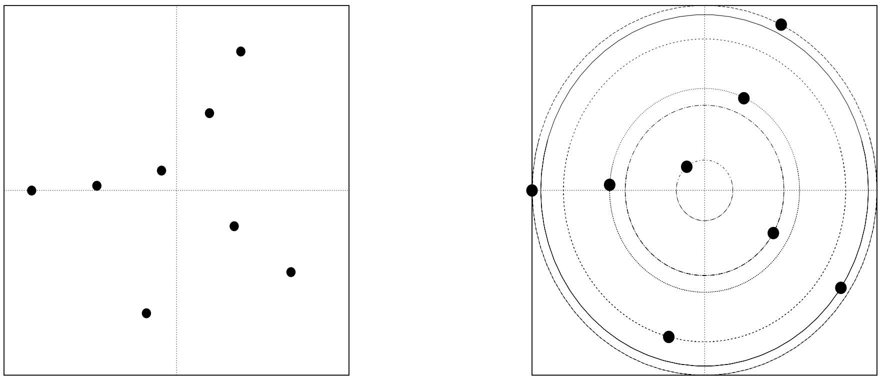


Figure D.9: Asymmetric configuration of eight vortices obtained from growing from co-rotating point A. The circles drawn in are for convenience in seeing which vortices share a common radius.

not get the identical configuration as obtained from A, nor any rotation of it. We rather get a reflection of the configuration obtained from A. While certainly similar, we see this is a unique configuration. The two configurations are given in figure D.10.

A B

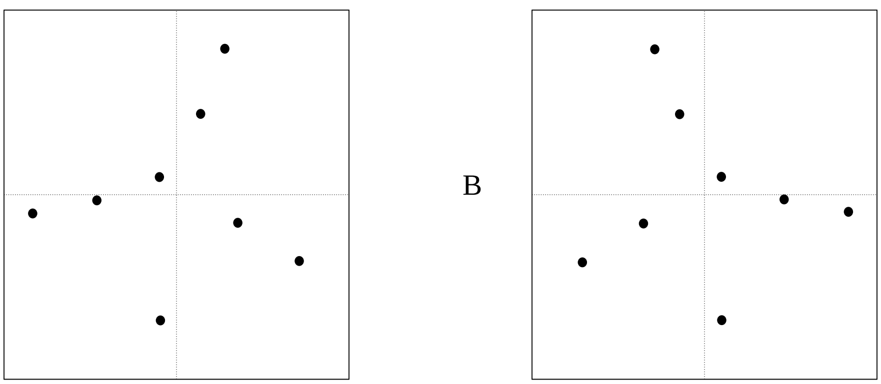


Figure D.10: Two unique asymmetric configurations found from co-rotating points A and B.

This case tells us several things. First, it shows us that we can no longer assume that a co-rotating point yields an identical configuration as its symmetric counterpart. This makes things more difficult still. However, this example also shows that asymmetric configurations must always come in pairs, as the reflection of any asymmetric configuration also seems to be a valid configuration. We also find (not surprisingly) that the co-rotating points for an asymmetric initial configuration seems to also be asymmetric. Finally, we find that growing an equilibrium configuration from an asymmetric configuration can yield rotationally and reflectionally asymmetric configurations, as well as configurations with symmetric properties. Clearly, asymmetric configurations considerably complicate our purpose of growing equilibrium configurations systematically. The co-rotating points for an asymmetric configuration of eight points, as well as examples of both asymmetric and symmetric configurations of nine points grown from the asymmetric configuration are given in figure D.11.

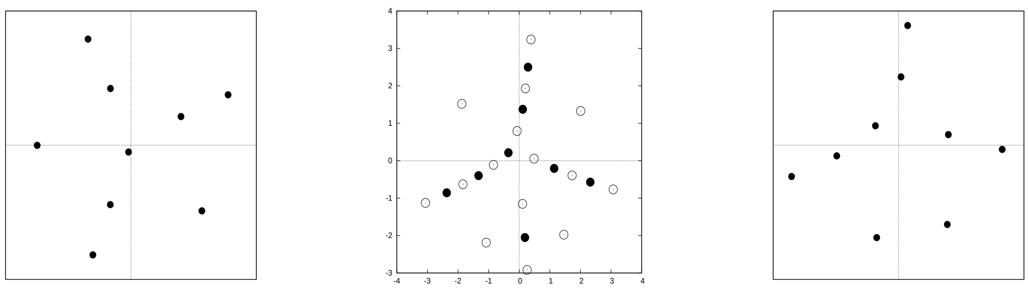


Figure D.11: Co-rotating points for asymmetric configuration of eight points with two possible resulting configurations. Notice that both asymmetric and symmetric configurations can be found.

D.6 Future Research

While at this point we have merely verified results found in previous research, the stage is now set to continue in a number of different directions. An obvious course of action is simply to continue growing equilibrium configurations for larger values of N beyond what is currently known. It is also possible to reverse this process, beginning with a system of N vortices and decreasing the strength of one of them from 1 to 0, finding new configurations of $N-1$ vortices. Through doing this, one could see if any new configurations are found, as well as compare with the growing process described in this article to note major differences and comparative effectiveness.

Another avenue could include comparing the stability of the configurations found with this method with stable configurations found in the past, as well as to see if any new stable configurations can be found through this method. Along similar lines, we can consider the energies of the systems and, more specifically, the change in energies when growing one configuration from another. If we consider the entropies of the systems in the same way, we can look specifically for cases where we see configurations being grown that have a lower energy, yet higher entropy. While we have started looking into this relatively untapped area of research, we have not yet obtained substantial enough results to present in this report.

Yet another interesting aspect to look at in future research would be to adjust the programs used for the method described in this report for growing a system of vortices on a sphere. While the method described here would essentially remain the same, we would primarily simply be changing the equations being considered for the vortex systems. Paul Newton gives the equations needed to consider a system of vortices on a sphere in [4]. His book also gives a very detailed overview of vortex systems in general as well as areas of research such as those mentioned here.

Though many of these areas have already been researched to a small extent, as in [1] and [3], there is still clearly a wealth of possibilities for new discoveries to be made. Indeed, vortex systems have seen a rich history of past research in a number of different applications, and it seems the topic will continue to provide such opportunities in the future.

D.7 Complete Work for Conversion to Real Variables

$$
\begin{aligned}
z_k^* &= \Psi_k + \Phi_k \\
\Psi_k &= \sum_{j=1}^{N} \frac{1}{z_k - z_j} \\
&= \sum_{j=1}^{N} \frac{1}{(x_{2k-1} - x_{2j-1}) + i(x_{2k} - x_{2j})} \cdot \frac{(x_{2k-1} - x_{2j-1}) - i(x_{2k} - x_{2j})}{(x_{2k-1} - x_{2j-1}) - i(x_{2k} - x_{2j})} \\
&= \sum_{j=1}^{N} \frac{(x_{2k-1} - x_{2j-1}) - i(x_{2k} - x_{2j})}{(x_{2k-1} - x_{2j-1})^2 + (x_{2k} - x_{2j})^2} \\
\Psi_{2k-1} &= \sum_{j=1}^{N} \frac{(x_{2k-1} - x_{2j-1})}{(x_{2k-1} - x_{2j-1})^2 + (x_{2k} - x_{2j})^2} \\
\Psi_{2k} &= -i \sum_{j=1}^{N} \frac{(x_{2k} - x_{2j})}{(x_{2k-1} - x_{2j-1})^2 + (x_{2k} - x_{2j})^2} \\
\Phi_k &= \frac{p}{z_k - z_{N+1}} \\
&= \frac{p}{(x_{2k-1} - x_{2N+1}) + i(x_{2k} - x_{2N+2})} \cdot \frac{(x_{2k-1} - x_{2N+1}) - i(x_{2k} - x_{2N+2})}{(x_{2k-1} - x_{2N+1}) - i(x_{2k} - x_{2N+2})} \\
&= p \frac{(x_{2k-1} - x_{2N+1}) - i(x_{2k} - x_{2N+2})}{(x_{2k-1} - x_{2N+1})^2 + (x_{2k} - x_{2N+2})^2} \\
\Phi_{2k-1} &= p \frac{(x_{2k-1} - x_{2N+1})}{(x_{2k-1} - x_{2N+1})^2 + (x_{2k} - x_{2N+2})^2} \\
\Phi_{2k} &= -ip \frac{(x_{2k} - x_{2N+2})}{(x_{2k-1} - x_{2N+1})^2 + (x_{2k} - x_{2N+2})^2}
\end{aligned}
$$

Recombining terms, we then find:

$$
\begin{aligned}
x_{2k-1} &= \sum_{j=1}^{N} \frac{(x_{2k-1}-x_{2j-1})}{h_j} + p\frac{(x_{2k-1}-x_{2N+1})}{d_k} \\
x_{2k} &= \sum_{j=1}^{N} \frac{(x_{2k}-x_{2j})}{h_j} + p\frac{(x_{2k}-x_{2N+2})}{d_k} \\
f_{2k-1} &= x_{2k-1} - \sum_{j=1}^{N} \frac{(x_{2k-1}-x_{2j-1})}{h_j} - p\frac{(x_{2k-1}-x_{2N+1})}{d_k} \\
f_{2k} &= x_{2k} - \sum_{j=1}^{N} \frac{(x_{2k}-x_{2j})}{h_j} - p\frac{(x_{2k}-x_{2N+2})}{d_k}
\end{aligned}
$$

where $h_j = (x_{2k-1}-x_{2j-1})^2 + (x_{2k}-x_{2j})^2$, $d_k = (x_{2k-1}-x_{2N+1})^2 + (x_{2k}-x_{2N+2})^2$, $k = 1, ..., N$, and $j \neq k$.

Similarly, we find

$$
\begin{aligned}
z^*_{N+1} &= \sum_{j=1}^{N} \frac{1}{z_{N+1}-z_j} \\
x_{2N+1} - ix_{2N+2} &= \sum_{j=1}^{N} \frac{1}{(x_{2N+1}-x_{2j-1}) + i(x_{2N+2}-x_{2j})} \cdot \frac{(x_{2N+1}-x_{2j-1}) - i(x_{2N+2}-x_{2j})}{(x_{2N+1}-x_{2j-1}) - i(x_{2N+2}-x_{2j})} \\
x_{2N+1} &= \sum_{j=1}^{N} \frac{(x_{2N+1}-x_{2j-1})}{\tilde{h}} \\
x_{2N+2} &= \sum_{j=1}^{N} \frac{(x_{2N+2}-x_{2j})}{\tilde{h}} \\
f_{2N+1} &= x_{2N+1} - \sum_{j=1}^{N} \frac{(x_{2N+1}-x_{2j-1})}{\tilde{h}} \\
f_{2N+2} &= x_{2N+2} - \sum_{j=1}^{N} \frac{(x_{2N+2}-x_{2j})}{\tilde{h}}
\end{aligned}
$$

where $\tilde{h}_j = (x_{2N+1}-x_{2j-1})^2 + (x_{2N+2}-x_{2j})^2$.

D.8 Acknowledgements

I would like to thank my advisor, Dr. Mike Jolly, without whose guidance this project would not have been possible. My thanks to Indiana University for hosting this mathematics REU program and especially Dr.

Kevin Pilgrim and Dr. Chris Judge for directing the program. Also, thank you to the National Science Foundation for funding the program.

Bibliography

1. Aref, Newton, and Stremler Vortex crystals, *Advances in Applied Mathematics* **39** (2002) 1-80

2. Aref and Vainchtein Point vortices exhibit asymmetric equilibria, *Nature* **392** (1998) 769-770

3. Campbell and Ziff Vortex patterns and energies in a rotating superfluid, *Physical Review B* **20** (1979)

4. Newton, Paul K. *The N-Vortex Problem*, Springer-Verlag, New York, 2001

5. Press, Teukolsky, Vetterling, and Flannery *Numerical Recipes in Fortran 77: The Art of Scientific Computing*, Cambridge UP, New York, 1992

The ends of pants complexes of small genus

Ted Spaide
Harvey Mudd College

Indiana University REU Summer 2008
Advisor: ADVISOR

E.1 Definitions

Denote the surface of genus g with b boundary components by $S_{g,b}$, or simply S_g if $b = 0$.

Definition E.1. A curve on a surface is *essential* if it cannot be contracted to a point, and *non-peripheral* if it is not isotopic to a boundary component.

Definition E.2. Say $3g + b > 3$, or $g = 0, b = 3$. A *pants decomposition* of $S = S_{g,b}$ is a set $\{C_1, \dots, C_n\}$ of essential, non-peripheral (isotopy classes of) curves such that $S \backslash (C_1 \cup \cdots \cup C_n)$ is a disjoint union of thrice-punctured spheres, or "pairs of pants".

The number of curves n in a pants decomposition is $3g + b - 3$ and the number of pairs of pants is $2g + b - 2$. There is only one pants decomposition of $S_{0,3}$ and that is the empty set.

Definition E.3. The *pants graph* of a surface S is a graph whose vertices are the pants decompositions of S and whose edges connect any two pants decompositions $\{C_1, \dots, C_n\}$ and $\{C'_1, \dots, C'_n\}$ such that:

1. $\{C_2, \dots, C_n\} = \{C'_2, \dots, C'_n\}$ (after some reordering of the indices), and
2. C_1 and C'_1 have minimal geometric intersection among all pairs of isotopy classes of essential, non-peripheral curves on the component of $S \backslash (C_2 \cup \cdots \cup C_n)$ they belong to.

In particular, if C_1 and C'_1 belong to a once-punctured torus, they intersect once, and if they belong to a four-punctured sphere they intersect twice. In the first case they are said to differ by an "S-move" and in the second by an "A-move".

One can add 2-cells to this graph to form a cell complex, called the *pants complex*; we will be unconcerned with these here. For a general overview of the pants complex see [1].

The pants graph of S will be denoted by $\mathcal{PG}(S)$ or simply $\mathcal{PG}_{g,b}$ if $S = S_{g,b}$.

Given two vertices v, w of a graph, we may define the distance between them to be the smallest number of edges in a path between them, or infinity if no such path exists. In [2], Hatcher and Thurston show that the pants graph is indeed connected, so the distance between two decompositions is finite. Let $B_r(v)$ be the ball of radius r aroud the vertex v.

Definition E.4. Let O be an arbitrary vertex of a graph G. Say a subgraph of G is *infinitely deep* or has *infinite depth* if it is unbounded when considered as a subset of G. Let n_r be the (possibly infinite) number of connected components of $G \backslash B_r(O)$ with infinte depth, and let $n = \lim_{r\to\infty} n_r$. G is said to have n *ends*.

Remark E.1. For a subgraph of a graph G to be infinitely deep, it is necessary for it to be unbounded (when considered as a graph by itself), but this is not sufficient.

Remark E.2. Clearly this definition is independent of the choice of basepoint O.

Note that if S and T have finite diameter, and $S \subseteq T$, then each infinitely deep component of $G \backslash S$ will contain at least one infinitely deep component of $G \backslash T$. Thus n_r will never decrease because $B_r(O) \subseteq B_{r+1}(O)$, and the limit certainly exists (or approaches infinity). Also note that the choice of basepoint does not matter in our definition.

In [4] Masur and Schleimer show that $\mathcal{PG}_g$ has one end for $g \geq 3$. Although they only show this for closed surfaces, their arguments appear to be valid for $PG_{g,b}$ with $g \geq 3$ and arbitrary $b \geq 0$. The pants graphs $\mathcal{PG}_{0,4}$ and $\mathcal{PG}_{1,1}$ are isomorphic and have infinitely many ends (in fact, there are already infinitely many components when we remove $B_1(O)$).

E.2 Possible number of ends

We show that a certain class of graphs, including the pants graphs, can only have 0, 1, 2, or ∞ ends. Since the pants graphs have infinite diameter, they can thus have 1, 2, or ∞ ends.

Lemma E.2.1. *Let G be a connected graph and say there exists an $R > 0$ such that for $v \in G$, $G \backslash B_R(v)$ has at least 3 infinitely deep components. Then G has infinitely many ends.*

Proof. Say G has $n < \infty$ ends. Pick some $O \in G$ and some r such that $G \backslash B_{r'}(v)$ has n infinitely deep components for $r' \geq r$. Choose some infinitely deep component of $G \backslash B_R(v)$ and a v' in that component such that $d(v, v') = D > r + R$; note that $B_r(v')$ is contained in that component. Then there are at least $n + 1$ infinitely deep components of $G \backslash (B_R(v) \cup B_r(v'))$: $n - 1$ infinitely deep components of $G \backslash B_R(v)$ not containing v', and at least 2 infinitely deep components of $G \backslash B_r(v')$ which do not contain v. $B_{D+r}(v)$ contains $B_R(v) \cup B_r(v')$, so each infinitely deep component of $G \backslash (B_R(v) \cup B_r(v'))$ restricts to at least one infinitely deep component $G \backslash B_{D+r}(v)$. But by our choice of r, $G \backslash B_{D+r}(v)$ has only n components, a contradiction. $\square$

Lemma E.2.2. *Let G be a connected graph whose automorphism group $Aut(G)$ acts on its vertices cofinitely, that is, there are finitely many orbits of the vertices of G under the action of $Aut(G)$. Then G has 0, 1, 2, or ∞ ends.*

Proof. Say G has strictly more than 2 ends, so there exist $v \in G$ and $r > 0$ such that $G \backslash B_r(v)$ has some $n > 2$ infinitely deep components. Choose representatives $v_1, \ldots, v_n$ of the orbits of G; since G is connected there exists some D such that $d(v, v_i) < D$ for $1 \leq i \leq n$. Then $B_r(v) \subseteq B_{r+D}(v_i)$ for each i, so $G \backslash B_{r+D}(v_i)$ has at least n components. But then for any $v' \in G$ we can pick $\phi \in Aut(G)$ such that $\phi(v_i) = v'$ for some i, and then $\phi(G \backslash B_{r+D}(v_i)) = G \backslash B_{r+D}(v')$ has at least n components as well. Thus the conditions of lemma E.2.1 hold and G has infinitely many ends. $\square$

Theorem E.2.3. *$\mathcal{PG}_{g,b}$ has 1, 2, or ∞ ends.*

Proof. Note that a pants decomposition is determined up to homeomorphism by the combinatorics of how its pairs of pants are connected; that is, for each pair of pants, we need only specify which pair of pants (if any) each of its boundary components is attached to. Since there are finitely many ways to determine this, the homeomorphisms of $S_{g,b}$ with itself act cofinitely on the pants decompositions of $S_{g,b}$, and these induce automorphisms of $\mathcal{PG}_{g,b}$. Thus lemma E.2.2 applies; as noted above $\mathcal{PG}_{g,b}$ has infinite diameter and therefore has at least one end. $\square$

E.3 End calculations using end strucures of curve complexes

In this section we prove that $\mathcal{PG}_{g,b}$ has at most as many ends as certain complexes of curves associated to $S_{g,b}$.

Definition E.5. The curve complex $\mathcal{C}(S)$ of a surface S is a simplicial complex whose vertices are the (isotopy classes of) curves on S and which has an n-simplex spanning vertices $\gamma_0, \ldots, \gamma_n$ if they can be realized by disjoint curves on S. If $S = S_{1,1}$ we instead require that each pair γ_i, γ_j intersect once, and if $S = S_{0,4}$ we require that they intersect twice.

If $g \geq 2$, or if $g = 1$ and $b \geq 1$, or if $g \geq 0$ and $b \geq 4$, then $\mathcal{C}(S_{g,b})$ is connected and unbounded (see [6] lemma 1.21, exercise 1.31, and corollary 2.25).

We will need a result of Masur and Minsky to relate the curve complex of a surface S to the pants graph S. If V is a subsurface of S and α is an arc in W such that $\partial\alpha \subseteq \partial V$, let N be a regular neighborhood of $\alpha \cup \partial V$. Then ∂N will be a set of curves in V; define the *curve surgery* of α to be the subset of these curves which are essential and non-peripheral. If γ is a curve on S which intersects V, define the *subsurface projection* $\pi_V(\gamma)$ of γ in V as follows: if $\gamma \subseteq V$, then set $\pi_V(\gamma) = \{\gamma\}$; otherwise, $\gamma \cap V$ is a disjoint union of arcs, and $\pi_V(\gamma)$ is the union of the curve surgeries of those arcs. If P is a pants decomposition on S, define $\pi_V(P)$ to be the union of all $\pi_V(\gamma)$ where γ ranges over the curve in P; $\pi_V(P)$ has diameter at most 2 in $\mathcal{C}(W)$ (see [3] lemma 2.3). All of these operations preserve homotopy so there is no ambiguity in defining them.

A curve γ *cuts* a subsurface V if $\pi_V(\gamma)$ is nonempty; equivalently, γ cuts V iff it is not isotopic to any curve carried by $S\backslash V$. A subsurface V is *essential* if each of its boundary components are essential curves, and it is not an annulus. If V is a non-pants essential subsurface and P is a pair of pants, then it contains at least one curve which cuts V (so $\pi_V(P) \neq \emptyset$).

Given two curves γ and γ' which cut V, define $d_V(P, P')$ to be the distance between the sets $\pi_V(P)$ and $\pi_V(P')$ in $\mathcal{C}(V)$, and similarly, if P and P' are pants decompositions, define $d_V(P, P')$ to be the distance between $\pi_V(P)$ and $\pi_V(P')$ in $\mathcal{C}(V)$. When there is no subscript, $d(P, P')$ still denotes distance in the pants complex. Let $[x]_C$ equal x if $x \geq C$ and 0 otherwise.

Then, there exists $C_0 = C_0(S) \geq 1$ such that if $C \geq C_0$, there exist $K = K(C) \geq C$ and $E = E(C) \geq 0$, such that for any pants decompositions P and P' on S,

$$\frac{1}{K}\sum_V [d_V(P, P')]_C - E \leq d(P, P') \leq K\sum_V [d_V(P, P')]_C + E, \tag{E.1}$$

where the sums range over (isotopy classes of) essential non-pants subsurfaces V of S (see theorem 6.12 and section 8 in [3]).

Lemma E.3.1. *Let S be a surface and fix a basepoint $O \in \mathcal{PG}(S)$. Given $R > 0$, there exists $R' > 0$ such that for any essential $W \subseteq S$, if γ cuts W, $d_W(\gamma, O) > R'$ and $P \in \mathcal{PG}(S)$ contains γ, then $d(P, O) > R$*

Proof. Fix some $C \geq C_0$ and some K, E satisfying (E.1), and set $R' = K(R + E + C) + 2$. Then we have

$$\begin{aligned} d(P, O) &\geq \frac{1}{K}\sum_V [d_V(P, O)]_C - E \\ &\geq \frac{1}{K}[d_W(P, O)]_C - E \\ &\geq \frac{1}{K} d_W(P, O) - (E + C). \end{aligned}$$

Thus, given $R > 0$, if $d_W(P, O) > K(R + E + C)$, we will have $d(P, O) > R$. If $d_S(O, \gamma) > R'$ and if $P \in \mathcal{PG}(S)$ contains γ, then since $\pi_W(P)$ has diameter 2, we have $d_S(P, O) > R' - 2 = K(R + E + C)$, so $d(P, O) > R$. ∎

Definition E.6. Let $O \in \mathcal{PG}(S)$ and $R > 0$. Call a curve $\gamma \in \mathcal{C}(S)$ *R-far* from O if any pants decomposition $P \in \mathcal{PG}(S)$ containing γ lies outside $B_R(O)$.

In this language, lemma E.3.1 states that for any $R > 0$, any γ which satisfies the conditions of the theorem is R-far from O.

For the sake of convenience we expand the definition of the pants graph slightly.

Definition E.7. If S is a disk or annulus, define $\mathcal{PG}(S)$ to be the graph with one vertex and no edges.

If X be a disjoint union of surfaces,

$$X = \cup_{i=1}^{n} S^{(i)},$$

then define

$$\mathcal{PG}(S) = \Pi_{i=1}^{n} \mathcal{PG}(S^{(i)}),$$

the cartesian product of the pants graphs of the $S^{(i)}$. Namely, the vertices of $\mathcal{PG}$ are the n-tuples $(v_1, \ldots, v_n)$ where $v_i \in \mathcal{PG}(S^{(i)})$ and an edge connects two vertices $(v_1, \ldots, v_k, \ldots, v_n)$ and $(v_1, \ldots, v_k', \ldots, v_n)$ if v_k and v_k' are connected by an edge in $\mathcal{PG}(S^{(k)})$

Remark E.3. As long as none of the $S^{(i)}$ are annuli or disks, this provides the logical definition of $\mathcal{PG}(S)$, since a pants decomposition of the whole surface is just a pants decomposition of each component, and a single move corresponds to moving in just one surface. The definitions for disks and annuli are only included to simplify things.

We note that the product of connected graphs is also connected, so $\mathcal{PG}(S)$ is still always connected.

Lemma E.3.2. *Let S be a surface. Choose a basepoint $O \in \mathcal{PG}(S)$ and $R > 0$. If $\gamma \in \mathcal{C}(S)$ is R-far from O, then given two pants complexes P, P' both containing γ, there exists a path from P to P' which lies outside $B_R(O)$*

Proof. We construct a path from P to P', each of whose vertices contain γ. γ divides S into at most two components, neither of which are a disk or annulus since γ is essential and non-peripheral. If S' is a component of $S \backslash \gamma$ then P defines a pants decomposition on $S \backslash \gamma$ by taking those elements of $P \backslash \{\gamma\}$ which lie in S'; similarly given an element $P_0 \in \mathcal{PG}(S \backslash \gamma)$, we can add γ to the curves of thus P_0 to define a pants decomposition of S which contains γ. $P \backslash \{\gamma\}$ thus determines an element of $)PG(S \backslash \gamma)$, and similarly so does $P' \backslash \{\gamma\}$. Since $PG(S \backslash \gamma)$ is connected there is a path from $P \backslash \{\gamma\}$ to $P' \backslash \{\gamma\}$. This defines a path from P to P' whose vertices all contain γ, so the path lies o utside $B_R(O)$ by assumption. $\square$

Corollary E.3.3. *Let S be a surface. Choose a basepoint $O \in \mathcal{PG}(S)$ and $R > 0$. If $\gamma_0, \gamma_1, \ldots, \gamma_n$ is a path in $\mathcal{C}(S)$ such that each γ_i is R-far from O, P is a pants complex containing γ_0, and P' is a pants complex containing γ_n, then there exists a path from P to P' which lies outside $B_R(O)$.*

Proof. For $0 \leq i < n$ let P_i be some arbitrary pants decomposition containing both γ_i and γ_{i+1}. By lemma E.3.2, there exist paths from P to P_0, P_i to P_{i+1} for $0 \leq i < n$, and P_{n-1} to P', all of which lie outside $B_R(O)$. Therefore there is a path from P to P' which lies outside $B_R(O)$. $\square$

Theorem E.3.4. *Let S be a surface. Choose a basepoint $O \in \mathcal{PG}(S)$. Given $R > 0$ there exists $R' > 0$ such that if P, P' are pants decompositions such that $\gamma \in P, \gamma' \in P'$, and γ and γ' belong to the same component of $\mathcal{C}(S) \backslash B_{R'}(O)$, then there is a path from P to P' in $\mathcal{PG}(S) \backslash B_R(O)$.*

Proof. Apply lemma E.3.1 to find R' such that $d(P, O) > R$ whenever P contains a γ satisfying $d_W(\gamma, O) > R'$. Say now P and P' are pants decompositions where $\gamma \in P, \gamma' \in P'$, $\gamma \subseteq W, \gamma' \subseteq W$, and $d_W(O, \gamma), d_W(O, \gamma') > R'$. Then by assumption there exists a path $\gamma = \gamma_0, \gamma_1, \ldots, \gamma_n = \gamma'$ in $\mathcal{C}(W)$ such that $d(\gamma_i, O) > R'$ for each i. Then each γ_i is R-far from O, so by corollary E.3.3 we are done. $\square$

Next we use the following result of Schleimer (see [7]):

Theorem E.3.5. *Let $g \geq 2$. For any vertex $\gamma \in \mathcal{C}(S_{g,1})$ and $r > 0$, the subcomplex $\mathcal{C}(S_{g,1})\backslash B_r(\gamma)$ is connected.*

Remark E.4. This is slightly stronger than the statement that $\mathcal{C}(S_{g,1})$ has one end, because $\mathcal{C}(S_{g,1})\backslash B_r(\gamma)$ has only one component, infinitely deep or not.

Given this, we then have:

Corollary E.3.6. *Let $g \geq 2$. Given $O \in \mathcal{PG}_{g,1}$ and $R > 0$, there exists $R' > 0$ such that if P_1, P_2 are pants decompositions with $\gamma_i \in P_i$ and $d_S(\gamma_i, O) > R'$ for $i = 1, 2$, then there is a path from P_1 to P_2 in $\mathcal{PG}_{g,1}\backslash B_R O$*

Next we want to show that we can always move pants decompositions far out enough in the curve complex. First we have a partial converse of lemma E.3.1.

Lemma E.3.7. *Let S be a surface and fix a basepoint $O \in \mathcal{PG}(S)$. Given $A > 0$, there exists $A' > 0$ such that if $d(P, O) > A'$ then $d_V(P, O) > A$ for some essential non-pants subsurface V of S.*

Proof. Pick some $C > A$ and find $E = E(C)$ and $K = K(C)$ satisfying (E.1). Set $A' = E$. Then if $d(P, O) > A'$ we have

$$K \sum_V [d_V(P, P')]_C + E \geq d(P, O) > E,$$

and so $[d_V(P, P')]_C$ must be strictly positive for some essential non-pants V. By definition this means that $d_V(P, P') \geq C > A$. ◻

We will also need the following result ([6] lemma 2.28).

Lemma E.3.8. *Suppose that V is an essential subsurface of S, and let $\{\gamma_0, \gamma_1, \ldots, \gamma_n\}$ be a path in $\mathcal{C}(S)$ such that every γ_i cuts V. Then $d_V(\gamma_0, \gamma_n) \leq 6n$.*

Theorem E.3.9. *Let $g \geq 2$ and $b \geq 0$, or let $g = 1$ and $b \geq 2$. Say $O \in \mathcal{PG}_{g,b}$, $R > 0$, and $A > 0$. Then there exists $A' > R$ such that if $d(O, P) > A'$ then there is a path in $\mathcal{PG}_{g,b}\backslash B_R(O)$ from P to some $P' \in \mathcal{PG}_{g,b}\backslash B_R(O)$ such that P' contains a curve γ satisfying $d_S(O, \gamma) > A$.*

Proof. Set $S = S_{g,b}$. By lemma E.3.1 there exists $R' > 0$ such that if $d_V(O, \gamma) > R'$ for essential $V \subseteq S$ which γ cuts, then γ is R-far from O. By lemma E.3.7 there exists $A' > 0$ such that if $d(O, P) > A'$ then $d_V(O, P) > R'$ for some non-pants essential $V \subseteq S$; without loss of generality we can take $A' > R$. Then we can find some $\beta \in P$ which cuts V, so that $d_V(O, \beta) > R'$; in particular β is R-far from O.

We show that without loss of generality β is nonseparating. Say that β separates $S = S_{g,b}$ into $S^{(1)}$, $S^{(2)}$. Since $g \geq 1$, one of the $S^{(i)}$ has nonzero genus; call it V'. V' is essential since its boundary is a union of β and boundary components of S, and is not a pair of pants or annulus since it has nonzero genus. Now, given any two curves λ_1 and λ_2 on V', if there exists a homeomorphism from $V'\backslash\lambda_1$ to $V'\backslash\lambda_2$, then we can extend such a homeomorphism onto λ_1 and λ_2; thus any curve on V' is determined by the homeomorphism class of $V'\backslash\lambda$. There are finitely many such classes, since if λ is nonseparating, then $V'\backslash\lambda$ is homeomorphic to $S_{g-1,b+2}$, and if it is separating, $V'\backslash\lambda$ is a disjoint union of S_{g_1,b_1} and S_{g_2,b_2}, where $g_1 + g_2 = g$ and $b_1 + b + 2 = b + 2$. Thus $\mathcal{C}(V')$ is cofinite under the action of homeomophisms of S with itself. In particular, any vertex of $\mathcal{C}(V')$ lies within a constant distance M of a nonseparating curve, so we can find a curve β' such that $d_{V'}(O, \beta') > (M + R')$, and then a nonseparating curve β'' within M of β', so that $d_{V'}(O, \beta') > R'$, and so β'' is R-far from O. Since β'' does not separate V', it does not separate S. Since β is R-far from O, we can use lemma E.3.2 to move P to some P' containing both β and β''. Then we can replace P with P'', β with β'', and V with V'. Thus without loss of generality β is nonseparating.

Let $N(\beta)$ be a regular neighborhood of β and set $W = S \backslash N(\beta)$. W is essential since its boundary components consist of two curves isotopic to β and the boundary components of S. W is homeomorphic to $S_{g-1,b+2}$, which is not an annulus or pair of pants because either $g-1>0$ or $b+2>3$. Then $\pi_W(O)$ is nonempty, and since $\mathcal{C}(W)$ is unbounded we can find a curve $\gamma_0 \subseteq W$ such that

$$d_W(O, \gamma_0) > R' + 12A + 18. \tag{E.2}$$

Choose then a pants decomposition P_0 containing β and γ_0. By lemma E.3.2 there is a path connecting P and P_0 which lies in $\mathcal{PG}(S) \backslash B_R(O)$.

Any curve which does not cut W is isotopic to a curve on $S \backslash W = N(\beta)$, so the only curve which does not cut W is β itself. Connect γ_0 to a vertex which is distance $2A+3$ away by a geodesic, that is, find a path $\gamma_0, \gamma_1, \ldots, \gamma_{2A+3}$ such that $d_S(\gamma_0, \gamma_{2A+3}) = 2A+3$.

We show now that without loss of generality, each of the γ_i cut W, that is, none of the γ_i are β. By definition, γ_0 and β are disjoint. If $\gamma_i = \beta$ for $i > 1$, then since $d_S(\gamma_0, \beta) = 1$, there is a path $\gamma_0, \gamma_i, \gamma_i + 1, \ldots, \gamma_{2A+3}$ with length strictly less than $2A+3$, contradicting the assumption that $d_S(\gamma_0, \gamma_{2A+3}) = 2A+3$. We are left with the case that $\gamma_1 = \beta$; in this case choose a curve λ which intersects β but not γ_0, and consider instead the path $D_\lambda(\gamma_0), D_\lambda(\gamma_1), \ldots, D_\lambda(\gamma_{2A+3})$, where D_λ denotes the Dehn twist about λ. $\lambda \cap \gamma_0 = \emptyset$, so $D_\lambda(\gamma_0) = \gamma_0$. However, λ and β intersect, so $D_\lambda(\gamma_1) \neq \gamma_1$. Finally, Dehn twisting preserves distance, so for the same reason as a bove we cannot have $D_\lambda(\gamma_i) = \beta$ for $i > 1$. In either case we have a geodesic of length $2A+3$ whose vertices cut W, so without loss of generality we can assume all of the γ_i cut W.

Now, suppose that $d_S(\gamma_0, O) \leq A$ and $d_S(\gamma_{2A+3}, O) \leq A$. Then γ_0 and γ_{2A+3} both lie within distance A of the set $\pi_S(O)$, which is a set with diameter 2. But then $d_S(\gamma_0, \gamma_{2A+3}) \leq 2A+2$, a contradiction. Thus either $d_S(\gamma_0, O) > A$ or $d_S(\gamma_{2A+3}, O) > A$. If $d_S(\gamma_0, O) > A$, then P_0 is our desired pants decomposition and we are done. Thus suppose $d_S(\gamma_{2A+3}, O) > A$. Since each γ_i cuts W, we have by lemma E.3.8 that $d_W(\gamma_0, \gamma_i) \leq 6i$ for $0 \leq i \leq 2A+3$. By (E.2), we have by the triangle inequality that $d_W(\gamma_i, O) > R' + 6(2A+3-i) \geq R'$. By our choice of R' this means that the γ_i are R-far from O. Then by corollary E.3.3 we can connect P_0 to some arbitrary pants decomposition containing γ_{2A+3}. Then $d_S(\gamma_{2A+3}, O) \geq d_S(\gamma_{2A+3}, O) > A$ an d we are done. □

Corollary E.3.10. *For $g \geq 2$, $\mathcal{PG}_{g,1}$ has one end.*

Proof. Fix a basepoint O and $R > 0$. Find an R' satisfying the conclusion of lemma E.3.1, and an A' satisfying the conclusion of theorem E.3.9 with $A = R'$. Say P_1 and P_2 lie in infinitely deep components of $\mathcal{PG}_{g,1} \backslash B_R(O)$. By definition, for $i = 1, 2$ we can find a path in $\mathcal{PG}_{g,1} \backslash B_R(O)$ from P_i to some P_i' such that $d(O, P_i') > A'$. By theorem E.3.9 we can find a path in $\mathcal{PG}_{g,1} \backslash B_R(O)$ connecting P_i' to some P_i'', where P_i'' contains a curve γ_i satisfying $d(\gamma_i, O) > R'$. By corollary E.3.6 there is a path in $\mathcal{PG}_{g,1} \backslash B_R(O)$ from P_1'' to P_2''. Thus P_1 and P_2 lie in the same component. □

Of course, the reason we can prove this for $\mathcal{PG}_{g,1}$ for $g \geq 2$ is due to theorem E.3.5; if we had an analogous theorem for some other $\mathcal{C}(S_{g,b})$ satisfying $g \geq 2$ or $g = 1$ and $b \geq 2$ then we could show that $\mathcal{PG}_{g,b}$ has one end as well. However, none of these results are known. Also, there are other proofs that $\mathcal{PG}_{g,b}$ has one end for $g = 2, b \geq 2$, or for $g \geq 3$, so the only cases for which this would be useful are $g = 1, b \geq 2$, or $g = 2, b = 0$.

E.4 End calculations using end structures of pants graphs of subcomplexes

In the previous section we used the one-endedness of a complex of curves associated with a surface to guarantee we stay away from a base pants decomposition, then used the connectedness of the pants graphs of subsurfaces to actually follow the paths we made in the curve complex. In this section we will still be using the complexes of curves to provide a general course and using the pants graphs to follow this course, but now we will use the one-endedness of the pants graphs to stay away from the basepoint.

Lemma E.4.1. *Let S be a surface. Fix disjoint curves $\gamma_1, \ldots, \gamma_n$ on S. Set $S' = S \backslash (\gamma_1 \cup \cdots \cup \gamma_n)$. Then there exist $A, B > 0$ such that for any pants decompositions P, P' such that $\{\gamma_1, \ldots, \gamma_n\} = \Gamma \subseteq P, P'$,*

$$A \cdot d_{\mathcal{PG}(S')}(P\backslash\Gamma, P'\backslash\Gamma) - B \leq d_{\mathcal{PG}(S)}(P, P') \leq d_{\mathcal{PG}(S')}(P\backslash\Gamma, P'\backslash\Gamma). \tag{E.3}$$

Proof. Let $S^{(1)}, \ldots, S^{(k)}$ be the components of S' and let P_i be the pants decomposition of $S^{(i)}$ induced by $P\backslash\Gamma$, and similarly for P'_i Pick some $C > C_0(S), C_0(S^{(1)}), \ldots, C_0(S^{(n)})$ (defined before (E.1)) and pick E_1, K_1 such that

$$\frac{1}{K_1} \sum_{V \subseteq S} d_V(P, P') - E_1 \leq d_{\mathcal{PG}(S)}(P, P') \tag{E.4}$$

where the V range over essential non-pants $V \subseteq S$.

Similarly pick E_2, K_2 large enough that

$$d_{\mathcal{PG}(S^{(i)})}((P_i), P'_i) \leq K_2 \sum_{V \subseteq S^{(i)}} d_V(P_i, P'_i) + E_2$$

for all $1 \leq i \leq k$. Summing over for all i yields

$$\begin{aligned} d_{\mathcal{PG}(S')}(P\backslash\Gamma, P'\backslash\Gamma) &= \sum_{i=1}^{k} d_{\mathcal{PG}(S^{(i)})}((P_i), P'_i) \\ &\leq K_2 \sum_{V \subseteq S^{(i)}} d_V(P_i, P'_i) + kE_2 \end{aligned} \tag{E.5}$$

where the sum now ranges over non-pants V which are essential subsurfaces of any $S^{(i)}$. Since the $S^{(i)}$ are themselves essential, so are all the subsurfaces V, so we have

$$\sum_{V \subseteq S^{(i)}} d_V(P_i, P'_i) \leq \sum_{V \subseteq S} d_V(P, P'). \tag{E.6}$$

Combining equations (E.4), (E.5), and (E.6) give the result

$$A \cdot d_{\mathcal{PG}(S')}(P\backslash\Gamma, P'\backslash\Gamma) - B \leq d_{\mathcal{PG}(S)}(P, P')$$

for some choice of A and B.

Clearly $d_{\mathcal{PG}(S)}(P, P') \leq d_{\mathcal{PG}(S')}(P\backslash\Gamma, P'\backslash\Gamma)$, since any path from $P\backslash\Gamma$ to $P'\backslash\Gamma$ in $\mathcal{PG}(S')$ defines a path of the same length from P to P' by adding Γ to each vertex. ∎

Lemma E.4.2. *Let S be a surface and let $O \in \mathcal{PG}(S)$ and $R > 0$. There exists R' such that if $d(O, P) > R'$ then P lies in an infinitely deep component of $\mathcal{PG}(S)\backslash B_R(O)$.*

Proof. Choose $A > 0$ be lemma E.3.1 such that if $d_V(\gamma, O) > A$ for essential nonpants $V \subseteq S$ then γ is R-far from O. Choose $R' > 0$ by lemma E.3.7 such that if $d(O, P) > R'$ then P contains a curve γ such that $d_V(\gamma, O) > A$ for some essential nonpants $V \subseteq S$.

Then let $d(O, P) > R'$; P contains some curve γ which is R-far from O. For any $N > 0$ use lemma E.3.1 to find some $\gamma_N \subseteq S\backslash\gamma$ which is N-far from O, and let P_N be some pants decomposition containing both γ and γ_N. By lemma E.3.2 there is a path from P to P_N which lies outside of $B_R(O)$. Thus whatever component of $\mathcal{PG}(S)\backslash B_R(O)$ contains P must contain pants decompositions which are arbitrarily far from O, and by definition has infinite depth.

Lemma E.4.3. *Let S be a surface and let $O \in \mathcal{PG}(S)$ and $R > 0$. There exists R' such that the following holds: if α and β are curves on S such that $\mathcal{PG}(S\backslash\alpha)$ and $\mathcal{PG}(S\backslash\beta)$ have one end, at least one component $S\backslash(\alpha \cup \beta)$ has an unbounded curve complex, and P, P' are pants complexes such that $\alpha \in P, \beta \in P'$ and $P, P' \in \mathcal{PG}(S)\backslash B_{R'}(O)$, then there is a path from P to P' lying in $\mathcal{PG}(S)\backslash B_R(O)$.*

Proof. Choose A by lemma E.3.1 so that if $d_V(O, \gamma) > A$ for some essential nonpants $V \subseteq S$ then γ is R-far from O. Choose R' by lemma E.3.1 so that any pair of pants Q satisfying $d(O, P) > R'$ contains a γ such that $d_V(O, \gamma) > A$ for some essential nonpants $V \subseteq S$, and in particular, is R-far from O.

Then let $P, P' \in \mathcal{PG}(S)\backslash B_{R'}(O)$ with $\alpha \in P$, $\beta \in P'$. Choose a component V of $S\backslash(\alpha \cup \beta)$ whose curve complex has infinite diameter; note that V is essential and not a pair of pants. Choose some $\lambda \in \mathcal{C}(V)$ such that $d_V(O, \lambda) > A'$, so λ is R-far from O. We move P to some P_1 containing λ and similarly P' to some P_1' containing λ. By our choice of R', P contains a curve which is R-far from O. If this curve is α, we can use lemma E.3.2 to move P to some P_1 containing both α and λ.

Thus assuume that some curve $\alpha' \in P$ is R-far from O and α is not. Since α is not R-far from O, there is some O' containing α such that $d(O, O') \leq R$. Choose A, B satisfying the results of lemma E.4.1, setting $\Gamma = \{\alpha\}$, $S' = S\backslash\alpha$. Then if Q is any other decomposition containing α such that $d(O, Q) \leq R$, we have by the triangle inequality that $d(O', Q) \leq 2R$, so by lemma E.4.1, $d_{\mathcal{PG}(S')}(O'\backslash\alpha, Q\backslash\alpha) \leq \frac{2R+B}{A} = M$.

We then find a path from P to some P_0 such that $\alpha \in P_0$ and $P_0\backslash\{\alpha\}$ lies in an infinitely deep component of $\mathcal{PG}(S')\backslash B_M(O'\backslash\{\alpha\})$, and this path stays outside of $B_R(O)$. By assumption $\mathcal{PG}(S\backslash\alpha)$ has one end, so it is not $S_{0,4}$ or $S_{1,1}$, and there is at least one component W of $S'\backslash(\alpha')$ whose curve complex $\mathcal{C}(W)$ has infinte diameter. By lemma E.4.2 there is some M' such that any pants decomposition at least M' away from $O'\backslash\alpha$ lives in an infinitely deep component of $\mathcal{PG}(S')\backslash B_M(O'\backslash\{\alpha\})$. Then by lemma E.3.1, we can choose some curve α'_M on W which is M'-far from $O'\backslash\{\alpha\}$. By lemma E.3.2 we can move P to some pants decomposition P_0 containing α, α', and α'_M wit hout entering $B_R(O)$.

Let P_1 be any pants decomposition containing both α and λ. Since λ was chosen to be R-far from O, we can use the same argument to find a path outside $B_R(O)$ from P_1 to some P_2 such that P_2 contains α and $P_2\backslash\{\alpha\}$ lies in an infinitely deep component of $\mathcal{PG}(S')\backslash B_M(O'\backslash\{\alpha\})$. Since S' is assumed to have one end, there is a path from $P_0\backslash\{\alpha\}$ to $P_2\backslash\{\alpha\}$ which stays outside $B_M(O'\backslash\{\alpha\})$. The path from P_0 to P_2 induced by adding α to each vertex then stays outside $B_R(O)$ because by our earlier discussion, if some Q on the path lay in $B_R(O)$, then $Q\backslash\{\alpha\}$ lies inside $B_M(O'\backslash\{\alpha\})$.

Thus there is a path outside of $B_R(O)$ from P to P_1 containing λ. Similarly P' can be moved to P_1' containing λ; since λ is R-far from O, we are done by lemma E.3.2. ∎

Theorem E.4.4. *Let S be a surface and $\mathcal{D}$ be some complex whose vertices are curves on S. Suppose that the following properties are true:*

1. *$\mathcal{D}$ is connected.*

2. *Given $O \in \mathcal{PG}(S)$, $R > 0$, each infinitely deep component of $\mathcal{PG}(S)\backslash B_R(O)$ contains some P containing a curve γ which lies in $\mathcal{D}$.*

3. *For each $\gamma \in \mathcal{D}$, $\mathcal{PG}(S\backslash\gamma)$ has one end.*

4. *If $\beta, \gamma \in \mathcal{D}$ are connected by an edge, then at least one component of $S\backslash(\beta \cup \gamma)$ has an unbounded curve complex.*

Then $\mathcal{PG}(S)$ has one end.

Proof. Suppose that P, P' lie in some infinitely deep components of $\mathcal{PG}(S)\backslash B_R(O)$. Choose R' as in lemma E.4.3. Move P far away from O to some P_0 in an infinitely deep component of $\mathcal{PG}(S)\backslash B_{R'}(O)$, and use property 2 to move this to some P_1 containing some $\gamma \in \mathcal{D}$ such that $d(O, P_1) > R'$; similarly move P' to some P_1' containing some $\gamma' \in \mathcal{D}$ such that $d(O, P_1') > R'$.

By property 1 there is some path $\gamma = \gamma_0, \ldots, \gamma_n = \gamma'$ in $\mathcal{D}$. For $0 < i < n$, $S\backslash\gamma_i \cup \gamma_{i+1}$ has a component with an unbounded curve complex, so in particular we can find an R'-far curve β_i which is disjoint from γ_i and we let Q_i be some pants complex which contains both β_i and γ_i. By properties 3 and 4 there are therefore paths outside of $B_R(O)$ connecting P_1 to Q_1, Q_i to Q_{i+1} for $0 < i < n-1$, and Q_{n-1} to P_1'. Thus there is a path in $\mathcal{PG}(S)\backslash B_R(O)$ from P to P_1 to P_1' to P', so $\mathcal{PG}(S)$ has one end. □

For the next result we will use an approach of Putman, at the heart of which is this lemma ([5] lemma 2.1):

Lemma E.4.5. *Let G be a group which acts on a simplicial complex X. Fix a basepoint $v \in X^{(0)}$ and a set S of generators of G. Assume the following hold:*

1. *For all $v' \in X^{(0)}$, the orbit Gv intersects the connected component of X containing v'.*

2. *For all $s \in S^{\pm 1}$, there is some path in X from v to $s \cdot v$.*

Then X is connected.

Theorem E.4.6. *Suppose that either*

1. *$g \geq 2$ and $b \geq 1$,*

2. *$g = 1$ and $b \geq 3$, or*

3. *$g = 0$ and $b \geq 6$.*

Then if $\mathcal{PG}(S_{g,b})$ has one end then so does $\mathcal{PG}(S_{g,b+1})$.

Proof. Let $K_1, \ldots, K_{b+1}$ are the boundary components of $S_{g,b+1}$ (see figure E.1 (a)). Consider the complex $\mathcal{D}$ of curves γ on $S = S_{g,b+1}$ which separate S into an homeomorphic copy of $S_{g,b}$ and a pair of pants whose boundary components are γ, K_b, and K_{b+1}. We show this is connected using lemma E.4.5.

Recall that the *pure mapping class group* of S is the group $PMod(S)$ of isotopy classes of homeomorphisms of S with itself which have the additional property that they send each K_i to itself as well. $PMod(S)$ is generated by Dehn twists about the curves in figure E.1 (b); see e.g. [1] section 4.4.5. Let γ, shown in figure E.1 (c) be our basepoint. We show $PMod(S)$ and $\mathcal{D}$ satisfy both of the properties of lemma E.4.5.

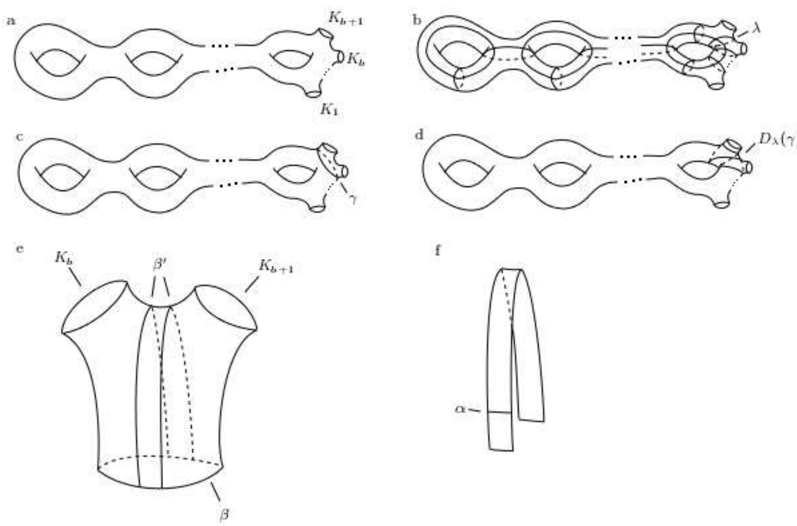


Figure E.1: a. The K_i b. The generators of PMod(S) c.γ d.$D_\lambda\gamma$ e. V, with the arcs of β' e. $V\backslash W$ and α

For the first, let $\gamma' \in X^{(0)}$; say that γ separates S into a subsurface V which is homeomorphic to $S_{g,b}$ and a pair of pants W whose boundary components are γ, K_b, and K_{b+1}, and γ separates S into V' which is homeomorphic to $S_{g,b}$ and W' whose boundary components are γ, K_b, and K_{b+1}. Send V onto V' by a homeomorphism sending each K_i to K_i and similarly W onto W' by a homeomorphism sending each K_i to K_i. Attach these homeomorphisms in a way that is consistent on γ; this is a homeomorphism of S which sends each K_i to K_i, so it is an element of $PMod(S)$, and it sends γ to γ'.

Note that none of the curves in figure E.1 (b) intersect γ except for the curve λ, so only the Dehn twist $D_\lambda^{\pm 1}$ does not leave γ fixed. But $D_\lambda(\gamma)$ intersects γ four times (see figure E.1 (c) and (d)) so there is a path from γ to $D_\lambda(\gamma)$ containing one edge; applying D_λ^{-1} to these curves show that γ and $D_\lambda^{-1}(\gamma)$ also only intersect four times. Thus the second property of lemma E.4.5 is satisfied, so $\mathcal{D}$ is connected, which is property 1 of lemma E.4.4.

As mentioned in the proof of theorem E.2.3, the group of homeomorphisms $S \to S$ acts on the pants graph cofinitely. If we instead look at how $PMod(S)$ acts on the pants graph, I claim the action is still cofinite. Let n be the number of pairs of pants in any pants decomposition of S. Define a *scheme* to be some specification, for each boundary component of each of n pairs of pants, of which other pair of pants that boundary component is attached to or which K_i that boundary component is. Given any two pants decomposition for which the pants are attached according to the same scheme, we can define homeomorphisms from each pair of pants in one to a pair of pants in the other in a way that respects this scheme; we extend these homeomorphisms to the curves of the pants decompositions to define an element of $PMod(S)$. Since the number of schemes is finite, the number of orbits of $PMod(S)$ on $\mathcal{PG}(S)$ is therefore als o finite. Then if we fix some scheme of attaching the pants for which one pair of pants has K_i and K_{i+1} as boundary components, some pants decompositions following this scheme will lie within some constant M of any other pants decomposition. Thus property 2 of lemma E.4.4 is satisfied.

Property 3 is satisfied by assumption, because for $\gamma \in \mathcal{D}$, we have $\mathcal{PG}(S_{g,b+1}\backslash\gamma) = \mathcal{PG}(S_{g,b})$.

Finally, say β and β' are connected by an edge in $\mathcal{D}$, so they intersect four times. Let V be the pair of pants cut of by β and W the pair of pants cut off by β'; β' then defines two arcs in V, each of which has both its endpoints on β. Therefore they are isotopic to the arcs shown in figure E.1 (e); note that β' then separates V into two annuli and a disk. Since the annuli both contain one of the K_i which W must contain, W contains the annuli. If W also contained the disk, then $V \subseteq W$, which is impossible since β and β' are not isotopic. Thus $V \backslash W$ is a single disk, which is the regular neighborhood in $S \backslash W$ of some arc α whose endpoints lie in β'. Consider now the subsurface $S \backslash (V \cup W)$. $S \backslash W$ is by assumption homeomorphic to $S_{g,b}$. $S \backslash (V \cup W) = (S \backslash W) \backslash (V \backslash W)$ is therefore some surface equal to $S_{g,b}$ minus the regular neighborhood of α. This will leave either $S_{g-1,b+1}$ or two surfaces S_{g_1,b_1} and S_{g_2,b_2} such that $g_1 + g_2 = g, b_1 + b_2 = b + 1$, and $b_1, b_2 \geq 1$. One can check that as long as the conditions of the theorem are satisfied, that in any case at least one of the components of $S \backslash (V \cup W)$, and therefore of $S \backslash (\beta \cup \beta')$, has a curve complex with infinite diameter, so property 4 is satisfied as well. Thus all the criteria of theorem E.4.4 are satisfied and we are done. ◻

In a similar vein, we have

Theorem E.4.7. *Say that $g \geq 2$ and $b \geq 0$. If $\mathcal{PG}_{g-1,b+2}$ has one end then so does $\mathcal{PG}_{g,b}$.*

Proof. We choose $\mathcal{D} = Nonsep(S_{g,b})$, the induced subcomplex of $\mathcal{C}(S_{g,b})$ containing only those vertices which are nonseparating curves on S. Property 1 of theorem E.4.4 is given by [5] theorem 1.2 (Putman proves this for $b = 0$, but the proofs for $b > 0$ are nearly identical). Property 2 follows as it did in theorem E.4.6. For $\gamma \in \mathcal{D}$, since γ is nonseparating, $S_{g,b} \backslash \gamma$ is homeomorphic to $S_{g-1,b+2}$, and $\mathcal{PG}_{g-1,b+2}$ has one end by assumption, giving us property 3. Finally if β and γ are connected by an edge, then $\gamma \in S_{g,b} \backslash \beta$, so $S_{g,b} \backslash (\beta \cup \gamma) = S_{g-1,b+2} \backslash \gamma$; this latter surface is either $S_{g-2,b+4}$ or a disjoint union $S_{g_1,b_1} \cup S_{g_2,b_2}$ where $g_1 + g_2 = g$, $b_1 + b_2 = b + 4$. In either case, at least one c omponent of $S_{g,b} \backslash (\beta \cup \gamma)$ has an unbounded curve complex, satisfying property 4. Thus theorem E.4.4 applies and we are done. ◻

Combining theorems E.3.10, E.4.6, and E.4.7 yields

Corollary E.4.8. *Let $g \geq 3, b \geq 0$ or $g = 2, b \geq 1$. Then $\mathcal{PG}_{g,b}$ has one end.*

As mentioned, this result is known for $g \geq 3$, but this gives an alternate proof.

Theorem E.4.9. *Say that $b \geq 3$. If $\mathcal{PG}_{0,b+2}$ has one end then so does $\mathcal{PG}_{1,b}$.*

Set $S = S_{1,b}$ and let $\mathcal{D}$ be the induced subcomplex of $\mathcal{C}(S)$ whose vertices are the nonseparating curves and separating curves β such that neither component of $S \backslash \beta$ is a pair of pants. Recall that $Mod(S)$ is the group of isotopy classes of homeomorphisms $S \to S$. Choose the base vertex γ of $\mathcal{D}$ (see figure E.2 (a)). We show that $\mathcal{D}$, γ, and $Mod(S)$ satisfy the conditions of lemma E.4.5. $Mod(S)$ acts cofinitely on $\mathcal{D}$ so the first condition is fulfilled. By the discussion in [1] 4.4.5, we can choose our generators of $Mod(S)$ to be the Dehn twists around the curves of E.1 (b) and certain maps which permute the boundary components of S. In particular these maps can be chosen so that they are constant on γ. Thus the only map which does not leave γ fixed is the Dehn twis t around λ (figure E.2 (b)), but there is an intermediate path from γ to γ' to $D_\lambda(\beta)$, and similarly $D_\lambda^{-1}(\beta)$ (figure E.2 (c) and (d)). Thus the second condition is fulfilled and $\mathcal{D}$ is connected, which is the property 1 of E.4.4.

Property 2 follows as it did in theorem E.4.6.

For $\gamma \in \mathcal{D}$, either γ is nonseparating, in which case $\mathcal{PG}(S \backslash \gamma) = \mathcal{PG}_{0,b+2}$ has one end by assumption, or γ separates S into surfaces $S^{(1)}$ and $S^{(2)}$ which are not pairs of pants, and in particular $\mathcal{PG}(S^{(i)})$ has infinite diameter; it is easy to show that the product of two graphs with infinite diameter has one end, so $\mathcal{PG}(S \backslash \gamma) = \mathcal{PG}(S^{(1)}) \times \mathcal{PG}(S^{(2)})$ has one end. Thus property 3 is true.

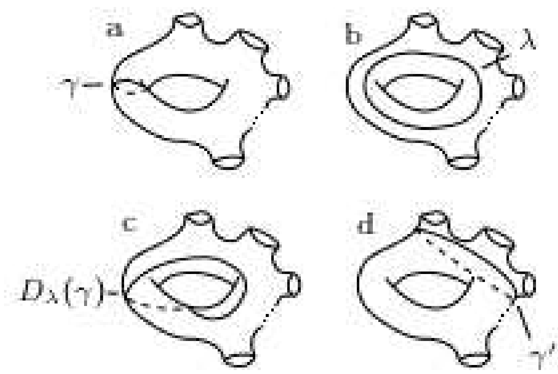


Figure E.2: a. γ b. λ c.$D_\lambda(\gamma)$ d. γ', which is disjoint from both γ and $D_\lambda(\gamma)$

Finally, say β and β' are connected by an edge in $\mathcal{D}$, that is, $\beta \cap \beta' = \emptyset$. If β is nonseparating then $S \backslash \beta = S_{0,b+2}$, so $S\backslash(\beta \cup \beta') = (S\backslash\beta)\backslash\beta'$ is two surfaces S_{0,b_1} and S_{0,b_2} where $b_1 + b_2 = (b+2) + 2 \geq 7$; it follows that at least one of the b_i is at least 4, so at least one of the components has a curve complex with infinite diameter. If β is separating, let S' be the component of $S\backslash\beta$ not containing β'; by assumption S' is not a pair of pants. Then $S'\backslash\beta' = S'$, so it is a component of $S\backslash(\beta \cup \beta')$ whose curve complex has infinite diameter. Thus property 4 is true and we are done. □

Bibliography

1. Farb and Margalit A primer on mapping class groups, `http://www.math.utah.edu/ margalit/primer/`

2. Hatcher and Thuston A presentation for the mapping class group of a closed orientable surface, *Topology* **19** (1980) 221-237

3. Masur and Minsky Geometry of the complex of curves II: Hierarchical structure, *Geometric and Functional Analysis* **10** (2000) 902-974

4. Masur and Schleimer The pants complex has only one end, arXiv:math/0312385v1 [math.GT] (2003)

5. Putman A note on the connectivity of certain complexes associated to surfaces, preprint

6. Schleimer Notes on the complex of curves, `http://www.math.rutgers.edu/ saulsch/math.html`

7. Schleimer The end of the curve complex, arXiv:math/0608505v1 [math.GT] (2006)

Parametric computation of 2-d invariant manifolds

S. Ulery
Bowdoin College

Indiana University REU Summer 2008
Advisor: ADVISOR

F.1 Motivation and Overview

Ordinary differential equations (ODEs) commonly arise in a variety of modeling contexts. Examples include periodically forced oscillators, which can model predator-prey scenarios and sleep cycles, among other systems, and the Lorenz system, which is a simple model of the atmosphere that exhibits interesting mathematical behaviors.

$$\frac{du}{dt} = F(u), \quad \in \mathbb{R}^n$$

In this paper, we consider the case of a steady-state u_* (i.e. $F(u_*) = 0$) where the Jacobian matrix $DF(u_*)$ has two positive and one negative eigenvectors. This gives rise to a two-dimensional unstable invariant manifold which we denote $\mathcal{U}(u_*)$. As a set $\mathcal{U}(u_*)$ consists of all initial conditions u_0 for which the solution to the ODE satisfies $u(t, u_0) \to u_*$ as $t \to -\infty$. It is tangent at u_* to the plane spanned by the eigenvectors associated with the positive eigenvalues. Analogously, the stable manifold $\mathcal{S}(u_*)$ is made up of all u_0 for which $u(t, u_0) \to u_*$ as $t \to +\infty$. As $\mathcal{S}(u_*)$ is essentially $\mathcal{U}(u_*)$ under a reversal of time, any algorithm to compute $\mathcal{U}(u_*)$ can be applied as well to $\mathcal{S}(u_*)$.

Unstable manifolds play an important role in long term dynamics. For dissipative system, i.e. one in which all trajectories eventually enter an *absorbing ball*, unstable manifolds form the backbone of the *global attractor*, the largest compact invariant set. Stable manifolds which have *co-dimension* one (i.e. are associated with $n-1$ negative eigenvalues for in system of n ODEs) form a *separatrix* which divides *phase space* ($\mathbb{R}^n$) into two portions. Trajectories with initial conditions on either side of the separatrix have different fates as $t \to \infty$. When the stable manifold for one fix point intersects with the unstable manifold of another, their intersection typically forms a curve that is an orbit connecting the two states. In general, the onset of such an intersection as a parameter is varied signals a global bifurcation, a dramatic change involving distinct elements of the global attractor. It is therefore useful to visualize these manifolds. Generally, these manifolds cannot be found analytically, so they must instead be "grown" or evolved from a local information.[9]

Previous methods for calculating these manifolds include approximation by geodesic level sets [8, 7], BVP continuation of trajectories [9], computation of fat trajectories [6], PDE formulation [4], and box covering [1, 2]. In this paper, we extend the process of approximation by level sets.

Consider, then, a closed curve of initial conditions parameterized by a variable α:

$$u_0(\alpha) = (x_0(\alpha), y_0(\alpha), z_0(\alpha) \ , \quad \alpha \in [0, \alpha_1] \ , \quad u_0(0) = u_0(\alpha_1) \ .$$

Without loss of generality, we assume that the positive eigenspace is the x, y-plane and the steady state is $u_* = 0$. The initial closed curve is then taken to be a small circle around the tangent point.

The evolution of this curve under the flow of the ODE over any finite time period produces the two-dimensional invariant manifold (with boundary), which we express as

$$u(\alpha, t) = (x(\alpha, t), y(\alpha, t), z(\alpha, t)), \quad t \in [t_1, t_2] \ .$$

The idea is demonstrated in Figure F.1.

If we evolve the points under the flow with no adjustment, however, the curve tends to elongate and not represent the manifold evenly. See Figure F.2. This shows points on trajectories of the ODE

$$\frac{dx}{dt} = 2x, \quad \frac{dy}{dt} = -y, \quad \frac{dz}{dt} = z \tag{F.1}$$

starting from initial conditions along the circle $x^2 + y^2 = 1, z = 0$. It is shown in two dimensions for clarity and because z does not change as the ring evolves.

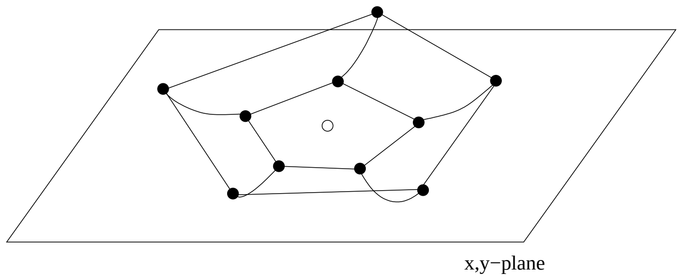


Figure F.1: One time-step from the original ring.

Since we are interested here only in the manifold, and not the dynamical process by which it is generated, we may adjust the flow under which the points are evolved as long as they stay on the manifold. Thus, we may change the component of the flow tangential to the curve as long as we preserve the normal and binormal components, because the entire ring is on the manifold, and the tangential component simply moves the point along the ring. One way to adjust the flow would be to set the tangential component of the flow to zero, thereby preserving locally geodesic flow. This is the approach taken in [5], and is recreated below.

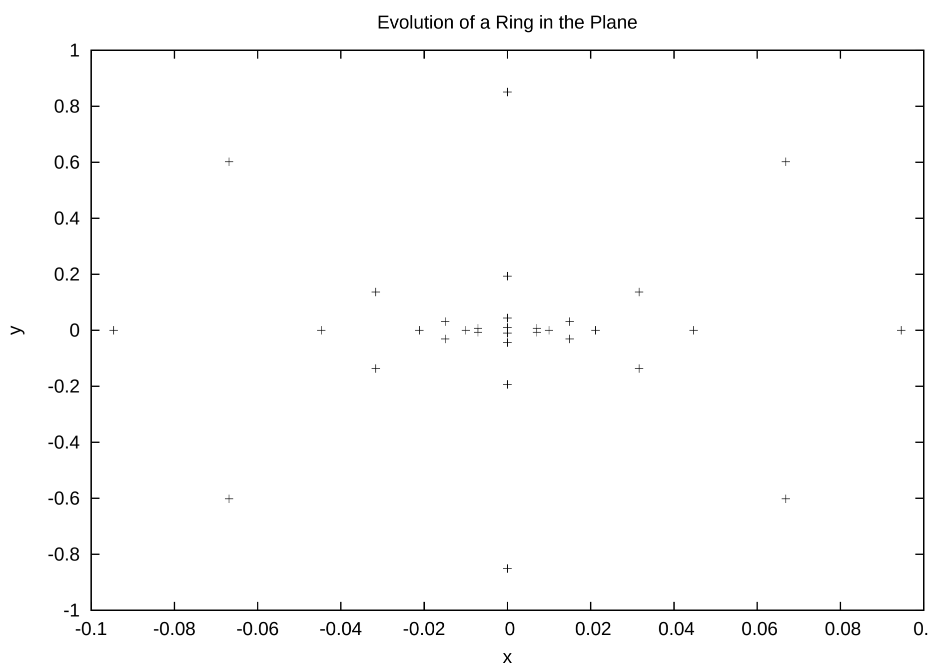


Figure F.2: The graph is distorted when points are unevenly spaced in arclength.

F.2 Zeroing Out the Tangential Component

Given a closed curve

$$\Gamma = (x(\alpha,t), y(\alpha,t), z(\alpha,t)) = u(\alpha,t) \ , \quad \alpha \in [0, \alpha_1] \tag{F.2}$$

we select the right hand coordinate system with unit tangent vector

$$w = \frac{1}{S_\alpha} u_\alpha \ , \quad \text{where} \quad S_\alpha = \sqrt{x_\alpha^2 + y_\alpha^2 + z_\alpha^2} = |u_\alpha| \ , \tag{F.3}$$

unit normal vector

$$n = \frac{1}{\kappa S_\alpha} w_\alpha \ , \quad \text{where} \quad \kappa = \frac{1}{S_\alpha} |w_\alpha| \tag{F.4}$$

and unit binormal vector

$$b = w \times n \ . \tag{F.5}$$

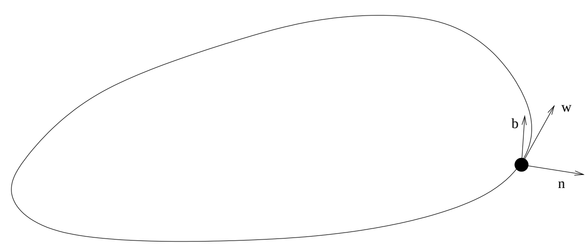


Figure F.3: The component unit vectors at a point.

The motion of the curve under local geodesic flow is given by

$$u_t = 0w + Un + Vb \tag{F.6}$$

where we keep

$$U = F \cdot n \quad \text{and} \quad V = F \cdot b \ . \tag{F.7}$$

We apply this approach to the ODE in F.1 and display the results in Figure F.4. The ring of points is increasingly elongated at each step, although the effect is much less severe than when the flow is unadjusted.

F.3 Adjusting the Tangential Component

Instead of zeroing out the tangential component, we directly calculate T to preserve equal distribution of points in arclength.

Choosing T so that the arclength spacing of a finite number of points on the curve remains constant (in α), is equivalent to satisfying, at each $t \in [t_1, t_2]$, the condition

$$S_\alpha(\alpha, t) = \frac{1}{2\pi} \int_0^{2\pi} S_\alpha(\tilde{\alpha}, t) d\tilde{\alpha} \ , \quad \text{for all} \quad \alpha \in [0, 2\pi] \ . \tag{F.8}$$

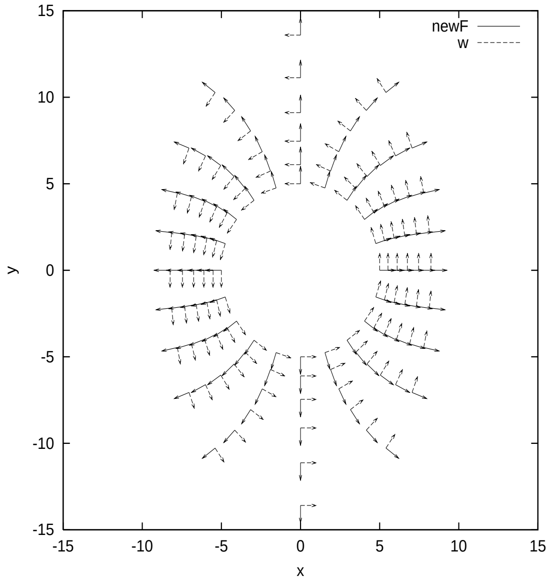


Figure F.4: Zeroing the tangential component does not preserve equal arclength distribution.

Suppose this holds at $t = 0$. We ensure it holds at all other t by choosing T such that

$$S_{\alpha t} = \frac{1}{2\pi}\int_0^{2\pi} S_{\alpha,t} d\tilde{\alpha} \tag{F.9}$$

Differentiating the second relation in (F.3) with respect to t, and then using (F.2) we find that

$$\begin{aligned} S_{\alpha t} &= \frac{1}{S_\alpha}(x_\alpha x_{\alpha t} + y_\alpha y_{\alpha t} + z_\alpha z_{\alpha t}) && \text{(F.10)} \\ &= u_{\alpha t} \cdot w && \text{(F.11)} \end{aligned}$$

We then write

$$u_{\alpha t} = T_\alpha w + T w_\alpha + U_\alpha n + U n_\alpha + V_\alpha b + V b_\alpha \ . \tag{F.12}$$

From (F.4) we have

$$w_\alpha = \kappa S_\alpha n \ . \tag{F.13}$$

It will turn out, due to a projection in the direction of w, that U_α, and V_α do not effect the calculation of T. The remaining quantities, n_α and b_α, are obtained by

Theorem F.3.1. *[10]*

$$n_\alpha = S_\alpha(\tau b - \kappa w) \ , \qquad b_\alpha = -\tau S_\alpha n \tag{F.14}$$

where the torsion τ is defined as the determinant

$$\tau = \kappa^2 \begin{vmatrix} \frac{1}{S_\alpha} x_\alpha & \frac{1}{S_\alpha} y_\alpha & \frac{1}{S_\alpha} z_\alpha \\ \frac{1}{S_\alpha}\partial_\alpha \frac{1}{S_\alpha} x_\alpha & \frac{1}{S_\alpha}\partial_\alpha \frac{1}{S_\alpha} y_\alpha & \frac{1}{S_\alpha}\partial_\alpha \frac{1}{S_\alpha} z_\alpha \\ \frac{1}{S_\alpha}\partial_\alpha \left[\frac{1}{S_\alpha}\partial_\alpha \frac{1}{S_\alpha} x_\alpha\right] & \frac{1}{S_\alpha}\partial_\alpha \left[\frac{1}{S_\alpha}\partial_\alpha \frac{1}{S_\alpha} y_\alpha\right] & \frac{1}{S_\alpha}\partial_\alpha \left[\frac{1}{S_\alpha}\partial_\alpha \frac{1}{S_\alpha} z_\alpha\right] \end{vmatrix} .$$

A proof of Theorem F.3.1 is found in [10].

Using the Frenet-Serret formulae in (F.12), we obtain

$$\begin{aligned} S_{\alpha t} &= w \cdot u_{\alpha t} \\ &= w \cdot [T_\alpha w + T\kappa S_\alpha n + U_\alpha n + U S_\alpha \tau b - U S_\alpha \kappa w + V_\alpha b - V\tau S_\alpha n] \\ &= T_\alpha - U\kappa S_\alpha \end{aligned}$$

We can now express (F.9) as

$$T_\alpha - U\kappa S_\alpha = \frac{1}{2\pi}\int_0^{2\pi} T_\alpha - U\kappa S_\alpha d\tilde{\alpha} \ ,$$

which by periodicity reduces to

$$T_\alpha = U\kappa S_\alpha - \frac{1}{2\pi}\int_0^{2\pi} U\kappa S_\alpha d\tilde{\alpha} \ . \tag{F.15}$$

Integrating both sides of (F.15), we arrive at

$$T(\alpha,t) = T(0,t) + \int_0^{\alpha} U\kappa S_\alpha d\tilde{\alpha} - \frac{\alpha}{2\pi}\int_0^{2\pi} U\kappa S_\alpha d\tilde{\alpha} \ . \tag{F.16}$$

Using (F.13), we can rewrite (F.16) as

$$T(\alpha,t) = T(0,t) + \int_0^{\alpha} F \cdot w_\alpha d\tilde{\alpha} - \frac{\alpha}{2\pi}\int_0^{2\pi} F \cdot w_\alpha d\tilde{\alpha} \ . \tag{F.17}$$

Later we will make a specific choice for the constant of rotation $T(0,t)$, but any choice would still preserve arclength parametrization. Note also that in a practical implementation, one need not even compute the vectors n, and b. Instead one may write (F.6) as

$$u_t = F(u) - [F(u) \cdot w - T]w \ . \tag{F.18}$$

F.4 Fourier Transform

Equation (F.17) is not simple to compute analytically, so we use a discrete Fourier transform (DFT) to compute the derivatives and antiderivatives. We used a pre-written Fast Fourier Transform algorithm. [3]. A DFT associates to a list of function values $u(\alpha_j)$ at equally spaced points a list of Fourier coefficients $\hat{u}_k$ fulfilling the following equation.

$$u(\alpha_j) = \sum_{k=\frac{-N}{2}}^{\frac{N}{2}} \hat{u}_k e^{i\alpha_j k}, \alpha_j = \frac{2\pi j}{N}$$

This is a useful transformation because it is simple to take the derivative and antiderivative of the right hand side.

$$\frac{du(\alpha)}{d\alpha} = \sum_{k=\frac{-N}{2}}^{\frac{N}{2}} ik\hat{u}_k e^{i\alpha k}$$

$$\int u(\alpha)d\alpha = \sum_{k=\frac{-N}{2}}^{\frac{N}{2}} \frac{\hat{u}_k}{ik} e^{i\alpha k}$$

The FFT uses at most N =the number of points on the ring different sine and cosine curves to approximate the curve represented by those points. Since there are a finite number of discrete points rather than a true curve input into the equation, the Fourier representation is necessarily an approximation.

F.5 Minimizing Error

The error involved in this approximation is large enough to be a practical barrier to creating a good graphical representation of the manifold in question, particularly because repeated applications of the FFT exaggerate the error. There are two types of error involved. The first error occurs when the FFT tries to be too smart, and in so doing gives too much weight to high frequency sines and cosines when approximating the curve, resulting in a "jumpy" representation. This is particularly problematic when taking derivatives. See Figure F.5 for an example.

To reduce the error, a process of filtering the data is introduced. To filter the Fourier representations of data points, we adjust

$$\hat{u}_k = \hat{u}_k e^{-10\frac{k}{N}^{10}}$$

The higher k is, the higher frequency $\hat{u}_k$ is contributing. Thus, our adjustment dampens the effect more as k increases. We don't throw away any data, but we limit the effect that disruptive, high frequency data influences our final result. It is also possible to filter data by setting to zero any component curve of frequency above some point, but for this application this form of filtering was not found to be useful.

Another problem occurs because creating an exact representation of the ring in Fourier space may require summing an infinite series:

$$u(\alpha_j) = \sum_{k=-\infty}^{\infty} \hat{u}_k e^{i\alpha_j k}, \alpha_j = \frac{2\pi j}{N}$$

Since we have only a finite number of points, this is impossible. This means that the components outside the limits of the sum get incorrectly moved into that range. This was not a large source of error for this application. We addressed it simply by taking a relatively large number of points on the ring (e.g. 256 or 512) to minimize the error from this source.

F.6 Computational Results

To test the parametric approach in various geometric situations we consider a couple examples of vector fields in $\mathbb{R}^3$ specified by

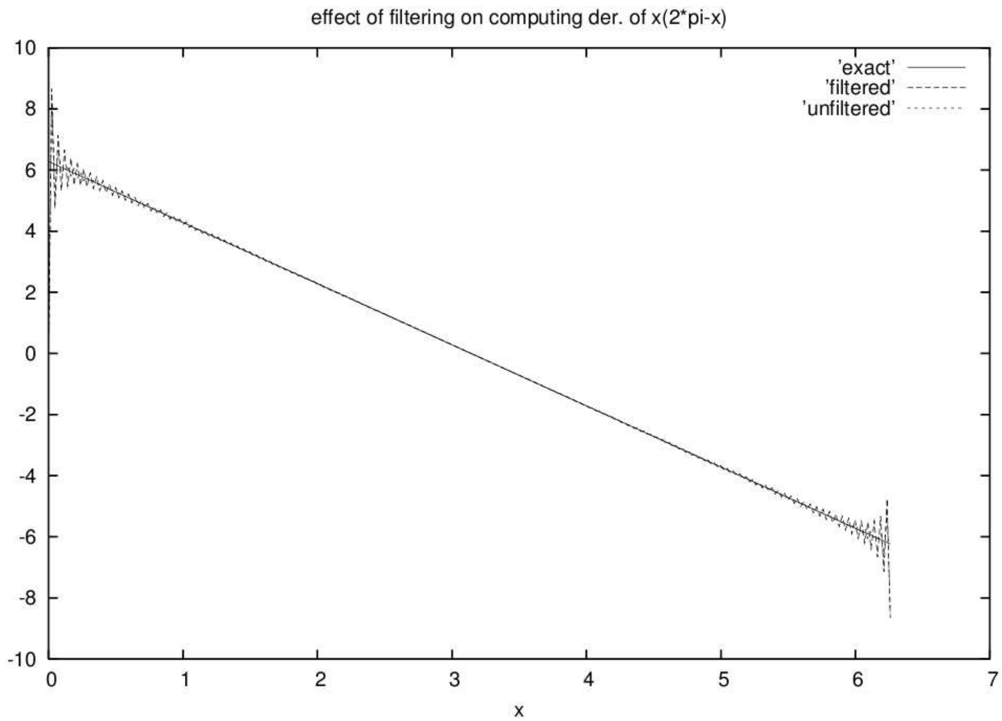


Figure F.5: The derivative of $x(2 * pi - x)$ computed using the FFT with and without filtering.

- (i) the (decoupled) flow in two-variables

$$\begin{aligned} \dot{x} &= f_1(x, y) \\ \dot{y} &= f_2(x, y) \ , \end{aligned} \tag{F.19}$$

- (ii) an (attracting) invariant manifold expressed as the graph of a function $z = \Phi(x, y)$, so that

$$\dot{z} = \nabla\Phi \cdot \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} - c(z - \Phi) \ , \tag{F.20}$$

for some positive number c.

F.6.1 The Dali

The first choice for the invariant manifold resembles a warped clock face from the works of the artist Dali. It is described in rectangular coordinates as the graph of

$$\Phi(x, y) = .1x^3 \ .$$

The vector field is then determined by

$$\begin{aligned} \dot{x} &= x \\ \dot{y} &= 2y \ , \end{aligned} \tag{F.21}$$

so that the “Dali” is the unstable manifold of the origin, which is associated with real and distinct eigenvalues.

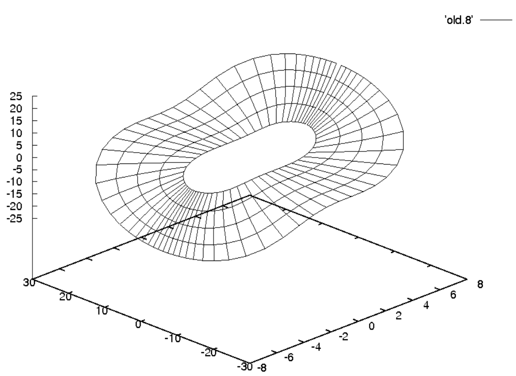


Figure F.6: We normalize the flow, but don’t adjust the tangential component.

To demonstrate the need to do something more than simply integrating a ring of initial conditions under the flow of the original vectorfield we plot in Figure F.6 the result of normalizing but not adjusting the direction of the flow. This is not a complicated manifold, so straightforward integration of a ring of initial data generates a reasonable representation of the manifold, but the data points are far from evenly spaced. Following the geodesic flow (see Figure F.7) is only a slight improvement. Following the parametric flow provides much a much better representation (see Figure F.8).

Figure F.9 demonstrates the differences in point spacing on the outermost ring between the three methods. The points are spaced much more evenly around the ring under parametric flow than either of the other two methods. Geodesic flow is only a slight improvement over no adjustment. On a more complicated manifold, these differences would cause a more severe problem in representation.

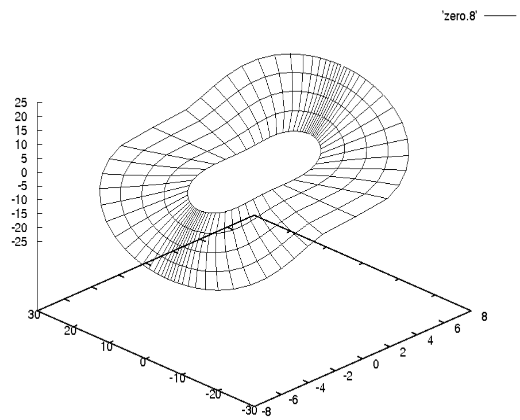


Figure F.7: We normalize the flow and set the tangential component to zero.

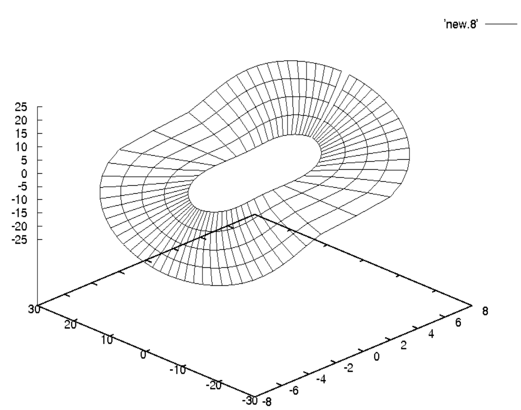


Figure F.8: We normalize the flow and adjust the tangential component for equal spacing.

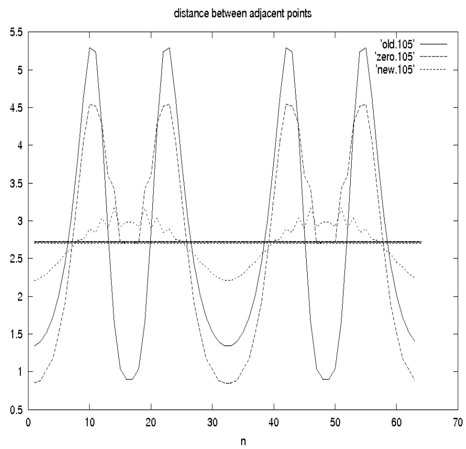


Figure F.9: The distance between adjacent points on the ring.

F.6.2 The Dial

The next choice for the invariant manifold resembles a sundial. It is a manifold with

$$\Phi(r,\theta) = \begin{cases} r^2(\cos(20\theta)+1) & \text{if } |\theta| \le \frac{\pi}{20} \\ 0 & \text{else} \end{cases}$$

$$\dot{z} = \nabla_{r,\theta}\Phi(r,\theta) \cdot \begin{bmatrix} \dot{r} \\ \dot{\theta} \end{bmatrix} - (z - \Phi(r,\theta))$$

We plot in Figure F.10 the result of using the original flow, neither normalizing nor adjusting the direction of the flow. This is a more complicated manifold, and in particular it has a "spike," so it is the type of manifold for which we expected parametric flow to be a considerable improvement over the original (see Figure F.11) or geodesic (see Figure F.12) flow. As it turns out, however, original flow appears to give the best representation. It is unclear whether this is the result of human error in carrying out the procedure described above, or whether this is a true artifact of using parametric flow on this manifold.

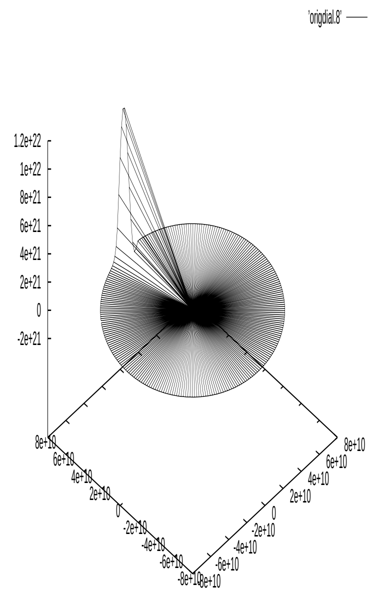


Figure F.10: We neither normalize the flow nor adjust the tangential component.

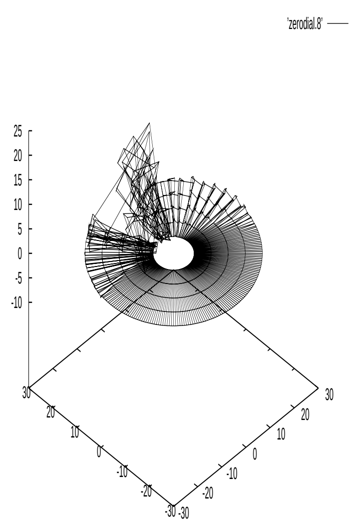


Figure F.11: We normalize the flow and set the tangential component to zero.

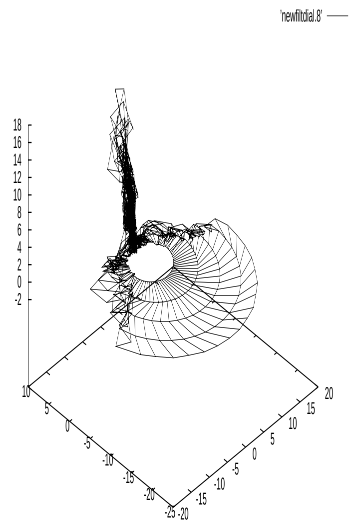


Figure F.12: We normalize the flow and adjust the tangential component for equal spacing.

F.7 Future Research

This project suggests that research could be undertaken to explore the possibility of adding points as the ring expands in order to prevent the points from becoming too far separated, even as the manifold grows. This would be another adaptation to prevent distortion of the general model of a ring evolving on a manifold. There could also be interest in applying these techniques to a 2-D manifold in 4-space, or generalizing to even higher dimensions. A systematic approach to deciding which of the many methods for growing manifolds found in the literature is best for specific types of problems would also be useful.

Bibliography

1. Dellnitz, M. and Hohmann, A. [1996] "The computation of unstable manifolds using subdivision and continuation," *Nonlinear Dynamical systems and Chaos PNLDE* 19, eds. Broer, H. W., Van Gils, S. A., Hoveign, I. and Takens, F. (Birkhauser, Basel), pp. 449-459.

2. Dellnitz, M. and Hohmann, A. [1997] "A subdivision algorithm for the computation of unstable manifolds and global attractors," *Numer. Math.* **75**, 293-317.

3. FFTW [Computer software]. [n.d.]. Retrieved from http://www.fftw.org

4. Guckenheimer, J. and Vladimirsky, A. [2004] "A fast method for approximating invariant manifolds." *SIAM J. Appl. Dyn. Syst.* **3**, 232-260.

5. Guckenheimer, J. and Worfolk, P. in *Bifurcations and periodic orbits of vector fields*, Kluwer Acad. Publ., 1992

6. Henderson, M. E. [2003] "Computing invariant manifolds by integrating fat trajectories," *SIAM J. Appl. Dyn. Syst.* **4**, 832-882.

7. Krauskopf, B. and Osinga, H. M. [2003] "Computing geodesic level sets on global (un)stable manifolds of vector fields," *SIAM J. Appl. Dyn. Syst.* **4**, 546-569.

8. Krauskopf, B. and Osinga, H. M. [1999] "Two dimensional global manifolds of vector fields," *Chaos* **9**, 768-774.

9. Krauskopf, B., Osinga, H. M., Doedel, E. J., Henderson, M. E., Guckenheimer, J., Vladimirsky, A., et al. [2005] "A survey of methods for computing (un)stable manifolds of vector fields," *International Journal of Bifurcation and Chaos* **15**, 763-791.

10. Oprea, J. [2007] *Differential geometry and its applications.* Washington, DC: Mathematical Association of America.

Ends and ideal spaces of negatively curved surfaces

John Ullman
Indiana University

Indiana University REU Summer 2008
Advisor: ADVISOR

G.1 Some Technical Lemmas

We begin with some general constructions which are used repeatedly in the next section.

Lemma G.1.1. *Given functions $r_1(z)$ for $z \leq a$ and $r_2(z)$ for $z \geq b > a$ which are C^2 and have positive second derivative, and such that $r_2'(b) > r_1'(a)$ and $r_1(a) + r_1'(a)(b-a) < r_2(b) < r_1(a) + r_2'(b)(b-a)$, there is a C^2 function $r(z)$ defined in [a,b] with positive second derivative such that $r(z) = r_1(z)$ for $z \leq a$ and $r(z) = r_2(z)$ for $z \geq b$.*

Proof. We will show there is a positive, piecewise linear function $g(z)$ defined on $[a, b]$ such that

1. $g(a) = r_1''(a)$, $g(b) = r_2''(b)$
2. $\int_a^b g(z)dz = r_2'(b) - r_1'(a)$
3. $\int_a^b \int_a^z g(s)dsdz = r_2(b) - r_1(a) - r_1'(a)(b-a)$.

Then we will define $r(z)$ by

$$r(z) = r_1(a) + \int_a^z \left(r_1'(a) + \int_a^s g(\tau)d\tau \right) ds$$

for $z \in [a, b]$. Then since $r'(z) = r_1'(a) + \int_a^z g(s)ds$ and $r''(z) = g(z)$ we will have matched function values and derivatives up to order two at a and b with $r''(z) > 0$. The function $g(z)$ will be of the form shown in Figure G.1 below. We will construct a family $\{g_\epsilon(z)\}$ for $\epsilon \in (0, 1)$ of such functions with properties 1 and

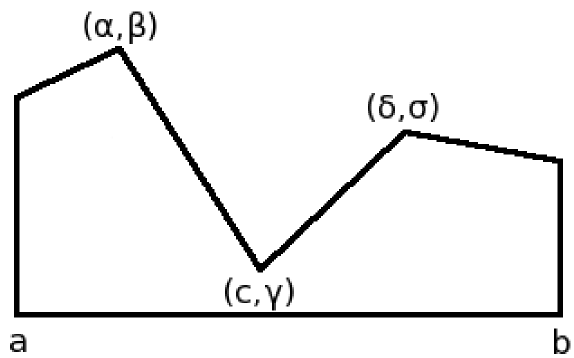


Figure G.1: The form of g(z)

2 above, then choose ϵ appropriately. First we let

$$c = a + (1-\epsilon)(b-a)$$

$$\gamma = \min\left\{ \frac{\epsilon(1-\epsilon)}{b-a}(r_2'(b) - r_1'(a)), \frac{1}{2}r_1''(a), \frac{1}{2}r_2''(b) \right\}$$

Then we find that

$$\int_a^c g_\epsilon(z)dz = \frac{1}{2}[(r_1''(a) - \gamma)\alpha + (c-a)\beta + (c\gamma - ar_1''(a))]$$

Since $\gamma < r_1''(a)$, we find

$$\int_a^c g_\epsilon(z)dz > \frac{1}{2}[(r_1''(a)-\gamma)a + (c-a)0 + (c\gamma - ar_1''(a))] \\ = \frac{1}{2}\gamma(c-a)$$

and $\frac{1}{2}\gamma(c-a) \le \frac{1}{2}(1-\epsilon)(b-a)\frac{\epsilon(1-\epsilon)}{b-a}(r_2'(b)-r_1'(a)) < \frac{1}{2}\epsilon(r_2'(b)-r_1'(a))$. Hence $\int_a^c g_\epsilon(z)dz < \epsilon(r_2'(b)-r_1'(a))$ for sufficiently small α and β. Now

$$\frac{1}{2}[(r_1''(a)-\gamma)\alpha + (c-a)\beta + (c\gamma - ar_1''(a))] = \epsilon(r_2'(b)-r_1'(a))$$

defines a line in (α,β) with $\frac{\Delta\beta}{\Delta\alpha} = -\frac{r_1''(a)-\gamma}{c-a} < 0$, so the α-intercept $\frac{2\epsilon(r_2'(b)-r_1'(a))-(c\gamma-ar_1''(a))}{r_1''(a)-\gamma}$ is bigger than a, since $(a,0)$ lies below the line. Hence we choose

$$\alpha = \min\left\{\frac{1}{2}\left(\frac{2\epsilon(r_2'(b)-r_1'(a))-(c\gamma-ar_1''(a))}{r_1''(a)-\gamma}+a\right), \frac{1}{2}(a+c)\right\}$$

Then, finally, we choose

$$\beta = \frac{2\epsilon(r_2'(b)-r_1'(a)) - (r_1''(a)-\gamma)\alpha - (c\gamma - ar_1''(a))}{c-a}$$

so that $\int_a^c g_\epsilon(z)dz = \epsilon(r_2'(b)-r_1'(a))$. Similar analysis leads to

$$\delta = \max\left\{\frac{1}{2}\left(\frac{2(1-\epsilon)(r_2'(b)-r_1'(a))-(br_2''(b)-c\gamma)}{\gamma-r_2''(b)}+b\right), \frac{1}{2}(b+c)\right\}$$

$$\sigma = \frac{2(1-\epsilon)(r_2'(b)-r_1'(a)) - (\gamma - r_2''(b))\delta - (br_2''(b)-c\gamma)}{b-c}$$

so that $\int_c^b g_\epsilon(z)dz = (1-\epsilon)(r_2'(b)-r_1'(a))$ and thus $\int_a^b g_\epsilon(z)dz = r_2'(b)-r_1'(a)$. Next note that $\int_a^z g_\epsilon(s)ds$ is strictly increasing, so that

$$(c-a)0 + (b-c)\epsilon(r_2'(b)-r_1'(a)) < \int_a^b\int_a^z g_\epsilon(s)dsdz < (c-a)\epsilon(r_2'(b)-r_1'(a)) + (b-c)(r_2'(b)-r_1'(a))$$

$$\epsilon^2(b-a)(r_2'(b)-r_1'(a)) < \int_a^b\int_a^z g_\epsilon(s)dsdz < (2\epsilon-\epsilon^2)(b-a)(r_2'(b)-r_1'(a))$$

From this it follows that

$$\lim_{\epsilon\to 0^+}\int_a^b\int_a^z g_\epsilon(s)dsdz = 0$$

$$\lim_{\epsilon\to 1^-}\int_a^b\int_a^z g_\epsilon(s)dsdz = (b-a)(r_2'(b)-r_1'(a))$$

We are given that $r_1(a)+r_1'(a)(b-a) < r_2(b) < r_1(a)+r_2'(b)(b-a)$ and hence $0 < r_2(b)-r_1(a)-r_1'(a)(b-a) < (b-a)(r_2'(b)-r_1'(a))$. Also, $\int_a^b\int_a^z g(s)dsdz$ is a continuous function of $\alpha,\beta,c,\gamma,\delta$ and σ, while these depend continuously on ϵ, so $\int_a^b\int_a^z g_\epsilon(s)dsdz$ is a continuous function of ϵ. Therefore by the intermediate value theorem there is some ϵ for which $\int_a^b\int_a^z g_\epsilon(s)dsdz = r_2(b)-r_1(a)-r_1'(a)(b-a)$. We choose $g = g_\epsilon$. $\square$

When we use this construction we will speak of "ramping" $r(z)$ from $r_1(z)$ to $r_2(z)$. Next we examine surfaces with circular cross sections of the form

$$(x - f(z))^2 + y^2 = r(z)^2$$

where $r(z) > 0$ and $r(z), f(z) \in C^2$, yielding a C^2 surface.

Lemma G.1.2. *A surface of the above form is negatively curved provided $r''(z) > |f''(z)|$.*

Proof. We parametrize the surface with

$$(x, y, z) = (r(z)\cos\theta + f(z), r(z)\sin\theta, z)$$

It follows that

$$K = -\frac{r'' + f''\cos\theta}{r(1 + r'^2 + \cos^2\theta f'^2 + 2r'f'\cos\theta)^2}$$

and therefore $K < 0$ identically if for all θ and z

$$r''(z) + f''(z)\cos\theta > 0$$

Now, for a given z, $f''(z)\cos\theta$ oscillates between $|f''(z)|$ and $-|f''(z)|$, so $K < 0$ identically if $r''(z) - |f''(z)| > 0$, i.e. if $r''(z) > |f''(z)|$. □

We will refer to the change $x \to x - f(z)$ as "skewing." Next we obtain a procedure for joining two hyperboloids. Let $a > 1$, and place the axes of the hyperboloids at distance $2a$ from one another. Then cut out the parts of each hyperboloid on the opposite side of the intersection, to obtain a surface of the form

$$(|x| - a)^2 + y^2 = 1 + z^2$$

We then smooth the intersection with a technique given in [1] (referred to here as the BVK construction). When $|x| < a - 1$ we have $(|x| - a)^2 - 1 > 0$ and so $z = \pm\sqrt{(|x| - a)^2 - 1 + y^2}$. This is smooth except at $x = 0$, so we seek to replace $(|x| - a)^2 - 1$ with $\phi(x)^2$ where $\phi(x) = c_0 + c_2x^2 + c_4x^4 > 0$ when $|x| \le b$ and b is some positive number less than $a - a^{1/3}$. We require

$$c_0 + c_2b^2 + c_4b^4 = \phi(b) = \sqrt{(a - b)^2 - 1}$$

$$2c_2b + 4c_4b^3 = \phi'(b) = -\frac{a - b}{\sqrt{(a - b)^2 - 1}}$$

$$2c_2 + 12c_4b^2 = \phi''(b) = -\frac{1}{((a - b)^2 - 1)^{3/2}}$$

to match function values and derivatives up to second order. These equations have the unique solution

$$c_0 = \phi(b) - \frac{5}{8}b\phi'(b) + \frac{1}{8}b^2\phi''(b)$$

$$c_2 = \frac{3}{4b}\phi'(b) - \frac{1}{4}\phi''(b) = \frac{1}{4b\phi(b)^3}(-3((a - b)^3 - (a - b)) + b)$$

$$c_4 = -\frac{1}{8b^3}\phi'(b) + \frac{1}{8b^2}\phi''(b)$$

Since $b < a - a^{1/3}$, we have $(a-b)^3 > a$, whence $(a-b)^3 - (a-b) > b$ and

$$c_2 < \frac{-3b + b}{4b\phi(b)^3} < 0$$

Then since $\phi''(x) = 2c_2 + 12c_4x^2$ and $\phi''(b)$ is also negative, it follows that $\phi''(x) < 0$ for all $x \in [-b, b]$. Also, $\phi(-b) = \phi(b)$ so we have $\phi(x) \geq \phi(b) > 0$ for all $x \in [-b, b]$. Finally, we compute

$$K = \frac{\phi^3\phi''}{(\phi^2 + \phi^2\phi'^2 + 2y^2)^2} < 0$$

Thus we have smoothed out the intersection while retaining negative curvature. Note that since affine transformations don't change the sign of the curvature, we may join two hyperboloids of the form $x^2 + y^2 = 1 + (cz)^2$ by using the above construction and then transforming the z-coordinate. Next we obtain our main theorems. Hyperboloids will be our building blocks in what follows. (N.B. by "complete" we mean complete in the extrinsic sense)

G.2 Main Theorems

Theorem G.2.1. *Let g, n_{circ}, and n_{cusp} be natural numbers such that $n_{circ} \geq 1$ if $g = 0$, and $n_{circ} \geq 2$ if $g > 0$. Then there exists a complete, negatively curved, C^2 surface embedded in $\mathbb{R}^3$ with genus g, n_{circ} circular ends and n_{cusp} cuspidal ends.*

Proof. We begin with the genus zero case. If $n_{circ} = 1$ and $n_{cusp} = 0$, we have the simple example $z = xy$, so we may assume that $n_{circ} + n_{cusp} \geq 2$. We will construct these surfaces by lining up hyperboloids and joining the bottoms using the BVK construction, then replacing the tops with narrower hyperboloids or cusps using Lemma G.1.1, and finally skewing the tops using Lemma G.1.2 so they don't hit one another. For definiteness we will use the minimum value of ϵ when applying Lemma G.1.1 (N.B. the set of possible ϵ's is compact) and $b = \frac{1}{2^k}(a - a^{1/3})$ when applying the BVK construction, where k is the smallest positive integer making b small enough for our needs. We begin with $n_{circ} + n_{cusp} - 1$ hyperboloids $\{H_j\}_{j=0}^{n_{circ}+n_{cusp}-2}$ with axes given by $x = x_j, y = 0$, where $x_0 = 0$ and the x_j's are increasing (the values to be chosen later). The top of H_j for $j = 0, \ldots, n_{circ} - 2$ will have a circular end, $j = n_{circ} - 1, \ldots, n_{circ} + n_{cusp} - 2$ a cuspidal end. For circular ends we ramp from $r_j(z) = \sqrt{1 + z^2}$ to

$$r_j(z) = \sqrt{1 + (\frac{1}{4^{j+1}}z)^2}$$

between $z = 0$ and $z = \frac{1}{6}$ using Lemma G.1.1. For cuspidal ends, we begin with $r_j(z) = \sqrt{1 + z^2}$. See the figure below. Since $r_j''(z) > 0$ we have

$$r_j(-\frac{1}{6}) + r_j'(-\frac{1}{6})(0 - (-\frac{1}{6})) < r_j(0) = 1 < r_j(-\frac{1}{6})$$

Also, $(0, 1)$ lies on the line $r_j(-\frac{1}{6}) + \frac{1 - r_j(-\frac{1}{6})}{0 - (-\frac{1}{6})}(z - (-\frac{1}{6}))$ so that

$$r_j(0) = 1 < r_j(-\frac{1}{6}) + \frac{1}{2}\frac{1 - r_j(-\frac{1}{6})}{0 - (-\frac{1}{6})}(0 - (-\frac{1}{6}))$$

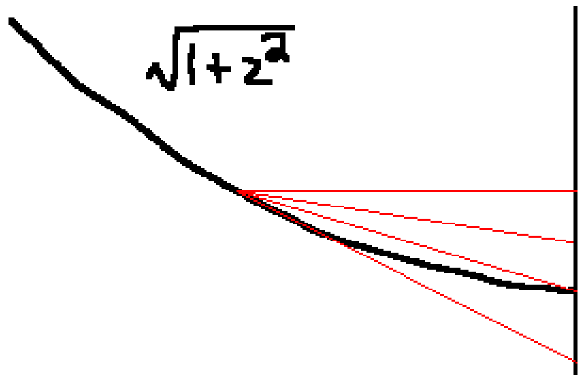


Figure G.2: Constructing the slope to ramp to

Thus we may apply Lemma G.1.1 to ramp $r_j(z)$ from $\sqrt{1+z^2}$ to $e^{3(1-r_j(-\frac{1}{6}))z}$ between $z=-\frac{1}{6}$ and $z=0$. Since $r_j'(z)<0$ at both points and $r_j''(z)>0$, it follows that $r_j(z)>0$ during the transition. Next we skew the surfaces H_j using functions $f_j(z)$ (in the notation of Lemma G.1.2). For each j and all $z<-\frac{1}{6}$ we require

$$|f_j''(z)| < r_j''(z) = \frac{1}{(1+z^2)^{3/2}}$$

First we define a continuous function $h(z)$ as follows: $h(z)=\frac{1}{(1+z^2)^{3/2}}$ between $z=-\frac{12}{5}$ and $z=-\frac{5}{12}$, $h(z)$ is affine for $z\in[a,-\frac{12}{5}]\cup[-\frac{5}{12},-\frac{1}{3}]$ and 0 elsewhere (there is a unique continuous function with these characteristics). The number a is chosen close enough to $-\frac{12}{5}$ so that $h(z)<\frac{1}{(1+z^2)^{3/2}}$ when $z<-\frac{12}{5}$ ($a=-2.5$ suffices). The form of $h(z)$ is shown in the figure below. The red curve is $\frac{1}{(1+z^2)^{3/2}}$, the blue is $h(z)$ (the region below the graph of $h(z)$ is shaded blue). Then

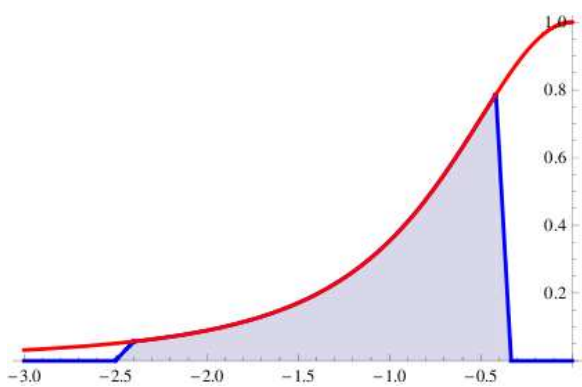


Figure G.3: The form of h(z)

$$\int_{-\infty}^{\infty} h(z)dz > \int_{-\frac{12}{5}}^{-\frac{5}{12}} \frac{1}{(1+z^2)^{3/2}}dz = \frac{7}{13} > \frac{1}{2}$$

Finally, we define the skewing functions:

$$f_j(z) = x_j + \frac{\frac{1}{2} - \frac{2}{4^{j+1}}}{\int_{-\infty}^{\infty} h(s)ds} \left(\int_{-\infty}^{z} \int_{-\infty}^{s} h(\tau) d\tau ds \right)$$

By the above we have

$$f_j''(z) = \frac{\frac{1}{2} - \frac{2}{4^{j+1}}}{\int_{-\infty}^{\infty} h(s)ds} h(z) \leq \frac{1/2}{7/13} h(z) < \frac{1}{(1+z^2)^{3/2}}$$

Then since $f_j''(z) = 0$ when $z > -\frac{1}{3}$, it follows that $|f_j''(z)| < r_j''(z)$ for all z, no matter what modifications were made to $r_j(z)$. Now we have negatively curved surfaces H_j described by

$$(x - f_j(z))^2 + y^2 = r_j(z)^2$$

Next we wish to choose the spacings $x_j - x_{j-1}$ large enough so that the H_j only intersect when $z < a$. To show this is possible, denote the least values of x of the circular cross sections of H_j by $left_j(z)$, the greatest by $right_j(z)$. That is to say, $left_j(z) = f_j(z) - r_j(z)$ and $right_j(z) = f_j(z) + r_j(z)$. When $z > \frac{1}{6}$ we compute $f_j'(z) = \frac{1}{2} - \frac{2}{4^{j+1}}$ and $r_j'(z) < \frac{1}{4^{j+1}}$ and so

$$\begin{aligned}(left_{j+1} - right_j)'(z) &= f_{j+1}'(z) - r_{j+1}'(z) - f_j'(z) - r_j'(z) \\ &> \frac{2}{4^{j+1}} - \frac{2}{4^{j+2}} - \frac{1}{4^{j+1}} - \frac{1}{4^{j+2}} = \frac{1}{4^{j+2}} > 0\end{aligned}$$

It follows that if $x_{j+1} - x_j$ is sufficiently large, the circular cross sections of H_{j+1} and H_j will not intersect when $z > \frac{1}{6}$. Hence, once x_j has been chosen, x_{j+1} may be chosen large enough so that they only intersect when $z < a - 1$. We make these choices, then finally join the H_j using the BVK construction, where the parameters b are chosen small enough that the surgeries only affect parts of the surfaces where $z < a$. This is now easily extended to the positive genus case. Since $n_{circle} \geq 2$, H_0 is a hyperboloid on bottom and a narrower hyperboloid on top. We simply line up g additional copies of H_0 with axes at $y = 0, x < 0$, space them far enough apart that they only intersect when $|z| > 1$, and join their tops and bottoms using separate applications of the BVK construction, making sure that the surgeries only affect parts of the surfaces where $|z| > \frac{1}{6}$. □

A typical such surface is illustrated below. The yellow indicates regions where smoothing has been performed, while grey dashes are on the back side. Note that we can continue adding to these surfaces in the same way to obtain surfaces with infinite g, n_{circ}, and/or n_{cusp}, although if we do so we will have a wild end as well.

Theorem G.2.2. *For any $g > 0$ there is a complete, negatively curved, C^2 surface embedded in $\mathbb{R}^3$ with genus g and one circular end.*

Proof. We begin with a hyperboloid and rip open one side of it. When $x \leq -\frac{1}{2}$ we have $x = -\sqrt{1 + z^2 - y^2}$ and $|y| \leq \sqrt{z^2 + \frac{3}{4}}$. We keep this part of the surface and replace the rest with the graph

$$x = y^4 + (-2z^2 - \frac{1}{2})y^2 + (z^4 + \frac{1}{2}z^2 - \frac{11}{16})$$

over the remainder of the yz plane. It is easily checked that function values and derivatives up to second order are matched, so we obtain a C^2 surface. The result is shown below. When $|y| > \sqrt{z^2 + \frac{3}{4}}$ we have

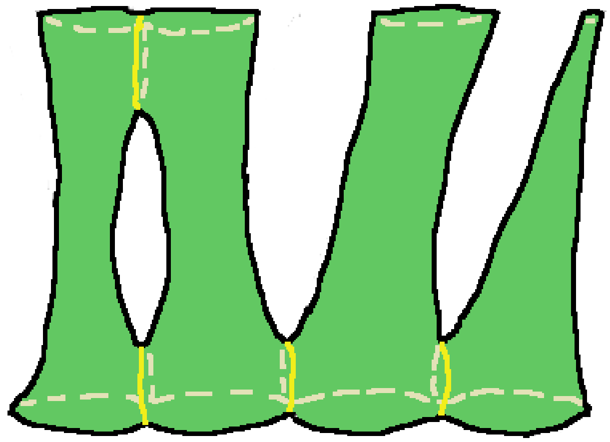


Figure G.4: A surface with genus 1, 3 circular ends and 1 cuspidal end

$$sgn(K) = sgn(x_{yy}x_{zz} - x_{yz}^2)$$

$$x_{yy}x_{zz} - x_{yz}^2 = -48y^4 + (96z^2 + 16)y^2 + (-48z^4 - 16z^2 - 1)$$

For fixed z, we find the roots of this polynomial in y:

$$y^2 = \frac{-96z^2 - 16 \pm 8}{-96} = z^2 + \frac{1}{4}, z^2 + \frac{1}{12}$$

$$y = \pm\sqrt{z^2 + \frac{1}{4}}, \pm\sqrt{z^2 + \frac{1}{12}}$$

All four roots are inside $\left[-\sqrt{z^2 + \frac{3}{4}}, \sqrt{z^2 + \frac{3}{4}}\right]$, so $x_{yy}x_{zz} - x_{yz}^2 \neq 0$ when $|y| > \sqrt{z^2 + \frac{3}{4}}$. Also, it is dominated by $-48y^4$, so in fact $x_{yy}x_{zz} - x_{yz}^2 < 0$ and thus $K < 0$. We now add genus by adding hyperboloids with axes $x = -3k, y = 0$ for $k = 1, \ldots, g$ and joining them using the BVK construction. □

A typical result of this construction is illustrated below.

Theorem G.2.3. *For any $g > 0$ and $n_{cusp} > 0$ there is a complete, negatively curved, C^2 surface embedded in $\mathbb{R}^3$ with genus g, one circular end and n_{cusp} cuspidal ends.*

Proof. We will skew the surfaces from the previous theorem, add cusps and skew them in the opposite direction. First we must gain some freedom to skew the ripped hyperboloid. Denote this surface by $x = x_0(y, z)$. We will try replacing this with

$$x = x_0(y, z) + f(z)$$

where $f(z) \in C^2$. Now we find that

$$x_{yy}x_{zz} - x_{yz}^2 = x_{0_{yy}}x_{0_{zz}} - x_{0_{yz}}^2 + f''(z)x_{0_{yy}}$$

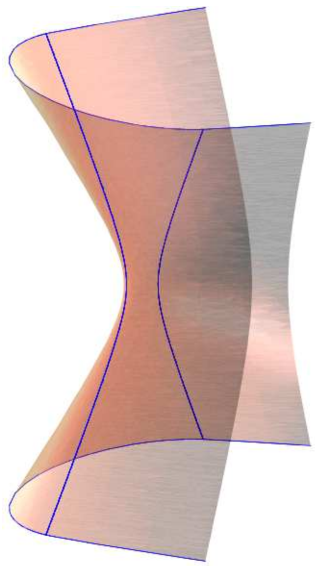


Figure G.5: A ripped-open hyperboloid with negative curvature

Also, it is easily checked that $x_{0_{yy}} > 0$, so $K < 0$ as long as

$$f''(z) < -\frac{x_{0_{yy}} x_{0_{zz}} - x_{0_{yz}}^2}{x_{0_{yy}}}$$

This must hold for all y, so we need to compute the minimum of the quantity on the right, which we denote by $m(y,z)$, for some values of z. By the above, when $|y| > \sqrt{z^2 + \frac{3}{4}}$ we have

$$m(y,z) = \frac{48y^4 - (96z^2 + 16)y^2 + (48z^4 + 16z^2 + 1)}{12y^2 - 4z^2 - 1}$$

$$m_y = y\frac{1156y^4 + (-776z^2 - 192)y^2 + (-380z^4 - 64z^2 + 8)}{(12y^2 - 4z^2 - 1)^2}$$

We view the numerator above as a quadratic in y^2; the discriminant is

$$2359296(z^4 + \frac{145}{576}z^2 - \frac{1}{18432})$$

This in turn has two imaginary roots and two real roots $z = \pm\epsilon$, where

$$\epsilon = \sqrt{\frac{1}{2}\left(-\frac{145}{576} + \sqrt{\left(\frac{145}{576}\right)^2 + 4\left(\frac{1}{18432}\right)}\right)} \approx .014674$$

Thus the discriminant is negative when $z \in (-\epsilon, \epsilon)$, so the numerator has no real roots. It is dominated by the positive term $1156y^4$, so it is positive. Then, due to the factor of y, $m_y < 0$ when $y < 0$ and $m_y > 0$

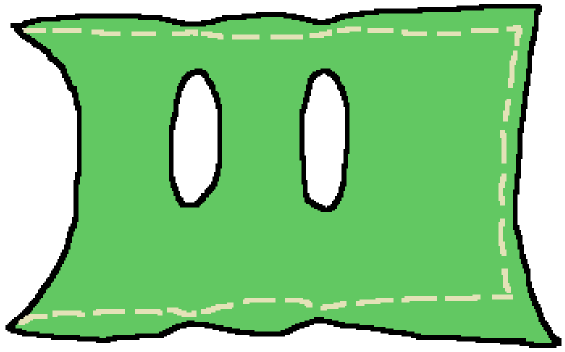


Figure G.6: A surface with genus 2 and 1 circular end

when $y > 0$. When instead $|y| \leq \sqrt{z^2 + \frac{3}{4}}$, we have

$$m(y, z) = \frac{1}{(1 + z^2)\sqrt{1 + z^2 - y^2}}$$

$$m_y = \frac{y}{(1 + z^2)(1 + z^2 - y^2)^{3/2}}$$

so for all y we have $m_y < 0$ when $y < 0$ and $m_y > 0$ when $y > 0$. Thus, if $z \in (-\epsilon, \epsilon)$, the minimum of m occurs at $y = 0$:

$$\min_{y \in \mathbb{R}} m(y, z) = m(0, z) = \frac{1}{(1 + z^2)^{3/2}}$$

Therefore, we have $K < 0$ if $0 \leq f''(z) < \frac{1}{(1+z^2)^{3/2}}$ when $z \in (-\epsilon, \epsilon)$ and $f''(z) = 0$ elsewhere. Since $\frac{d^2}{dz^2}\sqrt{1 + z^2} = \frac{1}{(1+z^2)^{3/2}}$, by Lemma G.1.2 we may skew hyperboloids by the same function and maintain negative curvature. We define a function $h(z)$ as follows: $h(z) = \frac{1}{2\epsilon}z + \frac{1}{2}$ when $z \in [-\epsilon, 0]$, $h(z) = -\frac{1}{2\epsilon}z + \frac{1}{2}$ when $z \in [0, \epsilon]$ and $h(z) = 0$ elsewhere. Finally, we define the skewing function:

$$f(z) = \int_{-\infty}^{z} \int_{-\infty}^{s} h(\tau) d\tau ds$$

Then, $f''(z) = h(z)$, so we maintain negative curvature. Next, we again add g hyperboloids with axes $x = -3k, y = 0$ for $k = 1, \ldots, g$ and skew them using the same function. We join the bottoms using the BVK construction, choosing the parameter b small enough so that the surgeries only affect parts of the surfaces where $z < -\epsilon$. Now, when $z > \epsilon$, $f'(z) = \int_{-\infty}^{z} h(s)ds = \frac{1}{2}\epsilon$, so the tops are simply hyperboloids which have been affinely skewed to one side. Thus we may join them, using an affinely transformed version of the BVK construction. Next, we add cusps; we must show that we can skew them quickly enough to escape the skewed hyperboloids. First, we note that

$$\int_{-\infty}^{0} \frac{1}{(1 + z^2)^{3/2}} dz = 1$$

so we may choose c_1, c_2 such that $c_1 < c_2 < 0$ and

$$\int_{c_1}^{c_2} \frac{1}{(1+z^2)^{3/2}} dz > 1 - \frac{1}{8}\epsilon$$

Recall the function $h(z)$ from the proof of Theorem G.2.1; we redefine $h(z)$ here to be a similar function, except that $h(z) = \frac{1}{(1+z^2)^{3/2}}$ when $z \in [c_1, c_2]$, a is chosen close enough to c_1 so that $h(z) < \frac{1}{(1+z^2)^{3/2}}$ when $z < c_1$, and $h(z)$ also goes to zero at $\frac{1}{2}c_2$. We find that

$$\int_{-\infty}^{\infty} h(z)dz > \int_{c_1}^{c_2} \frac{1}{(1+z^2)^{3/2}} dz > 1 - \frac{1}{8}\epsilon$$

Finally we define the skewing functions for the cusps ($k = 1, \ldots, n_{cusp}$):

$$p_k(z) = -\frac{1 - \frac{1}{8}\epsilon(1 + \frac{1}{2^{k-1}})}{\int_{-\infty}^{\infty} h(s)ds} \int_{-\infty}^{z} \int_{-\infty}^{s} h(\tau)d\tau ds$$

$$|p_k''(z)| = \frac{1 - \frac{1}{8}\epsilon(1 + \frac{1}{2^{k-1}})}{\int_{-\infty}^{\infty} h(s)ds} h(z) < \frac{1}{(1+z^2)^{3/2}}$$

When $z \geq \frac{1}{2}c_2$ we have the stronger condition $p_k''(z) = 0$. To make the cusps, we start with n_{cusp} hyperboloids: $r_k(z) = \sqrt{1+z^2}$. Recalling Figure G.2 and the associated argument, we may use Lemma G.1.1 to ramp to $r_k(z) = e^{(1-r_k(c_2/2))z/|c_2|}$ between $z = \frac{1}{2}c_2$ and $z = 0$. We then skew the resulting surfaces using the functions $p_k(z)$. Now let $left_k(z)$ and $right_k(z)$ be as in the proof of Theorem G.2.1 with these surfaces, and similarly let $Left(z)$ be the minimum values of x of the cross sections of the leftmost hyperboloid added in the first part of the proof. That is,

$$Left(z) = -3g + f(z) - \sqrt{1+z^2}$$

When $z > \epsilon$ we have

$$left_k'(z) = p_k'(z) - r_k'(z) > -(1 - \frac{1}{8}\epsilon(1 + \frac{1}{2^{k-1}}))$$

$$right_k'(z) = p_k'(z) + r_k'(z) < -(1 - \frac{1}{8}\epsilon(1 + \frac{1}{2^{k-1}}))$$

$$Left'(z) = f'(z) - \frac{z}{\sqrt{1+z^2}} > -(1 - \frac{1}{2}\epsilon)$$

From the above inequalities it follows that

$$(left_k - right_{k+1})'(z) > \frac{\epsilon}{8 \cdot 2^k} > 0$$

$$(Left - right_1)'(z) > \frac{1}{4}\epsilon > 0$$

Thus if we place our cusps with axes (before skewing) at $y = 0$ and small enough x (decreasing as k increases), they won't intersect one another or our other surface when $z > \epsilon$. Hence we may choose the spacings large enough that in fact they don't intersect except when $z < a - 1$. Finally, we join the bottoms using the BVK construction, choosing the parameters b small enough so that the surgeries only affect parts of the surfaces where $z < a$. □

Figure G.7: A surface with genus 1, 1 circular end and 2 cuspidal ends

A typical result of this construction is shown below. These results are summed up in the following theorem:

Theorem G.2.4. *Let g, n_{circ}, and n_{cusp} be natural numbers such that $n_{circ} \geq 1$. Then there exists a complete, negatively curved, C^2 surface embedded in $\mathbb{R}^3$ with genus g, n_{circ} circular ends and n_{cusp} cuspidal ends.*

G.3 Cusp Ends Only

This result raises the question of whether there are complete, negatively curved surfaces with only cusp ends. We begin by examining the following situation: we have a surface given in cylindrical coordinates by $r = r(\theta, z)$. The signed curvature of a cross section $z = const.$ is given by

$$k_s = \frac{2r_\theta^2 + r^2 - rr_{\theta\theta}}{(r^2 + r_\theta^2)^{3/2}}$$

so it is strictly convex ($k_s > 0$) if and only if $2r_\theta^2 + r^2 - rr_{\theta\theta} > 0$. The curvature of the surface is given by

$$K = \frac{-(2r_\theta^2 + r^2 - rr_{\theta\theta})rr_{zz} - (rr_{\theta z} - r_\theta r_z)^2}{(r^2 + r_\theta^2 + r^2 r_z^2)^2}$$

so $K < 0$ if the cross section is strictly convex and r_{zz} is positive. Next, suppose we have a corner at $z = 0$: for some $z_0 > 0$

1. $r \in C^2$ in $-z_0 < z \leq 0$ and $0 \leq z < z_0$
2. $r_{zz} > 0$ when $|z| \in (0, z_0)$ and $r_z(\theta, +0) > r_z(\theta, -0)$
3. the cross sections $z = const.$ are strictly convex

We have the following Lemma:

Lemma G.3.1. *In the above situation we may redefine $r(\theta, z)$ when $|z| \leq \epsilon$, where ϵ is an arbitrary positive number less than z_0, so that $r \in C^2$ and the curvature remains negative.*

We will refer to this as "ironing out." The proof is given in [2]. Note that, while the statement requires C^∞, the proof only uses C^2. In what follows, we will say that a tube is properly contracting at a certain cross section if $r_z < 0$ and $r_{zz} > 0$. As a final preliminary, consider a factorized form:

$$r(\theta, z) = \Theta(\theta)Z(z)$$

$$K = \frac{-(2\Theta'^2 + \Theta^2 - \Theta\Theta'')\Theta^2 Z''}{Z(\Theta^2 + \Theta'^2 + \Theta^4 Z'^2)^2}$$

where $\Theta, Z > 0$. We see that $K < 0$ if $\Theta(\theta)$ is strictly convex and $Z''(z) > 0$. We have the following Theorem, which rules out a certain type of counterexample to Milnor's conjecture:

Theorem G.3.2. *There is no complete, negatively curved, C^2 surface immersed in $\mathbb{R}^3$ whose ends are a finite collection of properly contracting cusps with strictly convex cross sections.*

Proof. Suppose that M is such a surface. For a particular end, choose coordinates so that $r = r(\theta, z)$ for z in some neighborhood of $z = 0$. Since r_z/r is continuous, it assumes a maximum on each cross section. Let $-b$ be the maximum of r_z/r when $z = 0$. Now we define

$$\Theta(\theta) = r(\theta, 0)$$

$$Z(z) = e^{-bz/2}$$

Finally, we redefine $r(\theta, z)$ to be $\Theta(\theta)Z(z)$ when $z \geq 0$. It is easy to see that the hypotheses of Lemma G.3.1 are satisfied, so we apply it to iron out the corner at $z = 0$. Then for all sufficiently large z,

$$K = \frac{-(2\Theta'^2 + \Theta^2 - \Theta\Theta'')\Theta^2}{(\Theta^2 + \Theta'^2 + \Theta^4 \frac{1}{4} b^2 e^{-bz})^2} \frac{b^2}{4}$$

This attains a (negative) maximum at each cross section. The only z dependence is in the denominator, and it decreases with increasing z, so the absolute value of K only increases with increasing z. Thus K is bounded away from zero on the end. Repeating this procedure for each end in turn, we obtain an immersed surface with curvature negative and bounded away from zero, which is precisely what is forbidden by Efimov's Theorem. □

Notice that the above argument only requires a single strictly convex cross section on each end where it is properly contracting. We now apply this Theorem to a classical example.

Example G.1 (Six-punctured Sphere). The equation for this surface, which is shown below, is

$$x^2y^2 + x^2z^2 + y^2z^2 = 3$$

It is easy (though tedious) to show that this surface is negatively curved everywhere except the eight marked points $x, y, z = \pm 1$. One might suppose that a bounded surgery could excise the points of zero curvature, but the above Theorem proves that this is not the case, since the ends are properly contracting with strictly convex cross sections.

It seems that convex cross sections may be something of a hindrance to having only cuspidal ends. We have the following example:

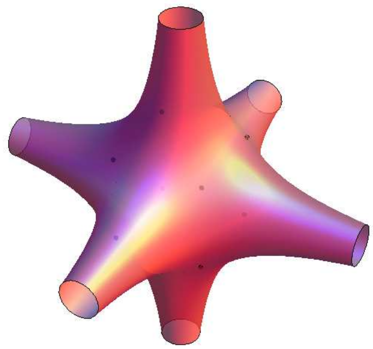


Figure G.8: The six-punctured sphere

Example G.2 (Vaigant's Surface). This surface, due to Vaigant[3], is given by

$$(z-u+v)^2(8+u+v)^2 - M^2[2-(u-1)(v-1)] = 0$$

where $u = \sqrt{1+x^2}$, $v = \sqrt{1+y^2}$ and $M \in (0, \frac{1}{2\sqrt{2}})$. It is negatively curved, and has four cuspidal ends. Note that the cross sections are not convex.

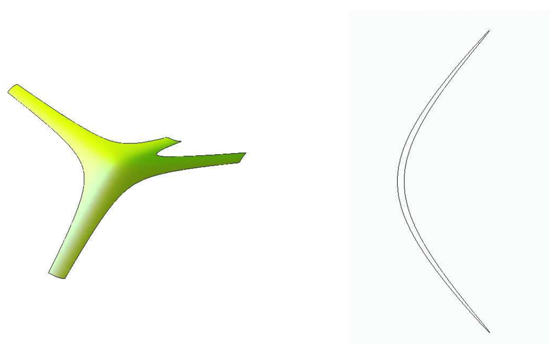


Figure G.9: Vaigant's surface and a (blown-up) cross section

G.4 Space at Infinity

Negatively curved surfaces may not be compact, so if one is complete (in the extrinsic sense) then it is unbounded. To study the behavior at large distances we compactify $\mathbb{R}^3$ in the following way: define the map Φ on $\mathbb{R}^3$ in polar coordinates by

$$\Phi((r,\theta,\phi)) = (\frac{2}{\pi}\tan^{-1} r, \theta, \phi)$$

With this identification of $\mathbb{R}^3$ with the open unit ball, we identify the points of the unit sphere with points at infinity. We now define the points at infinity of a surface to be the limit points on the unit sphere of the image of the surface under Φ. The figure below illustrates this for the hyperboloid. The red circles are not

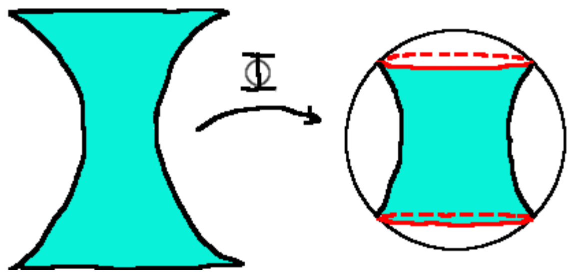


Figure G.10: Behavior of the hyperboloid at infinity

in the image of Φ; they are the points at infinity of the hyperboloid. We have the following Theorem about the points at infinity:

Theorem G.4.1. *The points at infinity are not contained in any open hemisphere.*

Proof. Suppose M is a negatively curved surface violating this. By rotational symmetry we may assume the points at infinity lie below the equator (at negative z). Now let

$$z_{max} = \sup\{z : \exists x \exists y (x, y, z) \in M\}$$

First suppose that $z_{max} = \infty$. Then we may choose a sequence of points (x_n, y_n, z_n) of M such that $z_n > n$. Then $\|(x_n, y_n, z_n)\| \to \infty$, so $\|\Phi((x_n, y_n, z_n))\| \to 1$. Since the unit ball is compact, we may choose a convergent subsequence $\Phi((x_{\nu(n)}, y_{\nu(n)}, z_{\nu(n)}))$, and

$$\lim_{n\to\infty} \Phi((x_{\nu(n)}, y_{\nu(n)}, z_{\nu(n)}))_z \geq 0$$

Thus, $\lim_{n\to\infty} \Phi((x_{\nu(n)}, y_{\nu(n)}, z_{\nu(n)}))$ is a point at infinity of M which is not below the equator, contradicting our assumption, so z_{max} must be finite. Now a negatively curved surface can not lie on one side of a plane that touches it, so for all $(x, y, z) \in M$ we have $z < z_{max}$. Next, we choose a sequence of points (x_n, y_n, z_n) of M such that $z_{max} > z_n > z_{max} - \frac{1}{n}$. Again we choose a convergent subsequence $\Phi((x_{\nu(n)}, y_{\nu(n)}, z_{\nu(n)}))$ of $\Phi((x_n, y_n, z_n))$. Let

$$p = \lim_{n\to\infty} \Phi((x_{\nu(n)}, y_{\nu(n)}, z_{\nu(n)}))$$

If p is on the unit sphere then $\lim_{n\to\infty} \|(x_{\nu(n)}, y_{\nu(n)}, z_{\nu(n)})\| = \infty$ so

$$p_z = \lim_{n\to\infty} z_{\nu(n)} \frac{\frac{2}{\pi}\tan^{-1}\|(x_{\nu(n)}, y_{\nu(n)}, z_{\nu(n)})\|}{\|(x_{\nu(n)}, y_{\nu(n)}, z_{\nu(n)})\|} = 0$$

Thus, p is a point at infinity of M which is not below the equator, again contradicting our assumption, so p must not be on the unit sphere, and hence is in the image of Φ. Since Φ is a homeomorphism, it follows that

$$\lim_{n\to\infty} (x_{\nu(n)}, y_{\nu(n)}, z_{\nu(n)}) = \Phi^{-1}(p)$$

so $\Phi^{-1}(p)$ is a limit point of M. Since M is complete, $\Phi^{-1}(p) \in M$. However,

$$\Phi^{-1}(p)_z = \lim_{n\to\infty} z_{\nu(n)} = z_{max}$$

so M lies to one side of a plane touching it, which is impossible. $\square$

As an application of this result, we derive a result concerning skewing in two different ways, the second more general. Consider a negatively curved horn with equation

$$x^2 + y^2 = r(z)^2$$

where $r'(z) \to -k$ as $z \to -\infty$ and $r'(z) \to 0$ as $z \to \infty$. How far can we move the point at infinity corresponding to the cusp by skewing the surface? If we maintain negative curvature with skewing function $f(z)$ (initially zero), we have

$$\begin{aligned}\lim_{z\to\infty} f'(z) &= \int_{-\infty}^{\infty} f''(z)dz < \int_{-\infty}^{\infty} r''(z)dz \\ &= \lim_{z\to\infty} [r'(z) - r'(-z)] = k\end{aligned}$$

so the point at infinity must lie within the circle opposite the other points at infinity (which form a circle at polar angle $\theta = \pi - \tan^{-1} k$), as shown in the figure below. Using the space at infinity we can put this more

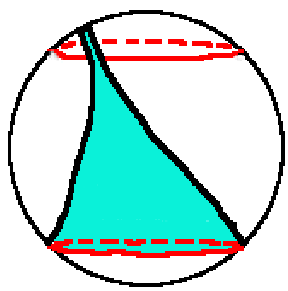


Figure G.11: Cusp skewing limit

generally:

Proposition G.4.2. *Let M be a negatively curved surface whose set of points at infinity is $S \cup \{p\}$. If S is contained in the (not great) circle C, then p is contained in its opposite circle C'.*

Proof. Suppose not. By rotational symmetry we may assume that C lies in a plane below and parallel to the xy plane, while p lies in the xz plane with positive x, as shown in the figure below. We then describe the great circle D (shown below in yellow) through the leftmost point of C and its antipodal point, the rightmost point of C', with normal vector in the xz plane. Now we simply tilt D slightly toward p, so that all the points at infinity lie on an open hemisphere (delimited by D), which contradicts Theorem G.4.1. $\square$

With this and Theorem G.4.1, we can also rule out negatively curved surfaces with one cusp end or two which are not antipodal, since cusp ends get a single point at infinity. Furthermore, we can also rule out three linearly independent cusps as follows: let v_1, v_2, v_3 be the three directions of the points at infinity. We simply

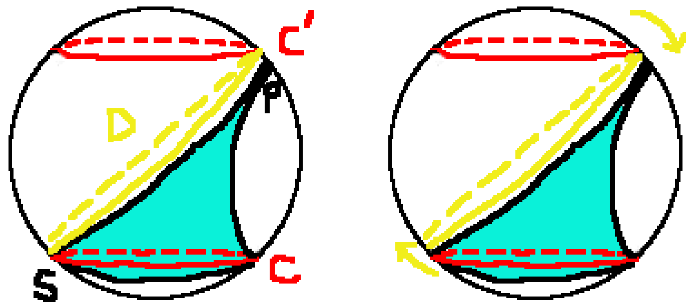


Figure G.12: More general cusp skewing limit

apply a linear transformation (which doesn't change the sign of the curvature) to move those directions to the standard basis vectors, and we have three points at infinity which fit on an open hemisphere. Finally, we give an argument which can be used to rule out various symmetric geometries.

Proposition G.4.3 (Symmetry Argument)**.** *If C is a curve in a plane of symmetry Π of a negatively curved surface M, then C is a principle curve.*

Proof. Let $p \in C$. By symmetry, the normal vector N at p is in Π, so the tangent plane to M is spanned by the tangent vector to C and the normal vector to Π. The principle directions (really lines) must still be principle after reflection across Π, and are distinguishable from one another by the sign of the normal curvature, so they must in fact be invariant under reflection across Π. Hence the tangent vector to C and the normal vector to Π are principle. Since this is true for every p in C, C is principle. □

From this it follows that C has no points of zero curvature, since this would imply normal curvature zero. Hence we can rule out certain symmetric arrangements by inspecting a curve of symmetry. In the following examples, the offending curves are shown in yellow. Approximate points of zero curvature are marked. Assume the drawings have the obvious symmetries.

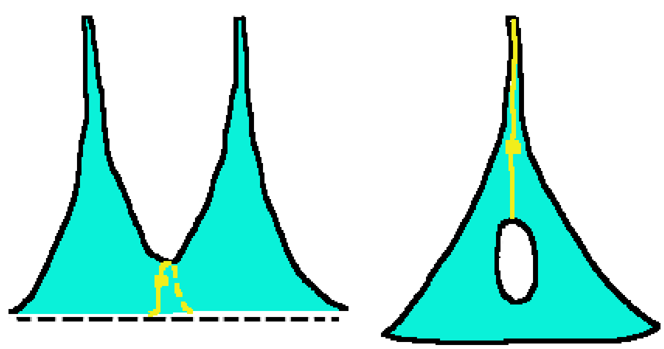
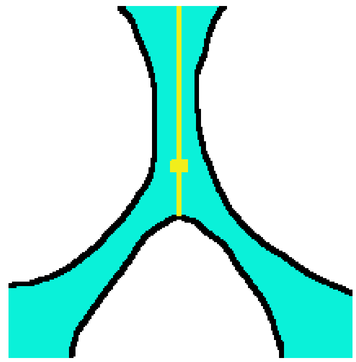


Figure G.13: Symmetric geometries (not negatively curved)



Bibliography

1. I. Ya. Bakel'man, A. L. Verner, B. E. Kantor, *Introduction to Differential Geometry in the Large*, Nauka, Moscow, 1973, Zbl.276.53093.

2. È. R. Rozendorn, *Weakly irregular surfaces of negative curvature*, Uspehi Mat. Nauk 21 (1966), no. 5 (131), 97-100; English transl. in Russian Math Surveys 21 (1966)

3. Yu. D. Burago, V. A. Zalgaller, eds, *Encyclopedia of Mathematical Sciences Vol. 48*, Springer-Verlag, New York, 1992.

A modification of periodic Golay sequence pairs

Adam Vollrath
Northwestern University

Indiana University REU Summer 2008
Advisor: ADVISOR

H.1 Preliminaries

Throughout our discussion, we will say that a sequence, A, is *binary* if its only entries are 1 and -1. Likewise, a matrix is *binary* if its only entries are 1 and -1. Given an $n \times n$ binary matrix, H, it is well known what the maximum attainable determinant is of that size. This bound depends not only on n, but also the value of $n \pmod 4$. We are particularly interested in the case that $n \equiv 0$ or $2 \pmod 4$. If $n \equiv 0 \pmod 4$, the maximum possible determinant is $n^{\frac{n}{2}}$; this bound is achieved if and only $HH^T = nI_n$ where I_n is the $n \times n$ identity matrix [10]. A matrix satisfying this property is said to be *Hadamard*, and it is conjectured that a Hadamard matrix exists for every multiple of 4. If $n \equiv 2 \pmod 4$, the maximum possible determinant is $2(n-1)(n-2)^{\frac{n-2}{2}}$ which Ehlich and Wojtas independently derived. However, the Ehlich/Wojtas bound can only be achieved if $n-1$ is a sum of two squares [6], [13]. Since there are many positive integers $n \equiv 2 \pmod 4$ with $n-1$ not equal to a sum of two squares, we may consider how close to the Ehlich/Wojtas bound the determinant of a binary matrix of this size can be. Before we can answer this question, however, we require the following preliminary results and definitions.

Definition H.1. Let $A = (a_0, a_1, \ldots, a_{l-1})$ be a complex-valued sequence of length l, then the k- *th periodic autocorrelation* of A is

$$P_A(k) = \sum_{i=0}^{l-1} a_i a_{i+k}$$

where $i + k$ is taken modulo l for every $k = 0, 1, \ldots, l-1$.

We can easily generalize this definition to any finite number of complex-valued sequences of length l. If $X = A_1, A_2, \ldots, A_n$ is a collection of sequences of length l, then its k- *th periodic autocorrelation* is simply the sum of the periodic autocorrelations of the A_i. That is,

$$P_X(k) = \sum_{i=1}^{n} P_{A_i}(k).$$

We are specifically interested in the case when there are two binary sequences, $A = (a_0, a_1, \ldots, a_{l-1})$ and $B = (b_0, b_1, \ldots, b_{l-1})$. Accordingly, we will assume these conditions hold whenever we discuss A and B in the future, unless specified otherwise. It is clear in this case that $\sum_{i=0}^{l-1}(a_i a_{i+k} + b_i b_{i+k}) = 2l$ for $k = 0$.

Definition H.2. Suppose there exists a $c \in \mathbb{Z}$ such that $P_A(k) + P_B(k) = c$ for every nonzero k, then A and B are said to be *compatible*. If $c = 0$, then A and B form a *periodic Golay pair*.

Example H.1. Let $A = (1, 1)$ and $B = (1, -1)$, then $P_A(1) + P_B(1) = (1 + 1) + (-1 - 1) = 0$. Thus, A and B form a periodic Golay pair.

Definition H.3. Let $A = (a_0, \ldots, a_{l-1})$, then the $l \times l$ matrix

$$M = \begin{bmatrix} a_0 & a_1 & \ldots & a_{l-1} \\ a_{l-1} & a_0 & \ldots & a_{l-2} \\ \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & \ldots & a_0 \end{bmatrix}$$

is called the *circulant matrix obtained from A*.

With this basic understanding, we are now able to modify our definition of a periodic Golay pair. This will create a new type of pair which has not been studied before and has applications to the maximal determinant problem described above.

H.2 Properties of Periodic and Alternating Golay Pairs

We now generalize the notion of a compatible pair by allowing the non-trivial periodic autocorrelations of a pair to take on multiple values. As we shall see, these generalized pairs share many of the same desirable properties present in periodic Golay pairs. Moreover, such pairs have been used to create binary matrices of record determinant, specifically of sizes 22, 34, 70, and 106 [11], [12].

Definition H.4. We say that the binary pair $A = (a_0, \ldots, a_{l-1})$ and $B = (b_0, \ldots, b_{l-1})$ is *alternating compatible with odd constant s and even constant t* if

$$\sum_{i=0}^{l-1}(a_i a_{i+k} + b_i b_{i+k}) = \begin{cases} 2l & \text{if } k = 0; \\ s & \text{if } k \text{ is odd}; \\ t & \text{if } k \text{ is even, } k \neq 0. \end{cases}$$

If $s = -4$ and $t = 0$, then we say A and B form an *alternating Golay pair.*

Example H.2. Let $A = (1, 1, -1, -1)$ and $B = (1, -1, 1, -1)$. Then, $P_A(1) + P_B(1) = -4 = P_A(3) + P_B(3)$ and $P_A(2) + P_B(2) = 0$, so A and B form an alternating Golay pair.

From the definition of alternating compatible sequences it is immediate that if s and t are distinct, then l must be even. This is because $P_A(k)+P_B(k) = P_A(k+l)+P_B(k+l)$, and the value of the nonzero periodic autocorrelation depends only on whether k is even or odd. It is worth noting, however, that if $s = t$ (that is, if A and B are compatible in the usual sense), then l need not be even.

Proposition H.2.1. *Suppose $A = (a_0, \ldots, a_{l-1})$ and $B = (b_0, \ldots, b_{l-1})$ form an alternating compatible pair with distinct odd constant s and even constant t. Then,*

$$\left(\sum_{i=0}^{l-1} a_i\right)^2 + \left(\sum_{i=0}^{l-1} b_i\right)^2 = \left(2 + \frac{s}{2} + \frac{t}{2}\right) l - t.$$

Proof. Straightforward calculations show that

$$\begin{aligned} \left(\sum_{i=0}^{l-1} a_i\right)^2 + \left(\sum_{i=0}^{l-1} b_i\right)^2 &= \sum_{i=0}^{l-1} \left({a_i}^2 + {b_i}^2\right) + \sum_{i=1}^{l-1} \left(P_A(i) + P_B(i)\right) \\ &= 2l + s\frac{l}{2} + t\left(\frac{l}{2} - 1\right) \\ &= \left(2 + \frac{s}{2} + \frac{t}{2}\right) l - t. \end{aligned}$$

□

Corollary H.2.2. *If A and B form an alternating Golay pair of length l, then precisely $\frac{l}{2}$ entries in both A and B are 1.*

Proof. Letting $s = -4$ and $t = 0$, Proposition H.2.1 shows that $\left(\sum_i a_i\right)^2 + \left(\sum_i b_i\right)^2 = 0$. Thus, the sum of the entries of A and B must both be zero, which occurs if and only if precisely $\frac{l}{2}$ entries of both A and B are 1. □

Suppose A and B form a periodic Golay pair of length l, and let α denote the sum of the components of A and β denote the sum of the components of B. Then, it is well known that $2l = \alpha^2 + \beta^2$ and hence $l = (\frac{\alpha+\beta}{2})^2 + (\frac{\alpha-\beta}{2})^2$ (this fact also follows from Proposition H.2.1, since the length of a periodic Golay pair must be even). The following theorem presents the analogous condition in the case of alternating Golay sequences.

Theorem H.2.3. *Suppose $A = (a_0, \ldots, a_{l-1})$ and $B = (b_0, \ldots, b_{l-1})$ form an alternating Golay pair of length l, then*

$$l = \left(\sum_{i=0}^{l/2-1} a_{2i}\right)^2 + \left(\sum_{i=0}^{l/2-1} b_{2i}\right)^2.$$

Proof. Let M and N be the $l \times l$ circulant matrices obtained from A and B respectively, and let

$$H = \begin{bmatrix} M & N \\ -N^T & M^T \end{bmatrix}.$$

We wish to compute the product HH^THH^T in two separate ways: first by calculating $(HH^T)(HH^T)$ and then by determining $H(H^TH)H^T$. The desired result comes from comparing the (1,1)- entries of these two products, which we denote by λ.

Let $\Delta = (\delta_{ij})$ be the $l \times l$ matrix defined by

$$\delta_{ij} = \begin{cases} 2l & \text{if } i = j; \\ -4 & \text{if } i+j \text{ is odd;} \\ 0 & \text{if } i+j \text{ is even, } i \neq j. \end{cases}$$

Because circulant matrices commute, we have

$$(HH^T)(HH^T) = \begin{bmatrix} \Delta & 0 \\ 0 & \Delta \end{bmatrix}^2 = \begin{bmatrix} \Delta^2 & 0 \\ 0 & \Delta^2 \end{bmatrix}.$$

Hence, if $v \in \mathbb{R}^l$ is $v = (2l, -4, 0, \ldots, -4)$, then

$$\lambda = v \cdot v = 4l^2 + 8l. \tag{H.1}$$

Let $E_A = \sum_{i=0}^{l/2-1} a_{2i}$ and $E_B = \sum_{i=0}^{l/2-1} b_{2i}$. Observe that

$$E_A + \sum_{i=0}^{l/2-1} a_{2i+1} = \sum_{i=0}^{l/2-1} a_{2i} + \sum_{i=0}^{l/2-1} a_{2i+1} = \sum_{i=0}^{l-1} a_i = 0$$

by Corollary H.2.2, and thus $\sum_{i=0}^{l/2-1} a_{2i+1} = -E_A$. Likewise, $\sum_{i=0}^{l/2-1} b_{2i+1} = -E_B$. We see

$$H(H^TH)H^T = \begin{bmatrix} M & N \\ -N^T & M^T \end{bmatrix} \begin{bmatrix} \Delta & 0 \\ 0 & \Delta \end{bmatrix} \begin{bmatrix} M^T & -N \\ N^T & M \end{bmatrix},$$

and hence

$$\begin{aligned}
\lambda &= a_0\left(2la_0 - 4\sum_{i=0}^{l/2-1} a_{2i+1}\right) + \cdots + a_{l-1}\left(2la_{l-1} - 4\sum_{i=0}^{l/2-1} a_{2i}\right)\\
&\quad + b_0\left(2lb_0 - 4\sum_{i=0}^{l/2-1} b_{2i+1}\right) + \cdots + b_{l-1}\left(2lb_{l-1} - 4\sum_{i=0}^{l/2-1} b_{2i}\right)\\
&= 2l\sum_{i=0}^{l-1}({a_i}^2 + {b_i}^2) - 4\left(2\sum_{i=0}^{l/2-1} a_{2i}\sum_{i=0}^{l/2-1} a_{2i+1} + 2\sum_{i=0}^{l/2-1} b_{2i}\sum_{i=0}^{l/2-1} b_{2i+1}\right)\\
&= 4l^2 + 8{E_A}^2 + 8{E_B}^2. \qquad \text{(H.2)}
\end{aligned}$$

Combining (H.1) and (H.2), we have that $4l^2 + 8l = 4l^2 + 8{E_A}^2 + 8{E_B}^2$ and therefore $l = {E_A}^2 + {E_B}^2$. □

The results of Theorem H.2.3 lead us to consider if there exists a relationship between the sum of the even and odd entries of a periodic Golay pair and its length. Suppose $A = (a_0, \ldots, a_{l-1})$ and $B = (b_0, \ldots, b_{l-1})$ form a periodic Golay pair, and let

$$E_A = \sum_{i=0}^{l/2-1} a_{2i}, \quad D_A = \sum_{i=0}^{l/2-1} a_{2i+1}, \quad E_B = \sum_{i=0}^{l/2-1} b_{2i}, \quad D_B = \sum_{i=0}^{l/2-1} b_{2i+1}. \qquad \text{(H.3)}$$

Consider the complex-valued polynomials $Q_A(x) = a_0 + a_1x + \cdots + a_{l-1}x^{l-1}$ and $Q_B(x) = b_0 + b_1x + \cdots + b_{l-1}x^{l-1}$, then for any l- th root of unity, ζ,

$$Q_A(\zeta)Q_A(\zeta^{-1}) + Q_B(\zeta)Q_B(\zeta^{-1}) = 2l.$$

Hence, if $\zeta_1 = 1$ and $\zeta_2 = -1$, then

$$\begin{aligned}
4l &= Q_A(\zeta_1)Q_A(\zeta_1^{-1}) + Q_B(\zeta_1)Q_B(\zeta_1^{-1}) + Q_A(\zeta_2)Q_A(\zeta_2^{-1})\\
&\quad + Q_B(\zeta_2)Q_B(\zeta_2^{-1})\\
&= (E_A + D_A)^2 + (E_B + D_B)^2 + (E_A - D_A)^2 + (E_B - D_B)^2\\
&= 2\left({E_A}^2 + {D_A}^2 + {E_B}^2 + {D_B}^2\right).
\end{aligned}$$

Thus,

$$2l = {E_A}^2 + {D_A}^2 + {E_B}^2 + {D_B}^2$$

which provides a nice property regarding the length of any periodic Golay pair. Notice that this property is equivalent to stating that

$$E_AD_A = -E_BD_B. \qquad \text{(H.4)}$$

This gives rise to the following result which greatly improves the efficiency of the search described in Section 4.

Theorem H.2.4. *Assume $A = (a_0, \ldots, a_{l-1})$ and $B = (b_0, \ldots, b_{l-1})$ form a periodic Golay pair of length $l = 2m$ for some odd $m \in \mathbb{Z}$, and assume further that there is a unique decomposition of l into a sum of two squares of nonnegative integers. Then, if E_A, D_A, E_B, and D_B are as in (H.3), then*

$$l = {E_A}^2 + {D_A}^2 = {E_B}^2 + {D_B}^2.$$

Proof. Let α and β be the respective sums of the entries of A and B, then $2l = \alpha^2 + \beta^2$. Furthermore, this decomposition into a sum of squares is unique, since the decomposition of l into a sum of squares is unique. By (H.4), $E_A D_A = -E_B D_B = \gamma$ for some $\gamma \in \mathbb{Z}$, so $\alpha = E_A + \frac{\gamma}{E_A}$ and likewise $\beta = E_B - \frac{\gamma}{E_B}$. Multiplying the first equation by E_A and the second by E_B gives

$$E_A = \frac{\alpha \pm \sqrt{\alpha^2 - 4\gamma}}{2}, \qquad E_B = \frac{\beta \pm \sqrt{\beta^2 + 4\gamma}}{2}. \tag{H.5}$$

Because E_A and E_B must both be integers, it follows that $\alpha^2 - 4\gamma = p^2$ and $\beta^2 + 4\gamma = q^2$ for some integers p and q. Therefore, $p^2 + q^2 = \alpha^2 + \beta^2 = 2l$, so either $p^2 = \alpha^2$ and $q^2 = \beta^2$ or vice versa by the uniqueness hypothesis.

Suppose $p^2 = \alpha^2$, then $E_A = 0$ or α. However, this is impossible because then either E_A or D_A is 0, which contradicts the assumption that m is odd. Thus, $p^2 = \beta^2$ and $q^2 = \alpha^2$. Substituting these values into (H.5) gives

$$E_A = \frac{\alpha \pm \beta}{2}, \quad D_A = \frac{\alpha \mp \beta}{2}, \quad E_B = \frac{\beta \pm \alpha}{2}, \quad D_B = \frac{\beta \mp \alpha}{2}.$$

Without loss of generality, we may assume $E_A = E_B = \frac{\alpha+\beta}{2}$ since shifting the entries of an individual sequence in a periodic Golay pair does not affect its periodic autocorrelation. Hence, $D_A = \frac{\alpha-\beta}{2}$ and $D_B = \frac{\beta-\alpha}{2}$, and the result follows from taking ${E_A}^2 + {D_A}^2$ and ${E_B}^2 + {D_B}^2$. □

Arasu and Xiang proved that if $l = p^{2t}u$ is the length of a periodic Golay pair for some prime $p \equiv 3 \pmod 4$ and positive integers t and u with u relatively prime to p, then $u \geq 2p^t$ [1]. Because of the close similarity between the properties of alternating Golay pair lengths and periodic Golay pair lengths, we conjecture a similar result holds for alternating Golay pairs.

H.3 Implementing Alternating Golay Pairs to Construct Binary Matrices of Record Determinant

With this basic understanding of the properties of alternating Golay pairs, we return to the issue of maximal determinants; our objective is to apply these pairs to create binary matrices of record determinant. The following is a well known result whose proof is omitted.

Theorem H.3.1. *Assume A and B form a periodic Golay pair of length l, and let M and N be the circulant matrices obtained from A and B, respectively. Then,*

$$H = \begin{bmatrix} M & N \\ -N^T & M^T \end{bmatrix}$$

forms a $2l \times 2l$ Hadamard matrix.

As mentioned previously, if $n \equiv 2 \pmod 4$, then for any binary $n \times n$ matrix H, $\det(H) \leq 2(n-1)(n-2)^{\frac{n-2}{2}}$ [6], [13]; we denote this bound by $\beta(n)$. We wish to use a similar construction of circulant matrices as in Theorem H.3.1 to generate $n \times n$ binary matrices of record determinant where $n \equiv 2 \pmod 4$ but with alternating Golay pairs. However, we first require the following lemma:

Lemma H.3.2. *Let $S = (s_{ij})$ be the $n \times n$ matrix defined by*

$$s_{ij} = \begin{cases} k & \text{if } i+j \text{ is even;} \\ -k & \text{if } i+j \text{ is odd} \end{cases}$$

for some $k \in \mathbb{C}$ and $1 \leq i, j \leq n$, and let $R = S + dI_n$ where I_n is the $n \times n$ identity matrix. Then, $\det(R) = (kn + d)d^{n-1}$.

Proof. Since S is not invertible, 0 is an eigenvalue of S. The dimension of its eigenspace is $n-1$, because every row of S is a multiple of its first row. Moreover, kn is an eigenvalue of S with eigenvector $v = (x_1, \ldots, x_n)$ where $x_i = (-1)^{i+1}$ for $i = 1, \ldots, n$. Thus, kn and 0 are the only eigenvalues of S with algebraic multiplicities 1 and $n-1$, respectively. By the Spectral Theorem, there exists and invertible $n \times n$ matrix Q such that $Q^{-1}SQ$ is the $n \times n$ matrix which has kn in its (1,1)- entry and zeroes everywhere else. Thus,

$$\begin{aligned} \det(R) &= \det(S + dI_n) = \det(Q^{-1}(S + dI_n)Q) = \det(Q^{-1}SQ + dI_n) \\ &= \det\left(\begin{bmatrix} kn & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 \end{bmatrix} + \begin{bmatrix} d & 0 & \ldots & 0 \\ 0 & d & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & d \end{bmatrix} \right) \\ &= (kn + d)d^{n-1}. \end{aligned}$$

∎

With this lemma, it is now possible to determine how close matrix constructions from alternating Golay pairs come to reaching the Ehlich/Wojtas bound.

Theorem H.3.3. *Suppose that A and B form an alternating Golay pair of length l. Let M and N denote the $l \times l$ circulant matrices obtained from A and B, respectively, and let X, J, and K be the following 2×2 and $2 \times l$ matrices:*

$$X = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad J = \begin{pmatrix} 1 & \ldots & 1 \\ 1 & \ldots & 1 \end{pmatrix}, \quad K = \begin{pmatrix} 1 & \ldots & 1 \\ -1 & \ldots & -1 \end{pmatrix}.$$

If H is the $2l + 2 \times 2l + 2$ matrix given by

$$H = \begin{bmatrix} X & J & K \\ J^T & M & N \\ K^T & -N^T & M^T \end{bmatrix},$$

then $\det(H) = 2(2l)^{l+1}$.

Proof. Observe that

$$HH^T = \begin{bmatrix} (2l+2)I_2 & P & Q \\ P^T & R & 0 \\ Q^T & 0 & R \end{bmatrix}$$

where I_2 is the 2×2 identity matrix, P and Q are the $2 \times l$ matrices

$$P = \begin{pmatrix} 2 & \ldots & 2 \\ 0 & \ldots & 0 \end{pmatrix}, \qquad Q = \begin{pmatrix} 0 & \ldots & 0 \\ 2 & \ldots & 2 \end{pmatrix},$$

and $R = (r_{ij})$ is the $l \times l$ matrix defined by

$$r_{ij} = \begin{cases} 2l+2 & \text{if } i = j; \\ -2 & \text{if } i+j \text{ is odd;} \\ 2 & \text{if } i+j \text{ is even, } i \neq j. \end{cases}$$

Let Y be the $2 \times 2l$ matrix obtained by gluing P and Q together, and let

$$Z = \begin{bmatrix} R & 0 \\ 0 & R \end{bmatrix},$$

then

$$\begin{aligned} \det(H)^2 &= \det\left(\begin{bmatrix} (2l+2)I_2 & Y \\ Y^T & Z \end{bmatrix}\right) \\ &= \det\left(\begin{bmatrix} (2l+2)I_2 & Y \\ Y^T & Z \end{bmatrix}\begin{bmatrix} I_2 & 0 \\ -Z^{-1}Y^T & I_{2l} \end{bmatrix}\right) \\ &= \det\left(\begin{bmatrix} (2l+2)I_2 - YZ^{-1}Y^T & Y \\ 0 & Z \end{bmatrix}\right) \\ &= \det(Z)\det\left((2l+2)I_2 - YZ^{-1}Y^T\right) \end{aligned}$$

It is straightforward to verify that

$$Z^{-1} = \begin{bmatrix} \Gamma & 0 \\ 0 & \Gamma \end{bmatrix}$$

where $\Gamma = (\gamma_{ij})$ is the $l \times l$ matrix given by

$$\gamma_{ij} = \begin{cases} \frac{2l-1}{4l^2} & \text{if } i = j; \\ \frac{1}{4l^2} & \text{if } i+j \text{ is odd;} \\ \frac{-1}{4l^2} & \text{if } i+j \text{ is even, } i \neq j. \end{cases}$$

Thus, $YZ^{-1}Y^T = 2I_2$, so $(2l+2)I_2 - YZ^{-1}Y^T = 2lI_2$. By Lemma H.3.2, $\det(R) = 2(2l)^l$, so $\det(Z) = \det(R)^2 = 2^2(2l)^{2l}$. Therefore,

$$\det(H) = \sqrt{\det(Z)\det(Z - YZ^{-1}Y^T)} = \sqrt{2^2(2l)^{2l}(2l)^2} = 2(2l)^{l+1}.$$

$\square$

If H is an $n \times n$ matrix constructed from an alternating Golay pair of length l as in the previous theorem, then $\frac{\det(H)}{\beta(n)} = \frac{n-2}{n-1}$ and therefore

$$\lim_{n\to\infty} \frac{\det(H)}{\beta(n)} = 1.$$

Hence, alternating Golay pairs are excellent for producing $n \times n$ binary matrices of large determinant for $n \equiv 2 \pmod 4$ whenever the Ehlich/Wojtas bound is not attainable. As mentioned before, the construction of H as in Theorem H.3.3 has been used to generate binary matrices of record determinant for $n = 22$, 34, 70, and 106 [11], [12].

H.4 The Search for Alternating and Periodic Golay Pairs

In 1997 and 1998 papers, Gysin and Seberry provide a method for locating D-optimal designs using generalized cyclotomy [8], [9]. We utilized this method to search for a variety of pairs. A description of the method is provided below.

Let x and l be positive integers with x relatively prime to l. Define *the i- th generalized coset of l with respect to x* to be

$$C_i = \bigcup_{k \in \mathbb{Z}/l\mathbb{Z}} ix^k \pmod{l}$$

for $i = 0, 1, \ldots, t$ for some $t \in \mathbb{Z}$. Note that these generalized cosets are only cosets in the usual algebraic sense if l is prime. In fact, if l is composite, then there will be nonzero generalized cosets of different lengths. They do, however, form a partition of $\mathbb{Z}/l\mathbb{Z}$ regardless of their lengths. Let $v_i \in \mathbb{Z}^l$ be the incidence vector corresponding to C_i; that is, the k- th entry of v_i is 1 if $k \in C_i$ and 0 otherwise for $k = 0, 1, \ldots, l-1$. Then, by choosing appropriate sums and differences of the v_i, we obtain a list of possible sequence pairs satisfying the desired periodic autocorrelation. By appropriate sums and differences, we refer to the linear combinations of the v_i with weights 1 and -1 which result in the requisite sums of the entries of the potential A and B sequences. If l is even, then each element in a given coset is either even or odd. This fact and Theorems H.2.3 and H.2.4 greatly reduce the number of potential alternating and periodic Golay pairs to be tested.

Example H.3. Suppose we would like to search for an alternating Golay pair, denoted by A and B, of length $l = 10$. If $x = 9$, then $C_0 = (0), C_5 = (5)$ and $C_i = (i, 10-i)$ for $i = 1, 2, 3,$ and 4. Since we are searching for alternating Golay pairs, exactly five entries of A and B must be 1 and the other five must be -1. In fact, Theorem H.2.3 requires that four of the even-ordered entries of A must be 1, and three of the even-ordered entries of B must be 1. Thus, the only possibility for A in this case is $A = -v_0 - v_1 + v_2 - v_3 + v_4 + v_5$. Likewise, B must be obtained by adding the incidence vectors of one even coset of size 2, one odd coset of size 2, and v_0, and subtracting off the remaining incidence vectors. Hence, $A = (-1, -1, 1, -1, 1, 1, 1, -1, 1, -1)$ and $B = (1, 1, -1, -1, 1, -1, 1, -1, -1, 1)$ form an alternating Golay pair of length 10. In this case, $B = v_0 + v_1 - v_2 - v_3 + v_4 - v_5$.

Due to the flexible nature of the program, it is possible to search for several different types of sequences with various periodic autocorrelations. However, we focused our search on three types of pairs: periodic Golay, alternating Golay, and generalized Legendre pairs. A *generalized Legendre pair* is two binary sequences whose entry sums are both 1 and whose periodic autocorrelations sum to -2. Generalized Legendre pairs are of special interest as the only known restriction of their lengths is that they be odd. Moreover, if l is the length of a generalized Legendre pair, then a $2l \times 2l$ Hadamard matrix exists. Therefore, if there exists a generalized Legendre pair for every positive, odd integer, then there exists a Hadamard matrix of size $4n \times 4n$ for every positive integer n.

The results of the search are provided below in addition to two previously known pairs which were not found via generalized cyclotomy; they are included for completeness[1]. Missing lengths in the generalized Legendre case up to 51 can be located in [7]. In the tables, we replace 1 by + and -1 by $-$ for clarity, and x denotes the relatively prime element used to generate the generalized cosets.

If $A = (a_0, \ldots, a_{l-1})$ and $B = (b_0, \ldots, b_{l-1})$ are two binary sequences, then they form a *Golay comple-*

[1]† indicates the pair is listed in [12]
‡ indicates the pair is listed in [2].

mentary pair if

$$\sum_{i=0}^{l-k-1} a_i a_{i+k} + b_i b_{i+k} = 0$$

for $k = 1, 2, \ldots, l-1$. It is straightforward to verify that every Golay complementary pair is a periodic Golay pair. In 1998 and 2007, Dokovic found periodic Golay pairs of lengths 34 and 50 which are known to possess no Golay complementary pairs [4], [3]. This is the first time a periodic Golay pair of length 82 has been found and only the third time that a periodic Golay pair has been discovered for a given length which does not admit any Golay complementary pairs [2].

Table H.1: Alternating Golay Pairs

Length	x	Pair
2	1	+−; +−
4	1	++−−; +−+−
8	1	+++−+−−−; ++−+−+−−
10	9	++−+−−−+−+; ++−−+−+−−+
16	1	+++++−+−−−+−+−−−; +++−+−−+−+−−++−−
20	9	++−+−+++−+−+−−+−−−−+; ++++−−−+−++−−+−−−++−
26	3	+++++++−−++−+−−+−−++−−−−−−; ++−+−+−+−+−+−−+++−−+−++−−−
34†	n/a	++−++−++++−+−+−+−−−+−−−−−+++++−−−−; ++−−+−+−++−−−−+−−+−+++−−+++−+−++−−
52	9	+++++−++−+−+++−−−−+−−+−+−−−+−+−−−+ ++++−−−−−+−−++−−++; ++−++−−−−+−−+−+++−−+−+++−−++−+−+−+ −+++−−+−++−−−−+−−+

Table H.2: Periodic Golay Pairs

Length	x	Pair
2	1	++; +−
4	1	+++−; +++−
8	1	++++++−−; ++−+−+−−
10	1	++++++−+−−; +++−+−++−−
16	7	++++−++++−+−−−+−; +++−−−+++−++−++−
20‡	n/a	++++−+−−−++−−++−+−−+; ++++−+++++−−−+−+−++−
26	3	+++++−++−++++−−−−++−−+−+−+; −+++−+++−+−+−−+++−++−++−−−
82	37	++++++−++++−−+−−++++++−+−+−−−+−++− −−++−++−+−−+++−+++−−−+−+−+−−+−+−++ −−−++−+−−−+−++; +++−−+++−−−+++−−+−++−−+−++−+++++−+ −+++−−−−−+++++−+++−++−++++−+−−−−+− −++−−+−−−−−++−

Table H.3: Generalized Legendre Pairs

Length	x	Pair
3	2	+ + −; + + −
5	4	+ + − − +; + − + + −
7	2	+ + + − + − −; + + + − + − −
9	1	+ + + + − + − − −; + + − + − + + − −
11	3	+ + − + + + − − − + −; + + − + + + − − − + −
13	3	+ + + + − + + − − + − − −; + + − + − − − + + + − + −
15	2	− + + + + − + − + + − − + − −; + + + + − + − + + − − + − − −
17	2	+ + + − + − − − + + − − − + − + +; + − − + − + + + − − + + + − + − −
19	4	+ + − − + + + + − + − + − − − − + + −; + + − − + + + + − + − + − − − − + + −
23	2	+ + + + + − + − + + − − + + − − + − + − − − −; + + + + + − + − + + − − + + − − + − + − − − −
29	4	+ + − − + + + + − + − − − + − − + − − − + − + + + + − − +; + − + + − − − − + − + + + − + + − + + + − + − − − − + + −
31	4	+ + + + + − + + + − − − + − + − + + − + − − − − + + − − + − −; + + + + + − + + + − − − + − + − + + − + − − − − + + − − + − −
37	3	+ + − + + − − + − + + + + − − − + − − − − + − − − + + + + − + − − + + −+; + − + − − + + − + − − − − + + + − + + + + − + + + − − − − + − + + − − +−
41	3	+ + + − + + − − + + + − − − − − + − + − + + − + − + − − + − − − + + + − − + + − + +; + − − + − − + + − − − + + + + + − + − + − − + − + − + + + + + − − − + + − − + −−;
43	4	+ + + − + − − + + − − + − − − − + − + − − + + − − − + + + + + − + − + + − − − + − ++; + + + − + − − + + − − + − − − − + − + − − + + − − − + + + + + − + − + + − − − + − ++
47	4	+ + + + + − + + + + − − + − + − + + + − − + − − + + − + + − − − + − + − + + − − − − + − − −−; + + + + + − + + + + − − + − + − + + + − − + − − + + − + + − − − + − + − + + − − − − + − − −−
49	18	+ + + + + + + + − − + + − + + − − − + + − − + + − − − − + − + + − + − − + − + − − + − − − − − ++; + + + − − − − + − − − + + + + − − + + − + + − − + − + + + − + − − − + + + − + − + − + − − + − +−
53	4	+ + − − + − + + − + + + − + − + + + − − − − − − + + − − + + − − − − − − + + + − + − + + + − + + − + − −+; + − + + − + − − + − − − + − + − − − + + + + + + − − + + − − + + + + + + − − − + − + − − − + − − + − + +−

Bibliography

1. Arasu, K.T. and Xiang, Q. On the Existence of Periodic Complementary Binary Sequences, *Designs, Codes and Cryptography* **2** (1992) 257-262

2. Borwein, P. B. and Ferguson, R. A. A Complete Description of Golay Pairs for Lengths up to 100, *Mathematics of Computation* **73** (2004) no. 246, 967-985 (electronic)

3. Dokovic, D. Z. Cyclic $((v; r, s; \lambda)$ Difference Families with Two Base Blocks and $v \leq 50$, (2007) arXiv:0707.2173v2 [math.CO]

4. Dokovic, D. Z. Note on periodic complementary sets of binary sequences, *Designs, Codes, and Cryptography* **13** (1998) 251-256

5. Dokovic, D. Z. Some New D-Optimal Designs, *Australasian Journal of Combinatorics* **15** (1997) 221-231

6. Ehlich, H. Determinantenabscähtzungen für binäre Matrizen, *Math. Z.* **83** (1964) 123-132

7. Fletcher, R. J., Gysin, M., and Seberry, J. Application of the Discrete Fourier Transform to the Search for Generalised Legendre Pairs and Hadamard Matrices, *Australasian Journal of Combinatorics* **23** (2001) 75-86

8. Gysin, M. New D-Optimal Designs via Cyclotomy and Generalised Cyclotomy, *Australasian Journal of Combinatorics* **15** (1997) 247-255

9. Gysin, M. and Seberry, J. An Experimental Search and New Combinatorial Designs via a Generalisation of Cyclotomy *JCMCC* **27** (1998) 143-160

10. Hadamard, J. Résolution dúne question relative aux déterminants, *Bull. Sci. Math.* **2** (1893) 240-246

11. Orrick, W. and Solomon B., comp. The Hadamard Maximal Determinant Problem, 18 Jan. 2007. Mathematics Dept., Indiana U. 07 August 2008 <http://www.indiana.edu/∼maxdet/>

12. Rokicki, T. New Determinant Records From Circulant Pairs, 07 August 2008 <http://tomas.rokicki.com/newrec.html>

13. Wojtas, M. On Hadamard's Inequality for the Determinants of Order Non-Divisible by 4, *Colloq. Math.* **12** (1964) 73-83

Completeness results in syllogistic logic

SAM ZIEGLER
Carnegie Mellon University

INDIANA UNIVERSITY REU SUMMER 2008
Advisor: Larry Moss

I.1 Introduction

The syllogistic fragment we deal with first, $\mathcal{L}(\mathit{all}, \mathit{most})$, requires the following definitions:

"Syntax" We start with a set of *variables* X, Y, ..., representing plural common nouns. These variables are used in sentences of the following form:

$$\mathit{All}\ X\ \mathit{are}\ Y, \mathit{Most}\ X\ \mathit{are}\ Y$$

These are the only types of sentences allowed, and there is no recursion whatsoever.

Notation If Γ is a set of sentences in $\mathcal{L}(\mathit{all}, \mathit{most})$, we write Γ_{all} for the subset of Γ containing only sentences of the form *All* X *are* Y. We do this for *Most* as well, writing Γ_{most}.

A proof in $\mathcal{L}(\mathit{all}, \mathit{most})$ consists of a two-dimensional proof tree.

Definition I.1. A *proof tree over* Γ is a finite tree $\mathcal{T}$ whose nodes are labeled with sentences in our fragment, with the additional property that each node is either an element of Γ or comes from its parent(s) by an application of one of the rules. $\Gamma \vdash S$ means that there is a proof tree $\mathcal{T}$ for over Γ whose root is labeled S.

Example I.1.1. Suppose we wanted to see

$\{\mathit{All}\ X\ \mathit{are}\ Y, \mathit{All}\ Y\ \mathit{are}\ Z, \mathit{Most}\ Y\ \mathit{are}\ X\} \vdash \mathit{Most}\ Y\ \mathit{are}\ Z$:

$$\dfrac{\mathit{Most}\ Y\ \mathit{are}\ X \qquad \dfrac{\mathit{All}\ X\ \mathit{are}\ Y \quad \mathit{All}\ Y\ \mathit{are}\ Z}{\mathit{All}\ X\ \mathit{are}\ Z}}{\mathit{Most}\ Y\ \mathit{are}\ Z}$$

Semantics One starts with a set M, a subset $[\![X]\!] \subseteq M$ for each variable X. This gives a *model* $\mathcal{M} = (M, [\![\ ]\!])$. We then define

$$\begin{array}{lll} \mathcal{M} \models \mathit{All}\ X\ \mathit{are}\ Y & \text{iff} & [\![X]\!] \subseteq [\![Y]\!] \\ \mathcal{M} \models \mathit{Most}\ X\ \mathit{are}\ Y & \text{iff} & |[\![X]\!] \cap [\![Y]\!]| > \frac{1}{2}\,|[\![X]\!]| \end{array}$$

We allow $[\![X]\!]$ to be empty, and in this case, recall that $\mathcal{M} \models \mathit{All}\ X\ \mathit{are}\ Y$ vacuously. Also, note that *Most* requires *strictly* more than half of one set to intersect with the other. And if Γ is a finite or infinite set of sentences, then we write $\mathcal{M} \models \Gamma$ to mean that $\mathcal{M} \models S$ for all $S \in \Gamma$.

Semantic definition $\Gamma \models S$ means that every model which makes all sentences in the set Γ true also makes S true. This is the relevant form of semantic entailment for this paper.

I.2 (Partial) Completeness of $\mathcal{L}(\mathit{all}, \mathit{most})$

Before we move on to the completeness of our system, we note that it is easily sound.

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

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

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

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

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

Figure I.1: The logic of *All* and *Most*.

Proof. The soundness of this system follows easily by induction on proof trees. The base case involves verifying the soundness of the rules of inference themselves. For instance, the soundness of the last rule in Figure I.1 can be seen as follows (we omit the semantic brackets- the difference between variables and sets is understood): given $|Z \cap Y| > \frac{1}{2}|Z|$, $Y \subseteq X$, and $X \subseteq Z$, we see $Y \subseteq Z$, so $|Z \cap Y| = |Y| = |X \cap Y|$, and also $|X| \leq |Z|$, giving $|X \cap Y| = |Z \cap Y| > \frac{1}{2}|Z| \geq \frac{1}{2}|X|$. The rest of the rules are straighforward. $\square$

We can split the proof of completeness for $\mathcal{L}(\textit{all}, \textit{most})$ into two separate proofs: one in which we prove $\Gamma \models \textit{Most } X \textit{ are } Y \implies \Gamma \vdash \textit{Most } X \textit{ are } Y$, and a similar result for $\textit{All } X \textit{ are } Y$. The latter case is much easier, and we will examine it first. Note that throughout, we're assuming that Γ is a finite set, and that all the models we use are finite (but perhaps an arbitrarily large size).

Theorem I.2.1. *Let* $\Gamma \subseteq \mathcal{L}(\text{all}, \text{most})$*. Then if* $\Gamma \models \text{All } X \text{ are } Y$*, then* $\Gamma \vdash \text{ All } X \text{ are } Y$

Proof. In this case, we note the following simple result from [1]: the fragment $\mathcal{L}(\textit{all})$ is complete. That is, the fragment which consists only of *All* statements is complete, and indeed, our current system includes the old system's syntax, semantics, and rules of inference. Therefore, if we were to show that $\Gamma_{all} \models S$, where $S = \textit{All } X \textit{ are } Y$, then citing this result, $\Gamma_{all} \vdash S$. Our definition of proof clearly allows for expansion of hypotheses, i.e. $\Gamma \vdash S$. So it suffices to show $\Gamma_{all} \models S$.

To see this, let $\mathcal{M} = (M, [\![\ ]\!])$ be a model of Γ_{all}. The main idea of this proof is that, for any two finite sets A and B, we can choose a finite set of elements Q, disjoint to A and B, so that $|(A \cup Q) \cap (B \cup Q)| > \frac{1}{2}|(A \cup Q)|$. Consider the collection $\mathcal{C} = \{[\![U]\!] : U \in V(\Gamma)\}$, where $V(\Gamma)$ is collection of variables which appear in Γ. We can pick a set Q, disjoint from $\bigcup \mathcal{C}$, whose size is sufficiently large. Now we construct a new model $\mathcal{M}' = (M', [\![\ ]\!]_{\mathcal{M}'})$ where $M' = M \cup Q$, and for any U, $[\![U]\!]_{\mathcal{M}'} = [\![U]\!]_{\mathcal{M}} \cup Q$. Given that we picked Q to be sufficiently large, for any sentence $\textit{Most } U \textit{ are } W \in \Gamma$, $\mathcal{M}' \models \textit{Most } U \textit{ are } W$. It is also clear that if $\mathcal{M} \models \textit{All } X \textit{ are } Y$, then $\mathcal{M}' \models \textit{All } X \textit{ are } Y$. So $\mathcal{M}' \models \Gamma$. Thus, $\mathcal{M}' \models S$. But S is an *All* statement, so since $\mathcal{M}'$ interprets all variables with the same added subset, as compared to $\mathcal{M}$, S must have held in $\mathcal{M}$. So $\mathcal{M} \models S$, and we have shown that $\Gamma_{all} \models S$. $\square$

The other half of the completeness result has proved more elusive. We present here the proof of a couple special cases, hopefully towards the proof of completeness.

Theorem I.2.2. *Let* $\Gamma \subseteq \mathcal{L}(\text{all}, \text{most})$*,* S *be* $\text{Most } X \text{ are } Y$*. Given that* $\Gamma \vdash \text{All } X \text{ are } Y$*, then* $\Gamma \models S \implies \Gamma \vdash S$.

Proof. We prove the implication by contrapositive: so assume, along with $\Gamma \vdash$ *All* X *are* Y, that $\Gamma \nvdash S$. We wish to show $\Gamma \not\models S$. To do this, we need to find a model of Γ which falsifies S.

Consider the model $\mathcal{M} = (\{*\}, [\![\ ]\!])$, with $[\![X]\!] = \emptyset$, and if $\Gamma \vdash$ *All* Z *are* X then $[\![Z]\!] = \emptyset$. Otherwise, $[\![Z]\!] = M = \{*\}$. To see that $\mathcal{M} \models \Gamma$, note that the only way we could have $\mathcal{M} \not\models$ *Most* U *are* W is if one of $[\![U]\!]$,$[\![W]\!]$ is $\emptyset$. If $[\![U]\!] = \emptyset$, then $\Gamma \vdash$ *All* U *are* X. But given *Most* U *are* W, our rules give us *Most* U *are* U. But with *All* U *are* X, this gives *Most* U *are* X, hence *Most* X *are* X, hence, from our assumption that $\Gamma \vdash$ *All* X *are* Y, we would end up with $\Gamma \vdash$ *Most* X *are* Y, which contradicts our contrapositive assumption. The case is similar if $[\![W]\!] = \emptyset$. Thus, if $T =$ *Most* U *are* W is a *Most* statement in Γ, then $\mathcal{M} \models T$.

Suppose $P =$ *All* U *are* W is a sentence in Γ and $\mathcal{M} \not\models P$. This is only possible if $[\![U]\!] = M = \{*\}$ and $[\![W]\!] = \emptyset$. But then we must have had $\Gamma \vdash$ *All* W *are* X. But then this combined with P gives us *All* U *are* X, from which it follows that $[\![U]\!] = \emptyset$, a contradiction. So $\mathcal{M} \models P$.

So we have shown that $\mathcal{M} \models \Gamma$. But obviously $\mathcal{M} \not\models S$, since $[\![X]\!] = \emptyset$. So we have show $\Gamma \not\models S$. □

Here is the proof of another subcase. We may now assume that $\Gamma \nvdash$ *All* X *are* Y.

Theorem I.2.3. *Let* $\Gamma \subseteq \mathcal{L}(\text{all}, \text{most})$*,* S *be* Most X are Y*. Given that* $\Gamma \vdash$ All Y are X*, then* $\Gamma \models S \Longrightarrow \Gamma \vdash S$.

Proof. Again, we're going to take a contrapositive approach to this proof. So assume that $\Gamma \nvdash$ Most X are Y. So we're looking for a model $\mathcal{M}$ which satisfies Γ but falsifies S. Let $M = \{1, 2, \ldots, 7\}$, $A = \{1, 2, 3\}$, $B = \{1, 2, 3, 4\}$. The variable assignments are as follows: let $[\![X]\!] = M$, $[\![Y]\!] = A$. Now for all other Z we case three cases: if $\Gamma \vdash$ *All* X *are* Z, then let $[\![Z]\!] = M$. If $\Gamma \vdash$ *All* Z *are* Y, then let $[\![Z]\!] = A$. Note that we can't have both, because then we'd have $\Gamma \vdash$ *All* X *are* Y, which we assumed we didn't. If neither of these hold for Z, then let $[\![Z]\!] = B$.

Now we want to see that this model satisfies Γ. There are only three ways that our model could falsify *All* U *are* V, and they are all impossible, given that *All* U *are* V is in Γ. For if $[\![V]\!] = A$, then $\Gamma \vdash$ *All* V *are* Y, so that then $\Gamma \vdash$ *All* U *are* Y, so that by definition $[\![U]\!] = A$. If $[\![V]\!] = B$, then we'd be in trouble if $[\![U]\!] = M$; but this is impossible, since this would require $\Gamma \vdash$ *All* X *are* U, and hence $\Gamma \vdash$ *All* X *are* V, so that $[\![V]\!] = M$ as well. So this shows that $\mathcal{M}$ satifies any *All* sentence in Γ.

The only way for our model to falsify *Most* U *are* V would be to have $[\![U]\!] = M$ and $[\![V]\!] = A$. But if *Most* U *are* V is in Γ this is impossible. To see this: in order to have those assignments, we must have $\Gamma \vdash$ *All* X *are* U and $\Gamma \vdash$ *All* V *are* Y. Now we also are given that $\Gamma \vdash$ *All* Y *are* X. From these, we get the following proof tree:

$$\frac{\dfrac{\textit{All } X \textit{ are } U \quad \textit{Most } U \textit{ are } V \quad \dfrac{\textit{All } V \textit{ are } Y \quad \textit{All } Y \textit{ are } X}{\textit{All } V \textit{ are } X}}{\textit{Most } X \textit{ are } V} \qquad \textit{All } V \textit{ are } Y}{\textit{Most } X \textit{ are } Y}$$

Thus, we would have $\Gamma \vdash$ *Most* X *are* Y, which is contrary to our assumption. Thus, $\mathcal{M} \models \Gamma$. But clearly $\mathcal{M} \not\models S$. So therefore $\Gamma \not\models S$, and we're done. □

I.3 Verbs: A Hilbert System

In his paper [2], Prof. Moss describes a syllogistic fragment which includes sentences of the form *NP V NP*, where any occurence *NP* is noun phrase of the form *All X* or *Some X*, and *V* is some verb which takes a direct object. In this paper, I outline a complete Hilbert style system which includes such sentences. We will refer to this system as $\mathcal{H}(all, some, verbs)$.

Syntax We have variables X, Y, Z, etc. representing plural nouns. The basic sentences in this fragment are of the following forms: *All X are Y*, *Some X are Y*, and $Q_1X\ V\ Q_2Y$, where Q_1 and Q_2 can take the form *Some* or *All*. The verb V, for the purposes of this paper, will only take form *see* and *not see*, where *not see* will act as a complement to *see*. When mixing existential and universal phrases, ambiguity may arise: for instance, we could take "All students write some paper" to mean either that each student writes his or her own paper, or to mean that there is some paper which all students helped to write. For sentences with verbs, and in which $Q_1 \neq Q_2$, we will include a notion of scope: the scope of a basic sentence will tell us how to read it. We denote subject wide scope with an sws tag, and object wide scope with an ows tag. For example, $(\textit{All X see some Y})_{\mathsf{sws}}$ should be read as "For each $x \in X$, there is a $y \in Y$ such that x sees y", whereas the ows version would be read as "There is some $y \in Y$ such that, for all $x \in X$, x sees y."

The language of the system $\mathcal{H}(all, some, verbs)$ consists of the boolean combinations of these basic sentences, using the usual boolean connectives $\wedge$, $\vee$, and $\neg$. It is understood that $A \to B$ is just convenient shorthand for $\neg A \vee B$.

Now we introduce some more shorthand. We will identify the verb sentences with symbols of the form $\sigma_{i,U,W}$, sometimes dropping the U and W when they are clear from context. The associations are as follows:

$$\begin{array}{rr}
\textit{All U see all W} \Rightarrow \sigma_{1,U,W} & (\textit{Some U see all W})_{\mathsf{sws}} \Rightarrow \sigma_{2,U,W} \\
(\textit{All U see some W})_{\mathsf{ows}} \Rightarrow \sigma_{3,U,W} & (\textit{Some U see all W})_{\mathsf{ows}} \Rightarrow \sigma_{4,U,W} \\
(\textit{All U see some W})_{\mathsf{sws}} \Rightarrow \sigma_{5,U,W} & \textit{Some U see some W} \Rightarrow \sigma_{6,U,W}
\end{array}$$

For the corresponding sentences with "not see", we put a bar over the σ symbol. For example, $\overline{\sigma}_{1,U,W}$ would stand for *All U not see all W*. Given this, we use a mneumonic for the negation of the σ's which arises from their natural interpretation: $\neg\sigma_1 \equiv \overline{\sigma}_6$, $\neg\sigma_2 \equiv \overline{\sigma}_5$, $\neg\sigma_3 \equiv \overline{\sigma}_4$, and then three more in which the roles of the subscripts are switched.

Just like any Hilbert style system, $\mathcal{H}(all, some, verbs)$ has axioms. As the reader may expect, there are a lot of axioms for this system. A full list is contained in the appendix. Here are a few examples:

$$\begin{array}{c}
\textit{All X are Y} \wedge \textit{All Y are Z} \to \textit{All X are Z} \\
(\textit{Some X see all Y})_{\mathsf{sws}} \to (\textit{Some X see all Y})_{\mathsf{ows}} \\
\textit{Some X are Y} \to \textit{Some Y are X}
\end{array}$$

Note that the system $\mathcal{L}(all, most)$ had many rules of inference and no axioms (one rule of inference happened to have no antecedents). The system $\mathcal{H}(all, some, verbs)$, on the other hand, has many axioms and just one rule of inference, and that is **Modus Ponens**. It states that, given A and $A \to B$, we can infer B. A *proof* from hypotheses Γ in the system $\mathcal{H}(all, some, verbs)$ would be a finite list of boolean formulas, say $\{\Delta_1, \Delta_2, \ldots, \Delta_n\}$, such that for all i, Δ_i is an axiom, $\Delta_i \in \Gamma$, or Δ_i is the result of an application of modus ponens on two formulas from the set $\{\Delta_1, \Delta_2, \ldots, \Delta_{i-1}\}$. We say that $\Gamma \vdash S$ if S is the last line in some proof from Γ. We can now define, for any variables U and V, $Th_\Gamma(U,V) := \{\pm\sigma_{i,U,V} : \Gamma \vdash \pm\sigma_{i,U,V}\}$, where $\pm\sigma_i$ can stand for σ_i or $\neg\sigma_i$.

Semantic notions As with the fragment $\mathcal{L}(all, most)$, the semantics is based on models. Given a model $\mathcal{M} = (M, [\![\ ]\!])$, we have $[\![see]\!] \subseteq M \times M$ is a binary relation on M, and we define $[\![not\ see]\!] = M \times M \backslash [\![see]\!]$. We can then define the following:

$$
\begin{array}{lll}
\mathcal{M} \models \textit{All X are Y} & \text{iff} & [\![X]\!] \subseteq [\![Y]\!] \\
\mathcal{M} \models \textit{Some X are Y} & \text{iff} & |[\![X]\!] \cap [\![Y]\!]| \neq \emptyset \\
\mathcal{M} \models \sigma_{1,X,Y} & \text{iff} & (\forall x \in [\![X]\!] \forall y \in [\![Y]\!])\ x[\![see]\!]y \\
\mathcal{M} \models \sigma_{2,X,Y} & \text{iff} & (\exists x \in [\![X]\!] \forall y \in [\![Y]\!])\ x[\![see]\!]y \\
\mathcal{M} \models \sigma_{3,X,Y} & \text{iff} & (\exists y \in [\![Y]\!] \forall x \in [\![X]\!])\ x[\![see]\!]y \\
\mathcal{M} \models \sigma_{4,X,Y} & \text{iff} & (\forall y \in [\![Y]\!] \exists x \in [\![X]\!])\ x[\![see]\!]y \\
\mathcal{M} \models \sigma_{5,X,Y} & \text{iff} & (\forall x \in [\![X]\!] \exists y \in [\![Y]\!])\ x[\![see]\!]y \\
\mathcal{M} \models \sigma_{6,X,Y} & \text{iff} & (\exists x \in [\![X]\!] \exists y \in [\![Y]\!])\ x[\![see]\!]y
\end{array}
$$

I.3.1 Basic Observations

From here forward, by **basic sentence** we mean any *All*, *Some*, or *Verb* sentence, or a negation of a sentence. It is the same idea as a "literal" in boolean algebra.

The following observations are key to the completeness of this fragment. First note that, since our system has negation, the proof of completeness is equivalent to a proof that any **consistent** set is satisfiable. A set is consistent if there is some formula which is doesn't prove. We may also invoke the **Lindenbaum lemma** (see [4]), which states that for any consistent set Δ, there exists a consistent, set Θ such that $\Delta \subseteq \Theta$ and, for every formula φ, either $\varphi \in \Theta$ or $\neg\varphi \in \Theta$. We call such a set Θ **complete**. So it suffices to show that every *complete* consistent set in $\mathcal{H}(all, some, verbs)$ is satisfiable.

We need one more observation: that is that any boolean combination has an equivalent boolean combination is disjunctive normal form. This is a standard result, so it will be assumed here.

Now for any complete consistent set Δ, let $\Gamma := \{S : S \in \Delta \text{ and } S \text{ basic}\}$.

Lemma I.3.1. $\Gamma \models \Delta$

Proof. Let $\mathcal{M} \models \Gamma$, and say $\varphi \in \Delta$. We know that φ has a disjunctive normal form, so let φ' be in DNF and $\varphi \equiv \varphi'$. To show $\mathcal{M} \models \varphi'$, we just need to see that one of it's disjuncts holds. Well, since Δ is complete and consistent, at least one disjunct, say ξ, is in Δ. So ξ is a conjunction of basic sentences. Thus each conjunct of ξ is in Δ, else Δ be inconsistent, and so furthermore, each conjunct of ξ is in Γ by definition. So $\mathcal{M} \models \xi$, and so therefore $\mathcal{M} \models \varphi'$. So $\mathcal{M} \models \Delta$. □

So finally, given a complete consistent set, we need only find a model for the basic sentences therein. Note that this set is also consistent, and complete in terms of basic sentences.

I.3.2 Completeness of $\mathcal{H}(\boldsymbol{all}, \boldsymbol{some}, \boldsymbol{verbs})$

Let Γ be our complete consistent set of basic sentences. We construct a model as follows.

Let $M = \{U_1, U_2, U_3 : \Gamma \vdash \exists U\} \cup \{\{A, B\} : \Gamma \vdash \textit{Some A are B}\}$. So we take three copies of each variable for which Γ proves the existence, and some sort of representative to satisfy *Some* sentences.

The semantics are as follows:

$W_i \in [\![U]\!]$	iff $\Gamma \vdash$ *All W are U*
$\{A, B\} \in [\![U]\!]$	iff $\Gamma \vdash$ *All A are U*, or $\Gamma \vdash$ *All B are U*
$U_i[\![V]\!]W_j$	iff $\Gamma \vdash \exists U, \Gamma \vdash \exists V, (U_i, W_j) \in R_{U,W,Th_\Gamma(U,W)}$
$\{U, Z\}[\![V]\!]W_2$	iff $\Gamma \vdash \sigma_{1,U,W}$ or $\Gamma \vdash \sigma_{1,Z,W}$
$\{U, Z\}[\![V]\!]W_1$	iff $\{U, Z\}[\![V]\!]W_2$, or $\Gamma \vdash \sigma_{3,U,W}$, or $\Gamma \vdash \sigma_{3,Z,W}$
$\{U, Z\}[\![V]\!]W_3$	iff $\{U, Z\}[\![V]\!]W_2$, or $\Gamma \vdash \sigma_{5,U,W}$ or $\Gamma \vdash \sigma_{5,Z,W}$
$U_2[\![V]\!]\{W, Z\}$	iff $\Gamma \vdash \sigma_{1,U,W}$ or $\Gamma \vdash \sigma_{1,U,Z}$
$U_1[\![V]\!]\{W, Z\}$	iff $U_2[\![V]\!]\{W, Z\}$, or $\Gamma \vdash \sigma_{2,U,W}$, or $\Gamma \vdash \sigma_{2,U,Z}$
$U_3[\![V]\!]\{W, Z\}$	iff $U_2[\![V]\!]\{W, Z\}$, or $\Gamma \vdash \sigma_{4,U,W}$, or $\Gamma \vdash \sigma_{4,Z,W}$
$\{A, B\}[\![V]\!]\{C, D\}$	iff $\Gamma \vdash \sigma_{1,A,C}$, or $\Gamma \vdash \sigma_{1,A,D}$, or $\Gamma \vdash \sigma_{1,B,C}$, or $\Gamma \vdash \sigma_{1,B,D}$

where $R_{U,W,Th_\Gamma(U,W)}$ refers to the subset of $\{U_1, U_2, U_3\} \times \{W_1, W_2, W_3\}$ determined by the set $Th_\Gamma(U, W)$ and its corresponding diagram in Figure I.2.

This is a full description of our model. Now we must see that our model satisfies the complete consistent set of basic sentences Γ.

Suppose *All X are Y* is in Γ. Then it is a simple monotonicity point that $\mathcal{M} \models$ *All X are Y*. If $Z_i \in [\![X]\!]$, then $\Gamma \vdash$ *All Z are X*. So using the axiom for transitivity of *All*, and modus ponens, we get $\Gamma \vdash$ *All Z are Y*, hence $Z_i \in [\![Y]\!]$. It is a similar point if $\{A, B\} \in [\![X]\!]$.

If $\neg$(*All X are Y*) is in Γ, we need to show that $[\![X]\!] \backslash [\![Y]\!] \neq \emptyset$. Axiom 15 yields that $\Gamma \vdash$ *Some X are X*, so we will have $[\![X]\!] \neq \emptyset$. Furthermore, we will have $X_1 \in [\![X]\!]$ but $X_1 \notin [\![Y]\!]$, because if it was, then $\Gamma \vdash$ *All X are Y*, and since Γ is complete, anything it proves is indeed already inside Γ. So then Γ would be inconsistent. So we conclude $[\![X]\!] \backslash [\![Y]\!] \neq \emptyset$.

It's easy to see that if *Some X are Y* is in Γ, then it will be satisfied by our model. This is due to the elements $\{A, B\}$. If $\neg$(*Some X are Y*) is in Γ, we must show $[\![X]\!] \cap [\![Y]\!] = \emptyset$. If $\Gamma \vdash$ *All Z are X* and $\Gamma \vdash$ *All Z are Y*, then we can't have $\Gamma \vdash$ *Some Z are Z*, otherwise we would be able to prove *Some X are Y*, which would contradict the consistency of Γ. Without *Some Z are Z*, $Z_i \notin M$, so Z won't add any common elements to $[\![X]\!]$ and $[\![Y]\!]$. In a similar way we can see that $\{A, B\} \in [\![X]\!]$ and $\{A, B\} \in [\![Y]\!]$ is impossible.

So we have seen that any *All* or *Some* basic sentence in Γ is satisfied by $\mathcal{M}$.

We want to show that, for any variables X and Y, sentences of the form $\pm\sigma_{i,X,Y} \in \Gamma$ are satisfied. Here, we are going to consider four cases.

Suppose *Some X are X* and $\neg$(*Some Y are Y*) are in Γ. Then from the axioms outlined in axiom scheme 16 (see appendix), we get that Γ proves $\sigma_1, \sigma_2, \neg\sigma_3, \sigma_4, \neg\sigma_5, \neg\sigma_6$ where each σ is understood to carry the subscripts $\{X, Y\}$ as well. Thus these are all in Γ, and they account for all the $\pm\sigma_{i,X,Y}$ in Γ by consistency. To see that our model satisfies all of them, note that since $\neg$(*Some Y are Y*) is in Γ, we will have $[\![Y]\!] = \emptyset$. Then it is easy to see that by the definition of our semantics, and the mneumonics $\neg\sigma_{3,X,Y} \equiv \overline{\sigma}_{4,X,Y} = (\textit{Some X not see all Y})_{\mathsf{ows}}$, that an empty $[\![Y]\!]$ and a nonempty $[\![X]\!]$ will satisfy this sentence: it is vacuously true. This also shows the soundness of the axioms outlined in axiom scheme 16.

The cases in which $\neg$(*Some X are X*) and *Some Y are Y* are in Γ and $\neg$(*Some X are X*) and $\neg$(*Some Y are Y*) are in Γ are examined similarly, and the reader is encouraged to consider the axiom schemes 17 and 18 to see that, in these cases, $\mathcal{M} \models \pm\sigma_{i,X,Y}$.

Now the only case we have left to consider is that both *Some X are X* and *Some Y are Y* are in Γ. This involves even more casework. Suppose that $\sigma_{1,X,Y} \in \Gamma$. Then from $\Gamma \vdash \exists X$ and $\Gamma \vdash \exists Y$ we see that $\Gamma \vdash \sigma_{i,X,Y}$ for each i (see axioms 2-5), which completely classifies the $\pm\sigma_{i,X,Y}$ in Γ.

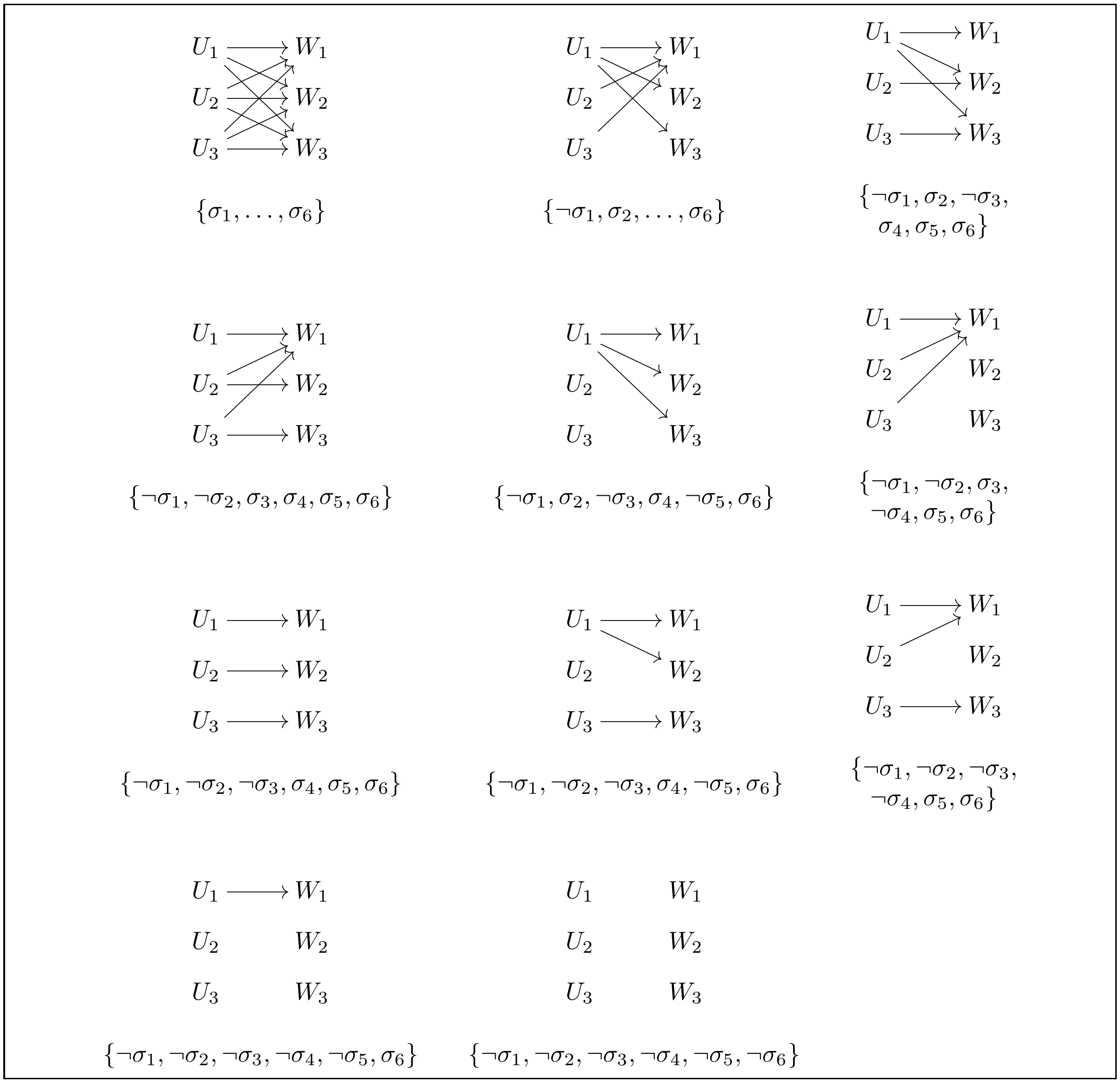


Figure I.2: These are the possible complete sets (given $\exists U$ and $\exists W$) and the definition of special relations used in the completeness proof.

Now suppose that $\neg\sigma_{6,X,Y} \in \Gamma$. We noted in the mneumonics that this is equivalent to $\overline{\sigma}_{1,X,Y}$. So, similar to above, it follows that $\overline{\sigma}_{i,X,Y}$ is in Γ for every i, since we have *Some* X *are* X and *Some* Y *are* Y. This is the same as saying $\neg\sigma_{i,X,Y} \in \Gamma$ for every i.

So these are two possible combinations of $\pm\sigma_{i,X,Y}$'s under the big case that *Some* X *are* X and *Some* Y *are* Y are in Γ. Now, with axioms 2-5, and $\exists X$, $\exists Y$, we have that σ_2 will prove σ_4, σ_3 will prove σ_5, and that all of these will prove σ_6. The same is true of the $\overline{\sigma}$'s, since these have precisely the same structure, but a different (complimentary) verb. I claim that in our consistent, complete Γ, there are nine other combinations of $\pm\sigma_{i,X,Y}$'s which will work under this last, big case. All of them include $\neg\sigma_{1,X,Y}$ and $\sigma_{6,X,Y}$, since the negation of either one would point us to two the combinations previously considered. All nine other combinations appear in Figure I.2, below a diagram.

These are the only combinations of the $\pm\sigma_{i,X,Y}$'s we can have given *Some* X *are* X and *Some* Y *are* Y. We can check that the model satisfies them case by case. I present here some important cases; note that, given non-empty assignments (which we do in this case), if a model satisfies σ_2, then it will satisify σ_4; likewise, if it satisfies σ_3, it will satisfy σ_5. Negate all the σ's and reverse the statements to obtain two more truths. Thus, we can save time while checking.

Case 1: If we have $+\sigma_{i,X,Y}$ for each i, then by inspection of the diagram, it is clear that $X_i[\![\textit{see}]\!]Y_j$ for all i and j. If $\{A, B\} \in [\![X]\!]$ and $\{C, D\} \in [\![Y]\!]$ then from the monotonicity axioms, we will get something like $\Gamma \vdash \sigma_{1,X,C}$ and $\Gamma \vdash \sigma_{1,A,Y}$, as well as $\Gamma \vdash \sigma_{1,A,C}$, where the A and C could have been B or D resp., and it would work all the same. So $\sigma_{1,X,Y}$ is satisfied, and since $[\![X]\!]$ and $[\![Y]\!]$ are non-empty, it is clear that the rest of the σ's will be satisfied as well.

Case 2: In this case, we have $\{\neg\sigma_1, \sigma_2, \neg\sigma_3, \sigma_4, \sigma_5, \sigma_6\}$. It is clear to see that $\neg\sigma_{1,X,Y}$ is satisfied by our model, as in the associated picture for this set, $X_3[\![\textit{not see}]\!]Y_3$. For σ_2, clearly $X_1[\![\textit{see}]\!]W_j$ for every $W_j \in [\![Y]\!]$. If $\{C, D\} \in [\![Y]\!]$, then it is a monotonicity point to see $\Gamma \vdash \sigma_{2,X,C}$, WLOG, since we have $\sigma_{2,X,Y}$. Therefore $X_1[\![\textit{see}]\!]\{C, D\}$. So σ_2 is satisfied. For $\neg\sigma_{3,X,Y} \equiv \overline{\sigma}_{4,X,Y}$, suppose that $W_j \in [\![Y]\!]$. So $\Gamma \vdash$ *All* W *are* Y. Suppose $j = 1$. Then if $X_i[\![\textit{see}]\!]W_1$ for all i, we must have $\Gamma \vdash \beta_{3,X,W}$, which gives $\sigma_{3,X,Y}$, which isn't possible. If $j = 2$ and all X_i saw W_j, the only way this is possible is if $\Gamma \vdash \sigma_{1,X,W}$, which from $\exists W$ lead tos $\sigma_{3,X,W}$, then $\sigma_{3,X,Y}$. If $j = 3$, then similarly $\Gamma \vdash \sigma_{1,X,W}$. Thus, W_j is not seen by some X_i.

As for $\{C, D\} \in [\![Y]\!]$, we will have that $X_2[\![\textit{not see}]\!]\{C, D\}$, for otherwise we would have something like $\Gamma \vdash \sigma_{1,X,C}$, and with $\Gamma \vdash$ *All* D *are* Y, and the fact that $\Gamma \vdash$ *Some* C *are* D, with axiom 3, we see that $\Gamma \vdash \sigma_{3,X,D}$ and monotonicity would give $\Gamma \vdash \sigma_{3,X,Y}$, which is a contradiction. So indeed, each element of $[\![Y]\!]$ is "not seen" by something in $[\![X]\!]$. As for σ_4, it is clear from the diagram that each Y_j is seen by some element of $[\![X]\!]$. For $\{C, D\}$, since we have $\sigma_{4,X,Y}$, we can easily get $\sigma_{4,X,C}$, or possibly with D instead of C, either of which would give $X_3[\![\textit{see}]\!]\{C, D\}$. So $\sigma_{4,X,Y}$ will be satisfied by our model. With $\sigma_{5,X,Y}$, again the only issue might be $\{A, B\} \in [\![X]\!]$, but monotonicity will give us $\{A, B\}[\![\textit{see}]\!]Y_3$ pretty easily. So indeed the model will satisfy $\sigma_{5,X,Y}$. It is trivial to see that the model satisfies $\sigma_{6,X,Y}$. So in this case, our model $\mathcal{M}$ satisfies all such $\pm\sigma_{i,X,Y}$.

Case 3: It would be instructive to now consider the case in which we have $\neg\sigma_{i,X,Y}$ for all i except, of course, σ_6. The idea that $\neg\sigma_3$ is satisfied is similar to the idea above, where we considered it in case 2. In fact, as noted above, showing that $\neg\sigma_{5,X,Y}$ holds implies that $\neg\sigma_{3,X,Y}$ holds anyway. To see that $\neg\sigma_{5,X,Y}$ holds, I claim that X_2 is a witness for $\overline{\sigma}_{2,X,Y}$. Suppose we had $X_2[\![\textit{see}]\!]W_j$ for *All* W *are* Y. If $j = 1$ then from the diagrams we can conclude that $\sigma_{5,X,W}$ is provable, from which monotonicity would give us

$\Gamma \vdash \sigma_{5,X,Y}$ which is impossible. If $j = 2$, we also get σ_5, which is impossible. And if $j = 3$, then the only way this is possible is if $\Gamma \vdash \sigma_{1,X,W}$, which will clearly give us $\sigma_{5,X,Y}$. So indeed $X_2[\![\textit{not see}]\!]W_j$.

If $\{C, D\} \in [\![Y]\!]$, then we will have $X_2[\![\textit{not see}]\!]\{C, D\}$, because otherwise we would have, say, $\sigma_{1,X,C}$, and with the existence of X and C and monotonicity (with possibly using *Some C are D*, we would arrive at $\sigma_{3,X,Y}$, which is impossible. So $\neg\sigma_{5,X,Y}$ is definitely satisfied. To see $\neg\sigma_{4,X,Y}$ would be satisfied, we show that Y_2 is a witness for this rule. Clearly from the diagram $X_i[\![\textit{not see}]\!]Y_2$ for any i. If $\{A, B\} \in [\![X]\!]$, then it will not "see" Y_2. For if it did, then WLOG $\Gamma \vdash \sigma_{1,A,Y}$, which with $\exists A$, $\exists Y$ gives $\sigma_{4,A,Y}$ (we might need to turn that A into B, which is possible from *Some A are B*), which by monotonicity leads to $\sigma_{4,X,Y}$, which is impossible. So $\neg\sigma_{4,X,Y}$ holds, and therefore so does $\neg\sigma_{2,X,Y}$. Since $\neg\sigma_{1,X,Y}$ and $\sigma_{6,X,Y}$ are obvious, this case is taken care of.

Case 4: Now let's consider the case in which we have $\neg\sigma_{1,X,Y}$ and $\neg\sigma_{2,X,Y}$, with the rest of the $\pm\sigma_{i,X,Y}$ being positive. To see that $\neg\sigma_{2,X,Y}$ is satisified, say $Z_i \in [\![X]\!]$. If it was the case that $Z_i[\![\textit{see}]\!]Y_j$ for each j: well if $i = 1$, then we must have $\sigma_{2,Z,Y}$ from the diagrams, which by monotonicity gives $\sigma_{2,X,Y}$, which is impossible. If $j = 2, 3$ then the diagrams tell us the only way this is possible is if $\Gamma \vdash \sigma_{1,Z,Y}$, from which we get $\sigma_{2,Z,Y}$ and so $\sigma_{2,X,Y}$. So there is some j such that $Z_i[\![\textit{not see}]\!]Y_j$. If $\{A, B\} \in [\![X]\!]$. Then if $\{A, B\}[\![\textit{see}]\!]Y_2$ we would have WLOG $\Gamma \vdash \sigma_{1,A,Y}$, which from which can easily obtain $\sigma_{2,X,Y}$. So clearly $\{A, B\}[\![\textit{not see}]\!]Y_2$, and so our model satisfies $\neg\sigma_{2,X,Y}$.

Now we will see that our model satisfies $\sigma_{3,X,Y}$. I claim our witness is Y_1. Say $Z_i \in [\![X]\!]$. Suppose $Z_i[\![\textit{not see}]\!]Y_1$. If $i = 1$, then we would have to have $\neg\sigma_{6,Z,Y}$ by the diagrams, which would give $\neg\sigma_{3,Z,Y}$, from which it follows $\neg\sigma_{3,X,Y}$ which is bad. Likewise, if $i = 2$ we get by inspection that Γ proves $\neg\sigma_{3,Z,Y}$, and again $\neg\sigma_{3,X,Y}$. The same goes for $i = 3$. So this shows that $Z_i[\![\textit{see}]\!]Y_1$ for each i. If $\{A, B\} \in [\![X]\!]$, then since $\Gamma \vdash \sigma_{3,X,Y}$, it is a simple monotonicity point to see that $\{A, B\}[\![\textit{in}]\!]Y_1$. So $\mathcal{M} \models \sigma_{3,X,Y}$.

In previous cases, we have examined the modelling of σ_4 and σ_5, and I claim that here there is no difference. As of now, we have seen all possible $\pm\sigma_{i,X,Y}$ verified in some context or another, and in any other of the 11 cases, the verification would be similar. We therefore leave it to the reader to confirm the rest of the cases, but we hope that this casework has been sufficiently convincing.

The preceding discussion is the essence of the proof of the following theorem:

Theorem I.3.1. *The system* $\mathcal{H}(\text{all}, \text{some}, \text{verbs})$ *is complete.*

I.4 A system for *in*

The third and final system I present in this paper is a syllogistic fragment with a transitive preposition. If A is in B, and B is in C, then we would say that A is in C. The idea for making a logic based on such a preposition came from the paper by Zwarts and Winter [5]. We denote this fragment by $\mathcal{L}(\textit{all}, \textit{some}, \textit{in})$.

Syntax As with the last two systems, we use a countable number of variables, X, Y, etc. The sentences of this fragment are of the following form:

$$\begin{array}{l} \textit{All } X \textit{ are } Y \\ \textit{Some } X \textit{ are } Y \\ \textit{All } X \textit{ are in all } Y \Longrightarrow \beta_{1,X,Y} \\ \textit{All } X \textit{ are in some } Y \Longrightarrow \beta_{2,X,Y} \\ \textit{Some } X \textit{ are in all } Y \Longrightarrow \beta_{3,X,Y} \\ \textit{Some } X \textit{ are in some } Y \Longrightarrow \beta_{4,X,Y} \end{array}$$

where the β's will serve as a notational short cut. Unlike in the verbs fragment, all sentences with *in* will be assumed to be subject wide scope. So *All X are in some Y* should be read as $\forall x \exists y$ such that x is in y.

Proofs in this system are the same as in the system $\mathcal{L}(\textit{all}, \textit{most})$, proof trees. The rules of inference for this system are listed on a separate page in the appendix. As a reminder, the root is labeled with the sentence to be proven, and the leaves are labeled with elements of some set of hypotheses Γ. If there is a proof tree with root S and leaves from Γ, then we write $\Gamma \vdash S$.

We make a similar to definition as in verbs: we define $Th_\Gamma(X,Y)$ to be the set $\{\beta_{i,X,Y} : \Gamma \vdash \beta_{i,X,Y}\}$. We define the **downward closure** of $Th_\Gamma(X,Y)$ to be $\{\beta_{i,X,Y} : \Gamma \cup \{\exists X, \exists Y\} \ \vdash \beta_{i,X,Y}\}$, that is the set of *in* sentences provable from Γ plus *Some X are X* and *Some Y are Y*. We denote this set by $\downarrow Th_\Gamma(X,Y)$. For example, if $\beta_1 \in Th_\Gamma(X,Y)$, we must have $\downarrow Th_\Gamma(X,Y) = \{\beta_1, \beta_2, \beta_3, \beta_4\}$.

Semantics As with the above two systems, in $\mathcal{L}(\textit{all}, \textit{some}, \textit{in})$ we will use a model based semantics. We interpret variables as being subsets of some "universe" (set), and we interpret *in* as $[\![in]\!]$, some *transitive* binary relation on $M \times M$, where M is our universe. So for $\mathcal{M}$ to be a model, if $x[\![in]\!]y$ and $y[\![in]\!]z$, then $x[\![in]\!]z$. We have already seen how *All* and *Most* sentences should be interpreted. As for *in* sentences, using the shorthand:

$$\begin{array}{lll} \mathcal{M} \models \beta_{1,X,Y} & \text{iff} & (\forall x \in [\![X]\!] \ \forall y \in [\![Y]\!]) \ x[\![in]\!]y \\ \mathcal{M} \models \beta_{2,X,Y} & \text{iff} & (\forall x \in [\![X]\!] \ \exists y \in [\![Y]\!]) \ x[\![in]\!]y \\ \mathcal{M} \models \beta_{3,X,Y} & \text{iff} & (\exists x \in [\![X]\!] \ \forall y \in [\![Y]\!]) \ x[\![in]\!]y \\ \mathcal{M} \models \beta_{4,X,Y} & \text{iff} & (\exists x \in [\![X]\!] \ \exists y \in [\![Y]\!]) \ x[\![in]\!]y \end{array}$$

I.4.1 Completeness of $\mathcal{L}(\textit{all}, \textit{some}, \textit{in})$

To prove the completeness of this fragment, the plan is to exhibit a model $\mathcal{M}$, in the spirit of the model for $\mathcal{H}(\textit{all}, \textit{some}, \textit{verbs})$, which depends on our set $\Gamma \subseteq \mathcal{L}(\textit{all}, \textit{some}, \textit{in})$, which will satisfy Γ. Then, given that $\Gamma \models S$, we would have $\mathcal{M} \models S$, and the nature of the model will allow us to conclude that $\Gamma \vdash S$.

Suppose we have a set $\Gamma \subseteq \mathcal{L}(\textit{all}, \textit{some}, \textit{in})$. Let $\mathcal{E}(\Gamma)$ be the set $\{S : \Gamma \vdash \textit{Some } X \textit{ are } X\}$ We will call a set of sentences $\mathcal{S}$ **closed** if $\mathcal{E}(\Gamma) \subseteq \mathcal{S}$, and if $\Gamma \vdash \textit{All } X \textit{ are in some } Y$, where $X \in \mathcal{S}$, then $Y \in \mathcal{S}$.

So given Γ and a closed set $\mathcal{S}$, we will define a model $\mathcal{M}(\Gamma, \mathcal{S})$ as follows: we let $M = \{U_1, U_2 : U \in \mathcal{S}\} \cup \{\{A, B\} : \Gamma \vdash \textit{Some } A \textit{ are } B\}$, i.e. two copies of every variable in $\mathcal{S}$, and a representative for sentences of the form *Some A are B*. The sets $R_{U,W,s} \subseteq \{U_1, U_2\} \times \{W_1, W_2\}$ come from the arrow diagrams, where the diagram is chosen based on s. The model assignments are as follows:

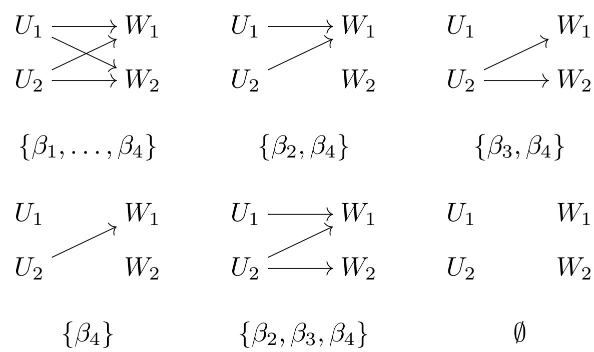


Figure I.3: These are the arrow diagrams for the model $\mathcal{M}(\Gamma, \mathcal{S})$. The $R_{U,W,s}$ in the model definition refer to the $[\![in]\!]$ relationship specified by the arrow diagram corresponding to the downward closed set s.

$W_i \in [\![U]\!]$	iff $\Gamma \vdash$ *All W are U*
$\{A, B\} \in [\![U]\!]$	iff $\Gamma \vdash$ *All A are U*, or $\Gamma \vdash$ *All B are U*
$U_i[\![in]\!]W_j$	iff $(U_i, W_j) \in R_{U,W,\downarrow Th_\Gamma(U,W)}$
$\{A, B\}[\![in]\!]W_2$	iff $\Gamma \vdash \beta_{1,A,W}$ or $\Gamma \vdash \beta_{1,B,W}$
$\{A, B\}[\![in]\!]W_1$	iff $\{A, B\}[\![in]\!]W_2$, or $\Gamma \vdash \beta_{2,A,W}$, or $\Gamma \vdash \beta_{2,B,W}$
$U_1[\![in]\!]\{C, D\}$	iff $\Gamma \vdash \beta_{1,U,C}$ or $\Gamma \vdash \beta_{1,U,D}$
$U_2[\![in]\!]\{C, D\}$	iff $U_1[\![in]\!]\{C, D\}$, or $\Gamma \vdash \beta_{3,U,C}$, or $\Gamma \vdash \beta_{3,U,D}$
$\{A, B\}[\![in]\!]\{C, D\}$	iff $\Gamma \vdash \beta_{1,A,C}$, or $\Gamma \vdash \beta_{1,A,D}$, or $\Gamma \vdash \beta_{1,B,C}$, or $\Gamma \vdash \beta_{1,B,D}$

Before we attempt to show that this model will satisfy Γ, we must first show that this model satisfies the transitivity requirement we place on models in $\mathcal{L}(\mathit{all}, \mathit{some}, \mathit{in})$.

Lemma I.4.1 (Transitivity). For any $\Gamma \subseteq \mathcal{L}(\mathit{all}, \mathit{some}, \mathit{in})$ and any set $\mathcal{S}$ which is closed with respect to Γ, $[\![in]\!]_{\mathcal{M}(\Gamma,\mathcal{S})}$ is transitive.

Proof. Due to the fact that both elements of the form Z_i and $\{A, B\}$ appear in our model, this proof has many cases. We will consider some of them here, so suffice it to say that the rest are fairly similar. So consider this a sketch.

Case: $Z_i[\![in]\!]W_k$ and $W_k[\![in]\!]V_j$. We want to see that $Z_i[\![in]\!]V_j$. So we just take subcases over the indices i, j, and k.

Suppose $i = j = k = 1$. By the construction of the model, and by looking at the diagrams, $Z_1[\![in]\!]W_1$ only if $\beta_{2,Z,W} \in\downarrow Th_\Gamma(Z, W)$. Similarly, $\beta_{2,W,V} \in\downarrow Th_\Gamma(W, V)$. Now, there are several possibilities. If $\Gamma \vdash \beta_{2,Z,W}, \beta_{2,W,V}$, then we could use the $\forall\exists - \forall\exists$ rule from the figure, that is the rule in the diagram whose antecedents are $\forall\exists - \forall\exists$, to conclude $\Gamma \vdash \beta_{2,Z,V}$. So then $\beta_{2,Z,V} \in\downarrow Th_\Gamma(Z, V)$ so by the model definition $Z_1[\![in]\!]V_1$.

We might also have that $\Gamma \vdash \beta_{1,Z,W}, \beta_{1,W,V}$. Then we would simply invoke the $\forall\forall - \forall\forall$ rule to get $\Gamma \vdash \beta_{1,Z,V}$, which by model definition gives $Z_1[\![in]\!]V_1$. Otherwise we could have either $\Gamma \vdash \beta_{1,Z,W}, \beta_{2,W,V}$ or $\Gamma \vdash \beta_{2,Z,W}, \beta_{1,W,V}$, in which case we invoke the $\forall\forall - \forall\exists$ rule and $\forall\exists - \forall\forall$ rule, respectively to arrive at similar conclusions. So in the case that $i = j = k = 1$, we have shown $[\![in]\!]$ is transitive.

Now suppose $i = k = 2$ and $j = 1$. So by looking at the diagrams we can conclude that $\beta_{3,Z,W} \in\downarrow Th_\Gamma(Z,W)$ and $\beta_{4,W,V} \in\downarrow Th_\Gamma(W,V)$. So there are many possibilities. Say $\Gamma \vdash \beta_{3,Z,W}$ and $\Gamma \vdash \beta_{4,W,V}$. Then using the $\exists\forall - \exists\exists$ rule we get $\Gamma \vdash \beta_{4,Z,V}$. Therefore $\beta_{4,Z,V} \in\downarrow Th_\Gamma(Z,V)$, so by our model construction and inspection of the diagrams, $Z_2[\![in]\!]V_1$, as required. To see that other possibilities work out, just note that either $\Gamma \vdash \beta_{3,Z,W}$ or $\Gamma \vdash \beta_{1,Z,W}$, and no matter which β_i we can show between W and V, each of $\beta_{3,Z,W}$ and $\beta_{1,Z,W}$ can be used to deduce that $\beta_{4,Z,V} \in\downarrow Th_\Gamma(Z,V)$.

The subcases in which we consider different values of i, j, k proceed in a similar fashion. Let us consider a different case: suppose $\{A,B\}[\![in]\!]\{C,D\}$, and $\{C,D\}[\![in]\!]\{E,F\}$. Then looking at our model construction, there are several β_1's which Γ might prove. Suppose we had $\Gamma \vdash \beta_{1,A,C}$ and $\Gamma \vdash \beta_{1,C,E}$. Then we can easily apply the $\forall\forall - \forall\forall$ rule to get $\Gamma \vdash \beta_{1,A,E}$ and so $\{A,B\}[\![in]\!]\{E,F\}$. However, we might only have $\Gamma \vdash \beta_{1,A,C}$ and $\Gamma \vdash \beta_{1,D,E}$. Now we cannot use a single application of a rule to get the result, but use the fact that $\Gamma \vdash$ *Some C are D* to see that $\Gamma \vdash \beta_{2,A,D}$. Now we can use the $\forall\exists - \forall\forall$ rule to get $\Gamma \vdash \beta_{1,A,E}$. So we still have $\{A,B\}[\![in]\!]\{E,F\}$. These two possibilites are representative of any β_1 situation we could have, thus we can safely conclude that transitivity holds.

In this way, we can proceed with the casework, the remainder of which is quite similar to the above. $\square$

Now that we've see that $\mathcal{M}(\Gamma, \mathcal{S})$ is definitely a valid model, we can now check that it will satisfy Γ.

Lemma I.4.2. For any $\Gamma \subseteq \mathcal{L}(all, some, in)$, and any set $\mathcal{S}$ which is closed with respect to Γ, $\mathcal{M}(\Gamma, \mathcal{S}) \models \Gamma$.

Proof. Suppose *All X are Y* is in Γ. If $Z_i \in [\![X]\!]$, then by the model construction, $\Gamma \vdash$ *All Z are X*. So $\Gamma \vdash$ *All Z are Y*, so that $Z_i \in [\![Y]\!]$. If $\{A,B\} \in [\![X]\!]$, it is a similar monotonicity point. So $\mathcal{M} \models$ *All X are Y*. If *Some X are Y* is in Γ, then this is clearly satisfied, via the element $\{X,Y\}$.

Suppose $\beta_{1,X,Y} \in \Gamma$. If either $[\![X]\!]$ or $[\![Y]\!]$ is empty, β_1 is satisfied vacuously, so we assume otherwise. Suppose $Z_i \in [\![X]\!]$, $W_j \in [\![Y]\!]$. So $\Gamma \vdash$ *All Z are X* and $\Gamma \vdash$ *All W are Y*. Then, using the monotonicity rules of our logic, we get $\Gamma \vdash \beta_{1,Z,W}$. Thus since $Th_\Gamma(Z,W) \subseteq\downarrow Th_\Gamma(Z,W)$, $\beta_{1,Z,W} \in\downarrow Th_\Gamma(Z,W)$, which by our model construction means that $Z_i[\![in]\!]W_j$. If $\{A,B\} \in [\![X]\!]$, WLOG say that $\Gamma \vdash$ *All A are X*. So $\Gamma \vdash \beta_{1,A,Y}$. Now, again, $\Gamma \vdash$ *All W are Y*, so $\Gamma \vdash \beta_{1,A,W}$, from which it follows by construction that $\{A,B\}[\![in]\!]W_j$. The cases in which $\{C,D\} \in [\![Y]\!]$ are quite similar. Thus with these cases, $\mathcal{M} \models \beta_{1,X,Y}$.

Now suppose $\beta_{2,X,Y} \in \Gamma$. We can assume that $[\![X]\!] \neq \emptyset$. Say $Z_1, Z_2 \in [\![X]\!]$. So $\Gamma \vdash$ *All Z are X*, which by monotonicity gives $\Gamma \vdash \beta_{2,Z,Y}$. Now by closedness, $Y \in \mathcal{S}$, so that $[\![Y]\!] \neq \emptyset$. Also, $\beta_{2,Z,Y} \in\downarrow Th_\Gamma(Z,Y)$, and looking at the diagrams, we see that we must have $Z_1[\![in]\!]Y_1$ and $Z_2[\![in]\!]Y_1$. If $\{A,B\} \in [\![X]\!]$ we can show similarly that $\{A,B\}[\![in]\!]Y_1$ by model construction. So $\mathcal{M} \models \beta_{2,X,Y}$.

Now suppose $\beta_{3,X,Y} \in \Gamma$. By our logic, $X \in \mathcal{S}$, and if $[\![Y]\!] = \emptyset$ then we're done. So assume $[\![Y]\!] \neq \emptyset$, i.e. $Y \in \mathcal{S}$. Let $W_j \in [\![Y]\!]$. Then $\Gamma \vdash$ *All W are Y*. So $\Gamma \vdash \beta_{3,X,W}$. So $\beta_{3,X,W} \in\downarrow Th_\Gamma(X,W)$, so by the model construction via the diagrams, $X_2[\![in]\!]W_1$ and $X_2[\![in]\!]W_2$. If $\{C,D\} \in [\![Y]\!]$ the case is similar. So $\mathcal{M} \models \beta_{3,X,Y}$.

Finally, suppose $\beta_{4,X,Y} \in \Gamma$. So both $X, Y \in \mathcal{S}$, and so given that $\beta_{4,X,Y} \in\downarrow Th_\Gamma(X,Y)$, it is clear from the diagrams that we'll have $X_2[\![in]\!]Y_1$. So $\mathcal{M} \models \beta_{4,X,Y}$. $\square$

Theorem I.4.1. *The system* $\mathcal{L}(\text{all}, \text{some}, \text{in})$ *is complete.*

Proof. Let $\Gamma \subseteq \mathcal{L}(all, some, in)$. We want to show for any sentence S that $\Gamma \models S \Longrightarrow \Gamma \vdash S$. We will do so by cases, depending on what kind of sentence S is.

First suppose that $S \in \mathcal{L}(all, some)$. Let $\Delta = \{S \in \mathcal{L}(all, some) : \Gamma \vdash S\}$. We will show that $\Delta \models S$. Let $\mathcal{N}$ be a model of Δ. So let's define a model $\mathcal{N}'$ which extends to *in* sentences by letting $[\![in]\!] = N \times N$. Now, if $\mathcal{N}' \models \Gamma$, then $\mathcal{N}' \models S$, and since $\mathcal{N}'$ is equal to $\mathcal{N}$ on variables, we would have $\mathcal{N} \models S$, as S does not involve "in." Note that if $[\![U]\!]$ and $[\![V]\!]$ are non-empty, then certainly any *in* sentence bewteen them will hold, by construction. So suppose $\beta_{1,U,V} \in \Gamma$. This sentence is satisfied trivially by $\mathcal{N}'$, because if either $[\![U]\!]$ or $[\![V]\!]$ is empty, the sentence is satisified, and if both are non-empty, then the "full" $[\![in]\!]$ will give $\mathcal{N}' \models \beta_{1,U,V}$. If $\beta_{2,U,V} \in \Gamma$, then things are a bit trickier; $[\![U]\!] = \emptyset$ is vacuous, so consider $[\![U]\!] \neq \emptyset$. We may then safely assume that $\Delta \vdash$ *Some U are U*. If it did not, then there would exist a model $\overline{\mathcal{N}}$ in which $[\![U]\!]$ was empty, and is equivalent to $\mathcal{N}$, in the sense that $\overline{\mathcal{N}} \models \Delta$ iff $\mathcal{N} \models \Delta$. So we assume $\Delta \vdash$ *Some U are U*. So along with $\beta_{2,U,V}$, this gives $\beta_{4,U,V}$, which in turn gives *Some V are V*. So $[\![V]\!] \neq \emptyset$, so $\beta_{2,U,V}$ will be satisfied. The case of $\beta_{3,U,V}$ is similar, and the case of $\beta_{4,U,V}$ is trivial. So $\mathcal{N}' \models \Delta$, and from above, $\mathcal{N} \models S$. So since $\Delta \models S$, we cite a previous result featured in Moss [1] which says that $\mathcal{L}(all, some)$ is complete. So $\Delta \vdash S$, hence $\Gamma \vdash S$.

The rest of the proof is split into cases.

Case: S is $\beta_{1,X,Y}$

$$\mathcal{S} = \{U : \Gamma \vdash \exists U\} \cup \{U : \Gamma \vdash \textit{All } X \textit{ are in some } U \text{ or } \Gamma \vdash \textit{All } Y \textit{ are in some } U\} \cup \{X, Y\}$$

We need to show that this $\mathcal{S}$ is closed for Γ so we can use it to build a model. Obviously $\mathcal{S}$ contains all U such that $\Gamma \vdash \exists U$. If $U \in \mathcal{S}$ and $\Gamma \vdash \beta_{2,U,Z}$, we want to see that $Z \in \mathcal{S}$. There are three ways that U could be in $\mathcal{S}$, looking at the above equation. If U is such that $\Gamma \vdash \exists U$, then we can use one of the rules to directly to derive $\beta_{4,U,Z}$, and so $\Gamma \vdash \exists Z$, so $Z \in \mathcal{S}$. If $U = X$ or $U = Y$, then Z fits into the second set in the union which defines $\mathcal{S}$, and so it in $\mathcal{S}$. So if U is such that $\Gamma \vdash \beta_{2,X,U}$, then we could use one of the rules of our logic to combine $\beta_{2,X,U}$ and $\beta_{2,U,Z}$ to get $\Gamma \vdash \beta_{2,X,Z}$, so $Z \in \mathcal{S}$. So $\mathcal{S}$ is closed, and we consider $\mathcal{M}(\Gamma, \mathcal{S})$.

As we know, from the lemma above, $\mathcal{M} \models \Gamma$, so $\mathcal{M} \models S$. Now by the definition of the model, using the $\mathcal{S}$ above, $[\![X]\!]$ and $[\![Y]\!]$ are non-empty. So since S is satisfied, we must have $X_1 [\![in]\!] Y_2$. Examining the diagrams, however, we see that this is only possible if $\beta_{1,X,Y} \in \downarrow Th_\Gamma(X, Y)$, which indeed implies that $\Gamma \vdash \beta_{1,X,Y}$. That is all.

Case: S is $\beta_{2,X,Y}$

$$\mathcal{S} = \{U : \Gamma \vdash \exists U\} \cup \{U : \Gamma \vdash \textit{All } X \textit{ are in some } U\} \cup \{X\}$$

The idea that $\mathcal{S}$ is closed is the same as the last case. So let's consider $\mathcal{M}(\Gamma, \mathcal{S})$. $[\![X]\!] \neq \emptyset$ by definition of the model, and the same is true of Y. So we know that X_1 is *in* something in Y. Say $X_1 [\![in]\!] Z_j$ where $\Gamma \vdash$ *All Z are Y*. If $j = 1$, then by looking at the diagrams, we must have $\beta_{2,X,Z} \in \downarrow Th_\Gamma(X, Z)$. So by definition of the downward closure there are two possibilities: either $\Gamma \vdash \beta_{2,X,Z}$ or perhaps we only have $\Gamma \vdash \beta_{1,X,Z}$. Well if $\Gamma \vdash \beta_{2,X,Z}$, then by monotonicity we would have $\Gamma \vdash \beta_{2,X,Y}$. If $\Gamma \vdash \beta_{1,X,Z}$, then since we know that $Z \in \mathcal{S}$, we have three possibilities: 1) $\Gamma \vdash \exists Z$: then $\Gamma \vdash \beta_{2,X,Z}$, and as above we are done. 2) $\Gamma \vdash$ *All X are in some Z*: again, we're done. 3) If $Z = X$, well one of our rules gives us that $\beta_{1,X,X}$ proves $\beta_{2,X,X}$, which with *All Z are Y* $=$ *All X are Y* gives $\Gamma \vdash \beta_{2,X,Y}$.

Now if $j = 2$, then by inspection of diagrams we see that $\beta_{1,X,Z} \in \downarrow Th_\Gamma(X, Z)$ i.e. $\Gamma \vdash \beta_{1,X,Z}$. Then the above discussion leads us to $\Gamma \vdash \beta_{2,X,Y}$.

The last possibility is that $X_1 [\![in]\!] \{C, D\}$. WLOG suppose that $\Gamma \vdash \beta_{1,X,D}$. If $\Gamma \vdash$ *All C are Y*, well we also have $\Gamma \vdash$ *Some C are D*,

$$\cfrac{\cfrac{\beta_{1,X,D} \quad \textit{Some } D \textit{ are } C}{\beta_{2,X,C}} \qquad \textit{All } C \textit{ are } Y}{\beta_{2,X,Y}}$$

So $\Gamma \vdash \beta_{2,X,Y}$.

Case: S is $\beta_{3,X,Y}$

$$\mathcal{S} = \{U : \Gamma \vdash \exists U\} \cup \{U : \Gamma \vdash \textit{All } Y \textit{ are in some } U\} \cup \{Y\}$$

Again, $\mathcal{S}$ is closed. $[\![Y]\!] \neq \emptyset$ by definition. Since our model satisfies S, we must have $[\![X]\!] \neq \emptyset$, and so $X \in \mathcal{S}$. Suppose X_1 is our witness. Then $X_1[\![in]\!]Y_2$. So $\beta_{1,X,Y} \in\downarrow Th_\Gamma(X,Y)$, i.e. $\Gamma \vdash \beta_{1,X,Y}$. Combined with $\Gamma \vdash \exists X$ this gives $\Gamma \vdash \beta_{3,X,Y}$.

Suppose X_2 is our witness. Then $X_2[\![in]\!]Y_2$, so by examining the diagrams we conclude $\beta_{3,X,Y} \in\downarrow Th_\Gamma(X,Y)$. If $\Gamma \vdash \beta_{3,X,Y}$ then we're done, so say $\Gamma \vdash \beta_{1,X,Y}$. Then as above we have $\Gamma \vdash \exists X$ which gives $\Gamma \vdash \beta_{3,X,Y}$.

Finally, suppose $\{A, B\}$ is our witness. Then $\{A,B\}[\![in]\!]Y_2$ Say WLOG that $\Gamma \vdash \beta_{1,A,Y}$. If $\Gamma \vdash$ *All* A *are* X, we can first use that $\Gamma \vdash \exists A$ (since $\Gamma \vdash$ *Some* A *are* B) to get $\Gamma \vdash \beta_{3,A,Y}$, then use monotonicity to get $\Gamma \vdash \beta_{3,X,Y}$. If $\Gamma \vdash$ *All* B *are* X, we can use $\Gamma \vdash$ *Some* A *are* B to get $\Gamma \vdash \beta_{3,B,Y}$ Then monotonicity will give us $\Gamma \vdash \beta_{3,X,Y}$.

Thusly, $\Gamma \vdash \beta_{3,X,Y}$.

Case: S is $\beta_{4,X,Y}$

$$\mathcal{S} = \{U : \Gamma \vdash \exists U\}$$

Clearly $\mathcal{S}$ is closed for Γ. Notice $[\![X]\!], [\![Y]\!] \neq \emptyset$. So we can conclude that $X, Y \in \mathcal{S}$.

Suppose that $Z_1 \in [\![X]\!]$ is our witness. Suppose further that $Z_1[\![in]\!]W_1$ is *in* relation we have. So $\Gamma \vdash$ *All* Z *are* X, *All* W *are* Y. By examining the diagram, we must have $\beta_{2,Z,W} \in\downarrow Th_\Gamma(Z,W)$. No matter if $\Gamma \vdash \beta_{2,Z,W}$ or $\Gamma \vdash \beta_{1,Z,W}$, the fact that $Z, W \in \mathcal{S}$ tells us that $\Gamma \vdash \exists Z, \Gamma \vdash \exists W$, which in either case will give us $\Gamma \vdash \beta_{4,Z,W}$, from which monotonicity yields $\Gamma \vdash \beta_{4,X,Y}$

Now, still supposing $Z_1 \in [\![X]\!]$ is our witness, assume $Z_1[\![in]\!]W_2$ is our in relation. The deal is similar to above, with the only possibility being that $\Gamma \vdash \beta_{1,Z,W}$, yielding $\Gamma \vdash \beta_{4,Z,W}$, monotoning to $\Gamma \vdash \beta_{4,X,Y}$.

Finally, still under the influence of Z_1 witnessing, suppose $Z_1[\![in]\!]\{C, D\}$ is our witnessing relation. We might very well have $\Gamma \vdash$ *All* C *are* Y, $\Gamma \vdash$ *All* Z *are* X, and $\Gamma \vdash \beta_{1,Z,D}$. Now, $\Gamma \vdash$ *Some* C *are* D, thus from $\beta_{1,Z,D}$, $\Gamma \vdash \beta_{2,Z,C}$, so by monotonicity $\Gamma \vdash \beta_{2,Z,Y}$; with $\Gamma \vdash \exists Z$, we get $\Gamma \vdash \beta_{4,Z,Y}$. Finally, monotonicity from *All* Z *are* X gives us $\Gamma \vdash \beta_{4,X,Y}$. All other possibilities regarding how $\{C, D\} \in [\![Y]\!]$ are similar.

And now we're catapulted into a veritable cornucopia of other cases. In the case that Z_2 is our witness, for some Z, the work is quite similar. Things also work out with $\{A, B\}$ the witness, but the author does not wish to bore the reader, and so will not list these cases. Suffice it to say that they are similar, and unremarkable.

In this way, $\Gamma \vdash \beta_{4,X,Y}$.

So indeed, we see that the very fact that $\Gamma \models S$, where S is any sentence at all in $\mathcal{L}(\textit{all}, \textit{some}, \textit{in})$, suffices to conclude that $\Gamma \vdash S$, via a well picked model. $\square$

I.5 Appendix

This is a list of the tons of axioms for $\mathcal{H}(\textit{all}, \textit{some}, \textit{verbs})$. Here, I abbreviate *All* X as $A(X)$, *Some* X as $S(X)$, and *All* X *are* Y and *Some* X *are* Y as $A(X,Y)$, $S(X,Y)$ respectively. Here, V indicates a verb which takes an object; for example, in the paper we consider *see* and its complement *not see*. There are also so axioms using the σ notation defined in the paper.

1. Tautologies of the propositional calculus
2. $((A(X)\ V\ NP))_{\mathsf{sws}} \wedge S(X,Y) \rightarrow ((S(Y)\ V\ NP))_{\mathsf{sws}}$
3. $((NP\ V\ A(X)))_{\mathsf{ows}} \wedge S(X,Y) \rightarrow ((NP\ V\ S(Y))_{\mathsf{ows}}$
4. $((S(X)\ V\ A(Y)))_{\mathsf{sws}} \rightarrow ((S(X)\ V\ A(Y)))_{\mathsf{ows}}$
5. $((A(X)\ V\ S(Y)))_{\mathsf{ows}} \rightarrow ((A(X)\ V\ S(Y)))_{\mathsf{sws}}$
6. $((S(X)\ V\ NP))_{\mathsf{sws}} \rightarrow S(X,X)$
7. $((NP\ V\ S(X)))_{\mathsf{ows}} \rightarrow S(X,X)$
8. $((A(X)\ V\ A(X)))_{\mathsf{sws}} \rightarrow ((A(X)\ V\ S(X)))_{\mathsf{sws}}$
9. $A(X,X)$
10. $A(X,Z) \wedge A(Z,Y) \rightarrow A(X,Y)$
11. $S(X,Y) \rightarrow S(Y,X)$
12. $S(X,Y) \rightarrow S(X,X)$
13. $A(Y,Z) \wedge S(X,Y) \rightarrow S(X,Z)$
14. $\neg S(X,Y) \wedge A(X,Y) \leftrightarrow \neg S(X,X)$
15. $\neg A(X,Y) \rightarrow S(X,X)$

    In the following, we use commas to abbreviate many axioms. The subscripts of the σ's are understood to include $\{X,Y\}$.

16. $S(X,X) \wedge \neg S(Y,Y) \rightarrow \sigma_1, \sigma_2, \neg\sigma_3, \sigma_4, \neg\sigma_5, \neg\sigma_6$
17. $\neg S(X,X) \wedge S(Y,Y) \rightarrow \sigma_1, \neg\sigma_2, \sigma_3, \neg\sigma_4, \sigma_5, \neg\sigma_6$
18. $\neg S(X,X) \wedge \neg S(Y,Y) \rightarrow \sigma_1, \neg\sigma_2, \neg\sigma_3, \sigma_4, \sigma_5, \neg\sigma_6$

    Monotonicity axiom schema:

19. $A(X^{\downarrow})\ V\ A(Y^{\downarrow})$
20. $(S(X^{\uparrow})\ V\ A(Y^{\downarrow}))_{\mathsf{sws}}$
21. $(A(X^{\downarrow})\ V\ S(Y^{\uparrow}))_{\mathsf{ows}}$
22. $(S(X^{\uparrow})\ V\ A(Y^{\downarrow}))_{\mathsf{ows}}$
23. $(A(X^{\downarrow})\ V\ S(Y^{\uparrow}))_{\mathsf{sws}}$
24. $S(X^{\uparrow})\ V\ S(Y^{\uparrow})$

The notation for the monotonicity axioms is taken from Johan van Benthem [3]. These six schema represent twelve rules. The up-arrow next to a variable $X^{\uparrow}$ indicates that we can take supersets of X, and a down arrow $X^{\downarrow}$ indicates we can take subsets. For example, $(S(X^{\uparrow})\ V\ A(Y^{\downarrow}))_{\mathsf{sws}}$ stands for

$$(S(X)\ V\ A(Y))_{\mathsf{sws}} \wedge A(X,Z) \rightarrow (S(Z)\ V\ A(Y))_{\mathsf{sws}}$$

$$(S(X)\ V\ A(Y))_{\mathsf{sws}} \wedge A(Z,Y) \rightarrow (S(X)\ V\ A(Z))_{\mathsf{sws}}$$

These are the rules for the fragment $\mathcal{L}(all, some, in)$. We use the following abbreviations: $\forall X \forall Y$ stands for *All X are in all Y*, $\forall X \exists Y$ stands for *All X are in some Y*, etc. In the top rules, we simplify even further by assuming that the variables involved in the top sentence of the antecedents are X then Y, and that the variables appearing in the second sentence of the antecedents are Y then Z. The conclusion features X then Z. The quantifiers with arrows are monotonicity rules, just as appear in the axioms for the *verbs* fragment.

$$\dfrac{\begin{matrix}\forall\forall\\ \forall\forall\end{matrix}}{\forall\forall} \qquad \dfrac{\begin{matrix}\forall\forall\\ \forall\exists\end{matrix}}{\forall\exists} \qquad \dfrac{\begin{matrix}\forall\forall\\ \exists\forall\end{matrix}}{\forall\forall} \qquad \dfrac{\begin{matrix}\forall\forall\\ \exists\exists\end{matrix}}{\forall\exists}$$

$$\dfrac{\begin{matrix}\forall\exists\\ \forall\forall\end{matrix}}{\forall\forall} \qquad \dfrac{\begin{matrix}\forall\exists\\ \forall\exists\end{matrix}}{\forall\exists} \qquad \dfrac{\begin{matrix}\exists\forall\\ \forall\forall\end{matrix}}{\exists\forall} \qquad \dfrac{\begin{matrix}\exists\forall\\ \forall\exists\end{matrix}}{\exists\exists}$$

$$\dfrac{\begin{matrix}\exists\forall\\ \exists\forall\end{matrix}}{\exists\forall} \qquad \dfrac{\begin{matrix}\exists\forall\\ \exists\exists\end{matrix}}{\exists\exists} \qquad \dfrac{\begin{matrix}\exists\exists\\ \forall\forall\end{matrix}}{\exists\forall} \qquad \dfrac{\begin{matrix}\exists\exists\\ \forall\exists\end{matrix}}{\exists\exists}$$

$$\dfrac{All\ X\ are\ Y \quad All\ Y\ are\ Z}{All\ X\ are\ Z} \qquad \dfrac{}{All\ X\ are\ X} \qquad \dfrac{All\ X\ are\ Y \quad Some\ Z\ are\ X}{Some\ Z\ are\ Y}$$

$$\dfrac{Some\ X\ are\ Y}{Some\ Y\ are\ X} \qquad \dfrac{Some\ X\ are\ Y}{Some\ X\ are\ X}$$

$$\forall^{\downarrow}\forall^{\downarrow} \qquad \exists^{\uparrow}\forall^{\downarrow}$$

$$\forall^{\downarrow}\exists^{\uparrow} \qquad \exists^{\uparrow}\exists^{\uparrow}$$

$$\dfrac{\exists X \exists Y}{Some\ X\ are\ X} \qquad \dfrac{\exists X \exists Y}{Some\ Y\ are\ Y} \qquad \dfrac{\exists X \forall Y}{Some\ X\ are\ X}$$

$$\dfrac{\forall X \exists Y \quad Some\ X\ are\ Z}{\exists Z \exists Y} \qquad \dfrac{\forall X \forall Y \quad Some\ X\ are\ Z}{\exists Z \forall Y} \qquad \dfrac{\forall X \forall Y \quad Some\ Y\ are\ Z}{\forall X \exists Z}$$

$$\dfrac{\exists X \forall Y \quad Some\ Y\ are\ Z}{\exists X \exists Z} \qquad \dfrac{\forall X \forall Y \quad All\ X\ are\ Y}{\forall X \exists Y}$$

Bibliography

1. Lawrence S. Moss. Completeness Theorems for Syllogistic Fragments. To appear in F. Hamm and S. Kepser (eds.) *Festschrift for Uwe Mönnich*, Mouton de Gruyter, 2008.

2. Lawrence S. Moss. Syllogistic Logic with Verbs. To appear in *Journal of Logic and Computation*, special issue of papers from a conference on *Order, Algebra, Logic*.

3. Johan van Benthem. *Essays in Logical Semantics*. Reidel, Dordrecht, 1986,

4. Ian Pratt-Hartmann and Lawrence S. Moss. Logics for the relational syllogistic. Ms., University of Manchester, 2008.

5. Joost Zwarts and Yoad Winter. Vector Space Semantics: a Modeltheoretic Analysis of Locative Prepositions. *Journal of Logic, Language and Information*, 2000