CALIFORNIA STATE UNIVERSITY, NORTHRIDGE

PROTEIN FOLDING: PLANAR CONFIGURATION SPACES

OF DISC ARRANGEMENTS AND HINGED POLYGONS:

PROTEIN FOLDING IN FLATLAND

A thesis submitted in partial fulfillment of the requirements
For the degree of Master of Science in Mathematics

By

Clinton Bowen

Spring 2014

The thesis of Clinton Bowen is approved:

Dr. John Dye — Date

Dr. Silvia Fernandez — Date

Dr. Bernardo Abrego — Date

Dr. Csaba Toth, Chair — Date

California State University, Northridge

DEDICATIONS

ACKNOWLEDGEMENTS

Table of Contents

ABSTRACT

PROTEIN FOLDING: PLANAR CONFIGURATION SPACES

OF DISC ARRANGEMENTS AND HINGED POLYGONS:

PROTEIN FOLDING IN FLATLAND

By

Clinton Bowen

Master of Science in Mathematics

Insert Abstract here

Abstract

We look into the decidability of whether a hinged configuration locks.

1 Introduction

We look into the decidability of continuity on planar configuration space using regular, unitary hexagonal polygons. These polygons can also represent unit disk configurations [3]

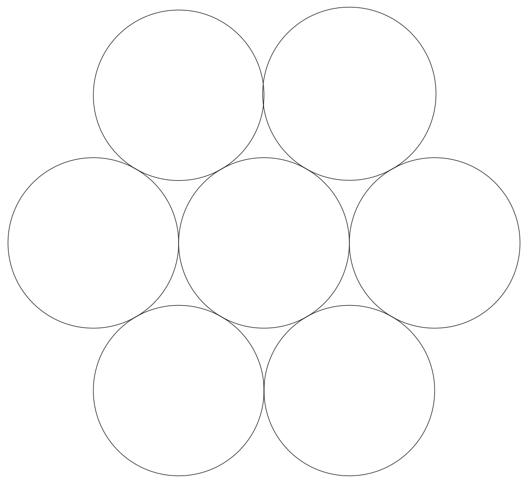

Figure 1: A locked 7 ball configuration

Motivation Protein folding, graphite, crystalline structures in metallurgy; disc packing; hexagonal configurations; Determine whether chemical structures are realizable.

Outline Section 2 covers the necessary mathematical concepts to understanding the problem. Section 3 explains the problem, Section 4 covers the results and findings about the problem. Section 5, the conclusion, offers final remarks on the problem.

2 Background

Here we review some of the necessary mathematics behind the problem. The definitions found in this chapter are those found in [7, 10, 6].

2.1 Linkages

Given a *graph*, an ordered pair $G = (V, E)$, comprising of a set V of vertices or nodes together with a set E of edges or lines, then a linkage of G is the realization (or embedding) of G in $\mathbb{R}^2$. For this paper, we focus on linkages that are simple planar graphs, i.e.:

(i) does not have multiple edges between any pair of vertices,

(ii) does not have edges that cross, or

(iii) have loops (i.e. $(v, v) \in E$).

2.1.1 Configuration Spaces of Linkages

To describe the types of motion that we are interested in linkages we must define the graph isomorphism. Two graphs $G = (V_1, E_1)$ and $\Gamma = (V_2, E_2)$, a bijection $f : V_1 \mapsto V_2$ such that for any two vertices $u, v \in V_1$ that are adjacent, i.e. $(u, v) \in E_1$, if and only if $(f(u), f(v)) \in E_2$.

Graph	Vertices	Edges
G	$\{a,b,c,d,e\}$	$\{(a,b),(b,c),(c,d),(d,e),(e,a)\}$
Γ	$\{1,2,3,4,5\}$	$\{(1,2),(2,3),(3,4),(4,5),(5,1)\}$

Table 1: Two graphs that are isomorphic with the alphabetical isomorphism $f(a) = 1$, $f(b) = 2$, $f(c) = 3$, $f(d) = 4$, $f(e) = 5$.

Next we add restrictions to our graph isomorphisms to narrow our focus:

(i) We focus on isomorphisms for simple planar graphs, and

(ii) the isomorphism preserves edge lengths, e.g. $d(u, v) = d(f(u), f(v))$.

With these restrictions of our isomorphisms, we can begin to describe a range of motion to transform a linkage. That range of motion is said to be the configuration space of that linkage. To expand on this concept, for given linkage, $L = (V, E)$, and for a given vertex $v \in V$, the set of points in which v can be realized in the plane would be the configuration space for that vertex, C_v. Defining some order of the vertices in L, i.e. $V = \{v_n\}_{i=1}^n$, then the *configuration space* for L is said to be the cartesion product of the configuration space of vertices:

$$C(L) = C_{v_1} \times C_{v_2} \times \cdots \times C_{v_n} \tag{1}$$

Some food for thought on configuration spaces and motions on linkages:

(i) A configuration space is said to be *connected* if there is a continuous mapping for any two planar realizations (linkages) of a graph in the plane. Otherwise it is said to be *disconnected*.

(ii) If the configuration space of a vertex, C_v, is a singleton set, then the vertex is said to be *pinned*. Otherwise it is said to be *free*.

(iii) The types of motions (mappings) that we refrain from using on linkages are translations.

2.1.2 Realizability of Linkages

Suppose we had two configurations of a linkage, $\mathscr{A}$ and $\mathscr{B}$. A question that can be posed is can we reconfigure $\mathscr{A}$ to $\mathscr{B}$ continuously while respecting simple planar graph conditions? The answer to this question is a yes or no. It has been shown that this problem can be posed as a planar satisfiability problem [3, 8] (Later on in this paper we'll cover satisfiability problems). This is the type of problem that we face in this paper. We will continue to explore this in a different manner, with circle packings.

2.2 Circle Packing

It turns out the circle packings are an equivalent way to to represent linkages and their corresponding problems. Before we establish the relation, we will cover some fundamental concepts of circle packings. A *circle packing*, P, embedded in a plane is a set of circles with disjoint interiors $\{C_i\}_{i=1}^n$ such that for any circle $C \in \{C_i\}_{i=1}^n$, C is tangent to a different circle of $\{C_i\}_{i=1}^n$.

Any circle embedded in a plane has a given center point and radius. This information of planar embedded circle packings allows us to establish the relationship to linkages with the following construction:

(i) let the centerpoints of the circle packing be a set of vertices V;

(ii) if two circles in a circle packing are tangent, we define an edge between their centerpoints. The distance of this edge is the sum the radii of the two tangent circles.

This construction establishes a relationship between linkages and circle packings. It begs questioning as to whether every connected simple planar graph has a circle packing. The question is answered in the following theorem.

Theorem 2.1 (Circle Packing Theorem)**.** *For every connected simple planar graph G there is a circle packing in the plane whose intersection graph is (isomorphic to) G.*

A proof of Theorem 2.1 is found in chapter 7 of [10]. Theorem 2.1 also gives us the ability to establish an equivalent definition of configuration spaces on circle packings and allows us to pose the same realizability problems found with simple planar graphs. To narrow the focus of the types of circle packing realizabilty problems that we are interested in, we add the following restriction: all circles in a circle packing have unit diameter.

2.2.1 Realizability Problems in Unit Disk Packings

In [3], it was shown that unit disk graph recognition is NP-Hard.

2.3 Area Packing Problem

2.3.1 Hinged Polygons

Definition 2.1 (Polygonal Chain)**.** A polygonal chain $P = (v_0, v_1, \ldots, v_{n-1})$ is a sequence of consecutively joined segments (or edges) $e_i = v_i v_{i+1}$ of fixed lengths $l_i = |e_i|$, in a plane. [2]

A chain is said to be closed if $v_{n-1} = v_1$, otherwise it is said to be open. Hinged polygons have been researched for decades and related to linkage problems [2, 4].

Consider the locked configuration of figure 2. We can configure the hexagons to be locked by placing hinged points as follows:

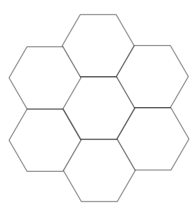

Figure 2: A locked 7 hexagonal configuration. (needs to modify picture by placing red points for hing points.)

2.3.2 Hinged Hexagons of Fixed Size

The Shapes Figure 3 is a locking shape: Figure 3 shall reside in the boundary of a lattice and have a hinge point at one vertex where the locking shape and boundary meet.

Junctions We define junctions to be the point three hexagons meet in a hexagonal lattice, e.g. Figure 4.

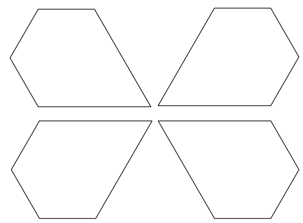

Figure 3: A locking shape in the lattice boundary's channel.

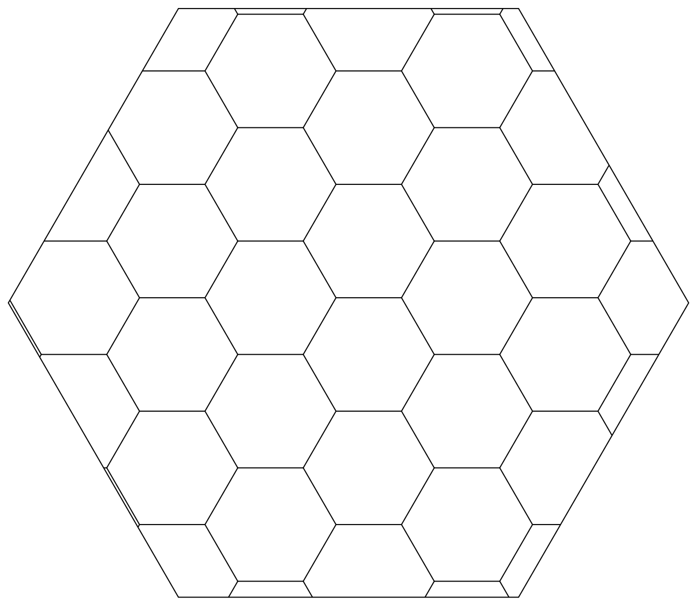

Figure 4: A portion of a hexagonal lattice.

Central Scaling

Junctions in Conjunctive Normal Form Explain the configurations we're interested in.

3 Configuration Spaces of Polygonal Chains

For any regular polygon, there exists a circle in which that polygon can be circumscribed.

3.1 Polygonal Linkages

3.1.1 Configurations and Locked Configurations

3.2 Dissections

Problem 3.1 (Polygonal Dissection). Given two polygons of equal area, P_1 and P_2, partition P_1 into smaller pieces,$\{P_{1,i}\}_{i=1}^{n}$, rearrange the pieces to form P_2. [6]

Theorem 3.1. *Any finite collection of polygons of equal area has a common hinged dissection. [1]*

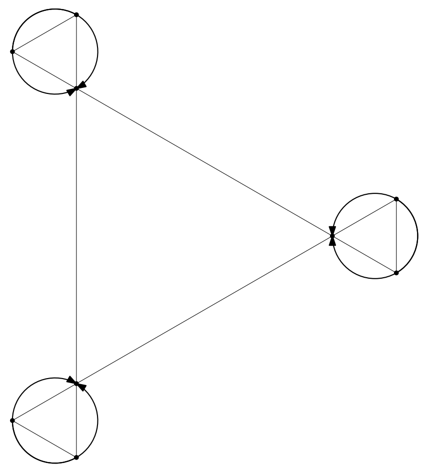

Figure 5: **Not sure how to describe this graph with free joints, pinned joints. Do I need to define another joint(i.e. axis of rotation joint)?**

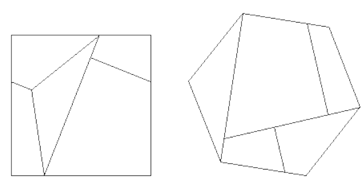

Figure 6: An axample of two polygons of equal area that can be rearranged into the other by the given partition.[5]

3.3 SAT Problems

Problem 3.2 (Satisfiability Problem). Let $\{x_i\}_{i=1}^n$ be boolean variables, and $t_i \in \{x_i\}_{i=1}^n \cup \{\bar{x}_i\}_{i=1}^n$. A *clause* is is said to be a disjuction of distinct terms:

$$t_1 \vee \cdots \vee t_{j_k} = C_k$$

Then the *satisfiability problem* is the decidability of a conjuction of a set of clauses, i.e.:

$$\wedge_{i=1}^m C_i$$

[9] A *3-SAT problem* is a SAT problem with all clauses having only three boolean variables.

Definition 3.1 (Planar 3-SAT Problem)**.** Given a boolean 3-SAT formula B, define the associated graph of B as follows:

$$G(B) = (\{\, v_x \,|v_x \text{ represents a variable in } B\,\} \cup \{\, v_C \,|v_C \text{ represent a clause in } B\,\}, \{\,(v_x, v_C)\,|x \in C \text{ or } \bar{x} \in C\,\}) \tag{2}$$

If $G(B)$ in equation (2) is planar, then B is said to be a *Planar 3-SAT Problem* [8].

4 Problem

4.1 Problem Statement

text

4.2 Decidability of Problem

test

4.3 Hexagonal Locked Configuration

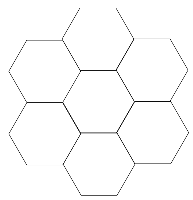

Figure 7: 7 hexagonal configuration

5 Conclusion

We conclude... WE REALLY NEED TO ADD THAT PLANAR SAT PROBLEM TO THE BIBLIOGRAPHY!!!! AND DISCUSS IT!!!!

References

[1] Timothy G Abbott, Zachary Abel, David Charlton, Erik D Demaine, Martin L Demaine, and Scott Duke Kominers. Hinged dissections exist. *Discrete & Computational Geometry*, 47(1):150–186, 2012.

[2] T. Biedl, E. Demaine, M. Demaine, S. Lazard, A. Lubiw, J. O'Rourke, M. Overmars, S. Robbins, I. Streinu, G. Toussaint, and S. Whitesides. Locked and Unlocked Polygonal Chains in 3D, 1999.

[3] Heinz Breu and David G. Kirkpatrick. Unit disk graph recognition is NP-hard. *Computational Geometry*, 9(1–2):3–24, 1998. Special Issue on Geometric Representations of Graphs.

[4] J. Canny. *The Complexity of Robot Motion Planning*. ACM doctoral dissertation award. MIT Press, 1988.

[5] David Eppstien. The Geometry Junkyard, November 2013.

[6] G.N. Frederickson. *Dissections: Plane and Fancy*. Cambridge University Press, 1997.

[7] J. Kleinberg and E. Tardos. *Algorithm Design*. Pearson Education, 2006.

[8] Wolfgang Mulzer and Günter Rote. Minimum-weight triangulation is np-hard. *Journal of the ACM (JACM)*, 55(2):11, 2008.

[9] S.S. Skiena. *The Algorithm Design Manual*. Springer, 2009.

[10] K. Stephenson. *Introduction to Circle Packing: The Theory of Discrete Analytic Functions*. Cambridge University Press, 2005.