Numerical Solution of Soap Film Dual Problems

Kenneth A. Brakke

CONTENTS

This work was partially supported by The Geometry Center of the University of Minnesota and the University of Massachusetts at Amherst.

The soap film problem is to minimize area, and its dual is to maximize the flux of a divergenceless bounded vector field. This paper discretizes the continuous problem and solves it numerically. This gives upper and lower bounds on the area of the globally minimizing film. In favorable cases, the method can be used to discover previously unknown films. No initial assumptions about the topology of the film are needed. The paired calibration or covering space model of soap films is used to enable representation of films with singularities.

1. INTRODUCTION

The soap film problem is to find the soap film that minimizes area subject to appropriate constraints. The two fundamental problems addressed by the numerical methods in this paper are finding a globally area-minimizing film for a given soap film problem, and proving that a given film is indeed a global minimum. A precise definition of what a soap film problem is will have to wait until Section 3, but an intuitive idea is enough to see the difficulties.

Loosely speaking, soap films are area-minimizing hypersurfaces, but their treatment in full generality is complicated by the fact that they are not always smooth manifolds, but may have singularities. For two-dimensional films in a three-dimensional ambient space, the possible singularities are a triple line, where three films meet at 120° along a curve, or a tetrahedral point, where six films and four triple lines meet at equal angles [Taylor 1976]. Further types of singularities are possible in higher dimensions [Brakke 1991; Sullivan 1995].

A computer program such as the Surface Evolver [Brakke 1992] can represent a surface as a set of flat triangles (or curved patches in more generality), and hence provide an upper bound for area.

1058-6458/96 $0.50 per page

But the topology of the surface has to be provided at the beginning, and although the topology may change during the evolution to minimum area, all one has at the end is a local minimum of a discrete problem. In very limited circumstances, one can show there is a smooth minimal surface nearby [Underwood 1993], but that is still only a local minimum. The topology of the global minimum may be entirely different. What is needed is a lower bound on the area of the global minimum. If one can get the upper and lower bounds equal, then one has a global minimum. Of course, the global minimum may not be unique. For example, a cubical wire frame bounds an apparent minimizer with a rounded square in the center, and that square can be parallel to any side of the cube.

Fortunately, it often happens that a minimization problem has a corresponding maximization problem whose optimum has the same value. Such problems are called *dual*. A simple example is minimizing the circumference of a given area, and maximizing the area bounded by a given length of circumference. A more relevant example here is the max-flow min-cut theorem of network theory: given a graph whose edges have fixed carrying capacities with some nodes designated sources and some designated sinks, find the maximum total flow from sources to sinks. The dual minimization problem is to find a minimal cut, a set of edges with minimum total capacity that separates the sources from the sinks. A continuous version of this that applies to orientable, nonsingular soap films was introduced by Federer [1969; 1974] and named *calibration* by Harvey and Lawson [1982]. The dual maximization problem is to find a divergenceless vector field of maximum magnitude 1 with maximum flux through the given boundary of the film. Intuitively, the vector field is the velocity of an incompressible fluid. The surface of minimal area is the bottleneck to the flow, so for maximum velocity 1 the maximum flux equals the minimum area.

This paper treats only the area minimization problem with boundary constraints. In particular, it does not treat soap bubble problems (with volume constraints), nor capillary problems (with gravitational energy), although in favorable circumstances these problems are susceptible to extensions of the methods of this paper.

Section 2 gives some preliminary background on surfaces and flows. Section 3 describes a more general calibration model that can handle soap film singularities. Section 4 describes discretization of the model into a form that is a standard optimization problem. Section 5 describes a particular implementation using the Surface Evolver and some custom programs to generate data that can be fed to optimization software. Section 6 discusses some results obtained so far. Briefly, there is the first known calibration of the network spanning the vertices of a regular hexagon, a novel solution found by computer of another plane problem, and some preliminary results on the conjectured minimal film spanning an octahedral frame. Section 7 discusses some works by others that have similarities to the approach presented here. Section 8 concludes and outlines some future prospects.

2. PRELIMINARIES

The only mathematical background necessary to the understanding of this article consists of the standard concepts of an advanced calculus course, such as surface integrals and the Divergence Theorem. Occasional references to more general concepts of integral geometry, such as currents and differential forms, are tossed in for the cognoscenti. This section explains just enough about currents and differential forms for our purposes. For fuller discussions, see [Federer 1969; Morgan 1995].

The overall domain will be N-dimensional Euclidean space $\mathbb{R}^N$, although the ideas extend naturally to any Riemannian manifold. Domains for particular problems are usually chosen to be compact convex sets, since a soap film is always contained in the convex hull of its boundary. All soap films will be $(N-1)$-dimensional. The types of integrals needed will be integrals of scalar functions over regions and integrals of vector fields over hypersurfaces. A region R may be defined by a characteristic function $\chi_R(x)$, with the integral of a scalar function $f(x)$ being

$$\int_R f(x)\, d^N x = \int_{\mathbb{R}^N} f(x)\chi_R(x)\, d^N x.$$

We will want to generalize the notion of region so that its characteristic function becomes instead a density function $Q(x)$ with real values instead of integer values. We still can define the integral of $f(x)$ as

$$\int_R f(x)\, d^N x = \int_{\mathbb{R}^N} f(x)Q(x)\, d^N x.$$

We call such an object a *smeared region*. Soap films will be viewed as the boundaries of regions. The boundary ∂R of a region R is defined precisely so as to make the Divergence Theorem true. If $\vec{u}(x)$ is a smooth vector field with compact support, then

$$\int_R \operatorname{div} \vec{u}(x)\, d^N x = \int_{\partial R} \vec{u}(x) \cdot \vec{N}(x)\, dA.$$

For an ordinary region R, the boundary ∂R is the surface of the region with outward unit normal $\vec{N}$. For a smeared region, the boundary is representable by the negative gradient of the density function:

$$\int_{\mathbb{R}^N} \operatorname{div} \vec{u}(x)Q(x)\, d^N x = -\int_{\mathbb{R}^N} \vec{u}(x) \cdot \operatorname{grad} Q(x)\, d^N x.$$

This may be derived by applying the ordinary Divergence Theorem to

$$\int_B \operatorname{div}(\vec{u}(x)Q(x))\, d^N x,$$

where B is some large ball containing the support of $\vec{u}(x)$. The boundary of a smeared region is a smeared surface.

Technically, objects one does k-dimensional integrals over are called *k-currents*, and the k-dimensional integrands are called *differential k-forms*. Thus our regions are N-currents, scalar functions are N-forms, vector fields are $(N-1)$-forms, and surfaces are $(N-1)$-currents. Smeared surfaces should still be regarded as $(N-1)$-dimensional objects, even though they are smeared out over all N dimensions. The particular class of currents we need are called *flat chains*, and the vector fields are the class of *flat cochains*. In particular, a *flow* will be a divergenceless flat $(N-1)$-cochain. It turns out to be relatively easy to characterize a flow. It need only be a measurable, bounded vector field, and have zero divergence in the weak sense. The latter means that if $\vec{v}$ is a vector field and f is a smooth function with compact support, then

$$\int \vec{v} \cdot \operatorname{grad} f = 0.$$

Flows need not be continuous, but at a surface of discontinuity the components normal to the surface on both sides must be equal. This is enough to guarantee the integrability of flows on all surfaces of interest. The flows in this paper will be piecewise linear and constructed to have zero divergence.

The surfaces in the continuous theory are flat $(N-1)$-chains. Flat chains are dual to the flat cochains. So a flat chain is anything one can integrate flat cochains over. A flat N-chain is simply an integrable scalar function. The boundary of a flat N-chain is automatically a flat $(N-1)$-chain, so the regions defined above with density functions Q are flat N-chains, and their boundaries are flat $(N-1)$-chains. Another way to form a flat chain is to take an oriented rectifiable set. Flat chains that are integer multiples of oriented rectifiable sets are called *integer flat chains*, and are what we usually think of as soap films. *Rectifiable flat chains* are rectifiable sets multiplied by real-valued densities. General flat chains are called *real flat chains* to emphasize the distinction. Hereafter we will often use *integer film* as a synonym for integer flat chain, and *real film* for real flat chain.

The equivalent of area for a flat chain T is its *mass* $M(T)$, defined as the maximum integral over flat cochains of maximum norm 1:

$$M(T) = \sup_{\vec{u}} \left\{ \int_T \vec{u} : \|\vec{u}(x)\| \leq 1 \text{ for all } x \right\}.$$

Note that the supremum is over all flat cochains, not just those with zero divergence. The mass need not be finite, although it will be for all our problems. The mass of an integer flat chain is just the area of the rectifiable set, times any multiplicities. If T is the boundary of a flat N-chain with density function Q, the mass of T is the total variation of Q:

$$M(T) = \int \|\operatorname{grad} Q(x)\| \, d^N x.$$

Hence, for finite surface mass, we may take Q to be any function of bounded variation. If Q is a characteristic function of a set, the boundary of the set is an integer flat chain. Defining the film by means of Q is the bounded variation approach to minimal surfaces of [Giusti 1984]. The flat $(N-1)$-chains in this paper will either be unions of $(N-1)$-dimensional simplices or be boundaries of piecewise linear functions.

3. THE CONTINUOUS MODEL

The mathematical model of soap films used in this paper is the paired calibration model [Lawlor and Morgan 1994] or the covering space model [Brakke 1995]. The paired calibration model will be described first, since it is perhaps a little clearer. Then the fully general covering space model will be defined.

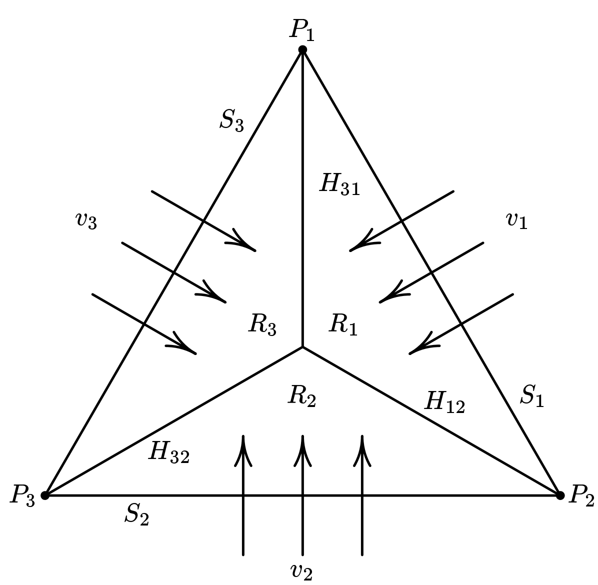


FIGURE 1. Paired calibration of the tripod.

The Paired Calibration Model

The *paired calibration model* regards a soap film as a set of interfaces between regions that partition the domain. Figure 1 is a diagram of the model for the problem of finding the shortest one-dimensional film joining the three vertices P_1, P_2, P_3 of an equilateral triangle.

Let the regions be denoted $R_1, \dots, R_s$. Let H_{ij} be the interface between regions R_i and R_j, regarded as an oriented surface, or flat chain to be more precise. Hence $H_{ij} = -H_{ji}$. Each region also has an outer boundary

$$S_i = \sum_j H_{ij} - \partial R_i,$$

regarded as fixed. S_i is called the *reference surface* for region R_i. The orientations are chosen here so that S_i and $\sum_j H_{ij}$ are homologous. Note that this makes S_i have inward normal. The minimization problem is:

$$\text{Minimize} \sum_{i<j} \operatorname{area}(H_{ij}).$$

A *paired calibration* is a set of flows $\vec{v}_i$, one per region, each defined over the entire domain, such that $\|\vec{v}_i(x) - \vec{v}_j(x)\| \le 1$ for all i, j and all x. The *total flux* F of a paired calibration is defined as

$$F = \sum_i \int_{S_i} \vec{v}_i \cdot d\vec{A}.$$

Theorem 3.1. *If $\{\vec{v}_i\}$ is a paired calibration for the set of surfaces $\{S_i\}$, then for any regions $\{R_i\}$ and corresponding interfaces $\{H_{ij}\}$, the total flux is at most the interface area*:

$$\sum_i \int_{S_i} \vec{v}_i \cdot d\vec{A} \le \sum_{i<j} \operatorname{area}(H_{ij}).$$

Proof. Since $\operatorname{div} \vec{v}_i = 0$, the Divergence Theorem and the flow difference bound imply, after some rearrangement:

$$\begin{aligned}\sum_i \int_{S_i} \vec{v}_i \cdot d\vec{A} &= \sum_i \int_{\sum_j H_{ij}} \vec{v}_i \cdot d\vec{A} \\ &= \sum_{i<j} \int_{H_{ij}} (\vec{v}_i - \vec{v}_j) \cdot d\vec{A} \\ &\leq \sum_{i<j} \int_{H_{ij}} 1\, dA \\ &= \sum_{i<j} \operatorname{area}(H_{ij}). \qquad \square\end{aligned}$$

Hence the total flux of a paired calibration provides a lower bound on the areas of possible soap films. If one can find a paired calibration whose flux is equal to the area of a soap film, one also has a proof that the soap film is a global minimum of area. We then say that the vector fields *calibrate* the film. The calibration is far from unique. The proof of Theorem 3.1 shows that the only serious constraint on it is that the difference of flows be a unit normal at the minimal films.

Note that if there are multiple global minimal films (as for the cubical wire frame), any calibration must calibrate all minimal films simultaneously. This follows immediately from the proof.

Example: Simplicial Cones

In Figure 1, the minimal film consists of three segments from the vertices P_1, P_2, P_3 to the center of the triangle. Thus it is a cone generated by the three vertices. I call this film the *tripod*. The simplest calibration consists of three vector fields, each constant over all of $\mathbb{R}^2$, of magnitude $1/\sqrt{3}$ and parallel to a segment. Note that it is the magnitude of the difference between vector fields that is bounded by 1, not the magnitudes of the individual vector fields. The same kind of calibration can be done to show that the $(N-1)$-dimensional cones over the $(N-2)$-dimensional skeletons of regular N-simplices are absolutely minimizing [Lawlor and Morgan 1994].

Existence

It follows from [Brakke 1995, Theorem 6.1] that, for any set of reference surfaces $\{S_i\}$, there always is a set of corresponding interfaces $\{H_{ij}\}$ that can be calibrated, but the notion of surface must be understood in the general sense of real flat chains. Soap films are usually imagined to be integer density surfaces, and the existence of global minimizers among integer density films may be proved by compactness. But there are cases where the real density minimum is different from the integer density minimum, and in that case the integer density minimum cannot be calibrated. A prime example is the single bubble problem: to find the minimum area enclosing a given volume. (Although we are not otherwise considering volume constraints in this paper, the techniques generalize.) The integer film minimum is just a sphere, but in the sense of real flat chains, a sphere of twice the radius and $\frac{1}{8}$ the density bounds the same volume but has only half the mass. Hence the real flat chain minimizer does not exist, as the radius goes to infinity and the density and mass go to zero.

The Covering Space Model

The *covering space model* is a generalization of the paired calibration model that can handle films that don't divide space into distinct regions, such as a Möbius band film, for example. Let $M \subset \mathbb{R}^N$ be a compact region, which will be the domain holding the film. Let $B \subset M$ be a closed set, meant to hold the boundary of the film. Let Y be a covering space of the complement $M - B$. Let W be the covering space of $M - B$ that, over each point of M, has one sheet for each oriented pair of sheets of Y. Locally, one can talk about sheet i of Y and sheet ij of W. For an oriented surface H in W, define the projection $\pi_\# H$ as the surface in Y obtained by copying from sheet ij of W a positive copy to sheet i of Y and a negative copy to sheet j of Y. The relation to paired calibrations is that regions correspond to sheets of Y, and sets of interfaces correspond to surfaces H in W. A reference surface S is defined to be a portion of the boundary of Y.

A particular area minimization problem is defined by a choice of Y, and a choice of S. The problem is:

$$\text{Minimize area}(H)$$
$$\text{with } \pi_{\#}H \text{ homologous to } S.$$

Intuitively, the idea is that the two sides of a soap film are oppositely oriented surfaces on different sheets of Y, and the covering space W is there to make sure that sides pair up. *Homologous* simply means that $\pi_{\#}H - S$ is the boundary of a region. If one takes the N-current Q such that $\partial Q = \pi_{\#}H - S$, then Q is a scalar density function times M. The current Q is intuitively the exterior of the film, and we will denote the density function as $Q(y)$ and call it the *exterior density function*, or just the *exterior function*. Q is required to be 1 on the reference surface S and zero on the rest of the boundary of Y.

A *calibrating flow* is a flow $\vec{v}$ in Y that at each point of M satisfies $\|\vec{v}_i - \vec{v}_j\| \le 1$, where i and j refer to sheets above the point. One can define the lift of $\vec{v}$ to a vector field $\pi^{\#}\vec{v}$ in W by

$$(\pi^{\#}\vec{v})_{ij} = \vec{v}_i - \vec{v}_j,$$

and then simply require $\|\pi^{\#}\vec{v}\| \le 1$. The dual maximization problem turns out to be:

$$\text{Maximize } \int_S \vec{v}\cdot d\vec{A}$$
$$\text{with } \operatorname{div}\vec{v} = 0 \text{ and } \|\pi^{\#}\vec{v}\| \le 1.$$

The same existence theorem for covering space films holds as for paired calibrations, with the same caveats. The surface S can be regarded as the source of the flow (in analogy with the network max-flow min-cut problem), and the rest of the boundary of Y as the sink.

Symmetries

If the problem has symmetries, the calibrating flow may be assumed to share those symmetries. The action of a symmetry transformation must be defined to include the permutation of regions in the paired calibration model and of sheets in the covering space model, so that reference surfaces get mapped to reference surfaces. To get a symmetric calibration, simply take any calibrating flow and average over all symmetric transformations of it. Thus if there is a mirror symmetry which maps a given region to itself, we may assume that the flow for that region has no flux across the mirror. For example, Figure 2 (left) shows the equilateral triangle divided into its six fundamental regions, labeled A through F. Use subscripts to denote sheets. By rotational symmetry, region A_1 is symmetric to C_2 and E_3. Including mirror symmetries, region A_1 is symmetric to B_1, D_2, and F_3. There are three symmetry classes, which can conveniently be represented either by a single stack of triangles A_1, A_2, A_3 on different sheets, or by triangles A_2, B_2, C_2 on a single sheet. The single stack is used for calculations, and the single sheet is useful for display, as in Figure 5 (page 281). The single sheet can be visualized as folding up into the single stack, with creases along the mirror lines. These lines (OP_2 and OJ in Figure 2, right) will be referred to as *fold lines*. The edge HP_2 is the reference surface for region R_2, that is, the source edge with $Q = 1$. The edges P_1J and JP_2 are sink edges with $Q = 0$. Since mirror symmetry along HP_1 maps region R_2 to itself, there is identically zero flux across edges HO and OP_1, and Q values are free to vary.

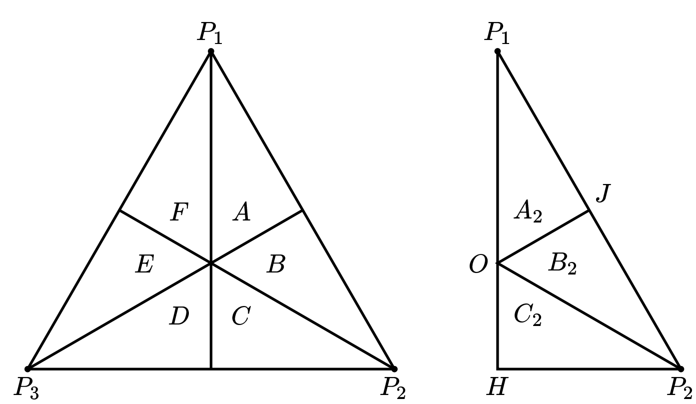


FIGURE 2. Left: Symmetric regions of tripod domain. Right: One unfolded stack.

Symmetries are very useful in cutting down the size of problems. Symmetry does not mean that all minimizing films will be symmetric (recall the case of the cubical frame). One can get a symmetric soap film by taking a symmetric average of a film, but the result will generally be a real film instead of an integer film. In general, the set of minimizing real films is a convex set, since any convex combinations of minimizers is trivially also a minimizer. It is the extreme points of the solution set which we are usually interested in.

4. DISCRETE MODELS

For numerical calculations, the continuous problems will be discretized into a standard optimization problem known as the "minimizing sum of norms" problem, or MSN for short. The goal of MSN is to minimize the sum of Euclidean norms of κ ν-dimensional affine transforms of an n-dimensional vector X, subject to p linear constraints. In our applications, ν will be the dimension of the ambient space ($\nu = 2, 3$), and X a large-dimensional vector representing the solution to the problem.

The primal problem is:

$$\begin{aligned} &\text{Minimize} \sum_{i=1}^{\kappa} \|A_i X - B_i\| \\ &\text{with } EX = C, \end{aligned} \tag{4.1}$$

where $X \in \mathbb{R}^n$, $A_i \in \mathbb{R}^{\nu \times n}$, $B_i \in \mathbb{R}^{\nu}$, $E \in \mathbb{R}^{p \times n}$, and $C \in \mathbb{R}^p$.

The dual problem is:

$$\begin{aligned} &\text{Maximize} \sum_i B_i^T Y_i + C^T Z \\ &\text{with} \sum_i A_i^T Y_i + E^T Z = 0 \text{ and } \|Y_i\| \le 1, \end{aligned} \tag{4.2}$$

where $Y_i \in \mathbb{R}^{\nu}$ and $Z \in \mathbb{R}^p$ are the variables. (The vectors Y_i are not to be confused with the covering space Y.) More general formulations are possible, but the software available to me uses the one above.

We will define three discretizations of the minimal surface problem:

- The first model finds a piecewise linear approximation to the exterior function $Q : Y \to \mathbb{R}$, and hence provides an upper bound.
- The second model finds a piecewise constant vector field, providing a lower bound on the total flux.
- The third model uses piecewise linear vector fields, providing better lower bounds.

The second model is included because it is a gentler introduction to the ideas involved, although the third model performs much better in practice. The models are phrased in terms of arbitrary dimension, but one-dimensional films in two-dimensional space illustrate all the ideas.

In all models, the N-dimensional space M is triangulated into N-dimensional simplices in a manner consistent with the boundary set B. Let the set of simplices be indexed by Greek subscripts $\alpha, \beta, \ldots$. Let the vertices be $V = v_1, \ldots, v_K$. All simplices will share the positive orientation of M. We will also need to refer to $(N-1)$-dimensional faces between simplices, and these will be indexed by ordered pairs $\alpha\beta$. Index 0 will be used for missing simplices outside M. Triangulations of Y and W are lifted from M. Let s be the number of sheets in the covering space Y. In general, it is not possible to assign sheet numbers to simplices of Y so adjacent simplices have the same sheet number, so we will not try. Instead, the simplices of Y over a simplex of M will arbitrarily labeled with sheet numbers, although in practice a useful assignment is made. Points of the film boundary B will be branch points of Y. We will require that any simplex has no more than $N-1$ branch points among its $N+1$ vertices.

The Upper Bound Model

We seek a piecewise linear scalar function Q on Y such that $Q(y) = 1$ for $y \in S$ and $Q(y) = 0$ for $y \in \partial Y - S$. The objective is to minimize the mass of a flat $(N-1)$-chain H in W that projects to the boundary of such a Q. The $(N-1)$-chain H need not be the boundary of any N-chain in W, but in this model we know H can be represented as a piecewise constant vector field $\vec{f}$ integrated over Lebesgue measure. The problem thus is:

$$\begin{aligned} &\text{Minimize } \int_W \|\vec{f}\| \\ &\text{with } \quad \pi_\# \vec{f} = \operatorname{grad} Q, \\ &\qquad Q(y) = \begin{cases} 1 & \text{for } y \in S, \\ 0 & \text{for } y \in \partial Y - S. \end{cases} \end{aligned}$$

We will define Q by its values q_k at the vertices v_k of Y, and to be piecewise linear on simplices. To evaluate the objective, introduce a vector $\vec{f}_{\alpha ij}$ for each pair of sheets ij over simplex α of M. Assume $\vec{f}_{\alpha ij} = -\vec{f}_{\alpha ji}$. Then the problem becomes:

$$\begin{aligned} &\text{Minimize } \sum_\alpha \sum_{i<j} \|\vec{f}_{\alpha ij}\| \operatorname{vol} \alpha \\ &\text{with } \operatorname{grad} Q = \sum_j \vec{f}_{\alpha ij} \text{ for each simplex } \alpha \\ &\qquad\qquad\qquad\qquad \text{and each sheet } i, \\ &\qquad Q(y) = \begin{cases} 1 & \text{for } y \in S, \\ 0 & \text{for } y \in \partial Y - S. \end{cases} \end{aligned}$$

The MSN vector X consists of all the q_k and all the $f_{\alpha ij}$.

Another formulation with fewer variables and constraints is to introduce vectors $\vec{F}_{\alpha ij}$ for $2 \le i < j \le s$, again with $\vec{F}_{\alpha ij} = -\vec{F}_{\alpha ji}$. Then the problem is

$$\begin{aligned} &\text{Minimize } \sum_\alpha \sum_{i<j} \left\| \frac{\operatorname{grad} Q_i - \operatorname{grad} Q_j}{s} + \vec{F}_{\alpha ij} \right\| \operatorname{vol} \alpha \\ &\text{with } \sum_i \operatorname{grad} Q_i = 0, \\ &\qquad Q(y) = \begin{cases} 1 & \text{for } y \in S, \\ 0 & \text{for } y \in \partial Y - S, \end{cases} \end{aligned}$$

where we are to understand

$$F_{\alpha 1j} = -\sum_{2 \le k} F_{\alpha kj}.$$

Note that the $F_{\alpha ij}$ are chosen to span the nullspace of $\pi_\#$. The relation between the two formulations is simply

$$f_{\alpha ij} = \frac{1}{s}(\operatorname{grad} Q_i - \operatorname{grad} Q_j) + \vec{F}_{\alpha ij}.$$

The linearity of Q is awkward around a branch point, since Q is changing very rapidly there. For this reason a modification is introduced whereby the value of Q at a branch point is not defined, but Q is defined on a simplex adjacent to a branch point by its values on the nonbranch vertices, by means of linear interpolation between them and level sets parallel to the faces containing all the branch points. That is, in barycentric coordinates with vertices $v_0, \dots, v_p$ on the branch set,

$$Q(\lambda_0 v_0 + \cdots + \lambda_N v_N) = \frac{\lambda_{p+1} q_{p+1} + \cdots + \lambda_N q_N}{\lambda_{p+1} + \cdots + \lambda_N}.$$

Because the piecewise linear Qs are a subset of all possible flat N-chains, solving the primal problem (4.1) gives an upper bound to the solution of the continuous problem. The dual problem (4.2) solved at the same time does not produce a feasible flow. It is only the dual problem of the discrete upper bound problem; it is not a discretization of the lower bound flow problem.

The Lower Bound Constant Vector Field Model

Here we assume a flow that has a constant value $\vec{u}_{\alpha i}$ on each simplex α on sheet i of Y. The divergenceless condition is trivial inside simplices, so we only need to require matching fluxes across faces. Also there is the bound on the difference of flows between pairs of sheets. The objective is to maximize flux through the reference surface S. The result is the dual MSN problem (4.2):

$$\begin{aligned} &\text{Maximize } \sum_{\text{faces } A_{\alpha 0 i} \in S} \vec{u}_{\alpha i} \cdot \vec{A}_{\alpha 0 i} \\ &\text{with } (\vec{u}_{\alpha i} - \vec{u}_{\beta i}) \cdot \vec{A}_{\alpha\beta i} = 0, \\ &\qquad\qquad \|\vec{u}_{\alpha i} - \vec{u}_{\alpha j}\| \le 1 \text{ for sheets } i, j. \end{aligned}$$

The notation gets a little awkward here, since sheet labelings are not consistent across faces.

Unfortunately, the software available to me at the moment can only handle bounds on the norms of single variables, not on linear combinations. So we have to introduce vectors $\vec{w}_{\alpha ij} = \vec{u}_{\alpha i} - \vec{u}_{\alpha j}$, and vectors $\vec{t}_\alpha = \sum_i \vec{u}_{\alpha i}$. Then

$$u_{\alpha i} = \frac{1}{s}\Big(\vec{t}_\alpha + \sum_j \vec{w}_{\alpha ij}\Big).$$

Hence the problem becomes:

$$\text{Maximize} \sum_{\text{faces } A_{\alpha 0i}\in S} \frac{1}{s}\Big(\vec{t}_\alpha + \sum_j \vec{w}_{\alpha ij}\Big)\cdot \vec{A}_{\alpha i0}$$

with

$$\Big(\frac{1}{s}\Big(\vec{t}_\alpha+\sum_j \vec{w}_{\alpha ij}\Big)-\frac{1}{s}\Big(\vec{t}_\beta+\sum_j \vec{w}_{\beta ij}\Big)\Big)\cdot\vec{A}_{\alpha\beta i} = 0,$$

$$\vec{w}_{\alpha 1i}+\vec{w}_{\alpha ij}+\vec{w}_{\alpha j1}=0 \text{ for } 2\le i<j<s, \qquad \|\vec{w}_{\alpha ij}\|\le 1.$$

The second set of constraints here is necessary and sufficient for the $\vec{w}_{\alpha ij}$ to be differences of $\vec{u}$'s. It is possible to eliminate the $\vec{t}$'s by subtracting $\vec{t}/s$ from each $\vec{u}$. But in practice we will want to delete large portions of Y that are not critical for the film, leaving Y to be an uneven covering space. Then the $\vec{t}$ elimination does not work. So, in practice, s is really s_α, depending on the simplex of M.

Solving the dual MSN problem (4.2) gives a feasible flow for the continuous problem, hence a lower bound on the continuous area. The simultaneous solution of the primal problem (4.1) that is generated is not a feasible film, but it should be an approximation of the optimal continuous film. The primal solution generates a vector for each simplex pair in W, whose magnitude is the mass of the film. For visualizing this approximate film, the film can be projected to Y with $\pi_\#$.

The Lower Bound Linear Vector Field Model

Here we assume a flow that is linear on each simplex of Y. The flow variables are a vector $\vec{u}_{\alpha i\tau}$ at each vertex of each simplex αi of Y. Here τ is the vertex index within a simplex. There is a separate vector for each simplex containing a given vertex. Inside the simplex, the flow is given by linear interpolation,

$$\vec{u} = \sum_{\tau=1}^{N+1} \lambda_\tau \vec{u}_{\alpha i\tau},$$

where the λ_τ are barycentric coordinates. The divergenceless condition inside a simplex turns out to be

$$\sum_\tau \vec{u}_{\alpha i\tau}\cdot\vec{A}_{\alpha i\tau} = 0,$$

where $\vec{A}_{\alpha i\tau}$ is the normal vector (proportional to area) of the face opposite vertex τ. Divergencelessness across faces can be guaranteed by requiring matching fluxes at each endpoint of each face. The flux through a face is the same as for a constant flow equal to the average of the flows at the vertices of the face. Thus the problem becomes:

$$\begin{aligned}\text{Maximize} \sum_{\text{faces } A_{\alpha 0i}\in S} \frac{1}{N}\sum_{\tau\in A_{\alpha 0i}} \vec{u}_{\alpha i\tau}\cdot\vec{A}_{\alpha 0i}\\ \text{with } \sum_\tau(\vec{u}_{\alpha i\tau}-\vec{u}_{\beta i\tau})\cdot\vec{A}_{\alpha\beta i\tau}=0,\\ (\vec{u}_{\alpha i\tau}-\vec{u}_{\beta i\tau})\cdot\vec{A}_{\alpha\beta i}=0,\\ \|\vec{u}_{\alpha i\tau}-\vec{u}_{\alpha j\tau}\|\le 1.\end{aligned}$$

Again, due to software limitations, we have to introduce vectors $\vec{w}_{\alpha ij\tau} = \vec{u}_{\alpha i\tau} - \vec{u}_{\alpha j\tau}$, and vectors $\vec{t}_{\alpha\tau} = \sum_i \vec{u}_{\alpha i\tau}$. Then

$$u_{\alpha i\tau} = \frac{1}{s}\Big(\vec{t}_{\alpha\tau} + \sum_j \vec{w}_{\alpha ij\tau}\Big).$$

Hence the problem becomes

$$\text{Maximize} \sum_{\text{faces } A_{\alpha i0}\in S} \frac{1}{sN}\sum_{\tau\in A}\Big(\vec{t}_{\alpha\tau}+\sum_j \vec{w}_{\alpha ij\tau}\Big)\cdot\vec{A}_{\alpha i0}$$

with

$$\Big(\sum_\tau\Big(\vec{t}_{\alpha\tau}+\sum_j \vec{w}_{\alpha ij\tau}\Big)-\sum_\tau\Big(\vec{t}_{\beta\tau}+\sum_j \vec{w}_{\beta ij\tau}\Big)\Big)\cdot\vec{A}_{\alpha\beta i}=0,$$

$$\Big(\frac{1}{s}\Big(\vec{t}_{\alpha\tau}+\sum_j \vec{w}_{\alpha ij\tau}\Big)-\frac{1}{s}\Big(\vec{t}_{\beta\tau}+\sum_j \vec{w}_{\beta ij\tau}\Big)\Big)\cdot\vec{A}_{\alpha\beta i}=0,$$

$$\vec{w}_{\alpha 1i\tau}+\vec{w}_{\alpha ij\tau}+\vec{w}_{\alpha j1\tau}=0 \text{ for } 2\le i<j<N,$$

$$\|\vec{w}_{\alpha ij\tau}\|\le 1.$$

As in the constant vector field model, we do not want to normalize $\vec{t}$ away since Y may be only an uneven cover of M.

As with the constant flow model, solving the dual MSN problem (4.2) gives a feasible flow for the continuous problem, hence a lower bound on the continuous area. The simultaneous solution of the primal problem (4.1) that is generated is not a feasible film, but it should be an approximation of the optimal continuous film. The primal solution generates a vector for each vertex of each simplex pair in W whose magnitude is the mass of the film. (See Figures 5 (bottom), 7, and 8, where the approximate film is shown projected to Y.)

5. IMPLEMENTATION

This section describes a particular implementation of the methods introduced in the preceding section, using the Surface Evolver [Brakke 1992], Knud Andersen's GOPT package [Andersen 1995] (used to solve MSN problems), and custom programs. All the programs, apart from the Evolver and GOPT, are still in their early stages, and are changing rapidly, so they are not described in detail here.

Two Dimensions

A typical run for a problem in two dimensions is schematically shown in Figure 3. The starting point is a file in Evolver data format that defines the basic simplicial structure for Y, listing the initial triangulation, source edges, sink edges, fold edges, and branch points. This structure is then suitably refined by the Evolver. Other programs could be used for this task; I use the Evolver because of my familiarity with it and its wide range of triangulation manipulation and visualization features.

After Y is refined, the script trimake.cmd, written in the Evolver command language, writes out the geometry in a custom format appropriate for input to the next stage, tri2mps. This latter program produces a file in the standard MPS format, for input to GOPT; it can produce MPS files for all three types of models. The GOPT program then processes the MPS file and creates report files containing solutions to both the primal and dual problems.

The report files generated by GOPT are merged with an Evolver dump file (created at the time of the initial Evolver processing) by the program out2flm, which produces an Evolver data file incorporating the GOPT solution. The output of out2flm depends on the model. For the upper bound model, the z-coordinate of each vertex is set to the value of the exterior function Q there, and a film mass for each triangle is derived from the dual solution. The dual solution actually has a mass for each pair of triangles in a stack (corresponding to the bound on the magnitude of the difference of the

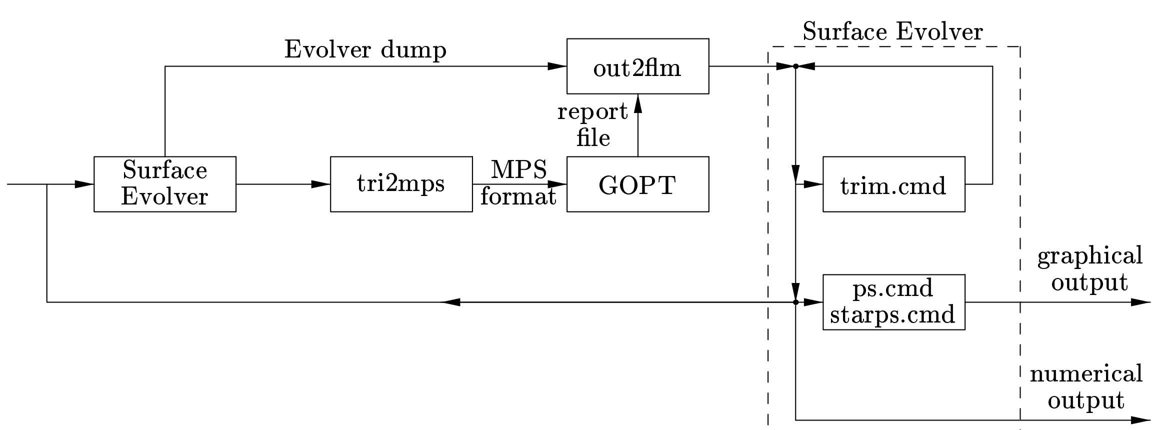


FIGURE 3. Schematic overview of programs used in two-dimensional experiments.

Q gradients), and out2flm gives this mass to each triangle of the pair. For the lower bound model, each triangle gets a flow vector and a film vector. For the linear flow model, each triangle gets the average of the three flows at its corners, and the total of the three film masses.

The output of out2flm can be loaded back into the Evolver, with optional *trimming* by an Evolver script, trim.cmd. Trimming means that zero mass triangles along the source and sink edges are removed: trim.cmd identifies such triangles, deletes them, and declares the new exposed edges to be source or sink edges as appropriate. The result is a covering space with an uneven covering of the base space, but the models are set up to be able to handle that. This is very useful in reducing the size of the numerical problem, particularly at higher refinements. Trimming can't hurt the upper bound, since any legal exterior function on the trimmed space can be trivially extended to the original space, but it can give an invalid lower bound, since it is not guaranteed to be possible to extend vector fields. Nonetheless, trimming is still very useful in the lower bound model when exploring to find an unknown film. The output shown in Figure 9 has undergone trimming.

The Evolver can also be used to generate graphical output in Postscript, through scripts ps.cmd and starps.cmd, which assemble the sheets of Y in different ways. Figures 5 and 7 show examples of output.

Higher Dimensions

There is a similar set of programs written for higher dimensions. We briefly discuss their use in a three-dimensional problem. They require the use of the simplex model of the Evolver. The main difference from the two-dimensional case is that here the counterpart of edges, namely faces, is not available. Hence vertices are labeled according to whether they are on the sink or source. Branch points can be on both. Which $(N-2)$-dimensional faces are sink or source faces has to be deduced from this information by sim2mps, the higher-dimensional analog of tri2mps. This program takes an input simplex file output by the Evolver and produces an MPS file for GOPT. At the moment, the only model supported is the lower bound linear flow model. The simplex file format is a simplified version of the triangle file format.

Next, out3flm reads the GOPT report, and can produce several types of output. One is an Evolver data file, with film and flow data merged into a previous Evolver dump file. Moreover, out3flm can produce files for the three-dimensional visualization program geomview [Phillips et al. 1993]: one file for the flow, one for the film, and one for the simplicial skeleton. Postscript files can then be obtained from geomview. Figure 11 shows an example of graphical output.

Reliability of Results

In numerical calculations, there is always the question of the accuracy of results, and when one should believe one has found the solution. In one sense, the calibration results are extremely reliable for estimating the mass of the minimizer. The upper and lower bound discrete models provide strict bounds theoretically. GOPT gives a "duality gap" for each discrete problem, so one has strict bounds on the true objective value of the discrete problems. Further, one can check that the solutions given by GOPT do indeed satisfy the necessary constraints.

On the other hand, the location of films found numerically is not so certain. There is no need for the current solving the discrete approximation to be near in flat norm to the true minimizer. I know of no theorems that restrict the location of minimizers, except the classical minimal surface barrier theorems [Morgan 1995, 10.4], which are usually not very informative for the covering space model. However, GOPT uses an interior point algorithm that tends to converge to the center of the solution set of currents. Hence the GOPT solution should be a superposition of all possible solutions. So far no situations have turned up where the numerical solution is misleading.

level	verts	tris	norms	vars	eqs	objective	gap	time
1	5	3	3	1	0	1.2159728	$1. \times 10^{-8}$	0.00
2	12	12	12	11	2	1.0399747	$1. \times 10^{-11}$	0.04
3	35	48	48	52	9	1.0187818	$1. \times 10^{-11}$	0.17
4	117	192	192	218	35	1.0087497	$3. \times 10^{-11}$	1.19
5	425	768	768	886	135	1.0040750	$1. \times 10^{-10}$	5.35

TABLE 1. Results obtained with the upper bound model for the tripod. The first three columns refer to Evolver output: refinement level and total numbers of vertices and triangles for all sheets. The remaining columns refer to GOPT output: number of Euclidean norms in the MSN problem, number of variables in the primal problem, number of equality constraints, value of the objective function (total flux or mass), gap between the primal and dual solutions of the MSN problem, and GOPT solution time in seconds on an SGI Indigo 2.

6. EXPERIMENTAL RESULTS

Tripod Upper Bound in Two Dimensions

The tripod ($H_{12} + H_{23} + H_{32}$ in Figure 1) is the minimal one-dimensional film connecting the three vertices of an equilateral triangle. For numerical purposes, the domain is taken to be the convex hull of the vertices, namely, the equilateral triangle. Further, symmetry is used to reduce the problem by a factor of six. There are three regions, hence three sheets. The full equilateral triangle is initially divided into its six symmetric fundamental regions, each a 30°-60°-90° triangle, making 18 triangles on all sheets together. By sixfold symmetry, each stack of triangles is equivalent, so we need to solve only one stack. The minimum of the objective in the continuous case is known to be 1. Experimental results are given in Table 1.

One conclusion that can be drawn here is that the upper bound model converges very slowly, at least as presently set up. The basic problem is that a continuous piecewise linear function is trying to approximate a step function. Faster convergence could probably be obtained by selective refinement of key regions. It may also be necessary to exert much more control over the directions of the edges in the triangulation, to permit Q to bend the way it wants. Due to its relatively poor performance, the upper bound model is neglected for the rest of this paper.

The Hexagon

The integer film joining the vertices of a regular hexagon is known to consist of five of the hexagon's sides, as shown in Figure 4, left. However, all attempts to calibrate it have previously failed. Does this very simple film have a complicated calibration, or is there some fractional density film of lower mass? The paired calibration setup for the hexagon has six regions, and the reference surfaces are the six sides, $S_1, \ldots, S_6$, shown in Figure 4, middle. Using the 12-fold symmetry of the problem, we reduce to one stack of six triangles, shown in Figure 4, right. The stack unfolds to a single sheet, shown in Figure 5. Edge AB is the source edge, edges BC, CD, DE, EF, FG are sink edges,

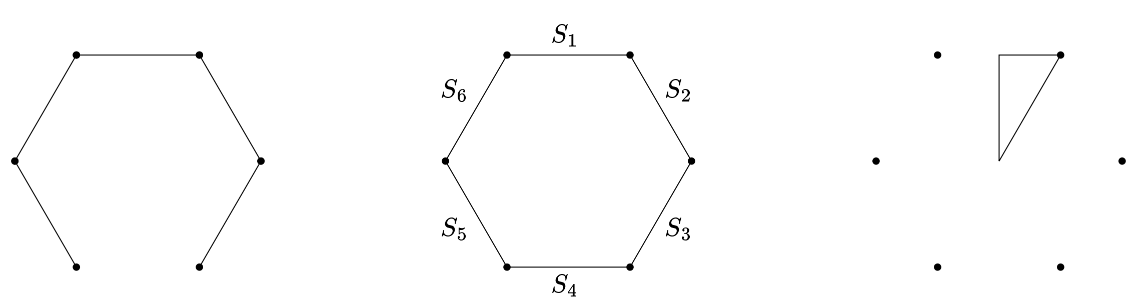
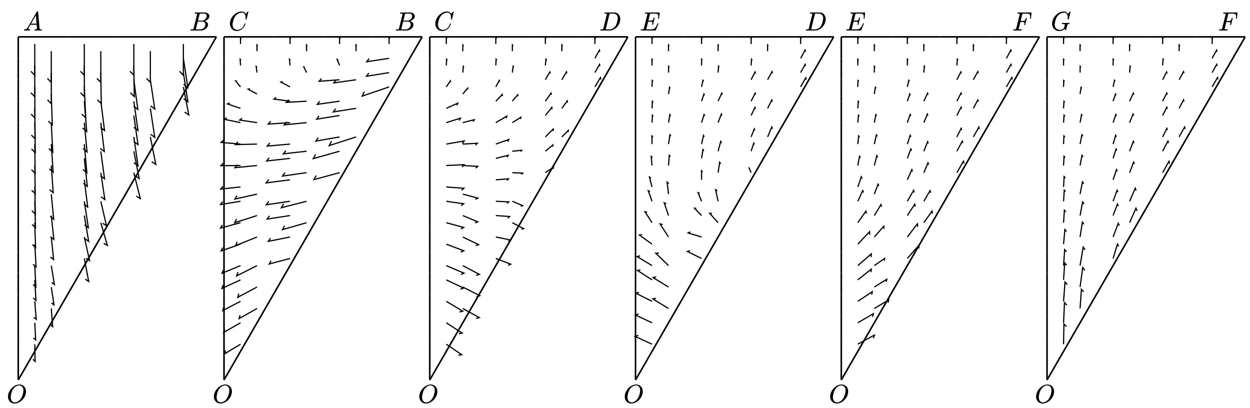
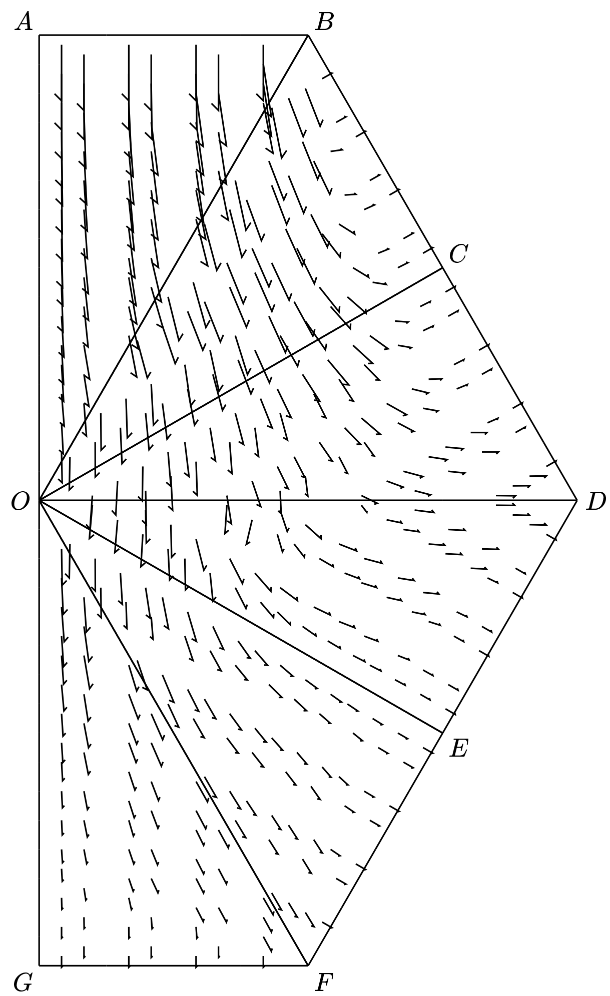


FIGURE 4. Left: Integer film spanning hexagon. Middle: Reference surfaces. Right: Fundamental region.





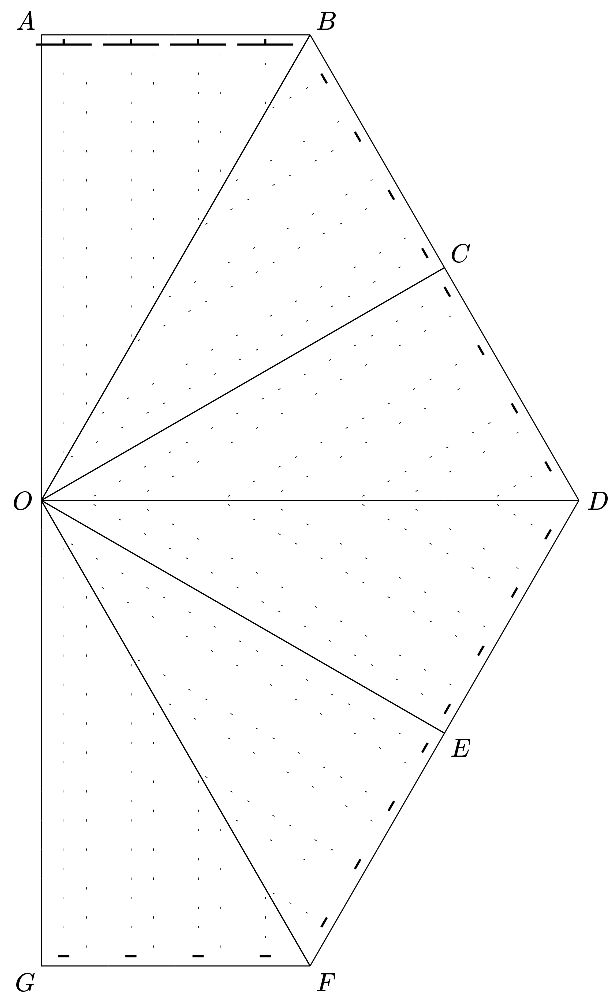


FIGURE 5. Top: Calibrating flow for the hexagon (from level 3 of the linear model). The flow is shown on separate sheets covering a fundamental region. Bottom left: Same flow laid out on one sheet. Bottom right: The corresponding symmetric film (the average of six integer films). It has density $\frac{5}{6}$, coming from five interface pairs of density $\frac{1}{6}$ each. We see the density $\frac{5}{6}$ along edge AB, and the other projections of the pairs give density $\frac{1}{6}$ along the sink edges.

			lower bound constant flow model						lower bound linear flow model					
level	verts	tris	norms	vars	eqs	objective	gap	time	norms	vars	eqs	objective	gap	time
1	20	24	60	117	8	0.3993794928	$2. \times 10^{-11}$	0.36	180	337	23	0.4140262160	$3. \times 10^{-11}$	1.30
2	63	96	240	466	32	0.4114536593	$2. \times 10^{-11}$	2.25	720	1346	94	0.4166362460	$8. \times 10^{-11}$	7.50
3	221	348	960	1860	128	0.4159658500	$1. \times 10^{-10}$	9.76	2880	1860	128	0.4166666667	$3. \times 10^{-10}$	43.73
4	825	1536	3840	7432	512	0.416663774	$3. \times 10^{-9}$	71.24						

TABLE 2. Results obtained with the lower bound models for the hexagon. See Table 1 for the meaning of the columns.

AO, OG are zero flux edges, and the other edges are folds.

The hexagon was run with both the lower bound constant flow and lower bound linear flow in order to compare the performance of the two models. Results are in Table 2. If the integer film is indeed the minimum, the limit value of the objective function should be $\frac{5}{12} = 0.41\bar{6}$. The right-hand column of the table certainly corroborates this conjecture.

The flow from level 3 of the linear model is shown in Figure 5, top and bottom left, and it is not something to design by hand. The same figure shows also the corresponding symmetric film, the average of all six integer films.

The linear flow model is obviously much more efficient than the constant flow model. Hereafter, all examples will be linear flow.

Crossed Tripods

This is an example where the minimal integer film is not the minimal real film. It is somewhat contrived, but it is the two-dimensional analog of what happens in the octahedron. Suppose the minimal real film joining the vertices of a regular hexagon were as shown in Figure 6. (This is in fact better than Figure 4 (left) for a large hexagon in the hyperbolic plane.) Any calibration would have to calibrate both of these films simultaneously. In particular, at the center it would have to calibrate two tripods simultaneously, in the configuration shown in Figure 7, left, called the *crossed tripods.*

The set of vertices of this configuration forms a regular hexagon, but each reference surface is a pair of sides, joining alternate vertices. One tripod setup has reference surfaces S_1, S_2, S_3, and the other has S_4, S_5, S_6 (Figure 7, right). The minimal integer film for this problem is presumably the crossed tripods of Figure 7, left. However, a couple of years ago, I found the fractional density film shown in Figure 8, left. It has all films of density $\frac{1}{2}$, and less mass than the crossed tripod.

However, an even better film, shown on the right in Figure 8, was obtained by running the lower bound linear flow model. More precisely, I ran the model on the data of Figure 7, right, modded out by the 12-fold symmetry, as in the previous example of the hexagon. Figure 9 shows the experimentally obtained flow and film. The source edges for

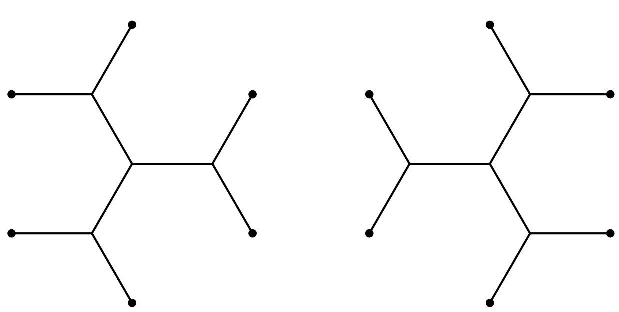


FIGURE 6. Hypothetical minimal films spanning the hexagon.

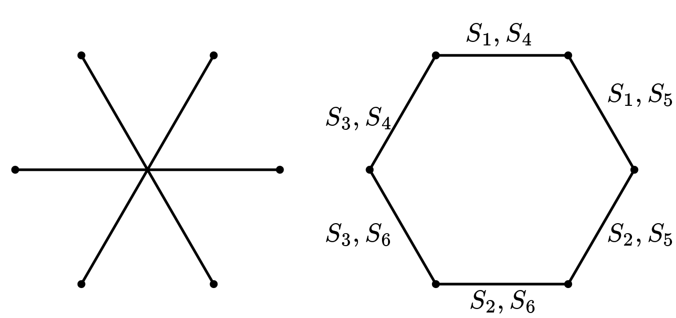


FIGURE 7. Left: Crossed tripods. Right: Reference surfaces for the crossed tripods.

level	verts	tris	norms	vars	eqs	objective	gap	time
1	8	6	45	79	5	0.49487172015	$2. \times 10^{-11}$	0.30
2	21	24	180	326	22	0.49487151606	$2. \times 10^{-11}$	1.65
3	65	96	720	1324	92	0.4954586682	$2. \times 10^{-10}$	6.02
4	225	384	2880	5336	376	0.4959077068	$6. \times 10^{-10}$	48.53
4*	78	119	429	836	245	0.49590770677	$3. \times 10^{-11}$	3.94
5*	238	412	1428	2885	903	0.4961183292	$2. \times 10^{-10}$	20.30
6*	679	1232	4416	8987	2719	0.496216250	$2. \times 10^{-9}$	118.54
7*	1981	3728	12060	25133	8655	0.496265	$1. \times 10^{-6}$	342.42

TABLE 3. Results for lower bound linear flow for crossed tripods. See Table 1 for the meaning of the columns. In the first column, an asterisk indicates that the domain was trimmed.

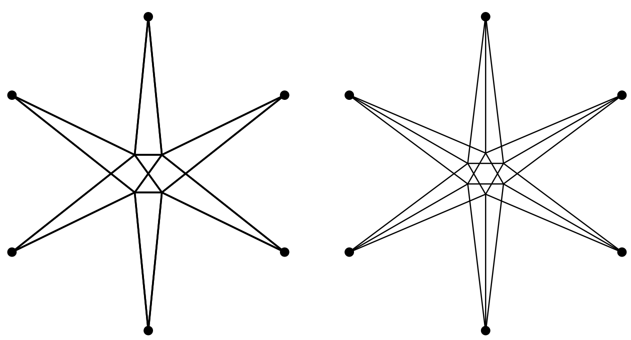


FIGURE 8. Left: Best film found by hand. Right: Best film found by computer.

the flow are ABC, the sink edges $CDEFG$, and the zero flux edges GOA. There are three equal films coming out of vertex C, so each has density $\frac{1}{3}$. By putting the whole film together from the piece that appears in Figure 9 one gets Figure 8, right. The value of the objective function, reported in Table 3, converges (when extrapolated to infinite refinement) to something consistent with 0.496324707689899, or $\frac{1}{12}$ of the total mass of the film of Figure 8, right, as reported by the Evolver. It seems therefore that this film is the absolute minimum. (The film appears somewhat smeared out in the bottom part of Figure 9, but that is probably due the discretization.)

The Octahedron

At least five different integer films can span an octahedral frame. Of these, the one with the smallest area is shown in two orientations in Figure 10. It consists of flat pieces, with a tetrahedral point in the center. The two views show the two different orientations possible for this film. Recall that any calibration would have to calibrate both these films simultaneously, hence calibrate two superposed tetrahedral points. This is the three-dimensional version of the crossed tripod problem. Calibrating by hand has failed, and attempts to find a three-dimensional analog to either film in Figure 8 have also failed. So we turn to the linear flow model. There are eight regions. Using the 48-fold symmetry, we need only compute the flow on $\frac{1}{6}$ of one face. The results are tabulated in Table 4, and the film is shown in Figure 11.

The corresponding mass of the integer density film is 0.235702260395516, but that is not close enough to claim calibration. The film in Figure 11

level	verts	tris	norms	vars	eqs	objective	gap	time
1	10	8	112	307	11	0.23231724054	$2. \times 10^{-11}$	7.24
2	63	96	896	2461	93	0.2341254003	$2. \times 10^{-10}$	126.30
3	165	512	7168	19702	758	0.235256244	$1. \times 10^{-9}$	1523.53

TABLE 4. Results for lower bound linear flow for crossed tetrahedra. See Table 1 for the meaning of the columns.

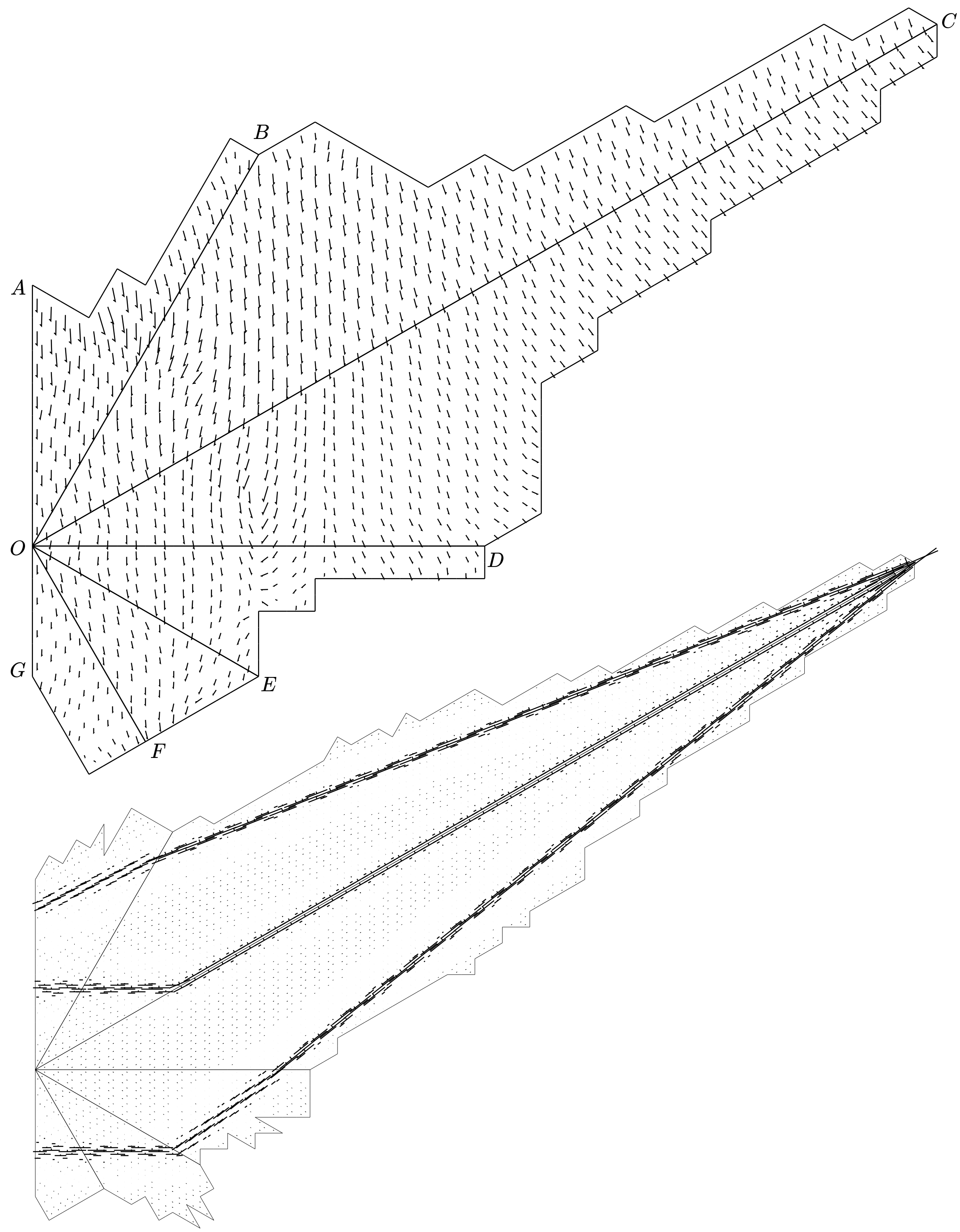


FIGURE 9. Top: Crossed tripod flow on one sheet of a trimmed domain (level 6$*$ in Table 3). Bottom: Fundamental region of minimal real film for crossed tripod problem (level 7$*$ in Table 3).

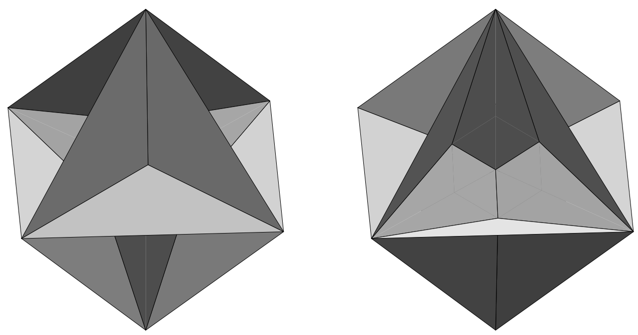


FIGURE 10. Two views of the presumed minimal film on octahedron.

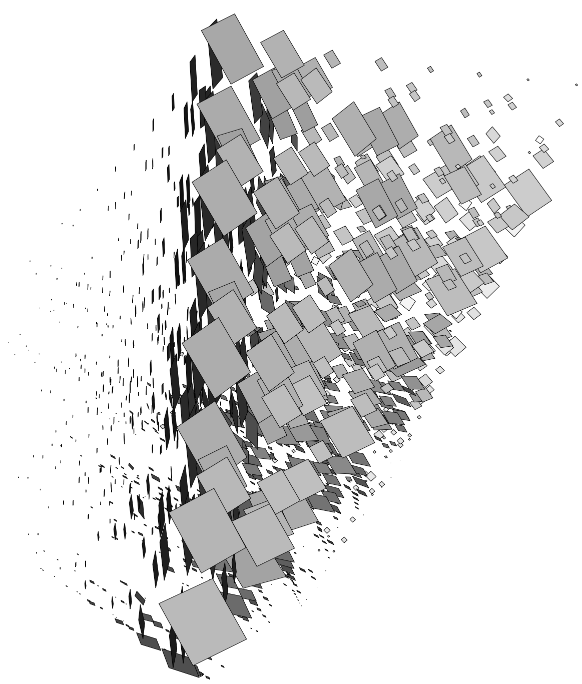


FIGURE 11. Unfolded fundamental piece of octahedral film (level 3 in Table 4).

looks like a cloud of surface bits. It might be approximating the integer film, but it might not.

This example does have an unexploited symmetry, namely homothetic symmetry. If there is a calibration, there will be one that is invariant radially, in the sense that $\vec{v}(x) = \vec{v}(cx)$ for any $c > 0$. This cuts the dimension of the problem from three to two, permitting higher refinements of the triangulation. I have implemented this conical model in a lower bound linear flow, the details of which are omitted here. The objective value for crossed tetrahedra is 0.353553390593274. Of all the refinement schemes tried, the best series of results is this:

level	objective
3	0.35354431624
4	0.35354631148
5	0.35354758418
6	0.35354843

This gets close to the crossed tetrahedra value, but one couldn't claim it is converging to it.

I have also tried a discretization of the upper bound model dual to the conical lower bound, but was not able to get below the crossed tetrahedral value.

The octahedron is the example that motivated all this numerical work, but the octahedron film must still be regarded as unresolved.

7. RELATED WORK

Many schemes have been proposed for numerically calculating minimal surfaces, but very few require no assumption on the topology of the surface, and of those few none can handle soap film singularities.

John Sullivan [1990] proposed a scheme in which the surface spanning a polygonal boundary is chosen from a large set of small polygonal surface elements of various orientations and locations. The actual selection of the set is done by a max-flow min-cut algorithm for a flow linking the boundary. Since an actual spanning surface is found, the scheme provides an upper bound on the area of the smooth absolute minimizer. The accuracy is limited by the available orientations of the set of surface elements, which is similar to some of the limitations in the upper bound schemes of this paper.

Harold Parks [1977; 1986; 1992] has developed a scheme in which minimal surfaces are the level sets of a function of bounded variation that minimizes the L_1 norm of its gradient in a convex domain with given boundary values. The implementation in [Parks 1992] represents the function as continuous piecewise linear on a simplicial decomposition of the domain. The discrete gradient norm minimization turns out to be the same minimizing-sum-of-norms problem that appears in this paper, although the algorithm and software used to solve it are different. The method finds a spanning surface, hence an upper bound on the true minimum area. The great advantage of this method is that by representing the desired surface as just one level set among many, an extra order of smoothness is gained. The orientation of the simplicial decomposition is not critical, since the level set can cut across a simplex in any orientation. The main limitation is that it can only handle a boundary curve on the boundary of a convex domain, and it cannot handle unoriented surfaces or soap film singularities, but within those limitations it would probably be my method of choice.

Harold Parks and Jon Pitts [1997] have another scheme to handle surfaces on arbitrary boundary curves, such as linked and knotted curves. The idea is to define some arbitrary reference surface that spans the boundary and then minimize the L_1 norm of the gradient of a function of bounded variation that has a jump of magnitude 1 across the reference surface. There will be a compensating jump across the minimal surface, and the area of the minimal surface is the L_1 norm of the gradient. In the discretization, the function is continuous and piecewise linear on a simplicial decomposition of some domain enclosing the boundary curve and reference surface, except for discontinuities across the reference surface. Again, the discrete minimization problem is minimizing a sum of norms. The accuracy of the method suffers in comparison with the convex domain scheme because here a discontinuous function with a jump of 1 at the minimal surface is being approximated by a continuous function. But there seems to be no obvious way to embed the surface in a foliation of minimal surfaces in order to gain smoothness of the bounded variation function. Essentially, this scheme is the same as the upper bound model of Section 4 of this paper, with two regions.

8. CONCLUSION

One would wish that every problem turned out as clearly as the film in Figure 9, but the examples show the need for many practical improvements. Even simple problems rapidly reach the limits of current computers. Improvements will undoubtedly be made to general optimization software such as GOPT. But cutting down the size of the problem will be far more important. The linear flow model is more efficient than the constant flow models, and higher order models should be even better. The immediate problem is how to fit higher order problems into the MSN framework. How does one guarantee a bound on the magnitude of a quadratic flow using Euclidean norms at a finite number of points? In the upper bound model, can a discrete function space be found which permits arbitrarily oriented step functions? Trimming is another technique that can be improved. Currently only simplices adjacent to source or sink edges are trimmed. As Figure 9 shows, there can be large interior areas with no film. These could be excised and replaced with a flux conservation constraint. There need to be methods of selective refinement. There also need to be "barrier theorems" that can restrict the film to a narrow region.

Although only soap films are discussed in this paper, the techniques extend to many related problems involving surfaces. The area objective function may be replaced with any positive definite quadratic form of the surface normal, and vector integrals over the surface may be included. This permits surfaces of different surface tensions, Riemannian metrics, gravitational energies, and contact angles on walls. Linear constraints may be added, volumes for example. All these will be added to the software in the future. However, a fundamental limitation will remain, that only real film minima can be treated. Thus one cannot use these techniques to solve bubble clusters, because the real film minimizer does not exist. But enough has been demonstrated to show the beginnings of a general "soap film technology."

REFERENCES

[Andersen 1995] K. D. Andersen, "Minimizing a sum of norms", Ph.D. Thesis, Odense University, 1995.

[Brakke 1991] K. A. Brakke, "Minimal cones on hypercubes", *J. Geom. Anal.* **1** (1991), 329–338.

[Brakke 1992] K. A. Brakke, "The Surface Evolver", *Experimental Mathematics* **1** (1992), 141–165.

[Brakke 1995] K. A. Brakke, "Soap films and covering spaces", to appear in *J. Geom. Anal.* Preprint available at http://www.geom.umn.edu/docs/preprints/online/GCG54.html.

[Federer 1969] H. Federer, *Geometric measure theory*, Grundlehren der mathematischen Wissenschaften **153**, Springer, Berlin, 1969.

[Federer 1974] H. Federer, "Real flat chains, cochains and variational problems", *Indiana U. Math. J.* **24** (1974), 351–407.

[Giusti 1984] E. Giusti, *Minimal surfaces and functions of bounded variation*, Birkhäuser, Boston, 1984.

[Harvey and Lawson 1982] R. Harvey and H. B. Lawson, Jr., "Calibrated geometries", *Acta Math.* **148** (1982), 47–157.

[Lawlor and Morgan 1994] G. Lawlor and F. Morgan, "Paired calibrations applied to soap films, immiscible fluids, and surfaces or networks minimizing other norms", *Pacific J. of Math.* **166** (1994) 55–83.

[Morgan 1995] F. Morgan, *Geometric Measure Theory: A Beginner's Guide*, 2nd ed., Academic Press, Boston, 1995.

[Parks 1977] H. R. Parks, "Explicit determination of area minimizing hypersurfaces", *Duke Math. J.* **44** (1977), 519–534.

[Parks 1986] H. R. Parks, "Explicit determination of area minimizing hypersurfaces II", *Mem. Amer. Math. Soc.* **342** (1986), 90 pp.

[Parks 1992] H. R. Parks, "Numerical approximation of parametric area-minimizing hypersurfaces", *SIAM J. Sci. Stat. Comput.* **13** (1992), 499–511.

[Parks and Pitts 1997] H. R. Parks and J. T. Pitts, "Computing least area hypersurfaces spanning arbitrary boundaries", to appear in *SIAM J. Sci. Comput.* (preprint available from authors at parks@math.orst.edu, jpitts@math.tamu.edu).

[Phillips et al. 1993] "Geomview: an interactive geometry viewer". *Notices Amer. Math. Soc.*, 985–988. See also Electronic Availability below.

[Sullivan 1990] J. M. Sullivan, *A crystalline approximation theorem for hypersurfaces*, Ph.D. Thesis, Princeton University, 1990.

[Sullivan 1995] J. M. Sullivan, "Convex deltatopes in all dimensions, and polyhedral soap films", preprint available from the author (sullivan@math.umn.edu).

[Taylor 1976] J. E. Taylor, "The structure of singularities in soap-bubble-like and soap-film-like minimal surfaces", *Ann. of Math.* **103** (1976), 489–539.

[Underwood 1993] A. Underwood, "Constructing barriers to minimal surfaces from polyhedral data", Ph.D. Thesis, Princeton University, 1993.

Software Availability

The GOPT package is available from Knud D. Andersen (kda@beta.dou.dk); there is no charge for academic use.

The Surface Evolver is available at http://www.geom.umn.edu/software/download/evolver.html. The distribution contains C source code, makefile, sample data files, and the manual. A Postscript version of the manual is also available at http://www.geom.umn.edu/docs/preprints/online/GCG71.html.

Geomview can be obtained at http://www.geom.umn.edu/software/download/geomview.html.

Other programs and files mentioned in this paper have not been publicly posted, but will be made available to interested parties.

Kenneth A. Brakke, Mathematics Department, Susquehanna University, Selinsgrove, PA 17870 (brakke@geom.umn.edu)

Received July 7, 1995; accepted in revised form October 5, 1995