SIMPLICIAL ENERGY AND SIMPLICIAL HARMONIC MAPS

JOEL HASS AND PETER SCOTT

ABSTRACT. We introduce a combinatorial energy for maps of triangulated surfaces with simplicial metrics and analyze the existence and uniqueness properties of the corresponding harmonic maps. We show that some important applications of smooth harmonic maps can be obtained in this setting.

1. INTRODUCTION

The energy of a map from one manifold to another is a measure of the total stretching of the map. Energy minimizing harmonic maps have found numerous applications in geometry, analysis, algebra and topology. The two survey papers of Eells and Lemaire [18, 19] give an introduction to the extensive literature of this subject.

One of the fundamental existence theorems states that a nontrivial map of a Riemannian manifold F to a negatively curved manifold M is homotopic to a unique harmonic map. When F is a surface, each conformal class of metrics on F gives rise to the same map, and smooth families of domains give rise to smooth families of harmonic maps [20, 22].

In this paper we introduce a new type of energy that leads to what we call a *simplicial harmonic map*. The key idea is to give a new, more combinatorial, definition of area, and then find a corresponding definition for energy. The payoff is that the existence and regularity of simplicial harmonic maps are simple to prove, but they retain enough of the features of smooth harmonic maps to be useful in applications. In this paper we focus on maps from surfaces into non-positively curved manifolds and spaces. We briefly discuss extensions to maps between manifolds of any dimension.

One motivation for studying energy from a more combinatorial point of view is to obtain numerical methods for computing minimal and harmonic surfaces. This was carried out in the work of Pinkall and Polthier [31]. See [30] for a survey of related work. The discrete harmonic maps of Pinkall and Polthier are maps of a triangulation of a domain surface into $\mathbb{R}^n$ that are linear on each 2–simplex. Their formula gives the natural energy for

1

a piecewise-linear approximation of a smooth map, and has found many uses in computational geometry [15], where harmonic maps give a preferred choice of surface map for purposes such as texturing and meshing. While the Pinkall-Polthier harmonic maps retain many of the useful features of smooth harmonic maps, particularly when the 2–simplices in the domain are acute angled, they do not in general satisfy the convex hull or mean value properties. (See section 6 for a discussion of these properties and their consequences.) An example of Polthier and Rossman [33] shows that a discrete harmonic map may fail to satisfy the convex hull property, with an interactive demonstration available at [32].

The simplicial harmonic maps we describe in this paper are defined for very general target spaces and do satisfy a convex hull property. A consequence is that if f is a simplicial harmonic immersion of a surface F into a Riemannian manifold M, then the induced curvature on F is more negative than the sectional curvature of M. When M itself is negatively curved, the Gauss-Bonnet theorem then implies that the area of f is bounded above by a constant times the Euler characteristic of F. When mapped into Euclidean space, simplicial harmonic maps also satisfy a mean value property and the maximum principle.

In this paper, we use the completeness and non-positive curvature of a Riemannian manifold M to conclude that each homotopy class of arcs in M contains a unique geodesic representative (rel boundary). There are other conditions that imply this property, and our construction then immediately applies. For example, Teichmuller space with either the Teichmuller metric or the Weil-Peterson metric has a unique geodesic connecting any two points [44], although the Weil-Peterson metric is not complete. Similarly, any two points in the interior of a hemisphere are connected by a unique geodesic arc. More generally, our methods work equally well if M is a metrized graph, a $CAT(0)$ space, or a path metric space with the property that any two points are joined by a unique geodesic in each homotopy class.

This paper is organized as follows. Initially we consider maps of triangulated compact surfaces into manifolds. In Section 2 we define the simplicial area of a map, and in Section 3 we define simplicial energy. In Section 4 we briefly discuss how to define simplicial area and energy of maps when the source surface is divided into polygons which need not be triangles. In Section 5 we define simplicial harmonic maps. In Section 6 we establish the mean value and convex hull properties of simplicial harmonic maps.

In Section 7 we consider the problem of deforming families of surfaces in 3–manifolds to surfaces of small area. Given any surface in a 3–manifold, it is possible to find a homotopy that deforms the surface to a collection of minimal surfaces joined by thin tubes. In a hyperbolic 3–manifold, there is a universal upper bound for the area of such a surface in terms of its Euler characteristic χ; a surface with $\chi < 0$ can be homotoped to have area at most $2\pi |\chi|$. But there is no canonical homotopy to such a small area surface. For many applications one would like to start with a family of surfaces and

continuously homotope the entire family of surfaces so that each surface in the family has small area. Area deformation techniques based on mean curvature and related flows are not well suited to such a process, due to the formation of singularities and the non-uniqueness of least area surfaces in a homotopy class. In [24] it was shown that smooth energy deformation techniques based on analytic results of Eells and Sampson [20] can be applied successfully in this context. We show that a simplified theory based on simplicial energy suffices to give a deformation of a family of surfaces to a family of small area surfaces. Then we extend this discussion to explain how the smooth energy deformation techniques in [24] can be replaced by simpler simplicial techniques.

In Section 8 we introduce the idea of the genus of an n–dimensional manifold, extending the notion of Heegaard genus from dimension three to all dimensions. We find connections between this n–dimensional genus and the areas of surfaces in a sweepout of an n–manifold.

The simplicial energy of a map depends on the choice of a simplicial metric on its domain. In Section 9 we show the existence of maps which are global minimizers of simplicial energy over the space of all simplicial metrics. In Section 10 we consider the problem of finding canonical parametrizations, or canonical triangulations of a surface. We solve this by showing that if a simplicial harmonic map from a closed surface of negative Euler characteristic to a non-positively curved Riemannian surface is homotopic to a homeomorphism, then it is itself a homeomorphism. Then we give applications of this result to study the space of straight triangulations of a compact surface F with a Riemannian metric of non-positive curvature. (A triangulation of F is called straight if each edge is a geodesic arc.) We show that the space of straight triangulations of F of a fixed combinatorial type is connected. We note that in the case when F is the disk with a flat metric, Bloch, Connelly and Henderson [6] proved much more, showing that this space is homeomorphic to some Euclidean space and so contractible. We thank Igor Rivin for telling us about this work, and for suggesting that our ideas could be useful for studying the space of straight triangulations of more general surfaces. Finally, in Section 11 we briefly discuss generalizations to maps of higher dimensional manifolds.

2. Simplicial area for surfaces

In this section we define the simplicial area of a compact 2–dimensional surface F. In place of a Riemannian metric on F, we specify a triangulation τ and a map l that assigns to each edge e_i, $1 \leq i \leq r$, of τ a length $l_i = l(e_i) > 0$, with the lengths l_i satisfying the triangle inequality for each triangle of τ. We call such an assignment a *simplicial metric* on F, and denote it by (F, τ, l), or just l when the context is clear. Note that we allow the possibility that for some triangles one of the triangle inequalities is an equality. We do not require the triangulation τ to be combinatorial. Thus

we allow triangulations where two, or even all three, vertices of a single triangle coincide in F, or where a pair of triangles intersect in two or three vertices. But we assume that τ has the structure of a Δ–complex [25]. This is not a serious restriction. This means that F is constructed from a disjoint union of triangles, each of whose vertices is ordered, by identifying faces of the triangles in an order preserving way. Thus the given orderings can be extended to an ordering of all the vertices of τ. For simplicity we will always discuss triangles in Δ–complexes as if they were embedded in F. If a triangle is not embedded in F, one needs to consider the source triangle in the initial disjoint union of triangles. Each simplex in a Δ–complex has a minimal (in the given order) vertex, which we call its preferred vertex. When considering limits of simplicial metrics, it is convenient to also allow edges in the triangulation that are assigned zero length. We call such an extension a *simplicial quasi-metric* on F.

We wish to assign a meaning to the area of (F, τ, l). A natural choice for the area of a triangle T with edge lengths a, b, c is the area of the Euclidean triangle with the same edge lengths. This is given by the classical Heron's formula, see [3] for example,

$$A(T) = \sqrt{\frac{(a+b+c)(a+b-c)(b+c-a)(a+c-b)}{16}}.$$

We will however use an alternate formula, one that is much simpler, but still useful. We define the *simplicial area* A_S of the triangle T by the formula

$$A_S(T) = ab + bc + ca.$$

We note that the Euclidean area A of a triangle is at most $xy/2$, for any pair x, y of edge lengths. So $A \leq ab/2$, $A \leq bc/2$ and $A \leq ca/2$. Equality occurs when an appropriate angle is a right angle. Adding these inequalities yields the inequality

$$A < A_S/6.$$

One defect of our definition of the simplicial area of a triangle, is that if we subdivide a triangle into smaller triangles, the simplicial area is not, in general, additive. There is however a special case where additivity holds. This is when we add a vertex at the midpoint of each edge of a Euclidean triangle T, and subdivide T into four similar triangles. We call this *conformal subdivision*.

Another defect of our definition is that the simplicial area of a triangle can only be zero when $a = b = c = 0$. For some purposes, this is not a problem. In the cases where this is a serious problem, we use quadrilaterals rather than triangles. See section 4.

Definition 2.1. *The* simplicial area *of* (F, τ, l) *is*

$$A_S(F, \tau, l) = \sum_{T \in \tau} A_S(T) = \sum_{T \in \tau} l_i l_j + l_j l_k + l_k l_i$$

where l_i, l_j, l_k *are the lengths of the edges of the 2–simplex T of τ.*

Unlike Heron's formula, this formula makes sense even when the triangle inequality does not hold. However the triangle inequality will prove useful, so we have built it into the definition of a simplicial metric. We will see that simplicial area retains enough similarities to the standard notion of area to be useful in many applications.

We now use this idea to define a simplicial analogue of the area of a map. Let $f : F \to M$ be a piecewise-smooth map of a metrized triangulated compact surface (F, τ, l) to a Riemannian manifold M and let L_i denote the length of $f|e_i$. Of course, this can only be defined when $f|e_i$ is rectifiable.

Definition 2.2. *The simplicial area of a map $f : (F, \tau, l) \to M$ is*

$$A_S(f) = \sum_{T \in \tau} L_i L_j + L_j L_k + L_k L_i,$$

where the sum is taken over all 2–simplices of τ.

The simplicial area of f does not depend on the simplicial metric l, but it does depend on the triangulation of F.

3. SIMPLICIAL ENERGY

The energy of a smooth map $f : (F, g) \to (M, h)$ of a Riemannian surface (F, g) to a Riemannian manifold (M, h) is

$$E(f) = \frac{1}{2} \int \int_F |\nabla f|^2 \ dA.$$

In local coordinates the energy can be obtained by integrating the 2–form

$$\frac{1}{2} g^{ij} h_{\alpha\beta} \frac{\partial f^\alpha}{\partial x^i} \frac{\partial f^\beta}{\partial x^j} dx_i dx_j.$$

If we take e_1, e_2 to be an orthonormal frame in a neighborhood on (F, g), then this expression simplifies to

$$\frac{1}{2} (||f_*(e_1)||^2 + ||f_*(e_2)||^2) dx_1 dx_2.$$

We now define a corresponding notion of simplicial energy. Suppose that f is a map from a metrized or quasi-metrized triangulated compact surface F to a Riemannian manifold M. If an edge e_i has $l_i > 0$ then we denote the ratio L_i/l_i by σ_i, and call it the *stretch factor* of the edge e_i under f. If e_i has length $l_i = 0$ then we define the stretch factor by $\sigma_i = 0$ if $L_i = 0$, and $\sigma_i = \infty$ if $L_i > 0$.

By analogy with the definition of smooth energy, we define the formula for the simplicial energy of a map on a triangle by multiplying the squares of stretch factors by the local contribution to the area.

Definition 3.1. *The* simplicial energy *of a piecewise-smooth map f of a metrized triangulated compact surface (F, τ, l) to a Riemannian manifold M*

is

$$E_S(f) = \frac{1}{2} \sum_{T \in \tau} (\sigma_i^2 + \sigma_j^2) l_i l_j + (\sigma_j^2 + \sigma_k^2) l_j l_k + (\sigma_k^2 + \sigma_i^2) l_k l_i,$$

where l_i, l_j, l_k are the edge lengths of the 2–simplex T of τ.

In the case where some edge e_i has length $l_i = 0$ in a quasi-metric and this edge is not mapped by f to a point, so that $L_i \neq 0$ and $\sigma_i = \infty$, we define the simplicial energy of f to be infinite. For all other cases the above formula gives a well defined finite energy.

Remark 3.2. *The simplicial energy of f is invariant under scale change of the domain simplicial metric. If the simplicial metric l is replaced by λl, resulting in multiplying each l_i by a constant $\lambda > 0$, then the energy is unchanged.*

It is sometimes convenient to rewrite the simplicial energy as a sum over the edges of τ. For a finite energy map $f : (F, \tau, l) \to M$, this gives

$$(1) \qquad E_S(f) = \frac{1}{2} \sum_{e_i \in \tau} \left(\frac{l_{i_1} + l_{i_2} + l_{i_3} + l_{i_4}}{l_i} \right) L_i^2$$

where the sum is over the edges $\{e_i\}$ of τ, and l_{i_1}, l_{i_2}, l_{i_3} and l_{i_4} are the lengths of the four edges of τ which are adjacent to e_i in the two triangles of which e_i is a face. See Figure 1.
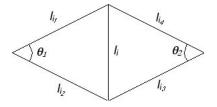

FIGURE 1. Edge lengths adjacent to e_i

The coefficient $w_i = (l_{i_1} + l_{i_2} + l_{i_3} + l_{i_4})/2l_i$ of L_i^2 is called the *weight* of e_i. Note that for any simplicial metric on (F, τ), each weight w_i is positive. Other approaches to combinatorial harmonic maps have studied a variety of ways of assigning such weights [31, 14, 26, 42]. For example the combinatorial energy of Pinkall and Polthier [31] assigns to the edge e_i the weight $w_i = (\cot \theta_1 + \cot \theta_2)/4$, where θ_1 and θ_2 are the two Euclidean angles opposite the edge e_i. See Figure 1. The cotangent formula goes back at least to Duffin [16]. Note that some of these weights will be negative if any of the triangles of τ has an obtuse angle. Our choice of weights was made because of its simplicity and its close connection to the previously defined simplicial area. For our purposes, it also turns out to be important that all our weights are positive. In fact, many of the results in this paper hold for any energy

functional of the form $\sum w_i L_i^2$, so long as all the weights are positive, though the connection between simplicial energy and area does not.

For our next result, it will be convenient to introduce the following notation. Let F be a surface with triangulation τ, and let l be a simplicial quasi-metric on (F, τ). Then we denote by τ_0 the subcomplex of τ which consists of all zero length edges.

Lemma 3.3. *Let $f : F \to M$ be a piecewise-smooth map of a connected triangulated compact surface (F, τ) to a Riemannian manifold M. Let l be a simplicial quasi-metric on (F, τ). Then*

$$E_S(f) \geq A_S(f).$$

Further equality holds if and only the stretch factors σ_i are constant on the edges of each component of $\tau - \tau_0$. In particular, if l is a simplicial metric, so that no l_i is zero, then equality holds if and only if all stretch factors σ_i take the same value on all edges of τ.

Remark 3.4. *This result is the combinatorial analogue of the fact that if F and M have Riemannian metrics and $f : F \to M$ is a smooth map, then the energy of f is always at least as large as its area, with equality if and only if f is conformal. The condition that the stretch factors σ_i take the same value on all edges of τ is a much weaker condition than conformality, but this may be an advantage for some purposes.*

Proof. Suppose that edges e_i and e_j lie on the boundary of a triangle in τ. If l_i and l_j are positive, we apply the inequality $a^2 + b^2 \geq 2ab$, with equality precisely when $a = b$. We obtain

$$(2) \qquad \frac{1}{2}(\sigma_i^2 + \sigma_j^2)l_i l_j \geq \sigma_i \sigma_j l_i l_j = \frac{L_i}{l_i} \frac{L_j}{l_j} l_i l_j = L_i L_j,$$

with equality precisely when $\sigma_i = \sigma_j$. If some l_i is zero, we have two cases. If $l_i = 0$, and $L_i > 0$, then σ_i is infinite, so that $\frac{1}{2}(\sigma_i^2 + \sigma_j^2)l_i l_j \geq L_i L_j$. If σ_j is finite, $l_i = 0$, and $L_i = 0$, then $\sigma_i = 0$ and the same inequality holds as both sides are zero. Thus this inequality holds in all cases. Summing over all triangles of F shows that $E_S(f) \geq A_S(f)$.

If $E_S(f) = A_S(f)$, and τ_0 is empty, the connectedness of the 1–skeleton of τ immediately implies that all stretch factors σ_i take the same value on all edges of τ. If τ_0 is non-empty, the connectedness of the 1–skeleton of the dual cellulation of τ implies that all the stretch factors σ_i are constant on the edges of each component of $\tau - \tau_0$. $\square$

We now give an example that shows some of the difficulties in achieving useful simplicial versions of area and energy. The example illustrates that it is not possible to capture all aspects of how a map stretches a surface using only information about the stretching of edges in a fixed triangulation.

Example 3.5. *Measuring stretching along the boundary of a triangle fails to capture Dirichlet energy.*

Let $f : T_1 \to T_2$ be a simplicial map from an isosceles triangle with edges of length $1, 1, 2 - \epsilon$ to an equilateral triangle with edge lengths all equal to one. The Euclidean area of T_1 is close to zero, while that of T_2 is $\sqrt{3}/4$. The affine map f that takes T_1 to T_2 stretches each of these edges by a factor of 1 or less. If we use a combinatorial definition of area that makes the area of T_1 close to zero, and a combinatorial definition of the energy of f which depends only on how much f stretches the edges of T_1, then the energy of f would also be close to zero. Thus the energy of f would be much less than its area, which is the area of the image triangle T_2, unlike the situation of Lemma 3.3. The problem is that measuring the stretching of f only on the boundary of T_1 ignores the large stretching by f of segments in T_1 orthogonal to the long edge.

We conclude that reasonable formulas defining simplicial area and energy cannot simultaneously satisfy all of the following:

(1) Simplicial energy is evaluated using stretching of edges and simplicial area.
(2) Simplicial energy is greater or equal to simplicial area.
(3) The simplicial area of a triangle contained in a line is zero.

As we pointed out in section 2 our definition of simplicial area only has the third property when every edge of a triangle has zero length. In section 7, we will see how using quadrilaterals instead of triangles removes some of these difficulties.

4. POLYGONAL DECOMPOSITIONS

In this section, we generalise the preceding two sections to the situation where a compact surface F is divided into polygons rather than just triangles. We again use τ to denote such a division. We again have a map l that assigns to each edge e_i, $1 \le i \le r$, of τ a length $l_i = l(e_i) > 0$. For each polygon of τ, we require that each edge has length no more than the sum of all the other edges of that polygon. This suffices to ensure the existence of a convex embedding of the polygon in the Euclidean plane. If the polygon has more than three edges, we cannot expect this embedding to be unique, but this is not important. Note that we allow the possibility that for some polygons one of these inequalities is an equality. We call such an assignment a *simplicial metric* on F, and denote it by (F, τ, l), or just l when the context is clear. As in the previous sections, we do not require that each polygon of τ be embedded in F. We will also extend this to the case where some edges of τ are assigned zero length. We call such an extension a *simplicial quasi-metric* on F.

Throughout this section, F will be a compact surface with a polygonal decomposition τ and a simplicial metric or quasi-metric l.

Let P be a n–gon, whose edges are $e_1, e_2, \ldots, e_n$ with lengths $l_1, l_2, \ldots, l_n$ reading round P. Thus l_i and l_{i+1} are lengths of adjacent edges, where we define l_{n+1} to equal l_1. We define the *simplicial area A_S* of this polygon by

the formula

$$A_S(P) = \sum_{i=1}^{n} l_i l_{i+1}.$$

As in the case of triangles, the simplicial area of polygons is not, in general, additive. However for quadrilaterals there is a natural analogue of the conformal subdivision of a triangle which we described near the start of section 2. But note that it may be that the process we describe does not correspond to subdividing a quadrilateral embedded in the plane. Given a quadrilateral P, there is a natural way to divide it into four sub quadrilaterals, by cutting along two lines which join interior points of opposite edges of P. We now assign lengths to the new edges as shown in Figure 2.
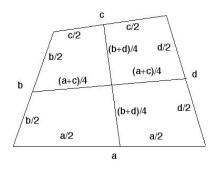

FIGURE 2. Conformal subdivision preserves area.

As for triangles, the area of a polygon can be non-zero even if several edges have length zero. However the following fact will be very useful. If P is a quadrilateral, and two opposite edges of P are assigned zero length, then the simplicial area of P is zero. For the applications in sections 7 and 8, this is all we will need.

Definition 4.1. *The* simplicial area *of* (F, τ, l) *is*

$$A_S(F, \tau, l) = \sum_{P \in \tau} A_S(P).$$

Let $f : F \to M$ be a piecewise-smooth map to a Riemannian manifold M and let L_i denote the length of $f|e_i$. Of course, this can only be defined when $f|e_i$ is rectifiable.

Definition 4.2. *For a n–gon P of τ, the* simplicial area *of a map $f : P \to M$ is $\sum_{i=1}^{n} L_i L_{i+1}$, which we denote by $A_S(f, P)$.*
The simplicial area *of a map $f : (F, \tau, l) \to M$ is*

$$A_S(f) = \sum_{P \in \tau} A_S(f, P).$$

Next we define the simplicial energy of f. As for triangles, we use the stretch factor $\sigma_i = L_i/l_i$, when $l_i > 0$. If $l_i = 0$ we define the stretch factor by $\sigma_i = 0$ if $L_i = 0$, and $\sigma_i = \infty$ if $L_i > 0$.

Definition 4.3. *For a n–gon P of τ, the* simplicial energy *of a map $f :$
$P \to M$ is $E_S(f, P) = \sum_{i=1}^{n} (\sigma_i^2 + \sigma_{i+1}^2) l_i l_{i+1}$, where σ_{n+1} equals σ_1.*
The simplicial energy *of a piecewise-smooth map f of (F, τ, l) to a Rie-
mannian manifold M is*

$$E_S(f) = \frac{1}{2} \sum_{P \in \tau} E_S(f, P).$$

As for triangles, the above formula gives a well defined finite energy, unless
some $\sigma_i = \infty$, when $E_S(f) = \infty$.

Remark 4.4. *The simplicial energy of f is invariant under scale change
of the domain simplicial metric. If the simplicial metric l is replaced by
λl, resulting in multiplying each l_i by a constant $\lambda > 0$, then the energy is
unchanged.*

It is sometimes convenient to rewrite the simplicial energy as a sum over
the edges of τ. For a finite energy map $f : (F, \tau, l) \to M$, this gives

$$(3) \qquad E_S(f) = \frac{1}{2} \sum_{e_i \in \tau} \left(\frac{l_{i_1} + l_{i_2} + l_{i_3} + l_{i_4}}{l_i} \right) L_i^2$$

where the sum is over the edges $\{e_i\}$ of τ, and l_{i_1}, l_{i_2}, l_{i_3} and l_{i_4} are the
lengths of the four edges of τ which are adjacent to e_i in the two polygons of
which e_i is a face. See Figure 1 for the case when the polygons are triangles.

If l is a simplicial quasi-metric on (F, τ), then we denote by τ_0 the sub-
complex of τ which consists of all zero length edges. The following result is
the natural generalisation of Lemma 3.3, and is proved in the same way.

Lemma 4.5. *Let $f : (F, \tau, l) \to M$ be a piecewise-smooth map of F to a
Riemannian manifold M. Then*

$$E_S(f) \geq A_S(f).$$

*Further equality holds if and only the stretch factors σ_i are constant on
the edges of each component of $\tau - \tau_0$. In particular, if l is a simplicial
metric, so that no l_i is zero, then equality holds if and only if all stretch
factors σ_i take the same value on all edges of τ.*

5. Simplicial harmonic maps

We call a map of a surface F into a manifold M *trivial* if the image of
$\pi_1(F)$ is contained in a cyclic subgroup of $\pi_1(M)$. When M is non-positively
curved, or more generally when $\pi_2(M) = 0$, this is equivalent to the map
factoring through a map of the circle into M.

In this section we show that any homotopy class of maps from a metrized
triangulated closed surface (F, τ, l) into a closed non-positively curved Rie-
mannian manifold M contains a simplicial-energy minimizing map. This
map is unique if the map is nontrivial and the curvature of M is strictly

negative. Note that nontrivial maps never exist if F is the sphere or projective plane. Further if the metric on M is strictly negatively curved, then nontrivial maps do not exist if F is the torus or Klein bottle.

Since M is geodesically complete, any map $f_0 : F \to M$ can be homotoped, fixing the vertices of F, to a new map that sends each edge of τ to a geodesic arc in M. If M is non-positively curved, such an arc is unique. After picking an identification of each edge e_i of τ with the interval $[0, l_i]$ of the real line, this yields a canonical map from the 1–skeleton of τ to M. The resulting map f_1 has energy no greater than the original map f_0, and it suffices to consider such maps in searching for an energy minimizer. The choice of an extension of f_1 to the 2–simplices of F does not affect its energy, so we are free to make this choice in any convenient way. To uniquely describe an extension, we make some further choices. Since the edge lengths of τ satisfy the triangle inequality, we can identify each triangle of τ with the Euclidean triangle having the same edge lengths. As discussed at the start of Section 2, each triangle of τ has a preferred vertex, and we fix the natural cone structure on the triangle with this vertex as cone point. There is a unique extension of f_1 to a map of F that sends each cone line in a triangle to a constant speed geodesic arc in M.

If τ is a polygonal decomposition of F, we proceed in the same way. The only difference is that we need to choose a preferred vertex for each polygon, choose a convex embedding of each polygon in the Euclidean plane, and then use the natural cone structure on the polygon with this vertex as cone point.

We make the following definition.

Definition 5.1. *A map $f : (F, \tau, l) \to M$ from a triangulated surface F with a simplicial metric l to a Riemannian manifold M is* simplicial *if each edge of τ and each cone line in each 2–simplex of τ is mapped as a geodesic arc.*

We can use the same definition and terminology if τ is a polygonal decomposition of F.

The preceding discussion shows that if M is non-positively curved, then any map $f_0 : F \to M$ can be homotoped, leaving the vertices fixed, to a unique simplicial map f_1 and that the simplicial energy of f_1 is no greater than that of f_0. Thus to establish the existence of simplicial-energy minimizing maps into a non-positively curved manifold, we need only consider simplicial maps. Given a simplicial map f we define a *variation* of f to be a 1–parameter family of maps f_t, $-\epsilon < t < \epsilon$ with $f_0 = f$. Note that f_t need not be simplicial when $t \neq 0$.

Definition 5.2. *A map $f : (F, \tau, l) \to M$ of a closed surface with a simplicial metric or simplicial quasi-metric to a Riemannian manifold M is* simplicial harmonic *if it is simplicial and is a critical point of simplicial energy under all variations f_t of f.*

If F is compact with non-empty boundary, then f is simplicial harmonic *if it is simplicial and is a critical point of simplicial energy under all variations f_t of f which leave ∂F fixed.*

The map f is simplicial conformal *if the stretch factors σ_i are constant on the edges of each component of $\tau - \tau_0$, where τ_0 denotes the subcomplex of τ which consists of all zero length edges.*

Remark 5.3. *Lemmas 3.3 and 4.5 tell us that if f is a simplicial map, then $E_S(f) = A_S(f)$ if and only if f is simplicial conformal.*

Remark 5.4. *The above definition of f being simplicial harmonic is equivalent to the one obtained by restricting attention to variations in which each map f_t is simplicial.*

For most of this paper, we will consider only the case when the surface F is closed, but in the next section it will be important to consider the case when F has non-empty boundary.

Proposition 5.5. *Let (F, τ, l) be a closed triangulated surface with a simplicial metric. Let $f : (F, \tau, l) \to M$ be a continuous map of F to a non-positively curved, closed Riemannian manifold M. Then*

(1) *f is homotopic to a simplicial map f_0 that minimizes simplicial energy among all maps homotopic to f, and hence is simplicial harmonic.*

(2) *If f is nontrivial and M is negatively curved, then f_0 is the unique simplicial harmonic map in the homotopy class of f.*

Furthermore, if l is a simplicial quasi-metric and f is homotopic to a map with finite energy, then (1) and (2) continue to hold.

Proof. (1) Any homotopy of a map $f : F \to M$ that decreases some image length L_i without increasing any L_j reduces the simplicial energy (see Equation 1). It follows that a map that minimizes simplicial energy sends each edge of τ to an arc in M that minimizes length in its relative homotopy class.

A simplicial map is not determined by the images of the vertices of τ, since distinct relative homotopy classes of arcs connecting two points will yield distinct geodesic arcs connecting those points. However a simplicial map f is determined uniquely by the images of the vertices and the homotopy class (rel boundary) of f on each edge of τ.

We select a vertex $v \in F$ as a basepoint and also fix a fundamental region W in the universal cover $\widetilde{M}$ of M. Since M is compact, so is W. Let $\mathcal{F}$ be the set of simplicial maps homotopic to f, let $E_0 > 0$ be a constant and let $\mathcal{F}_0$ be the subset of $\mathcal{F}$ consisting of maps with simplicial energy less than or equal to E_0. We will show that $\mathcal{F}_0$ is compact.

Let f be a map in $\mathcal{F}_0$, let P denote $f(v)$ and let $\widetilde{P}$ be a point of W projecting to P. Let $\widetilde{F}$ denote the universal cover of F, equipped with the natural triangulation lifted from τ, and let $\widetilde{v}$ denote a point of $\widetilde{F}$ above v.

There is a unique map $\widetilde{f} : \widetilde{F} \to \widetilde{M}$ that covers f and sends $\widetilde{v}$ to $\widetilde{P}$. As $\widetilde{M}$ is simply connected, the simplicial map $\widetilde{f}$ is determined completely by the images of the vertices of $\widetilde{F}$.

Each edge e_i in τ is part of a simple arc γ_i contained in the 1–skeleton of τ and connecting e_i to the base point v. The restriction of f to γ_i lifts to a path $\widetilde{\gamma}_i$ in $\widetilde{M}$ which starts at $\widetilde{P}$.

Let l_m be the length of the shortest edge in the given simplicial metric on τ, let l_M be the length of the longest edge and let L_i be the length of the restriction of f to e_i. As f lies in $\mathcal{F}_0$, we have

$$(L_i/l_i)^2 l_m^2 \leq (L_i/l_i)^2 l_i l_j = \sigma_i^2 l_i l_j \leq 2E_0,$$

so that

$$L_i \leq \sqrt{2E_0}(l_i/l_m) \leq \sqrt{2E_0}(l_M/l_m).$$

Let L_M denote the quantity $\sqrt{2E_0}(l_M/l_m)$, so that $L_i \leq L_M$, for all i. Since γ_i is simple it contains at most r edges, where r is the total number of edges in τ. Hence the restriction of f to γ_i has length at most rL_M. Note that this bound applies for any element f of $\mathcal{F}_0$. As the point $\tilde{P}$ lies in the compact fundamental domain W, there is a ball B in $\widetilde{M}$ such that each $\widetilde{\gamma}_i$ lies in B, for any $f \in \mathcal{F}_0$. It follows that the simplicial map $\widetilde{f}$ is completely determined by the finite number of points in B that are vertices of the $\widetilde{\gamma}_i$'s. The set of all such is compact, therefore a minimum energy map exists, which establishes part (1).

Note that the above argument does not really use the non-positive curvature assumption on M. Without that assumption, one can still find an energy minimizing map homotopic to f, but may not be able to arrange that this map is simplicial as defined in Definition 5.1.

(2) Now suppose that f is nontrivial, and that f_0 and f_1 are simplicial harmonic maps in the homotopy class $\mathcal{F}$ of f. Further suppose that M is negatively curved. As f_0 and f_1 are homotopic, there is a "straight" homotopy f_t between them, in which f_t is a simplicial map for each $0 \leq t \leq 1$ and the images of vertices of τ move along geodesic arcs in M at constant speed.

We will make use of the following key fact. Let $\alpha(t)$ and $\beta(t)$ denote two geodesic arcs in the negatively curved manifold M and let $\gamma(t)$ be the shortest geodesic arc from $\alpha(t)$ to $\beta(t)$. Then the length $L(t)$ of $\gamma(t)$ is a convex function of t. Further it is strictly convex except possibly when the arcs $\alpha(t)$, $\beta(t)$ and $\gamma(t)$ are all contained in a single geodesic [41]. Since each $L_i(t)$ is convex we see that the simplicial energy function

$$E_S(f_t) = \sum w_i L_i(t)^2, \ w_i > 0,$$

given in Equation 1 is also a convex function of t.

Now suppose that f_0 and f_1 do not coincide. Then there is a vertex v of τ such that the homotopy f_t from f_0 to f_1 moves v along a nontrivial geodesic arc μ. We let λ denote the unique complete geodesic of M which

contains μ. As f is nontrivial, there is an edge e_i of τ whose image under f_0 or f_1 is not contained in λ. First we suppose that v is a vertex of e_i. In this case, the function $L_i(t)$ is a strictly convex function of t, and it follows immediately that $E_S(f_t)$ is also strictly convex, and so can have only one critical point. Since both f_0 and f_1 are critical points, they must coincide, proving part (2) of the proposition in this case. However it may be the case that every edge of τ incident to v has image which remains in λ under each map f_t. Here we need a slightly refined version of the above key fact. Let $\alpha(t)$ and $\beta(t)$ denote two geodesic arcs in the negatively curved manifold M which are contained in a single geodesic λ, and suppose, in addition, that α and β each has constant speed. Let $\gamma(t)$ be the shortest geodesic arc from $\alpha(t)$ to $\beta(t)$. Then the length $L(t)$ of $\gamma(t)$ is a linear function of t. Hence $L(t)^2$ is a strictly convex function of t, unless $L(t)$ is constant.

Now we return to the situation where v is a vertex of τ such that the homotopy f_t moves v along a nontrivial geodesic arc contained in λ, there is an edge e_i of τ whose image under f_0 or f_1 is not contained in λ, and v is not a vertex of e_i. Choose a simple path in the 1-skeleton of τ that connects v to e_i, let e_j denote the first edge of this path and let w denote the other vertex of e_j. The preceding discussion shows that $L_j(t)^2$ is a strictly convex function of t unless $f_t(w)$ stays in λ for all t, and $L_j(t)$ is constant. In the first case, it follows as before that $E_S(f_t)$ is a strictly convex function of t. In the second case, it follows that w, like v, is a vertex of τ such that the homotopy f_t from f_0 to f_1 moves w along a geodesic arc contained in λ. Further this geodesic arc is nontrivial because $L_j(t)$ is constant. As w is joined to e_i by a shorter edge path than v, a simple inductive argument completes the proof that $E_S(f_t)$ is a strictly convex function of t in all cases. As above this implies that f_0 and f_1 coincide, which completes the proof of part (2) of the proposition.

Finally, we consider the case when l is a quasi-metric. It may no longer be true that a harmonic map homotopic to f exists. For example, l might assign zero length to each edge of τ, which would mean that all maps homotopic to f have infinite energy unless f is homotopic to a point. However if f is homotopic to a map with finite energy, then we can take l_m to be the smallest positive length of the edges in τ. All zero length edges must be mapped to a point if the total energy is finite, and so they make no contribution to the energy. The previous argument then applies and establishes that parts (1) and (2) of the theorem continue to hold in this case.   □

6. Mean value and convex hull properties

In this section, we show that simplicial harmonic maps share some of the classical properties of smooth harmonic maps.

Let F be a Riemannian manifold. A map $f : F \to \mathbb{R}^n$ has the *mean value property* if for any point $x \in F$, and any ball B of radius r centered at x, the value of f at x is the average of the values of f over ∂B. If f is harmonic it

has the mean value property. Harmonic maps to $\mathbb{R}^n$ also have the following *convex hull property*: for D any compact connected submanifold of F, the image $f(D)$ is contained in the convex hull of $f(\partial D)$. Note that if F is closed, then any harmonic map from F to $\mathbb{R}^n$ must be constant, so the discussion in this section is only of interest when F has non-empty boundary.

Finally a real valued function f on F is said to satisfy the *maximum principle* if, for any compact submanifold D of F, the restriction of f to D attains its upper and lower bounds only on ∂D. Non-constant harmonic functions also satisfy the maximum principle.

As mentioned in the introduction, the discrete harmonic maps of Pinkall and Polthier [31] do not satisfy the convex hull or mean value properties. See [32] for an interactive demonstration of an example found by Polthier and Rossman [33] which shows that discrete harmonic maps do not satisfy the convex hull property.

Next we show that simplicial harmonic maps do satisfy the convex hull and mean value properties. Let v be a vertex of a surface triangulation τ, and let D denote the closed star of v, so that D is the union of all the simplices of τ which contain v. We let $v_1, \ldots, v_k$ denote the other vertices of D.

Lemma 6.1. *(Euclidean Mean Value and Convex Hull Properties) Let f be a simplicial harmonic map of a connected metrized triangulated compact surface (F, τ, l) into $\mathbb{R}^n$. Then $f(v)$ lies in the convex hull of $f(v_1), \ldots, f(v_k)$. Further $f(v)$ is the centre of mass of a collection of positive weights placed at $f(v_1), \ldots, f(v_k)$.*

Proof. Equation (1) shows that the simplicial energy $E_S(f)$ has the form $\sum w_i L_i{}^2$, for some positive constants w_i. As f is harmonic, it is a critical point of the functional E_S. We will apply this fact to variations of f through simplicial maps which equal f on all vertices apart from v. For $1 \leq j \leq k$, let e_j denote the edge joining v to v_j, and consider the terms $\sum_{j=1}^{k} w_j L_j(t)^2$ of $E_S(f_t)$. Let $(x_1, \ldots, x_n)$ be the coordinates of $f(v)$, and let $(x_{1j}, \ldots, x_{nj})$ be the coordinates of $f(v_j)$, $1 \leq j \leq k$. Thus $\sum_{j=1}^{k} w_j L_j(t)^2 = \sum_{j=1}^{k} w_j \sum_{i=1}^{n} (x_i - x_{ij})^2$. As f is a critical point of E_S, we know that, for $1 \leq i \leq n$, the partial derivative of this summation with respect to x_i must be zero. Hence, for $1 \leq i \leq n$, we have $\sum_{j=1}^{k} w_j 2(x_i - x_{ij}) = 0$, so that $x_i \sum_{j=1}^{k} w_j = \sum_{j=1}^{k} w_j x_{ij}$. As each w_j is positive, x_i is a linear combination of the x_{ij}'s with coefficients which lie between 0 and 1. It follows that $f(v)$ lies in the convex hull of $f(v_1), \ldots, f(v_k)$, and that $f(v)$ is the centre of mass of weights $w_1, \ldots, w_n$ placed at $f(v_1), \ldots, f(v_k)$. $\square$

The special case of this result when $n = 1$ gives the following form of the Maximum Principle.

Lemma 6.2. *(Maximum Principle) Let f be a non-constant simplicial harmonic map of a connected metrized triangulated compact surface (F, τ, l)*

into $\mathbb{R}$. *Then for any subsurface E of F which is also a subcomplex, the restriction of f to E attains its upper and lower bounds only on ∂E.*

The convex hull of a subset of a general Riemannian manifold M may be rather complicated. For simplicity we consider only the case of non-positive curvature and will restrict attention to convex balls in M. In this setting we obtain the following result.

Lemma 6.3. *Let f be a simplicial harmonic map of a connected metrized triangulated compact surface (F, τ, l) into a convex ball B with a non-positively curved Riemannian metric. Let v be an interior vertex of τ and let D denote the closed star of v. Let $v_1, \ldots, v_k$ denote the other vertices of D. Then $f(v)$ lies in the convex hull of $f(v_1), \ldots, f(v_k)$.*

Proof. If $f(v)$ lies outside the convex hull of $f(v_1), \ldots, f(v_k)$, then there is a shortest geodesic arc α from $f(v)$ to the convex hull. Take the tangent plane P at $f(v)$ perpendicular to α. The geodesics joining $f(v)$ to $f(v_j)$ have tangent vectors at $f(v)$ that all lie on the same side of P. Moving $f(v)$ away from P towards this side will reduce each L_j and so reduce the energy of f, contradicting the assumption that f is simplicial harmonic. □

Lemma 6.3 implies a universal area bound for homotopy classes of maps of a fixed surface into a negatively curved manifold. We first recall the well-known corresponding bound for smooth maps.

Lemma 6.4. *Let M be a hyperbolic manifold, let F be the closed orientable surface of genus $g > 1$, and let $f : F \to M$ be any map. Then f can be homotoped to a map f_0 with $Area(f_0) \leq 2\pi(2g - 2)$.*

If M has sectional curvature bounded above by $-a < 0$ then we have $Area(f_0) \leq 2\pi(2g - 2)/a$.

Proof. Suppose that M has sectional curvature bounded above by $-a < 0$. If f sends a simple loop on f to a trivial loop in M, we compress f. Otherwise f is homotopic to a least area immersion in its homotopy class, using the existence theorems for minimal surfaces due to Schoen and Yau [40], and Sacks and Uhlenbeck [39]. Note that as $\pi_2(M) = 0$, the least area maps obtained in [40] and [39] must be homotopic to f.

The normal curvature k_n of a minimal immersion is non-positive. Its sectional curvature $K \leq -a + k_n \leq -a$. By the Gauss-Bonnet Theorem, with $F_0 = f_0(F)$,

$$2\pi(2 - 2g) = 2\pi\chi(F_0) = \int_{F_0} K \; dA \leq \int_{F_0} (-a) \; dA = -a\mathrm{Area}(F_0)$$

So $\mathrm{Area}(F_0) \leq 2\pi(2g - 2)/a$.

The homotopy class of the original uncompressed surface map can be recovered by adding tubes of zero area. □

We reprove this result using simplicial harmonic maps. In doing so we no longer need to use the existence results for minimal surfaces. Since our results apply equally well to non-orientable surfaces, we use the Euler number rather than the genus of a surface.

Lemma 6.5. *Let M be a hyperbolic manifold, let F be a closed surface with Euler characteristic $\chi < 0$, and let $f : F \to M$ be any map. Then f can be homotoped to a map f_0 with $Area(f_0) \leq 2\pi |\chi|$.*

If M has sectional curvature bounded above by $-a < 0$ then we have $Area(f_0) \leq 2\pi |\chi| / a$.

Proof. Suppose that M has sectional curvature bounded above by $-a < 0$. Choose a triangulation and a simplicial metric on F, and let f_0 be a simplicial harmonic map in the homotopy class of f. If f is trivial, then f_0 has image contained in some geodesic of M and so has zero Riemannian area. In this case the required inequality is obvious. Otherwise f is nontrivial, so that f_0 is unique and at least one 2–simplex of τ does not have image contained in a geodesic. The simplicial map f_0 is a ruled smooth immersion on each such 2–simplex. Thus the Riemannian metric induced on F has curvature $\leq -a$ on the interior of such 2–simplices, and edges are sent to geodesics. At the vertices, the convex hull property in Lemma 6.3 shows that the angle sum around each vertex is greater than 2π, and hence that the vertices contribute negative curvature to the total curvature when using the Gauss-Bonnet formula for simplicial complexes. The argument now follows as in the smooth case. $\square$

Remark 6.6. *The preceding lemma can be proved in essentially the same way if F is divided into polygons. For then each polygon has a cone structure, and so has a division into subcones which are triangles. Again each triangle either has image contained in a geodesic, or is mapped by a ruled smooth immersion into M.*

7. Families of maps

Given an immersion of a surface into a Riemannian manifold, the mean curvature flow gives an area decreasing flow that pushes the surface in the direction that decreases its area as rapidly as possible. However singularities develop as the surface evolves under mean curvature, making it difficult to extend the flow beyond the time of singularity formation. Additional difficulties arise in extending mean curvature flow to families of surfaces, and these have proved an obstruction to topological applications. Harmonic maps offer an alternative approach to constructing area decreasing flows, a fact exploited in [24].

In this section we show how to deform a multi-parameter family of simplicial maps to a family of simplicial harmonic maps. Under this deformation the simplicial energy of each map in the family decreases monotonically. The simplicial area may not be monotonically decreasing, but the simplicial area

of each surface is at all times bounded above by its initial value. Although some of the results in this section work for triangulations, we will use only decompositions into quadrilaterals. This is because of the applications we have in mind in this and the next section.

For these applications, it is important that the Riemannian area of a simplicial map is bounded by its simplicial energy. Note that no such bound can exist for general maps, as the simplicial energy of a map is determined solely by the restriction of that map to the 1–skeleton of the given triangulation.

Lemma 7.1. *Let F be a compact surface with a decomposition τ into quadrilaterals and a simplicial quasi-metric l. Let $f : F \to M$ be a simplicial map to a non-positively curved Riemannian manifold M. Then the Riemannian area, Area(f), satisfies the inequality*

$$Area(f) \leq E_S(f)/2.$$

Proof. Let P be a quadrilateral of τ, with edges $e_1, \ldots, e_4$ cyclically ordered. Recall that there is a preferred vertex v of P, that we use a cone structure on P with v as cone point, and that a simplicial map from F to M sends each cone line in this cone structure to a geodesic in M. The diagonal of P that has one end at v divides P into two triangles, and the restriction of f to each triangle is a ruled 2–simplex in M. Suppose that v is the vertex $e_1 \cap e_4$, so that one triangle contains e_1 and e_2, and the other triangle contains e_3 and e_4. As M is non-positively curved, the induced curvature on each triangle is also non-positive and so the induced Riemannian area is no greater than the Euclidean area of a triangle with the same edge lengths. The Euclidean areas of the two triangles are at most $L_1 L_2/2$ and $L_3 L_4/2$. We conclude that the Riemannian area of the restriction of f to P is at most $(L_1 L_2 + L_3 L_4)/2$, which is clearly at most one half of the simplicial area of the restriction of f to P. Now Lemma 4.5 implies that this is at most one half of the simplicial energy of the restriction of f to P. Summing over all the quadrilaterals of τ immediately yields the inequality Area(f) $\leq E_S(f)/2$, as required. $\square$

Let F be a closed surface with a decomposition τ into quadrilaterals and a simplicial quasi-metric l. Let $f : (F, \tau, l) \to M$ be a continuous map of F to a negatively curved, closed Riemannian manifold M. If f is nontrivial, then Proposition 5.5 states that f is homotopic to a unique simplicial harmonic map f_0.

When f itself is a simplicial map, there is a canonical way to deform it to the simplicial energy minimizing map f_0. The simplicial energy function is determined locally by the images under f of the vertices of τ, and thus is a real valued function on a finite dimensional manifold. The gradient of this function determines a direction of deformation for the vertices of $f(F)$. This gives rise to an energy decreasing flow f_t whose derivative at each t equals the negative of the gradient of the simplicial energy function. This flow moves the vertex images of $f_t(F)$ so as to decrease simplicial energy as quickly as possible. We define the *simplicial energy gradient flow* to be this

flow of the map f. Since there is a unique critical point for simplicial energy in any homotopy class, the gradient flow converges to f_0, the minimizing simplicial harmonic map homotopic to f.

We now show that the map f_0 obtained by minimizing energy among simplicial maps homotopic to an initial map f depends continuously on the map f and the simplicial metric l. This is a straightforward consequence of the uniqueness of simplicial harmonic maps within a homotopy class. It was shown in [17] that smooth harmonic maps depend continuously on the domain metric.

We use the term *continuous family of maps* $f_s : F \to M$ *of a surface F to a manifold M parametrized by a space $\mathcal{S}$* to refer to a continuous function $H : \mathcal{S} \times F \to M$, with $f_s(x) = H(s, x)$, for each $s \in \mathcal{S}$ and $x \in F$.

Proposition 7.2. *Let F be a closed surface with a decomposition τ into quadrilaterals, and let M be a negatively curved closed Riemannian manifold. Let $f_s : (F, \tau, l_s) \to M$ be a continuous family of nontrivial simplicial maps parametrized by a subspace $\mathcal{S}$ of $\mathbb{R}^n$, where l_s is a continuous family of simplicial quasi-metrics. Then the gradient flow applied to each f_s yields a continuous family of simplicial maps $f_{s,t} : (F, \tau, l_s) \to M$ parametrized by $\mathcal{S} \times [0, \infty)$, such that, for each s, the map $f_{s,0}$ equals f_s, and the family of maps $f_{s,t}$ converges to a simplicial harmonic map $f_{s,\infty}$. The simplicial harmonic maps $f_{s,\infty}$ depend continuously on the initial map f_s and on the initial simplicial metric l_s.*

Proof. For fixed l_s, the simplicial harmonic map homotopic to $f_s : (F, \tau, l_s) \to M$ is unique, so doesn't vary when f_s is changed by a homotopy. The path of maps $f_{s,t} : (F, \tau, l_s) \to M$ defined by the gradient flow, for each fixed s, varies continuously with the map, since each point is following the trajectory of a solution to a first order differential equation whose initial conditions vary continuously.

It remains to show that the simplicial harmonic maps $f_{s,\infty}$ depend continuously on the choice of simplicial metric l_s. Let $\{s_n\}$ be a sequence of points in $\mathcal{S}$ converging to s in $\mathcal{S}$, let l_n denote the metric associated to s_n, and let l_∞ denote the metric associated to s. Thus the sequence of simplicial quasi-metrics l_n converges to l_∞. Arguing as in the proof of the first part of Proposition 5.5 shows that the corresponding simplicial harmonic maps f_n converge to a limiting map f_∞ whose simplicial energy is equal to the limiting value of the simplicial energies of f_n. With respect to the simplicial metric l_∞, the map f_∞ has the same simplicial energy as $f_{s,\infty}$, which is the unique harmonic map in its homotopy class with simplicial metric l_∞. Therefore f_∞ must be equal to $f_{s,\infty}$. This establishes that the resulting simplicial harmonic maps vary continuously with the simplicial metric, as claimed. □

Now we can apply the above result to show that we can homotope a family of maps from F to M to arrange that there is a uniform bound

on the Riemannian area of the maps in the new family, where the bound depends only on the Euler characteristic of F, and the curvature of M.

Theorem 7.3. *Let F be a closed surface with Euler characteristic $\chi < 0$, and with a subdivision τ into quadrilaterals. Let $f_s : F \to M$, $s \in \mathcal{S}$, be a continuous family of simplicial maps of F to a Riemannian manifold M with sectional curvatures bounded above by $-a < 0$, where $\mathcal{S}$ is a path connected subset of $\mathbb{R}^n$. Further suppose that some, and hence every, f_s is nontrivial. Then, for each $s \in \mathcal{S}$, there is a simplicial quasi-metric l^s on F which varies continuously with s, and a continuous family of maps $g_{s,t} : (F, \tau, l^s) \to M$ parametrized by $\mathcal{S} \times I$, satisfying*

(1) *$g_{s,0} = f_s$,*
(2) *$g_{s,1}$ is a simplicial harmonic map,*
(3) *the simplicial energy of $g_{s,t}$ is non-increasing with t,*
(4) *$Area(g_{s,1}) \leq 2\pi |\chi| / a$, for every s.*

Proof. For each simplicial map f_s in the given family, we pick a simplicial quasi-metric l^s on τ by setting the edge length l_i^s of e_i equal to the length L_i^s of $f_s|e_i$. For this quasi-metric, all the stretch factors σ_i are equal to 1 on all edges with $l_i^s > 0$. If $l_i^s = 0$, then f_s must map e_i to a point. Thus f_s has finite simplicial energy. Now we apply Proposition 7.2. This yields a continuous family of simplicial maps $f_{s,t} : (F, \tau, l^s) \to M$ parametrized by $\mathcal{S} \times [0, \infty)$, such that $f_{s,0} = f_s$. Further, for each s, the function $E_S(f_{s,t})$ is a non-increasing function of t, and the family of maps $f_{s,t}$ converges to a simplicial harmonic map $f_{s,\infty}$. Now we use a homeomorphism between $[0, \infty)$ and $[0, 1)$ to obtain a new family of simplicial maps $g_{s,t} : (F, \tau, l^s) \to M$ parametrized by $\mathcal{S} \times I$, such that $g_{s,0} = f_{s,0} = f_s$, and $g_{s,1} = f_{s,\infty}$ is simplicial harmonic. This family is continuous by Proposition 7.2. In particular, parts 1)-3) of the theorem hold.

Since $g_{s,1}$ is simplicial harmonic, the proof of Lemma 6.5 shows that $Area(g_{s,1}) \leq 2\pi |\chi| / a$ for every s, so that part 4) of the theorem holds. $\square$

The above theorem and its proof are very similar to what occurs in the smooth setting. The crucial difference is that in the smooth setting, if $f : F \to M$ is a map to a Riemannian manifold M, one can only define the Riemannian metric on F induced by f when f is an immersion. In particular the energy of f must be non-zero. But in the above theorem maps with zero energy cause no problems, as we simply induce a quasi-metric on F. This greatly simplifies applications, as we will see later in this section and in the next section.

If we start with a family f_s of smooth or piecewise-smooth maps from F to M, we need to approximate it by a homotopic family of simplicial maps before we can apply the above result. This is easy to do by just choosing a subdivision τ of F into quadrilaterals, and then replacing each f_s by the unique simplicial map which agrees with f_s on the vertices of τ.

Theorem 7.3 generalizes from 1–parameter to k–parameter families certain results of Thurston [41], Bonahon [7], Canary [9], Minsky [28] and Wolf [43], that utilized pleated surfaces and simplicial hyperbolic surfaces to control the area and diameter of a 1–parameter family of surfaces. See also the recent use of bounded area 1–parameter families of surfaces by Agol [1] and Calegari-Gabai [8] in proving the tameness conjecture for ends of negatively curved 3–manifolds. Area bounds for 2–parameter families of surfaces obtained using smooth harmonic maps were used to find counterexamples to the stabilization conjecture for Heegaard splittings in [24]. See the discussion at the end of this section.

Pitts and Rubinstein [29] gave a construction of unstable minimal surfaces that utilizes a minimax argument applied to a family of surfaces called a sweepout. For later applications we give below a modified definition of a sweepout.

Given a closed orientable surface F, a closed orientable 3–manifold M, and a map $h : (F \times I, F \times \partial I) \to M$ with the property that $\text{Area}(h(F, 0)) = \text{Area}(h(F, 1)) = 0$, there is a natural homomorphism

$$\phi_h : H_3(F \times I, F \times \partial I) \to H_3(M).$$

We now explain the construction of ϕ_h.

Since $H_3(F \times I, F \times \partial I)$ is a cyclic group generated by the fundamental class μ of $(F \times I, F \times \partial I)$, it suffices to define the action of ϕ_h on μ.

Let α denote a fundamental 3–chain for $F \times I$, so that $\partial \alpha = \beta_1 - \beta_0$, where β_i denotes a fundamental 2–cycle for $F \times \{i\}$, $i = 0, 1$. Thus each $h_\#(\beta_i)$ is a singular 2–cycle in M with zero volume. A theorem of Federer implies that, for $i = 0, 1$, there is a singular 3–chain c_i in M such that $\partial c_i = h_\#(\beta_i)$, and c_i also has zero volume [21]. Thus the 3–chain $c_0 + h_\#(\alpha) - c_1$ is a cycle. We want to define the homology class represented by this cycle to be $\phi_h(\mu)$. We need to know that this is independent of the choices of the c_i's. Suppose that c_0 and c_0' are zero volume singular 3–chains in M such that $\partial c_0 = \partial c_0' = h_\#(\beta_i)$. Then $c_0 - c_0'$ is a zero volume 3–cycle in M, and a theorem of Federer implies this cycle must be null homologous. It follows that our definition of $\phi_h(\mu)$ is independent of the choice of 3–chain c_0. Similarly it is also independent of the choice of 3–chain c_1. It follows that $\phi_h(\mu)$ is well defined.

Definition 7.4. *Let M be a closed orientable 3–manifold, and let F be a closed orientable surface. A sweepout of M by F is a map $h : F \times I \to M$ with $\text{Area}(g(F, 0)) = \text{Area}(g(F, 1)) = 0$, and inducing an isomorphism*

$$\phi_h : H_3(F \times I, F \times \partial I) \to H_3(M).$$

Remark 7.5. *One can equally well define a sweepout using homology with $\mathbb{Z}_2$ coefficients.*

There is a natural way to associate a sweepout to a Heegaard splitting of a closed orientable 3–manifold M. In a handlebody H, we can choose a

homotopy which shrinks the boundary surface F onto a 1–dimensional spine Γ of the handlebody, so that the homotopy is through embeddings except at the end of the homotopy. Doing this for each of the two handlebodies of the Heegaard splitting of M yields a sweepout in which the surface F is the Heegaard surface and, as $t \to 0$, the homotopy crushes F down to the chosen spine of one handlebody, and, as $t \to 1$, the homotopy crushes F down to the chosen spine of the other handlebody. Pitts and Rubinstein [29] considered a closed 3–manifold with a strongly irreducible Heegaard splitting. They then constructed a minimal surface that has maximal area in some sweepout derived from the splitting. In a hyperbolic manifold M, a minimal surface of genus g has area less than $2\pi(2g - 2)$, and this implies the same area bound for each surface in the sweepout.

The existence of a sweepout by bounded area surfaces has implications for the geometry of M. It gives upper bounds on the injectivity radius, as noted in Rubinstein [35] and in Bachman-Cooper-White [4].

We will show that one can use harmonic maps to obtain the same area bound given by Pitts and Rubinstein, but without proving the existence of an unstable minimal surface and without needing to assume that the Heegaard splitting is strongly irreducible. The minimal surface obtained by Pitts and Rubinstein is embedded, whereas harmonic maps need not yield embedded surfaces, but embeddedness is not needed to establish results such as a bound on injectivity radius.

The natural way to apply the theory of harmonic maps to a sweepout $h : F \times I \to M$ is to apply Theorem 7.3 or its smooth analogue in order to obtain a sweepout by bounded area surfaces. In the smooth case the difficulty is that at the ends of the sweepout, i.e. at $t = 0$ and $t = 1$, the maps from F to M have 1–dimensional image and so there is no way to define the induced Riemannian metric on F and one cannot directly use the gradient flow. In [24], the authors showed how to work round this difficulty. In the simplicial setting, this is not a problem, but a new difficulty arises. It is not automatic that the maps from F to M at the ends of the sweepout have zero energy. If they do not, then the gradient flow applied to such maps may yield maps of non-zero Riemannian area, so that we no longer have a sweepout. We resolve this difficulty in the following way.

First we consider maps from surfaces to graphs in general. Let $q : S^1 \times I \to I$ denote projection onto the second factor. Let F be a closed surface, and let Γ be a graph.

Definition 7.6. *A map $f : F \to \Gamma$ is of* standard type *if there is a compact subsurface N of F, such that the following conditions hold:*

  (1) *Each component of N is an annulus.*
  (2) *Each component of the closure of $M - N$ is sent by f to a vertex of Γ.*

(3) *Each component annulus A of N is sent by f to an edge of Γ, and $f \mid A = S^1 \times I$ is the composite of q with the attaching map of the edge.*

Any map from F to a graph Γ is homotopic to a map of standard type. For future reference, we set out this result and its proof.

Lemma 7.7. *Let F be a closed surface, and let Γ be a graph. Then any map from F to Γ is homotopic to a map of standard type.*

Proof. We denote the edges of Γ by e_λ, $\lambda \in \Lambda$. The first step is to choose a point x_λ in the interior of each edge e_λ of Γ, and to homotope the map $f : F \to \Gamma$ to be transverse to each x_λ. Let C_λ denote $f^{-1}(x_\lambda)$, so that C_λ is a finite collection of disjoint simple closed curves in F. Further there is a closed interval E_λ in the interior of e_λ, and a homeomorphism of $f^{-1}(E_\lambda)$ with $C_\lambda \times I$, such that the restriction of f to each $S^1 \times I$ component of $f^{-1}(E_\lambda)$ is the composite of q, projection of $S^1 \times I$ onto the second factor, with a homeomorphism onto E_λ. Now consider a map $\phi : \Gamma \to \Gamma$ which, for each edge e_λ of Γ, maps E_λ to e_λ by a homeomorphism on their interiors and maps each component of $\Gamma - \cup E_\lambda$ to the vertex of Γ in that component. The composite $\phi \circ f : F \to \Gamma$ is of standard type. As ϕ is clearly homotopic to the identity, it follows that there is a homotopy of f to a map of standard type, as required. $\square$

Next we consider constructing subdivisions into quadrilaterals. Pick a subdivision of the circle S^1 into intervals. Taking the product with I yields a subdivision of the annulus $S^1 \times I$ into quadrilaterals, which we call a standard subdivision.

Now given a map $f : F \to \Gamma$ of standard type, we choose a subdivision τ of F into quadrilaterals by starting with a standard subdivision of each component of N, and then extending in any way to the rest of F. We will say that such a subdivision is *compatible* with f.

Next let l be the simplicial quasi-metric on (F, τ) induced by f, and suppose that Γ has a path metric. Thus we can consider the simplicial energy $E_S(f)$ of $f : (F, \tau, l) \to \Gamma$. Let P be a quadrilateral of τ. If P is not contained in N, it is mapped by f to a point and so contributes zero to $E_S(f)$. If P is contained in N, two of its edges are mapped to vertices of Γ and the other two edges are mapped onto edges of Γ. Further the two edges mapped to vertices are non-adjacent edges of P. It follows that in this case also P contributes zero to $E_S(f)$. We conclude that $E_S(f) = 0$.

Before returning to our discussion of sweepouts, we need one further refinement of Lemma 7.7. We consider a closed surface F, a finite collection of graphs $\Gamma_1, \ldots, \Gamma_n$, and maps $f_i : F \to \Gamma_i$. Lemma 7.7 tells us that, for each i, we can homotope f_i to be of standard type and can then find a compatible subdivision τ_i of F into quadrilaterals. The following lemma shows how to find a single subdivision of F into quadrilaterals which is compatible with each f_i.

Lemma 7.8. *Let F be a closed surface, and, for $1 \leq i \leq n$, let Γ_i be a graph, and let $f_i : F \to \Gamma_i$ be a continuous map. Then there is g_i homotopic to f_i, and a single subdivision of F into quadrilaterals which is compatible with each g_i.*

Proof. By applying Lemma 7.7, we can assume that each f_i is of standard type. Choose one point in the interior of each edge of Γ_i, and let X_i denote the union of all these points. Thus $f_i^{-1}(X_i)$ is a finite collection of disjoint simple closed curves on F. We can slightly homotope each f_i to a new map g_i, also of standard type, to arrange that the $g_i^{-1}(X_i)$'s intersect each other transversely and that there are no triple points. We let $C_i = g_i^{-1}(X_i)$. For each i, we choose a regular neighborhood N_i of C_i so that the following conditions hold:

(1) The intersection of three distinct N_i's is empty.
(2) For distinct i and j, the boundaries of N_i and N_j meet transversely.
(3) For distinct i and j, each component of $N_i \cap N_j$ is a disc which contains exactly one point of $C_i \cap C_j$, exactly one sub-arc of C_i and exactly one sub-arc of C_j. The picture must be as shown in Figure 3.
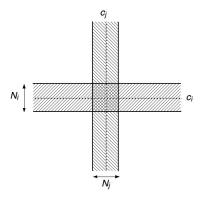

FIGURE 3. Intersection of N_i and N_j

Now we can construct a single subdivision τ of F into quadrilaterals which is simultaneously compatible with each f_i. For distinct i and j, consider a component of $N_i \cap N_j$. As in Figure 3, there is a natural way to regard this as a quadrilateral, with two edges in ∂N_i and two in ∂N_j. We start by choosing all these quadrilaterals. Next we consider a component of N_i. This is an annulus, and some pieces of this annulus have already been declared to be quadrilaterals. It is now trivial to subdivide the rest so as to obtain a standard subdivision of the annulus. Having made these choices for each i, we will have a decomposition of the union of all the N_i's. We now complete this to a subdivision of F in any way. It is easy to see that, after a further homotopy, τ is compatible with each g_i, as required. $\square$

Next we apply Theorem 7.3 to sweepouts.

Lemma 7.9. *If M is a closed orientable hyperbolic 3–manifold with a genus g Heegaard splitting, then M has a sweepout by genus g surfaces with Riemannian area bounded above by $2\pi(2g-2)$.*

Proof. Let F denote the Heegaard surface of M, and let H_0 and H_1 denote the two handlebodies into which F cuts M. For $i = 0, 1$, pick a 1–dimensional spine Γ_i for H_i. By subdividing the edges of these graphs, we can arrange that each edge is a geodesic segment in M. Then there is a sweepout $g : F \times I \to M$ of M by F such that g_0 collapses F to Γ_0, and g_1 collapses F to Γ_1. For each $i = 0$ or 1, there is a map $f_i : F \to \Gamma_i$, which is of standard type, and is homotopic to g_i. We homotope the given sweepout g so that $g_0 = f_0$ and $g_1 = f_1$. The above discussion yields a subdivision τ of F into quadrilaterals which is compatible with both g_0 and g_1, perhaps after further small homotopies of g_0 and g_1. As g_0 and g_1 are now of standard type, and the edges of Γ_0 and Γ_1 are geodesic, it follows that g_0 and g_1 are simplicial. Now we simultaneously homotope each g_t to the simplicial map which agrees with g_t on the vertices. This will not alter g_0 and g_1, so these maps still have zero simplicial energy with respect to the induced simplicial quasi-metric. Finally we can apply the proof of Theorem 7.3 to this sweepout by simplicial maps to obtain a 1–parameter family of simplicial harmonic maps. The crucial fact is that as the energies of g_0 and g_1 are each zero, the gradient flow leaves them unchanged. Thus we have homotoped the original sweepout, through sweepouts, to a sweepout by simplicial harmonic maps, each of which has Riemannian area less than $2\pi(2g-2)$, by Lemma 7.1. $\square$

The Scharlemann-Thompson genus of M can be used in place of the Heegaard genus g in Corollary 7.9, and may be smaller [38].

We end this section by discussing how to replace the use of smooth harmonic maps by simplicial harmonic maps in the work of Hass, Thompson and Thurston [24]. In that paper, the authors considered a particular closed hyperbolic 3–manifold M with a Heegaard splitting $M = H_0 \cup H_1$ and Heegaard surface F. They showed that there could not be an ambient isotopy of M which interchanged H_0 and H_1, and a key role in their argument was played by applying the gradient flow to a 2–parameter family of smooth maps from F to M. We will describe how to replace this part of their argument by the work in this paper.

As in the proof of Lemma 7.9, for $i = 0, 1$, we pick a 1–dimensional spine Γ_i for H_i, so that each edge of Γ_i is a geodesic segment in M. As in that lemma, we can find a sweepout $h : F \times I \to M$ by simplicial harmonic maps such that, for $i = 0, 1$, the map h_i sends F to Γ_i and is of standard type. In particular h_0 and h_1 each has zero simplicial energy.

Next suppose that $\Phi : M \times I \to M \times I$ is an ambient isotopy of M which interchanges H_0 and H_1. Thus Φ is a homeomorphism, Φ_0 is the identity map of M, and $\Phi_1(H_0) = H_1$ and $\Phi_1(H_1) = H_0$. Let $p : M \times I \to M$ denote projection. Then the composite map $F \times I \times I \xrightarrow{h \times 1} M \times I \xrightarrow{\Phi} M \times I \xrightarrow{p} M$

is a 2–parameter family of maps $h_{s,t}$ from F to M, with $h_{s,0} = h_s$, for all s. As h_0 collapses F to the spine Γ_0 of H_0, the map $h_{0,t}$ collapses F to the graph $\Phi_t(\Gamma_0)$ embedded in M. Further each map $h_{0,t}$ is of standard type. Similarly each map $h_{1,t}$ collapses F to the graph $\Phi_t(\Gamma_1)$ embedded in M, and is of standard type. Note that as Φ_1 interchanges H_0 and H_1, we have $\Phi_1(\Gamma_0)$ is a spine of H_1 and $\Phi_1(\Gamma_1)$ is a spine of H_0.

The next step is to replace each map $h_{s,t}$ by the unique simplicial map which agrees with $h_{s,t}$ on the vertices of τ. As each h_s is already simplicial, the maps $h_{s,0} = h_s$ do not change. The maps $h_{0,t}$ and $h_{1,t}$ may change, but they will continue to have zero simplicial energy. This is because any edge of τ which was mapped to a point before the change will continue to be mapped to a point after the change. Thus the gradient flow will leave all the maps $h_{0,t}$ and $h_{1,t}$ unchanged. Hence when we apply Theorem 7.3 to this 2–parameter family of maps, we will obtain a family $f_{s,t}$ of harmonic maps, each with Riemannian area less than $2\pi\,|\chi(F)|$, such that all the maps $f_{0,t}$ and $f_{1,t}$ have zero simplicial energy and hence zero Riemannian area.

In [24], the authors considered hyperbolic 3–manifolds of fixed Heegaard genus which have arbitrarily high volume. This implies that a sweepout $h : F \times I \to M$ of M by F in which each surface has Riemannian area less than $2\pi\,|\chi(F)|$, must be "long" in a certain sense. They show that this implies that in any 2–parameter family that flips the orientation of the Heegaard splitting, there must be a surface of large area. This in turn was used to show that it is not possible to connect such a Heegaard sweepout to the same sweepout with opposite orientation without stabilizing until the genus is doubled.

8. Sweepouts in higher dimensions

In this section we define the genus of an n–dimensional closed orientable manifold, extending the idea of the Heegaard genus of a 3–manifold to closed manifolds of arbitrary dimension. To do so we extend the definition of a sweepout of a 3–manifold by a 1–parameter family of surfaces to a sweepout of a closed n–dimensional manifold M by an $(n-2)$–parameter family of surfaces. We replace the unit interval I in a Heegaard sweepout with an arbitrary compact $(n-2)$–dimensional manifold which parametrizes the surface maps.

For this construction we let X be the orientable total space of a bundle over a compact $(n-2)$–manifold B with fiber a closed orientable surface F, with projection map $\pi : X \to B$, and with ∂B possibly empty. We will consider an appropriately defined degree one map $h : X \to M$ such that, for all $b \in \partial B$,

$$Area(h(\pi^{-1}(b))) = 0.$$

As before, we need to define a homomorphism

$$\phi_h : H_n(X, \partial X) \to H_n(M)$$

to make sense of the notion of degree for such maps. Again note that we could also use homology with $\mathbb{Z}_2$ coefficients.

Since $H_n(X, \partial X)$ is a cyclic group generated by the relative fundamental class μ, a homomorphism ϕ_h to $H_n(M)$ is determined by the action of ϕ_h on μ. So long as h satisfies some mild smoothness condition, the $(n-1)$–volume of $h(\partial X)$ will equal zero because $Area(h(\pi^{-1}(b))) = 0$ for $b \in \partial B$. So $h(\partial X)$ gives a zero volume $(n-1)$–chain in M which in turn is the boundary of a zero volume n–chain c in M. Thus $h_\#(\mu) + c$ is a n–cycle in M whose homology class $\alpha \in H_n(M)$ is independent of the choice of c. We set $\phi_h(\mu) = \alpha$. This leads to the following definition.

Definition 8.1. *A* sweepout *of a closed orientable n–manifold M by a closed orientable surface F is a map $h : X \to M$ from a bundle X with orientable total space, fiber F, and base space a compact $(n-2)$–manifold B, whose boundary may be empty, such that*

   (1) *$Area(h(\pi^{-1}(b))) = 0$, for all $b \in \partial B$, and the $(n-1)$–volume of $h(\partial X)$ also equals zero, and*
   (2) *h induces an isomorphism*

$$\phi_h : H_n(X, \partial X) \to H_n(M).$$

Definition 8.2. *If M is a closed orientable n–manifold which admits a sweepout by genus g surfaces, for some g, then the* genus *of M is the smallest such number g.*

For 3–manifolds this number is clearly less than or equal to the Heegaard genus, since the above definition allows non-embedded and non-disjoint surfaces to sweep over the 3–manifold. In addition, it allows the consideration of sweepouts by bundles over a circle, as well as over an interval. We do not know whether or not equality always holds for a given B.

Here are some examples of n–manifolds which admit a sweepout by surfaces. Our first examples are simply the product of a 3–manifold with another manifold. Let Y be a closed $(n-3)$–dimensional manifold, let Z be a closed 3–manifold, and let $M = Y \times Z$. Choose a sweepout $h : I \times F \to Z$ of Z by a closed surface F, let $B = Y \times I$, and let $X = B \times F$. Then the map $X = B \times F = Y \times (I \times F) \overset{(1,h)}{\to} Y \times Z = M$ is a sweepout of M by F.

Our next examples will play an important role in our discussion below. Let B be a compact $(n-2)$–dimensional manifold, and let F be a closed orientable surface of genus g. Let X be a bundle over B with fiber F whose restriction to ∂B is trivial. Choose a homeomorphism f from F to the boundary of a handlebody H, and let g denote the homeomorphism $f \times 1 : F \times \partial B \to \partial H \times \partial B$. We form a closed n–manifold M by attaching X to $H \times \partial B$ using the homeomorphism of their boundaries $\partial X = F \times \partial B \overset{g}{\to} \partial H \times \partial B = \partial(H \times \partial B)$. As in the previous section, there is a homotopy which shrinks the boundary surface F of the handlebody H onto a 1–dimensional spine Γ of H, such that the homotopy is through embeddings except at the end of the homotopy. Taking the product of this homotopy with the identity

map of B yields a map $(I \times \partial B) \times F \to \partial B \times H$ which, for each b in ∂B, shrinks $\{b\} \times F$ onto $\{b\} \times \Gamma$. Together with the identity map of X, this yields a map from $X \cup ((I \times \partial B) \times F)$ to M which is clearly a sweepout of M by F. Of course $X \cup ((I \times \partial B) \times F)$ is homeomorphic to X, so we can and will regard this sweepout as a map from X to M. This sweepout has the special property that for all interior points of B, the associated map from F to M is an embedding, and that for all boundary points of B, the associated map from F to M has image a copy of the graph Γ. If ∂B is not connected, one can make this construction slightly more general by separately choosing a homeomorphism from F to the boundary of the handlebody H for each component of ∂B.

When B is the $(n-2)$-ball, the bundle must be trivial over ∂B. These *n-dimensional Heegaard splittings* therefore fall into this class of examples.

Now suppose that we have a sweepout $h : X \to M$ of a closed orientable n–manifold M by a closed orientable surface F. In addition suppose that M has a metric of negative curvature. We would like to apply the ideas of the previous section to show that M admits a sweepout by small area maps from F to M. As in that section, the key to doing this is being able to find a sweepout such that, for each b in ∂B, the associated map from F to M has zero energy. The result we obtain is the following.

Theorem 8.3. *Let M be a closed Riemannian n–manifold with sectional curvatures bounded above by $-a < 0$, let F be a closed orientable surface of genus $g \geq 2$, and let τ be a subdivision of F into quadrilaterals. Suppose that $h : X \to M$ is a sweepout of M by F such that, for each b in ∂B, the associated map $F \to M$ is simplicial and has zero simplicial energy with respect to the induced quasi-metric on (F, τ). Then h can be homotoped, fixing ∂X, to a new sweepout so that, for each b in B, the associated map from F to M has Riemannian area at most $2\pi(2g-2)/a$.*

Proof. We will use the same ideas as in the proof of Lemma 7.9 in the case of dimension 3.

If X is a trivial bundle over B, we can simply apply Theorem 7.3 to homotope the sweepout to a family of simplicial harmonic maps from F to M. This homotopy will be fixed for fibers over ∂B, so still gives a degree one map from X to M, and is therefore still a sweepout. Theorem 7.3 also tells us that in this new sweepout the area of each surface is less than $2\pi(2g-2)/a$, as required.

If X is a nontrivial bundle over B, it is still locally trivial, and that suffices to apply Theorem 7.3. $\square$

Given a sweepout, it is not obvious that one can find a subdivision into quadrilaterals that satisfies the hypotheses of the above theorem. But it is easy to do this in the cases discussed just before the theorem. As in the second set of examples, we let B be a compact $(n-2)$–dimensional manifold, let F be a closed orientable surface of genus g, and let X be a bundle over B with fiber F whose restriction to ∂B is trivial. Finally let M

be a closed n–manifold formed by attaching X to $H \times \partial B$ using a product homeomorphism of their boundaries. Then there is a sweepout $h : X \to M$ such that, for all interior points of B, the associated map from F to M is an embedding, and, for all boundary points of B, the associated map from F to M has image a graph which depends only on the component of ∂B. Further, for a component C of ∂B, the maps from F to the graph Γ associated to the points of C are all equal. As ∂B has only finitely many components, Lemma 7.8 tells us that after a suitable homotopy of h, there is a single subdivision τ of F into quadrilaterals which is compatible with all of these maps. Now it is immediate that, for each b in ∂B, the associated map from F to M has zero simplicial energy with respect to the induced quasi-metric on (F, τ). If we assume that M has negative curvature, then we can homotope this sweepout to the sweepout by simplicial maps which agrees with the original sweepout on the vertices of τ. Thus, for each b in ∂B, the associated simplicial map from F to M has zero simplicial energy with respect to the induced quasi-metric on (F, τ), verifying all the hypotheses of the above theorem.

We can generalize the previous discussion as follows. Let M be as in the previous paragraph, but do not assume that M has negative curvature. Suppose that there a map k of degree 1 from M to some negatively curved n–manifold M'. Then the composite map $k \circ h : X \to M'$ is a sweepout such that, for each b in ∂B, the associated map from (F, τ, l) to M' has zero simplicial energy. Further, we can homotope this sweepout to the sweepout by simplicial maps which agrees with the original sweepout on the vertices of τ, again verifying all the hypotheses of the above theorem.

9. Varying the simplicial metric

In this section we examine global minimizers of simplicial energy over the space of all simplicial metrics for a fixed triangulation. It was known to Courant and Rado that one way to find a least area map in a homotopy class is to find a map that minimizes energy not just for a given metric on the domain but among all possible metrics [12, 34]. We now present this argument in the simplicial setting, where it becomes considerably easier.

Lemma 9.1. *Let $f : F \to M$ be a map of a triangulated closed surface F to a closed manifold M with a Riemannian metric. Let $E_S(f, l)$ denote the simplicial energy of f with respect to a simplicial quasi-metric l on F.*

(1) *If $A_S(f)$ has least possible value among all maps homotopic to f, then there is a simplicial quasi-metric l on F such that $E_S(f, l)$ has the least possible value among all maps homotopic to f and over all simplicial quasi-metrics on F. Further $E_S(f, l) = A_S(f)$.*

(2) *If l is a simplicial quasi-metric on F, and f is a map such that $E_S(f, l)$ has least possible value among all maps homotopic to f and over all simplicial quasi-metrics on F, then $A_S(f)$ also has the least possible value, and $E_S(f, l) = A_S(f)$.*

Proof. (1) Let l denote the simplicial quasi-metric on (F, τ) induced by f, i.e. for each edge e_i of τ , we set $l_i = L_i$. For this quasi-metric, all the stretch factors σ_i are equal to 1 on all edges with $l_i > 0$ and therefore $E_S(f, l) = A_S(f)$, by Lemma 3.3. The hypothesis that $A_S(f)$ has least possible value implies that, for any map f' homotopic to f, we have $A_S(f') \geq A_S(f)$. Thus for any simplicial quasi-metric k on F, we have the inequalities

$$E_S(f', k) \geq A_S(f') \geq A_S(f) = E_S(f, l).$$

It follows that $E_S(f, l)$ has the least possible value over all maps homotopic to f and over all simplicial quasi-metrics on F, and that $E_S(f, l) = A_S(f)$, as claimed.

(2) Suppose that $A_S(f)$ does not have the least possible value. Thus there is a map f' homotopic to f with $A_S(f') < A_S(f)$. We define a simplicial quasi-metric k on F by setting k_i equal to L_i'. For this quasi-metric and the map f', the stretch factors σ_i are equal to 1 on all edges with $l_i > 0$. Thus, by Lemma 3.3, we have

$$E_S(f', k) = A_S(f') < A_S(f) \leq E_S(f, l)$$

which contradicts the hypothesis that $E_S(f, l)$ has least possible value.

We conclude that $A_S(f)$ must have the least possible value. Now part (1) implies that there is a simplicial quasi-metric k on F such that $E_S(f, k)$ has the least possible value among all maps homotopic to f and over all simplicial quasi-metrics on F. Further $E_S(f, k) = A_S(f)$. As $E_S(f, l)$ also has least possible value, it follows that $E_S(f, l)$ must equal $E_S(f, k)$, so that $E_S(f, l) = A_S(f)$, as required.   □

The proof of the above lemma is much simpler than in the smooth setting. This is because any map of F into M for which one can measure the lengths of the edges of τ induces a simplicial quasi-metric on F, whereas in the smooth setting, one needs f to be an immersion in order to induce a Riemannian metric on F.

In Section 5 we saw that a finite energy map from a surface with a fixed simplicial metric or quasi-metric to a manifold of non-positive curvature can be homotoped to a simplicial harmonic map that minimizes simplicial energy. We now examine what happens when the simplicial metric on the domain surface is allowed to change. We will show that the class of maps homotopic to a given map $f : F \to M$ contains a map with minimal simplicial energy among all maps homotopic to f and all possible simplicial quasi-metrics (F, τ, l). In the smooth setting this gives a least area map in the homotopy class of f. The existence of a least area map in that setting was shown by Sacks and Uhlenbeck [39] and by Schoen and Yau [40], with the hypotheses that the surface is orientable with genus at least 1, and its fundamental group injects into the fundamental group of M.

It would seem more natural to consider only simplicial metrics rather than allow quasi-metrics, but the following discussion explains why this is unreasonable. In fact this is why we introduced the idea of a simplicial

quasi-metric. Consider a map $f_1 : F \to M$ from a triangulated surface (F, τ) to a Riemannian manifold and let

$$\mathcal{I} = \inf\{E_S(f)\}$$

where the infimum is taken over all maps f homotopic to f_1 and over all simplicial metrics on τ. We would like to establish the existence of a simplicial harmonic map homotopic to f_1 with simplicial energy equal to $\mathcal{I}$. There is a sequence l^n of simplicial metrics on F, and of homotopic simplicial maps $f_n : F \to M$ such that $E_S(f_n) \to \mathcal{I}$, as $n \to \infty$. We call such a sequence a *simplicial-energy minimizing* sequence. The main difficulty we face is that even if the sequence l^n converges to a function l^0 on the edges of τ, it may be that $l^0(e_i) = 0$, for some edge e_i of τ. It is clear that $l^0(e_i) \geq 0$, for all i, and that l^0 satisfies the triangle inequality for each triangle of τ. Thus l^0 is a simplicial quasi-metric on F. There is a further difficulty with this approach which is that it is possible that $l^0(e_i) = 0$, for every edge e_i of τ. For example, suppose that l is a simplicial metric on F, and $f : F \to M$ is a simplicial map such that $E_S(f) \to \mathcal{I}$. For each $n \geq 1$, we define the simplicial metric l^n by $l^n(e_i) = l(e_i)/n$, for each edge e_i of τ, and the simplicial map f_n to equal f. The scale invariance of simplicial energy implies that $E_S(f_n) = \mathcal{I}$, for each $n \geq 1$, but the sequence l^n converges to the simplicial quasi-metric l^0 such that $l^0(e_i) = 0$, for every edge e_i of τ. The analogous phenomenon occurs in the smooth setting, and as in that setting we avoid this problem by changing our initial choice of minimizing sequence.

The following lemma shows that allowing quasi-metrics does not change the minimal energy for a homotopy class.

Lemma 9.2. *Let $f_1 : F \to M$ be a map of a triangulated closed surface (F, τ) to a closed Riemannian manifold M. Let $\mathcal{I}$ denote the infimum of simplicial energies over the family $\mathcal{F}$ of all maps f homotopic to f_1 and over all simplicial metrics on τ, and let $\mathcal{I}_0$ denote the infimum of simplicial energies over the family $\mathcal{F}_0$ of all finite energy maps f homotopic to f_1 and over all simplicial quasi-metrics on τ.*

Then $\mathcal{I} = \mathcal{I}_0$.

Proof. As $\mathcal{F} \subset \mathcal{F}_0$, it is immediate that $\mathcal{I} \geq \mathcal{I}_0$.

Now let l^0 be a simplicial quasi-metric on F, and let $f : (F, \tau, l^0) \to M$ be a finite energy map homotopic to f_1. Consider the family of edge length functions l^t, defined for $t > 0$ by $l^t(e_i) = l^0(e_i) + t$. As l^t is positive and satisfies all triangle inequalities in τ, it follows that l^t is a simplicial metric for each $t > 0$.

Let $E_S(f, l^t)$ denote the energy of the map $f : F \to M$ computed using the simplicial quasi-metric l^t on τ. For $t > 0$ we have

$$E_S(f, l^t) = \frac{1}{2} \sum \left(\frac{L_i^2}{(l^t(e_i))^2} + \frac{L_j^2}{(l^t(e_j))^2} \right) l^t(e_i) l^t(e_j).$$

If either of $l^0(e_i)$ or $l^0(e_j)$ is zero then so is L_i or L_j, so that

$$\lim_{t \to 0} E_S(f, l^t) = E_S(f, l^0).$$

It follows immediately that $\mathcal{I} \leq \mathcal{I}_0$, so that $\mathcal{I} = \mathcal{I}_0$ as required.   □

Next we show the existence of a simplicial quasi-metric and a map which minimizes energy in a homotopy class. Note that, unlike the smooth case, this theorem places no requirements on the given map f_1 or on its homotopy class. In the smooth case, an incompressibility condition is required.

Theorem 9.3. *Let $f_1 : F \to M$ be a map of a triangulated closed surface (F, τ) to a closed non-positively curved manifold M. Then there is a simplicial quasi-metric l^0 on F, and a simplicial map f_0 homotopic to f_1, such that f_0 minimizes simplicial energy among all maps homotopic to f_1 and all choices of simplicial quasi-metric on (F, τ). Further f_0 minimizes simplicial area among all maps homotopic to f_1.*

Proof. Let $\mathcal{I}$ denote the infimum of simplicial energies over the family of all finite energy maps f homotopic to f_1 and over all simplicial quasi-metrics on τ.

Consider a sequence l^n of simplicial quasi-metrics on (F, τ) and finite-energy simplicial maps $f_n : F \to M$, each homotopic to f, such that $\lim_{n \to \infty} E_S(f_n) = \mathcal{I}$. For each n, let L_i^n denote the length of $f_n|e_i$ and define a new simplicial quasi-metric k^n on (F, τ) by $k^n(e_i) = L_i^n$. For this quasi-metric, the map f_n is a simplicial isometry, with $l_i^n = L_i^n$ and $E_S(f_n, k^n) = A_S(f_n)$. Now $A_S(f_n)$ is independent of the simplicial quasi-metric on F and $A_S(f_n) \leq E_S(f_n, l^n)$ by Lemma 3.3. Hence $E_S(f_n, k^n) \leq E_S(f_n, l^n)$, and replacing each simplicial quasi-metric l^n by k^n retains the property that $\lim_{n \to \infty} E_S(f_n) = \mathcal{I}$, so that f_n is still an energy minimizing sequence. In particular, the sequences $E_S(f_n)$ and $A_S(f_n)$ are bounded.

We now claim that, after passing to a subsequence, the sequence L_i^n is bounded for each i. First note that by passing to a subsequence we can arrange that each sequence L_i^n is either convergent or unbounded. Now consider an edge e_0 and suppose that $\{L_0^n\}$ is unbounded. By passing to a further subsequence we can assume that $\{L_0^n\}$ is increasing and unbounded. As e_0 meets one triangle of τ on each of its two sides, there are four additional edges of τ which belong to a triangle containing e_0. We denote these edges by e_1, e_2, e_3, e_4 (they may not all be distinct in F). The terms in $A_S(f_n)$ which involve L_0^n are $\sum_{i=1}^4 L_0^n L_i^n = L_0^n \sum_{i=1}^4 L_i^n$, and this sequence of sums is bounded, as we noted above. As L_0^n is increasing and unbounded, the sum $\sum_{i=1}^4 L_i^n$ must approach 0 as $n \to \infty$. Since each $L_i^n \geq 0$, we deduce that $L_i^n \to 0$ as $n \to \infty$, for $i = 1, 2, 3, 4$. Now the triangle inequality implies that L_0^n must also converge to 0 as $n \to \infty$, a contradiction. This establishes that L_i^n is bounded for each i, as claimed. It follows that there is a subsequence of f_n for which each sequence L_i^n is convergent. As M is compact, we can find a further subsequence such that for each vertex v of τ, the sequence $\{f_n(v)\}$

converges. As $k^n(e_i) = L_i^n$, the sequence $\{k^n\}$ also converges and the limit is a simplicial quasi-metric k^0 on F. So the sequence of finite-energy simplicial maps f_n converges to a finite-energy simplicial map f_0 with $E(f_0) = \mathcal{I}$. This completes the proof of the theorem.   □

We say that a map of a surface into M is *simply essential* if it sends nontrivial simple loops on the surface to nontrivial elements of $\pi_1(M)$. For two-sided embedded surfaces in a 3–manifold this is equivalent to being π_1–injective, but for singular surfaces in a 3–manifold this equivalence is unknown, and is called the simple loop conjecture. The result below shows that if we assume that the map $f_1 : F \to M$ in Theorem 9.3 is simply essential, then we can refine the result to assert the existence of a simplicial metric (not just quasi-metric) on F and an energy minimizing map.

Corollary 9.4. *Let $f_0 : F \to M$ be a simply essential simplicial map from a triangulated closed surface (F, τ) to a closed non-positively curved manifold M, and suppose that l is a simplicial quasi-metric on (F, τ) such that $E(f_0) = \mathcal{I}$. Then there is a new triangulation τ' of F, obtained by collapsing edges and triangles of τ, a new simplicial map f_0' homotopic to f_0, and a simplicial metric on τ' for which $E(f_0') = \mathcal{I}$.*

Proof. As f_0 is simplicial, it maps any edge of τ to a point or to a geodesic arc, and it maps any triangle of τ to a point or to a geodesic arc or to a triangle. If an edge e of τ is zero length, i.e. $l(e) = 0$, the fact that f_0 has finite energy implies that it must map e to a point. If a triangle σ of τ has $\partial\sigma$ consisting entirely of zero length edges, the fact that f_0 is simplicial implies that it must map σ to a point. If $\partial\sigma$ has two zero length edges, the fact that l satisfies the triangle inequality implies that all three edges of $\partial\sigma$ are zero length. Finally if $\partial\sigma$ has one zero length edge, the fact that l satisfies the triangle inequality implies that the other two edges of $\partial\sigma$ have the same l–length, and the fact that f_0 is simplicial implies that these edges have the same image under f_0.

Now we let Δ denote the subcomplex of τ consisting of all zero length edges of τ together with all triangles of τ whose boundary consists of zero length edges. Thus f_0 maps each component of Δ to a point. As f_0 is simply essential, if N denotes a regular neighborhood of a component of Δ, all except one of the components of the closure of $F - N$ must be a disc, and the union of N with these discs must itself be a disc. Suppose there is a disc component B of the closure of $F - N$, and let R denote the component of $F - \Delta$ which contains B. Then ∂R is contained in Δ, so is mapped to a point by f_0. The facts that f_0 maps ∂R to a point, is simplicial, and has least energy implies that it must map all of R to a point. Thus R is contained in Δ, which is a contradiction. This shows that no component of the closure of $F - N$ can be a disc, so that N itself must be a disc.

We form the quotient F'' of F by collapsing every simplex of Δ to a point. Note that f_0 must factor through a map $f_0'' : F'' \to M$. The preceding

discussion implies that F'' is homeomorphic to F. The image of τ in F'' need not be a triangulation. For if σ is a triangle of τ such that $\partial\sigma$ has one zero length edge, then the image of σ in F'' is a 2–gon. But the image of σ in M is a geodesic arc, and so we can take a further quotient F' of F'', by collapsing each such 2–gon to an arc, and f_0 must factor through a map $f_0' : F' \to M$. Again F' must be homeomorphic to F, and now the image of τ in F' is naturally a triangulation τ' of F'. (Note that even if τ is a simplicial triangulation of F, the triangulation τ' of F' may not be simplicial.)

Distinct edges of τ cannot map to the same edge in F''. Thus if distinct edges e_1 and e_2 of τ map to the same edge in F', it must be because certain 2–gons in F'' were collapsed, so that $l(e_1) = l(e_2)$. It follows that there is a well defined function l' on the edges of τ', given by $l'(e') = l(e)$, where e' is any edge of τ' and e is any edge of τ which maps to e'. Clearly l' is never zero and so is a simplicial metric on τ'. Further, as f_0 factors through the simplicial map $f_0' : F' \to M$, it follows that $E(f_0') = \mathcal{I}$. Finally, as F' is obtained from F by collapsing disjoint contractible subsets, f_0' must be homotopic to f_0. $\square$

A consequence of Corollary 9.4 is that any potential counterexample for the simple loop conjecture in a hyperbolic 3–manifold can be chosen to be geometrically controlled; it can be homotoped to be simplicial with area less than $2\pi |\chi|$. Since the simple loop conjecture has been proven for Seifert fiber spaces [23] and for graph manifolds [36], the hyperbolic case is central to this problem. Note that in [10], Cooper and Manning gave an example of a representation of a surface group into $PSL_2(\mathbb{C})$ which has nontrivial kernel, but no element of the kernel is represented by an essential simple closed curve. However this is not a counterexample to the simple loop conjecture as their representation is not known to be discrete, nor to have image in the fundamental group of a hyperbolic 3–manifold.

10. Maps between surfaces and canonical triangulations

In this section we consider a metrized triangulated closed surface (F_1, τ, l), and a closed surface F_2 with Euler characteristic $\chi < 0$, and with a Riemannian metric of non-positive curvature. We will consider maps from F_1 to F_2 which are homotopic to a given homeomorphism. We show that in the homotopy class of such a map there is a simplicial harmonic map which is a homeomorphism. If F_2 is negatively curved, this simplicial harmonic map is then unique in its homotopy class. Thus the simplicial harmonic map never contains unnecessary singularities or folds. Hence if τ is a fixed triangulation of F_1, this map gives a canonical triangulation of F_2 that varies continuously as we vary the metric h. Note that if F_1 and F_2 were orientable, with F_1 of genus 3, and F_2 of genus 2, and if f was a map of degree 1 from F_1 to F_2, then f would be homotopic to a simplicial harmonic map g, by Proposition 5.5, but g could not be a homeomorphism nor could it be a covering

map. Thus in general simplicial harmonic maps between surfaces must have some type of singularity. See Sampson [37] for an analysis of the possible singularities in the smooth case.

In order to prove the results in this section, we need to slightly restrict the type of triangulations we allow.

Definition 10.1. *A triangulation τ of a compact surface F is* good *if no edge of τ forms a null homotopic loop, and no union of two edges of τ forms a null homotopic loop.*

Clearly if a triangulation of F is not good and F has a metric of non-positive curvature, it is impossible to have a homeomorphism of F which sends every edge to a geodesic.

For the rest of this section, we fix a metrized triangulated surface (F_1, τ, l), a closed surface (F_2, h) with Euler characteristic $\chi < 0$, and with a non-positively curved Riemannian metric h, and a homeomorphism $f : F_1 \to F_2$. Let $I(f)$ denote the infimum of the simplicial energies of all homeomorphisms isotopic to f, and let τ^1 denote the 1–skeleton of τ. Note that we can isotope f to be piecewise geodesic on τ^1, so that the simplicial energy is now defined. In particular, the isotopy class of f always contains homeomorphisms whose simplicial energy is defined. We will consider an energy minimizing sequence of homeomorphisms $\{f_n\}$ from F_1 to F_2 each of which is isotopic to f, and will show that there is a subsequence which yields a simplicial homeomorphism $(F_1, \tau, l) \to (F_2, h)$ whose energy equals $I(f)$. The special case where τ is a 1–vertex triangulation and h is negatively curved turns out to be much simpler and will be treated separately after the following lemma. Note that this first lemma does not require that τ be good, nor that the metric on F_2 be non-positively curved.

Lemma 10.2. *Suppose that F_1 and F_2 are closed surfaces with negative Euler number. Let $f : (F_1, \tau, l) \to (F_2, h)$ be a homeomorphism from F_1, with simplicial metric l, to F_2 with Riemannian metric h. Let $\{f_n\}$ be a sequence of homeomorphisms from F_1 to F_2, each of which is isotopic to f and whose simplicial energies approach $I(f)$. Then there is a subsequence of the f_n's whose restrictions to τ^1 converges to a map $f_0 : \tau^1 \to F_2$. Further any map f_0 obtained in this way maps each edge of τ to a point or to a geodesic arc.*

Proof. As f_n has finite simplicial energy, it must map each edge of τ to a rectifiable arc. As the sequence of simplicial energies of the f_n's is convergent, it is bounded. As F_2 is compact, Ascoli's Theorem shows that the sequence of f_n's has a subsequence whose restriction to τ^1 converges to a continuous function f_0. If e is an edge of τ, then the sequence of lengths of the arcs $f_n \mid e$ is also bounded, so that $f_0 \mid e$ must have finite length. Thus f_0 maps each edge of τ to a rectifiable arc. Further for any extension of f_0 to all of F_1, we have $E_S(f_0) = I(f)$.

We claim that if z is any point of $f_0(\tau^1)$ which is not the image of a vertex of τ, then there must be a disc D centered on z such that $f_0(\tau^1)$ meets D in geodesic arcs. Assuming this claim, the fact that there can be only finitely many images of vertices of τ immediately implies that f_0 must map each edge of τ to a point or to a geodesic arc, as required.

To prove our claim, suppose that z is a point of $f_0(\tau^1)$ which is not the image of a vertex of τ. Then suitably small neighborhoods of z will also not contain the image of a vertex of τ. As f_0 maps each edge of τ to a rectifiable arc, it follows that there is a small convex disc D centered on z such that $f_0^{-1}(\partial D)$ consists of a finite number of points $w_1, \ldots, w_k$. As D does not contain the image of a vertex, the intersection of the image of $f_0(\tau^1)$ with D consists of isolated points in ∂D together with arcs joining the images of the remaining w_i's in pairs. We replace these arcs by geodesic arcs joining the images of the w_i's in pairs in the same way, to obtain a new map f_0' of τ^1 into F_2. As f_0 is the limit of the homeomorphisms $\{f_n\}$, we could make the analogous changes to f_n, for all large enough values of n. The geodesic arcs obtained when we do this for f_n will all be disjoint. We conclude that there is a new sequence f_n' of homeomorphisms such that the limit of $f_n' \mid \tau^1$ is f_0'. As no edge of τ is lengthened by this construction, we have $E_S(f_n') \le E_S(f_n)$, for each n. If there is any edge of τ whose image under f_0 meets D in a non-geodesic arc, our construction of f_0' will strictly shorten the image of this edge, so that the limit of the sequence $\{E_S(f_n')\}$ is strictly less than $I(f)$, which is a contradiction. We deduce that $f_0(\tau^1)$ must meet D in geodesic arcs, which completes the proof of the claim. $\square$

Before proceeding further, we consider the special case when τ is a 1–vertex triangulation of F_1, and F_2 has a Riemannian metric of negative curvature.

Theorem 10.3. *Let F_1 and F_2 be closed surfaces with negative Euler number. Let τ be a good 1–vertex triangulation of F_1, let h be a negatively curved Riemannian metric on F_2, and let $f : (F_1, \tau, l) \to (F_2, h)$ be a homeomorphism. Then the following hold:*

    (1) *f is isotopic to a simplicial homeomorphism.*
    (2) *The unique least energy simplicial map homotopic to f is a simplicial homeomorphism and is isotopic to f.*
    (3) *Any homeomorphism from F_1 to F_2 which is homotopic to f must be isotopic to f.*

Remark 10.4. *By taking F_1 equal to F_2 in part 3), we see that we have a new proof of the classical fact that homotopic homeomorphisms of F are isotopic, so long as F has negative Euler number.*

Proof. 1) Lemma 10.2 tells us that there is a sequence $\{f_n\}$ of homeomorphisms from F_1 to F_2 each of which is isotopic to f, and whose simplicial energies approach $I(f)$, such that the restrictions of the f_n's to τ^1 converge to a map f_0 which sends each edge of τ to a point or to a geodesic arc. Let

v denote the vertex of τ. As τ is a good 1–vertex triangulation, f_0 cannot send an edge of τ to a point, and it must send distinct edges of τ to distinct geodesics in F_2. In particular, the restriction of f_0 to each edge of τ is an immersion. If f_0 fails to embed τ^1 in F_2, there must be an edge e of τ with an interior point w such that $f_0(w)$ equals $f_0(v)$. See Figure 4. If U is a suitably small neighborhood of $f_0(v)$ in F_2, the image of τ^1 in U consists of the image of a neighborhood W of w in e, together with the image of a neighborhood V of v in τ^1. As the f_n's are all homeomorphisms, $f_0(V)$ lies on one side of $f_0(W)$, as shown in Figure 4.
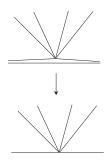

FIGURE 4. A vertex limiting to meet the interior of an edge.

Note that it is possible that some edges of $f_0(V)$ embed in $f_0(W)$. In all cases, we will obtain a contradiction, implying that f_0 must embed τ^1 in F_2. Assuming this, as the f_n's are all homeomorphisms, and f_0 is the limit of the sequence $f_n \mid \tau^1$, there is a unique extension of f_0 to a simplicial homeomorphism g from (F_1, τ, l) to (F_2, h). The Alexander Trick applied to each triangle then implies that f is isotopic to this simplicial homeomorphism, as required.

To obtain a contradiction, we first consider the case where each edge of $f_0(V)$ meets $f_0(W)$ only in $f_0(v)$. Now consider a homotopy of f_0 which moves $f_0(v)$ a small distance orthogonally to $f_0(W)$ towards this side of $f_0(W)$. Let f_0' denote the map so obtained. We further choose f_0' to be equal to f_0 outside V, and so that $f_0'(V)$ consists of a union of geodesic arcs joining $f_0'(v)$ to $f_0(V) \cap \partial U$. This homotopy will shorten the image of every edge of τ, so that f_0' has strictly less simplicial energy than f_0. As f_0' embeds τ^1 in F_2, it can be extended to a homeomorphism which is isotopic to f. As the energy of f_0' is less than $I(f)$, this is the required contradiction.

Now suppose that some edges of $f_0(V)$ embed in $f_0(W)$. This implies that a triangle of which e is an edge must be collapsed onto $f_0(e)$. No other triangles can collapse onto $f_0(e)$, as f_0 sends distinct edges of τ to distinct geodesics in F_2. Now we apply the preceding construction to obtain a map f_0' with strictly less simplicial energy than f_0. Of course f_0' does not embed τ^1 in F_2, but by homotoping very slightly the edges collapsed onto $f_0(e)$, we can obtain an embedding f_0'' of τ^1 in F_2 with strictly less simplicial energy than f_0. Note that if e' is a geodesic segment in F, and we move one end P

of e' along a geodesic orthogonal to e', then the derivative of the length of e' with respect to the distance moved by P is zero. This is why f_0'' has strictly less simplicial energy than f_0, if we move $f_0(v)$ a suitably small distance. As above this yields the required contradiction.

2) By construction, g is a homeomorphism which is isotopic to f, and g minimizes E_S among all homeomorphisms isotopic to f. In particular, g is a critical point for E_S under all variations through homeomorphisms. As any small deformation of g must still be a homeomorphism, this implies that g is a critical point for E_S under all variations. As h is negatively curved, it follows that g is the unique least energy simplicial map homotopic to f, as required.

3) Let f_1 denote a homeomorphism from F_1 to F_2 which is homotopic to f. By parts 1) and 2), f_1 is isotopic to an energy minimizing simplicial homeomorphism which must be g. In particular f_1 must be isotopic to f, as required. $\square$

Now we return to the case of a general triangulation. The result we obtain is similar to previous results obtained in the smooth and orientable setting. Sampson [37] and Schoen and Yau [40] showed that there is a least energy harmonic diffeomorphism in each homotopy class of diffeomorphisms from a Riemannian surface to a surface of non-positive curvature. Later it was shown by Jost and Schoen [27] that there is a least energy harmonic diffeomorphism in each homotopy class of diffeomorphisms between two Riemannian surfaces of genus at least 2, even if the target surface has some positive curvature. In [11], Coron and Helein showed that, in all cases, any harmonic diffeomorphism between two Riemannian surfaces of genus at least 2, must be energy minimizing in its homotopy class, and is the unique energy minimizing map in that homotopy class.

Theorem 10.5. *Suppose that F_1 and F_2 are closed surfaces with negative Euler number. Let $f : (F_1, \tau, l) \to (F_2, h)$ be a homeomorphism from F_1, with simplicial metric l, to F_2 with a Riemannian metric of non-positive curvature. Further assume that τ is a good triangulation. Then f is isotopic to a simplicial harmonic homeomorphism $g : (F_1, \tau, l) \to (F_2, h)$ that minimizes E_S among all homeomorphisms isotopic to f. Further, if (F_2, h) is negatively curved, then g is the unique simplicial harmonic map in the homotopy class of f, and so minimizes energy in this homotopy class.*

Proof. Recall that $I(f)$ denotes the infimum of the simplicial energies of all homeomorphisms isotopic to f, and that τ^1 denotes the 1–skeleton of τ.

Lemma 10.2 tells us that there is a sequence $\{f_n\}$ of homeomorphisms from F_1 to F_2 each of which is isotopic to f, and whose simplicial energies approach $I(f)$, such that the restrictions of the f_n's to τ^1 converge to a map f_0 which sends each edge of τ to a point or to a geodesic arc. In Lemma 10.12 we will show that f_0 must be an embedding. As the f_n's are all homeomorphisms, and f_0 is the limit of the sequence $f_n \mid \tau^1$, there is an extension of f_0 to a simplicial homeomorphism g from (F_1, τ, l) to (F_2, h). By

construction, g is a homeomorphism which is isotopic to f, and g minimizes E_S among all homeomorphisms isotopic to f. In particular, g is a critical point for E_S under all variations through homeomorphisms. As any small deformation of g must still be a homeomorphism, this implies that g is a critical point for E_S under all variations. Thus g is simplicial harmonic. This completes the proof of the first part of the theorem.

If (F_2, h) has negative curvature, Proposition 5.5 tells us that the homotopy class of f contains only one simplicial harmonic map, and that this map minimizes E_S in the homotopy class. It follows that g is this map, which completes the proof of the theorem subject to giving the proof of Lemma 10.12. $\qquad\square$

In order to show that f_0 must be an embedding, we need to analyze how this might fail. First we consider what happens if an edge of τ is collapsed to a point by f_0. Note that, as τ is good, this cannot happen for an edge which is a loop. Before dealing with the general case, we consider the following special case.

Lemma 10.6. *Suppose that F_1 and F_2 are closed surfaces with negative Euler number. Let $f : (F_1, \tau, l) \to (F_2, h)$ be a homeomorphism from F_1, with simplicial metric l, to F_2 with Riemannian metric h, of non-positive curvature. Let $f_0 : \tau^1 \to F_2$ be a map which is the limit of the restriction to τ^1 of a sequence of homeomorphisms from F_1 to F_2, each of which is isotopic to f, and whose simplicial energies approach $I(f)$.*

Suppose that there is an edge e of τ which is mapped to a point by f_0, such that none of the edges incident to e is mapped to a point. Then there is a geodesic λ through $f_0(e)$ such that every edge of τ incident to e has image contained in λ. Hence f_0 maps every edge in the star of e into λ.

Proof. Let v and w denote the vertices of e. Let C denote the unit circle in the tangent space to F_2 at $f_0(e)$. Each edge e' of τ which is incident to e determines a point of C which corresponds to the tangent vector to the image of e'. As f_0 is a limit of homeomorphisms, we can cut C at two points into two closed intervals A and B such that one contains the points determined by edges of τ incident to v, and the other contains the points determined by edges of τ incident to w.

Suppose that these endpoints of A and B are not diametrically opposite points of C. Then we can cut C at a pair of diametrically opposite points into two closed intervals L and R one of which contains one of A or B in its interior. Without loss of generality, we can suppose that L contains A in its interior.

Now consider a homotopy of f_0 which moves $f_0(v)$ a small distance towards the midpoint of L, and does not move any other vertices of τ, and let f_0' denote the map so obtained. If e' is a geodesic segment in F, and we move one end P of e' along a geodesic making an angle θ with e', then the derivative of the length of e' with respect to the distance moved by P is $\cos\theta$. In particular, if e' is an edge of τ other than e, which is incident

to v, it must be shortened by such a homotopy, as θ is not equal to $\pi/2$. The homotopy does increase the length of e, but the contribution of the length of e to the energy of f_0' remains zero to first order. It follows that f_0' has strictly less simplicial energy than f_0, if we move $f_0(v)$ a suitably small distance. Further f_0' is a limit of homeomorphisms f_n'. But this contradicts the energy minimizing property of f_0. Hence the endpoints of A and B must be diametrically opposite in C.

If the interior of A or the interior of B contains any points corresponding to edges incident to e, we can again obtain a contradiction by essentially the same argument. Note that, as we move the vertex v, any edge e' of τ incident to v whose tangent vector is an endpoint of A will not change in length to first order, as the angle θ is equal to $\pi/2$ for such e'. We conclude that there are only two tangent vectors determined by edges of τ incident to e and that these two points of C are diametrically opposite. If λ denotes the geodesic through $f_0(e)$ with these tangent vectors, then it follows that every edge incident to e has image contained in λ. If e' is an edge in the star of e which is not incident to e, then there are two edges of the star of e which are incident to e, and which together with e' bound a triangle of τ. Hence each vertex of e' is mapped into λ by f_0, and $f_0(e')$ is homotopic into λ fixing the vertices of e'. Now the non-positive curvature of F_2 implies that f_0 must map e' into λ, which completes the proof of Lemma 10.6. $\square$

In the above result, the boundary $\partial\sigma$ of any triangle σ in the star of e is mapped by f_0 into the geodesic λ, and $\partial\sigma$ is not mapped to a point. Intuitively, the picture is that as $n \to \infty$, the sequence f_n is collapsing the entire triangle σ into the geodesic λ. Unfortunately, by itself the condition that f_0 maps the boundary of σ into λ is not enough to ensure that there is an extension of f_0 to a map of σ into λ. For example if λ were a simple closed curve, f_0 could send $\partial\sigma$ to an essential loop in λ, and if λ were singular, f_0 need not even send $\partial\sigma$ to a loop in λ. In order to clarify the situation, we use the fact that if Λ is a geodesic in the universal cover of F_2 which lies above λ, then Λ is an embedded line. This is because of the non-positive curvature of F_2. This leads us to the following definition.

Definition 10.7. *Let λ be a geodesic in F_2. Then a triangle σ in τ is λ–degenerate under f_0 if*

(1) *f_0 maps $\partial\sigma$ into λ, and*
(2) *the restriction of f_0 to $\partial\sigma$ lifts to a map of $\partial\sigma$ into Λ.*

Note that if σ is λ–degenerate under f_0, then $f_0 \mid \partial\sigma$ extends to a map of σ into λ. The proof of the preceding lemma shows that if σ is a triangle in the star of e, then σ is λ–degenerate under f_0, but $\partial\sigma$ is not mapped to a point by f_0.

Next we consider the general case when edges adjacent to e may also be collapsed to the point $f_0(e)$. We consider the subcomplex of τ consisting

of all edges which are mapped to this point by f_0 together with each 2–simplex σ such that $f_0(\partial\sigma) = f_0(e)$. As the image of f_0 cannot be a point, this subcomplex is not equal to τ. Let K denote the component of this subcomplex which contains e. Thus K is not a single point. Lemma 10.6 is a special case of the following result, and much of the proof is similar. We proved Lemma 10.6 separately to clarify the argument. Note that f_0 is only defined on the 1–skeleton of K, but in this setting we will abuse notation and denote $f_0(e)$ by $f_0(K)$.

Lemma 10.8. *Using the above notation, there is a geodesic λ through $f_0(K)$ such that every triangle in the star of K is λ–degenerate under f_0.*

Proof. As f_0 is homotopic to the restriction of a homeomorphism, the inclusion map of K into F_1 must induce the trivial map of fundamental groups. Thus if N denotes a regular neighborhood of K, all except one of the components of the closure of $F_1 - N$ must be a disc, and the union of N with these discs must itself be a disc. Suppose there is a disc component B of the closure of $F_1 - N$, and let R denote the component of $F_1 - K$ which contains B. Then ∂R is contained in $\tau^1 \cap K$, so is mapped to $f_0(K)$ by f_0. Hence for n large, ∂R is mapped arbitrarily close to $f_0(K)$ by f_n. As B is a disc and f_n is a homeomorphism, it follows that the same holds for any edge of τ^1 which is contained in R, so f_0 must map each such edge to $f_0(K)$. It follows that R itself must be contained in K, which is a contradiction. We conclude that no component of the closure of $F_1 - N$ can be a disc, so that N itself is a disc. In particular, K must be simply connected. Now let ∂K denote the subcomplex in which K intersects the closure of $F_1 - K$. As K is simply connected, a vertex of K separates K if and only if it locally separates K. As K is finite, there must be at least two vertices of ∂K which are "extreme" and so do not separate K. As any such vertex does not locally separate K, it must have connected link in K. Let v and w denote two such vertices of ∂K. Now we will argue very much as in the proof of Lemma 10.6. See Figure 5.
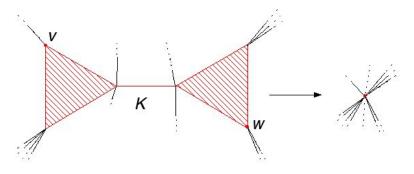

FIGURE 5. A subcomplex collapses to a point.

Let C denote the unit circle in the tangent space to F_2 at $f_0(K)$. Note that each edge of $\tau - K$ which is incident to K cannot be mapped to a point

by f_0. Thus each such edge e determines a point of C which corresponds to the tangent vector to the image of e. As f_0 is a limit of homeomorphisms, and v and w have connected link in K, we can cut C at two points into two closed intervals A and B such that one contains the points determined by edges of $\tau - K$ incident to v, and the other contains the points determined by edges of $\tau - K$ incident to w. Suppose that these endpoints of A and B are not diametrically opposite points of C. Then we can cut C at a pair of diametrically opposite points into two intervals L and R one of which contains one of A or B in its interior. Without loss of generality, we can suppose that L contains A in its interior. Now a homotopy of f_0 which moves $f_0(v)$ a suitably small distance towards the midpoint of L, and does not move any other vertices of τ, will yield a map f_0' with strictly less simplicial energy than f_0. For each edge of $\tau - K$ which is incident to v is shortened, and although edges of K which are incident to v increase in length, the contribution of the length of each such edge to the energy of f_0' remains zero to first order. Also all other edges are unchanged. Further f_0' is a limit of homeomorphisms f_n'. As before, this contradicts the energy minimizing property of f_0, so that the endpoints of A and B must be diametrically opposite in C. But if the interior of A or the interior of B contains any points corresponding to edges of $\tau - K$ incident to v or w respectively, we can obtain a contradiction by essentially the same argument. Note that any edge of $\tau - K$ incident to v whose tangent vector is an endpoint of A will not change in length to first order. We conclude that there are only two tangent vectors determined by edges of $\tau - K$ incident to v or w and that these two points of C are diametrically opposite. If λ denotes the geodesic through $f_0(K)$ with these tangent vectors, then it follows that every edge of $\tau - K$ incident to v or to w has image contained in λ. As f_0 is a limit of homeomorphisms, it follows that every edge of $\tau - K$ incident to K has image contained in λ. As in the proof of Lemma 10.6, it follows that f_0 maps every edge in the star of K into λ. It also follows that every triangle in the star of K is λ–degenerate under f_0, which completes the proof of Lemma 10.8. $\qquad\square$

Note that, in the above result, if σ is a triangle in the star of K which does not lie in K, then σ is λ–degenerate under f_0, but $\partial\sigma$ is not mapped to a point by f_0.

Now we simply suppose that we have a geodesic λ in F_2 and a triangle σ of τ such that σ is λ–degenerate under f_0, but $\partial\sigma$ is not mapped to a point by f_0. We consider the subcomplex of τ which consists of the union of all triangles which are λ–degenerate under f_0, and let L be the component of this subcomplex which contains σ. Let ∂L denote the intersection of L with the closure of $F_1 - L$.

Lemma 10.9. *Using the above notation, if v is a point of ∂L, and e is an edge of τ which is incident to v, then f_0 does not collapse e to a point.*

Proof. If f_0 collapses e to a point, we consider the subcomplex of τ consisting of all edges mapped to this point together with all 2–simplices whose boundary is mapped to this point, and let K denote the component of this subcomplex which contains e. Lemma 10.8 tells us that f_0 maps every edge in the star of K into some geodesic μ. As $\partial\sigma$ is not collapsed to a point, it follows that K does not contain some edge E of $\partial\sigma$. Pick a path in L which joins K to E. Some edge E' of this path must lie in the star of K but not in K. This edge E' is not collapsed to a point by f_0. As E' lies in L and in the star of K, the image under f_0 of E' must lie in λ and in μ. It follows that $\lambda = \mu$. As the star of e is contained in the star of K, it follows that every triangle of the star of e is λ–degenerate under f_0, which contradicts the fact that v lies in ∂L. This contradiction completes the proof of the lemma. $\square$

Of course L itself need not be a surface, but it will fail to be a surface precisely at those vertices whose link is not connected. We temporarily remove all such vertices and let L_0 denote the closure of the component of the resulting object which contains σ. Thus L_0 also need not be a surface but is obtained from a connected surface U by (possibly) identifying certain vertices in its boundary. Now we pick a boundary component C of U. Note that any edge of C also lies in ∂L, and so is not collapsed to a point by f_0. In what follows we will assume that C embeds in F_1. If this is not the case, we can push C slightly into the interior of U to obtain an embedded circle C' in F_1, and then apply the following arguments using C' in place of C.

Let $\widetilde{F_1}$ and $\widetilde{F_2}$ denote the universal covers of F_1 and F_2 respectively, where $\widetilde{F_1}$ has the triangulation $\widetilde{\tau}$ induced from τ and $\widetilde{F_2}$ has the Riemannian metric induced from h. Let $\widetilde{\tau}^1$ denote the 1–skeleton of $\widetilde{\tau}$, let $\widetilde{f_0} : \widetilde{\tau}^1 \to \widetilde{F_2}$ be a map which covers f_0, and let Λ be a geodesic in $\widetilde{F_2}$ above λ. Recall that, as F_2 is non-positively curved, Λ is an embedded line. Let $\widetilde{C}$ denote a component of the pre-image in $\widetilde{F_1}$ of C, chosen so that $\widetilde{f_0}(\widetilde{C}) \subset \Lambda$, and let $\widetilde{U}$ denote the component of the pre-image in $\widetilde{F_1}$ of U which has $\widetilde{C}$ as a boundary component. Note that $\widetilde{C}$ is a simple closed curve or a line embedded in $\widetilde{F_1}$. As each triangle in L is λ–degenerate under f_0, it follows that the restriction of $\widetilde{f_0}$ to $\widetilde{C}$ maps $\widetilde{C}$ into Λ.

Lemma 10.10. *Using the above notation, one of the following holds:*
 (1) *The map $\widetilde{f_0} : \widetilde{C} \to \Lambda$ is a homeomorphism, or*
 (2) *$\widetilde{C}$ is a circle, and the map $\widetilde{f_0} : \widetilde{C} \to \Lambda$ has precisely two critical points.*

Proof. To prove this, choose a homeomorphism f_n, for large n, lift to the universal cover and consider the induced map from $\widetilde{C}$ to $\widetilde{F_2}$. This map embeds $\widetilde{C}$ very close to Λ, and edges of $\widetilde{C}$ are almost geodesic. At each vertex v of $\widetilde{C}$, either $\widetilde{C}$ is almost straight, or it turns through an angle of approximately π. The vertices of $\widetilde{C}$ where the turn angle is approximately π correspond to the critical points of $\widetilde{f_0}$. At such a vertex, $\widetilde{U}$ must lie on

the side of $\widetilde{C}$ where the internal angle is almost zero. For any triangle of $\widetilde{\tau}$ which is incident to v and on that side of $\widetilde{C}$ must be mapped to Λ by $\widetilde{f}_0$ and hence must lie in $\widetilde{U}$.

If $\widetilde{C}$ is a simple closed curve, we know it must have total curvature close to 0 or to $\pm 2\pi$. Now each sharp bend contributes essentially $\pm\pi$, and every other vertex and every edge contributes essentially zero. Further the contributions of the sharp bends all have the same sign, as the side of $\widetilde{C}$ where the internal angle is almost zero always lies in $\widetilde{U}$. It follows immediately that $\widetilde{f}_0$ has exactly zero or two critical points. In the first case, $\widetilde{f}_0$ would map $\widetilde{C}$ to Λ by an immersion which is impossible. Thus part 2) of the lemma must hold.

If $\widetilde{C}$ is a line, it must be properly embedded, so for any subinterval I of Λ, there is a subinterval J of $\widetilde{C}$ which "starts at one end of I and ends at the other". For such an interval J, the total curvature must be close to zero. Again the contributions of the sharp bends all have the same sign, so it follows immediately that J has no critical points. Hence $\widetilde{f}_0$ has no critical points, so that $\widetilde{f}_0$ maps $\widetilde{C}$ to Λ by a homeomorphism, showing that part 1) of the lemma holds. This completes the proof of the lemma. $\qquad\square$

We can now describe U and L_0 as follows. See Figure 6.

Lemma 10.11. *Using the above notation, one of the following holds:*

(1) L_0 *is an annulus,* λ *is a simple closed curve, and* f_0 *maps each component of* ∂L_0 *to* λ *by a homeomorphism.*

(2) L_0 *is a Moebius band,* λ *is a simple closed curve, and* f_0 *maps* ∂L_0 *to* λ *by a double covering.*

(3) U *is a disc, and* f_0 *maps* $\partial U = C$ *to* λ *with exactly two critical points. Further, either* L_0 *is equal to* U, *or* L_0 *is obtained from* U *by identifying the two critical points of* C. *In the case when* L_0 *is not equal to* U, *the geodesic* λ *must be either a simple closed curve or cross itself at the image of the critical points.*
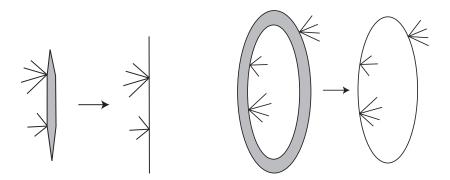

FIGURE 6. Two cases in which L_0 collapses to a geodesic

Proof. We apply Lemma 10.10 to each boundary component of U.

Suppose that some boundary component C of U is such that the map $\widetilde{f_0} : \widetilde{C} \to \Lambda$ is a homeomorphism. As C is a simple closed curve, it follows that λ is a simple closed curve in F_2, and that $f_0 \mid C : C \to \lambda$ is a covering map. As C is embedded in F_1, it follows that $f_0 \mid C : C \to \lambda$ must be a homeomorphism or a double covering. Further in the double covering case, C must bound a Moebius band in F_1. It follows that the inclusion of U into F_1 maps $\pi_1(U)$ to an infinite cyclic subgroup of $\pi_1(\lambda)$, and hence that U must be homeomorphic to an annulus with some discs removed, or to a Moebius band with some discs removed. As in the proof of Lemma 10.8, it follows that no discs are removed so that U is an annulus or Moebius band. Now it follows that L_0 must equal U as it is not possible to have any points of ∂U identified in F_1. Thus we have cases 1) or 2) of the lemma.

Now suppose that every boundary component of U maps to λ with exactly two critical points. We will show that we have case 3). Each boundary component of U is mapped to a null homotopic loop in F_2, and so must be null homotopic in F_1. Hence each boundary component of U bounds a disc in F_1. As in the proof of Lemma 10.8, it follows that U is a disc. Further, it follows that L_0 is equal to U, or L_0 is obtained from U by identifying the two critical points of C, as required. When L_0 is not equal to U, the image of L_0 in λ is a simple closed curve, so that either λ is a simple closed curve or λ crosses itself at the image of the critical points. This completes the proof that we have case 2). □

Now we are ready to prove that f_0 must be an embedding.

Lemma 10.12. *Suppose that F_1 and F_2 are closed surfaces with negative Euler number. Let $f : (F_1, \tau, l) \to (F_2, h)$ be a homeomorphism from F_1, with simplicial metric l, to F_2 with a Riemannian metric of non-positive curvature. Further assume that τ is a good triangulation. Let $f_0 : \tau^1 \to F_2$ be a map which is the limit of the restriction to τ^1 of a sequence of homeomorphisms from F_1 to F_2 each of which is isotopic to f, and whose simplicial energies approach $I(f)$. Then f_0 is an embedding.*

Proof. Recall from Lemma 10.2 that f_0 maps each edge of τ to a point or to a geodesic arc. Suppose that f_0 is not an embedding. As f_0 is a limit of homeomorphisms, it must fail to embed the boundary of some triangle of τ. Thus one of the following cases occurs.

(1) Some vertex is mapped to the image of the interior of an immersed edge of τ,
(2) some edge of τ is collapsed to a point by f_0, or
(3) some triangle of τ is mapped to a geodesic segment, not a point.

In case 1), there is a vertex v and an immersed edge e of τ with an interior point w such that $f_0(w)$ equals $f_0(v)$. The proof of part 1) of Theorem 10.3 provides a contradiction if no edge of τ incident to v is mapped into $f_0(e)$

or to a point. See Figure 4. It follows that if case 1) occurs, we must also be in case 2) or case 3).

In case 2), if an edge e of τ is collapsed to a point by f_0, we define the subcomplex K as we did just before Lemma 10.8. That lemma implies that there is a geodesic λ such that if σ is a triangle in the star of K which does not lie in K, then σ is λ–degenerate under f_0, but $\partial\sigma$ is not mapped to a point by f_0. Thus in all cases, there is a geodesic λ in F_2 and a triangle σ of τ such that $\partial\sigma$ is not collapsed to a point by f_0, and σ is λ–degenerate under f_0.

As in the preceding three lemmas, we define the subcomplexes L and L_0 of F_1, and the surface U. Lemma 10.11 tells us that U is an annulus, Moebius band or disc. In each case, we claim there is at least one vertex v of ∂U which is not a critical point of the map $f_0 \mid \partial U : \partial U \to \lambda$. This claim is trivial in the first two cases. In the third case when U is a disc, our hypothesis that τ is a good triangulation implies that ∂U must contain at least three vertices, so that at least one of these vertices is not a critical point, as claimed.

Let v be any vertex of a component C of ∂U which is not a critical point, and let $star(v)$ denote the star of v in F_1. As C is a boundary component of U, the two edges of C incident to v together divide $star(v)$ into two discs, one of which is contained in U, and hence is contained in L_0. We denote this disc by D, and let D' denote the other disc. Thus $star(v)$ equals $D \cup D'$, and $D \cap D'$ equals two edges of C. We claim that no edge of $\tau - D$ which is incident to v can lie in L. Let e be an edge of $\tau - D$ which is incident to v. Thus e splits D' into two subdiscs D_1 and D_2. Now suppose that e lies in L, so that $f_0(e) \subset \lambda$. Recall from Lemma 10.9, that f_0 does not collapse e to a point. As f_0 is a limit of homeomorphisms, it follows that f_0 maps all the edges in one of D_1 or D_2 into λ. If $f_0(D_i) \subset \lambda$, then D_i must be contained in L. Hence $D \cup D_i$ is contained in L. But this implies that the interior of an edge of C lies in the interior of L which contradicts the definition of C. It follows that no edge of $\tau - D$ which is incident to v can lie in L, as claimed. Recall again that Lemma 10.9 shows that every edge incident to v is not mapped to a point by f_0. As f_0 is a limit of homeomorphisms, all the edges of $\tau - D$ which are incident to v must be mapped by f_0 to the same side of λ. Note that this statement makes crucial use of the fact that v is not a critical point of the map from ∂U to λ.

Now consider the sequence of homeomorphisms f_n whose restriction to τ^1 converges to f_0. As $n \to \infty$, L_0 collapses down to the geodesic λ. Thus every edge of L_0 either collapses to a point or to a geodesic segment along λ. Consider a small homotopy of f_0 that only moves vertices of L_0, moves these vertices in a direction orthogonal to λ, and does not move the critical points of ∂U, if any. The contribution to the energy from the edges of L_0 is constant to first order for sufficiently small such deformations. If such a deformation moves points of ∂L_0 away from L_0, then it strictly shortens each edge of τ that does not lie in L_0 and is incident to L_0 at a noncritical

point of ∂L_0. The deformation will therefore shorten these edges at a rate bounded below by a linear function of the distance moved. If ∂U has no critical points, we conclude that for large enough values of n, isotoping f_n by moving a component of ∂L_0 away from L_0 yields a homeomorphism f_n' whose simplicial energy is strictly smaller than that of f_0. This contradicts the energy minimizing property of f_0. In the remaining cases when U is a disk and ∂U has two critical points, they divide ∂U into two intervals, and we apply the same argument to an isotopy of f_n which moves one of these intervals while fixing the endpoints. This contradiction completes the proof of Lemma 10.12. $\square$

We will end this section by discussing an application of the preceding arguments. But first we note that the arguments of this section work perfectly well if F_1 and F_2 have non-empty boundary, and in this case, one need not insist that the metrics be strictly negatively curved to obtain uniqueness. The result we obtain is the following.

Theorem 10.13. *Suppose that F_1 and F_2 are compact surfaces with non-empty boundary. Let $f : (F_1, \tau, l) \to (F_2, h)$ be a homeomorphism from F_1, with simplicial metric l, to F_2 with a non-positively curved Riemannian metric h such that ∂F_2 is locally convex. Further assume that τ is a good triangulation. Then f is isotopic rel ∂F_1 to a simplicial harmonic homeomorphism $g : (F_1, \tau, l) \to (F_2, h)$ that minimizes E_S among all homeomorphisms isotopic to f rel ∂F_1. Further, g is the unique simplicial harmonic map in the homotopy class of f, and so minimizes energy in this homotopy class.*

Now let F be a compact surface with a Riemannian metric h, and let τ denote a good triangulation of F. We will say that a triangulation of F is *straight* if each edge is a geodesic arc. We will be interested in the number and homotopy type of the components of the space of straight triangulations of F of a fixed combinatorial type. Note that Theorems 10.5 and 10.13 tell us that for any good triangulation τ of F, the space of straight triangulations of F modelled on τ is non-empty. We will show that, with certain curvature restrictions, there is a natural bijection between the components of this space and the group of isotopy classes of homeomorphisms of F. When F is the 2–disk with a flat metric, Bloch, Connelly and Henderson [6] proved this, and they proved, in addition, that the components of this space are homeomorphic to some Euclidean space. In related work, Awartani and Henderson [2] considered the problem of showing that when F is the 2–sphere, the analogous space has the homotopy type of $O(3)$. And Bloch [5] showed that the space of embeddings with geodesic edges and convex image of a triangulated 2–disk into the plane has the homotopy type of $O(2)$.

If we have a simplicial metric l on τ, and if g is any simplicial homeomorphism from (F, τ, l) to (F, h), then $g(\tau)$ is a straight triangulation of F. Conversely a straight triangulation of (F, h) by τ determines a unique simplicial homeomorphism from (F, τ, l) to (F, h). Thus, after fixing τ and

its simplicial metric, we can identify the space of straight triangulations of F by τ with the space of simplicial homeomorphisms of F. Note that any triangulation τ of F admits a simplicial metric l by specifying that l takes the value 1 on every edge of τ.

Having fixed a good triangulation τ of F and its simplicial metric l, let $H(\tau)$ denote the space of simplicial homeomorphisms from (F, τ, l) to (F, h). This is a subset of the space $Homeo(F)$ of all homeomorphisms from F to F. The components of $Homeo(F)$ naturally correspond to the isotopy classes of homeomorphisms of F.

Theorem 10.14. *Let F be a compact surface with a Riemannian metric h of non-positive curvature such that ∂F is locally convex. If F is closed we suppose that h has negative curvature. Let τ denote a good triangulation of F, with simplicial metric l. Using the above notation, each component of $Homeo(F)$ contains exactly one component of $H(\tau)$.*

Proof. Theorems 10.5 and 10.13 imply that any homeomorphism $f : (F, \tau, l) \to (F, h)$ is isotopic rel ∂F to a unique simplicial harmonic homeomorphism $g : (F, \tau, l) \to (F, h)$ that minimizes E_S among all homeomorphisms isotopic to f. In particular, each component of $Homeo(F)$ contains a point of $H(\tau)$. Note that the fact that $H(\tau)$ is non-empty is already nontrivial to prove.

Next we concentrate on homeomorphisms isotopic to f. Let H denote a component of the space of simplicial homeomorphisms from (F, τ, l) to (F, h) which are isotopic to f. We will show that H contains the energy minimizing simplicial harmonic homeomorphism g. As this holds for all such components, it will follow that the space of simplicial homeomorphisms from (F, τ, l) to (F, h) which are isotopic to f must be connected, which will prove the first part of the theorem.

Now H is contained in the space, $Map(f)$, of all simplicial maps from (F, τ, l) to (F, h) which are homotopic to f. The gradient flow of the energy functional on $Map(f)$ yields a deformation retraction of $Map(f)$ to the unique energy minimizer g, which we know to be a homeomorphism. It would be extremely convenient if this flow induced a deformation retraction of H to this same energy minimizer. One would need to show that if one starts with a simplicial homeomorphism from (F, τ, l) to (F, h), then the result of following the gradient flow remains a homeomorphism for all time. Unfortunately this is not the case. See the end of this section for a discussion and example. This means that we are unable to show that the components of $H(\tau)$ are contractible, although we believe this to be the case. Instead we consider the closure $\overline{H}$ of H in $Map(f)$. As H is open in $Map(f)$, it follows that $\overline{H} - H$ is closed in $Map(f)$. This implies that the restriction of the energy functional to $\overline{H} - H$ attains a minimum. To see this, recall that the proof of Proposition 5.5 showed that given any constant K, the subspace of $Map(f)$ consisting of maps with energy $\leq K$ is compact. Thus the same holds for the intersection of this subspace with $\overline{H} - H$. Now let g_0 be a

point of $\overline{H} - H$ of least energy. Thus g_0 is not a homeomorphism, and is a limit of simplicial homeomorphisms f_n in H. We make the following claim.

Claim 10.15. *There is n such that we can isotope f_n to a simplicial homeomorphism f'_n whose energy is strictly less than that of g_0.*

Of course, such f'_n also lies in H. Assuming this claim, consider the gradient flow starting at f'_n. Any map in this flow has energy less than that of f'_n, and so less than that of g_0. Hence no map in this flow can lie in $\overline{H} - H$. Hence the flow stays in H and so determines an isotopy of f'_n to the unique energy minimizer g. In particular it follows that H contains g, which completes the proof of the theorem.

It remains to prove the above claim. We would like to prove this in much the same way as we proved Lemma 10.12. Specifically we want to use Lemmas 10.8, 10.9, 10.10, 10.11 and 10.12. As stated, these lemmas do not apply to our situation, as g_0 is not an energy minimizer for all homeomorphisms. But we can apply the arguments in these lemmas as we now describe.

Recall that all these lemmas consider a map $f_0 : \tau^1 \to F_2$ which is the limit of the restriction to τ^1 of a sequence of homeomorphisms from F_1 to F_2, each of which is isotopic to f, and whose simplicial energies approach $I(f)$. Lemma 10.2 shows that any such map sends each edge of τ to a point or a geodesic segment. In the present situation, g_0 is a limit of simplicial homeomorphisms, so it is immediate that it sends each edge of τ to a point or a geodesic segment, but it need not have the minimal energy $I(f)$.

Lemma 10.8 used the least energy assumption on f_0 to show that if f_0 collapsed some edge e to a point, then it must map the star of some complex K which contains e into some geodesic λ. Now suppose that g_0 collapses some edge e to a point. The arguments in the proof of Lemma 10.8 show that either we get the same conclusion for g_0, or we can isotope some f_n to a simplicial homeomorphism f'_n whose energy is strictly less than that of g_0. An alternative way of putting this is to say that either the claim holds or the result of Lemma 10.8 holds for g_0.

Lemmas 10.9, 10.10 and 10.11 depend on Lemma 10.8 but do not otherwise use the energy minimizing assumption on f_0. Thus again we have that either the claim holds or their results also hold for g_0.

Finally the arguments in the proof of Lemma 10.12 use the conclusions of Lemmas 10.9, 10.10 and 10.11 to show that if f_0 is not a homeomorphism we can isotope f_n to a simplicial homeomorphism f'_n whose energy is strictly less than that of f_0. As g_0 is not a homeomorphism, when we apply these arguments to g_0, it follows that the claim must hold. This completes the proof of the claim, and hence of Theorem 10.14. $\qquad\square$

We end this section by discussing the fact that the gradient flow need not preserve the property of being a homeomorphism. Recall from section 3 the formula for the simplicial energy of a map. If we consider varying a single vertex P, only those edges incident to P can contribute to any change

in the energy. To each such edge e we associate a vector v_e in the tangent
space at P so that v_e is tangent to e, has magnitude equal to the length of
e, and points away from P. Suppose that every edge of τ has length 1 in our
simplicial metric. Then a simple calculation shows that under the gradient
flow, the flow vector of P is simply the sum of the vectors v_e. Thus P is
"pulled" much more by long edges than by short ones. If a configuration
like the one in Figure 7 is part of a triangulation of a surface, it will quickly
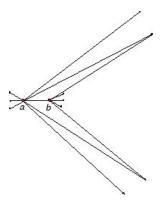cease to be embedded under the gradient flow.

FIGURE 7. A configuration which ceases to be embedded un-
der the gradient flow. The vertex a moves rightward faster
than vertex b.

11. HIGHER DIMENSIONS

In the previous sections we considered maps of a triangulated closed sur-
face F into a non-positively curved closed manifold M whose dimension
was unrestricted. In this section we observe that all the preceding work
can be generalized to maps of a triangulated closed k–manifold F into a
non-positively curved manifold M, for any $k \geq 2$. We do this by restricting
attention to the 2–skeleton $\tau^{(2)}$ of the triangulation τ of F. This seems
reasonable as the homotopy class of a map of F into M is determined by
its restriction to $\tau^{(2)}$, as M is aspherical. In the smooth setting, much work
has been done in this setting. See [18][19].

For simplicity we discuss only the case when $k = 3$. As in the 2–
dimensional case, we specify a triangulation τ of the closed 3–manifold
F, and a map l that assigns to each edge e_i, $1 \leq i \leq r$, of τ a length
$l_i = l(e_i) > 0$, with the lengths l_i realizable by a Euclidean tetrahedron for
each tetrahedron of τ. We call such an assignment a *simplicial metric* on
F, and denote it by (F, τ, l), or just l when the context is clear. Note that
we allow the possibility that for some triangles one of the triangle inequali-
ties is an equality. As before we order the vertices of τ using a Δ-complex
structure.

Define the *simplicial 2-volume* of (F, τ) to be the sum $V_2 = \Sigma l_i l_j$, where the sum is taken over all corners of all the 2–simplices of τ.

Given a map $f : F \to M$, let L_i denote the length of the restriction of f to the edge e_i, and as before define the i–th stretch factor to be $\sigma_i = L_i/l_i$. The *simplicial 2-volume of f* is defined to be $V_2(f) = \Sigma L_i L_j$, and the *simplicial 2-energy of the map f* to be $E_2(f) = \frac{1}{2}\Sigma(\sigma_i^2 + \sigma_j^2)l_i l_j$, where the sum is taken over all corners of all the 2–simplices of τ.

Next we need to define when a map from F to M is simplicial. As before we insist that each edge of the triangulation τ is mapped to M as a geodesic arc. In order to determine f canonically on the remainder of F, we use the ordering of the vertices of τ. For each triangle of τ this determines a cone structure, with cone point the minimal vertex of the triangle, and as before this determines a canonical extension of f to the 2–skeleton of F. We give each tetrahedron T of τ the cone structure from its minimal vertex v. Note that all the faces of T which meet v also have cone structure with v as the cone point. Thus the canonical extension of f to the tetrahedra of τ yields a well defined simplicial map to M.

As in Lemma 3.3, for any simplicial metric on F, we have the inequality $E_2(f) \geq V_2(f)$ with equality if and only if all the stretch factors are equal. We say that f is *2-simplicial harmonic* if it is a critical point of the 2-energy functional. As in Proposition 5.5, if M is closed, the energy functional has a minimum among all maps homotopic to f, and we can choose any minimizing map to be simplicial. Further if M is negatively curved and f is nontrivial, i.e. f cannot be homotoped to have image contained in a closed geodesic, then this minimum energy simplicial map g is unique.

As in Section 7, the fact that any map $f : F \to M$ can be homotoped to a unique simplicial map without moving the vertices allows one to construct families of harmonic maps with uniform area bounds.

In the preceding paragraphs, we prefixed 2 to our definitions, because there seem to be some natural alternative definitions. For example, we could define the *simplicial 3-volume* of (F, τ) to be the sum $V_3 = \Sigma l_i l_j l_k$, where the sum is taken over all corners of all the 3–simplices of τ. Then we would define the *simplicial 3-volume of f* to be $V_3(f) = \Sigma L_i L_j L_k$, and the *simplicial 3-energy of f* to be $E_3(f) = \frac{1}{2}\Sigma(\sigma_i^2 + \sigma_j^2 + \sigma_k^2)l_i l_j l_k$. In general, if the dimension of F equals d, then one could similarly define the simplicial k–volume of F, and the simplicial k–energy of f, for any k such that $2 \leq k \leq d$. However one cannot expect there to be any connection between simplicial k–volume and simplicial k–energy, except when $k = 2$.

REFERENCES

[1] I. Agol, *Tameness of hyperbolic 3-manifolds*, preprint, arxiv.org/abs/math/0405568.
[2] M. Awartani, D. W. Henderson, *Spaces of geodesic triangulations of the sphere*, Trans. Amer. Math. Soc. 304 (1987), no. 2, 721–732.
[3] R. C. Alperin, *Heron's Area Formula*, The College Mathematics Journal, 1987.

[4] D. Bachman, D. Cooper, M. White, *Large embedded balls and Heegaard genus in negative curvature*, Algebr. Geom. Topol. 4 (2004), 31–47 (electronic).

[5] E.D. Bloch, *Strictly convex simplexwise linear embeddings of a 2-disk*, Trans. Amer. Math. Soc. 288 (1985), no. 2, 723–737.

[6] E.D. Bloch, R. Connelly, D. W. Henderson, *The space of simplexwise linear homeomorphisms of a convex 2-disk*, Topology 23 (1984), no. 2, 161–175.

[7] F. Bonahon, *Bouts des variétés hyperboliques de dimension 3*, Annals of Math. 124 (1986), 71–158.

[8] D. Calegari and D. Gabai, *Shrinkwrapping and the taming of hyperbolic 3-manifolds*, J. Amer. Math. Soc. 19 (2006), no. 2, 385–446.

[9] R. Canary, *A covering theorem for hyperbolic 3-manifolds and its applications*, Topology 35 (1996), 751–778.

[10] D. Cooper and J. F. Manning, *Non-faithful representations of surface groups into SL(2,C) which kill no simple closed curve*, preprint, arxiv.org/abs/1104.4492.

[11] J.-M. Coron and F. Hélein, *Harmonic diffeomorphisms, minimizing harmonic maps and rotational symmetry*, Compositio Math. 69 (1989), no. 2, 175–228.

[12] R. Courant, *Dirichlet's Principle, Conformal Mapping, and Minimal Surfaces*. Appendix by M. Schiffer. Interscience Publishers, Inc., New York, N.Y., 1950.

[13] M. Culler and K. Vogtmann, *Moduli of graphs and automorphisms of free groups*, Invent. Math. 84 (1986), no. 1, 91–119.

[14] G. Daskalopoulos and C. Mese, *Harmonic maps from a simplicial complex and geometric rigidity*, J. Differential Geom. 78 (2008), 269–293.

[15] M. Desbrun, M. Meyer, P. Schroder, A. Barr, *Implicit fairing of irregular meshes using diffusion and curvature flow*, Proceedings of the 26th annual conference on Computer graphics and interactive techniques, (1999), 317–324.

[16] R. J. Duffin, *Distributed and lumped networks*, J. Math. Mech. 8 (1959), 793–826.

[17] J. Eells and L. Lemaire, *Deformations of metrics and associated harmonic maps*, Proc. Indian Acad. Sci. 90 (1) (1981), 33–45.

[18] J. Eells, Jr. and L. Lemaire, *A report on harmonic maps*, Bull. London Math. Soc. 10 (1978), 1-68. *Two reports on harmonic maps*, 1–68, World Sci. Publ., River Edge, NJ, 1995.

[19] J. Eells, Jr. and L. Lemaire, *Another report on harmonic maps*, Bull. London Math. Soc. 20 (1988), 385–524. *Two reports on harmonic maps*, 69–208, World Sci. Publ., River Edge, NJ, 1995.

[20] J. Eells, Jr. and J. H. Sampson, *Harmonic mappings of Riemannian manifolds*, Amer. J. Math. 86 (1964), 109-160.

[21] H. Federer, *Geometric measure theory,* Springer-Verlag, Berlin, Heidelberg, New York, 1969.

[22] P. Hartman, *On homotopic harmonic maps*, Canad. J. Math. 19 (1967), 673- 687.

[23] J. Hass, *Minimal Surfaces in Manifolds with S^1 Actions and the Simple Loop Conjecture for Seifert Fibered Spaces*, Proc. A.M.S., 99, (1987), 383-388.

[24] J. Hass, A. Thompson and W. P. Thurston, *Stabilization of Heegaard splittings*, Geometry & Topology 13 (2009), 2029–2050.

[25] A. Hatcher, *Algebraic Topology*, Cambridge University Press, Cambridge, 2002.

[26] H. Izeki and S. Nayatani, *Combinatorial Harmonic Maps and Discrete-group Actions on Hadamard Spaces*, Geometriae Dedicata 114 (2005), 147-188.

[27] J. Jost and R. Schoen, *On the existence of harmonic diffeomorphisms between surfaces*, Inventiones Mathematicae 66 (1982), 353–359.

[28] Y. Minsky, *Harmonic Maps into Hyperbolic 3-Manifolds*, Transactions of the American Mathematical Society 332 (1992), 607-632.

[29] J. Pitts and J. H. Rubinstein, *Applications of minimax to minimal surfaces and the topology of 3-manifolds*, Proc. Centre Math. Applic., Australian National University 12 (1987), 137-170.

[30] K. Polthier, *Computational Aspects of Discrete Minimal Surfaces*, Proceedings of the Clay Summer School on Global Theory of Minimal Surfaces (D. Hoffman, Ed.) (2002).

[31] U. Pinkall and K. Polthier, *Computing Discrete Minimal Surfaces and Their Conjugates*, Experimental Mathematics 2 (1993), 15–36.

[32] K. Polthier and W. Rossman, *Counterexample to the Maximum Principle of Discrete Minimal Surfaces*, http://www.eg-models.de/models/Surfaces/MinimalSurfaces/2000.11.040/PaMaxPrinciple.html

[33] K. Polthier and W. Rossman, *Discrete constant mean curvature surfaces and their index*, J. Reine Angew. Math. 549 (2002), 47–77.

[34] T. Rado, *On the problem of Plateau*, Bull. Amer. Math. Soc. 40 (1934), 194–196.

[35] J. H. Rubinstein, *Minimal surfaces in geometric 3-manifolds*, *Global theory of minimal surfaces*, 725-746, Clay Math. Proc., 2, Amer. Math. Soc., Providence, RI, 2005.

[36] J. H. Rubinstein and S. Wang, π_1–*injective surfaces in graph manifolds*, Comment. Math. Helv. 73 (1998), no. 4, 499–515.

[37] J.H. Sampson, *Some properties and applications of harmonic mappings*, Annales scientifiques de l'école Normale Supérieure Sec 4, 11 no. 2 (1978), p. 211-228.

[38] M. Scharlemann and A. Thompson, *Thin position for 3-manifolds*, Geometric topology (Haifa, 1992), Contemp. Math. 164, Amer. Math. Soc., Providence, RI (1994).

[39] J. Sacks and K. Uhlenbeck, *Minimal Immersions of Closed Riemann Surfaces*, Transactions of the American Mathematical Society, Vol. 271, No. 2 (Jun. 1982), pp. 639-652.

[40] R. Schoen and S.-T. Yau, *Existence of incompressible minimal surfaces and the topology of three-dimensional manifolds with nonnegative scalar curvature*, Ann. of Math. (2) 110 (1979), 127-142.

[41] W. P. Thurston, *The geometry and topology of 3-manifolds*, Princeton University Notes, available from Math. Sci. Res. Inst., Berkeley, library.msri.org/books/gt3m/.

[42] M. T. Wang, *A fixed point theorem for discrete group actions on Riemannian manifolds*, J. Diff. Geom 50 (1998), 249-267.

[43] M. Wolf, *Measured Foliations and Harmonic Maps of Surfaces*, J. Differential Geom., 49 (1998), 437-467.

[44] S. A. Wolpert, *Geodesic Length Functions and the Nielsen Problem*, J. Differential Geometry 25 (1987), 275-296.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, DAVIS, CALIFORNIA 95616

E-mail address: hass@math.ucdavis.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MICHIGAN, ANN ARBOR, MICHIGAN 48109

E-mail address: pscott@umich.edu