

Article

From Normal Surfaces to Normal Curves to Geodesics on Surfaces

Eli Appleboim

Faculty of Electrical Engineering, Technion, Haifa 3200, Israel; eliap@ee.technion.ac.il; Tel.: +972-4-8294723

Received: 23 March 2017; Accepted: 25 August 2017; Published: 20 September 2017

Abstract: This paper gives a study of a two dimensional version of the theory of normal surfaces; namely, a study of normal curves and their relations with respect to geodesic curves. This study results with a nice discrete approximation of geodesics embedded in a triangulated orientable Riemannian surface. Experimental results of the two dimensional case are given as well.

Keywords: 3-manifold; least area surface; normal surface; geodesic; normal curve

1. Introduction

Motivated by the topological theory of normal surface we give in this paper a complete study of the relations between geodesic curves and normal curves embedded in a triangulated Riemannian surface.

Normal surface theory is a topological piecewise linear ($p\ell$ for short) counterpart of the differential geometric theory of minimal surfaces. This theory studies the ways surfaces intersect with a given triangulation of a 3-manifold. A surface is normal if it intersects the tetrahedra of a triangulation in a fairly simple manner. It was proved by Haken in [1] that every incompressible (i.e., topologically non trivial) surface embedded in a triangulated 3-manifold can be continuously deformed to a normal surface with respect to any given triangulation of the manifold.

Having the theory of normal surfaces, it is then desirable to interpret minimal surfaces in terms of normal surfaces. In this direction Jaco and Rubinstein presented in [2] a $p\ell$ version, that is based on normal surfaces, of minimal and least area surfaces in a triangulated 3-manifold. It is shown in [2] that $p\ell$ minimal and $p\ell$ least area surfaces share many properties of classical minimal and least area surfaces. These results are the first to make precise the analogy between minimal and normal surfaces. However, the extent at which this analogy holds is far from being fully understood. For instance, the following two questions are very natural in this context, yet both of them are still open.

(I) Is it possible to subdivide a given triangulation of a 3-manifold arbitrarily fine, to obtain sequences of $p\ell$ -minimal surfaces that converge in some suitable way to classical minimal surfaces?

An affirmative answer to the above question gives an alternative topological-combinatorial proof of many classical existence results of least area incompressible surfaces in 3-manifolds, such as those obtained for example in [3], using partial differential equations. Moreover, the use of existing algorithms for finding normal surfaces makes this topological-combinatorial result of computable algorithmic nature.

The second question is the following:

(II) Can every minimal surface in the smooth sense, be presented as a limit surface of a sequence of $p\ell$ -minimal surfaces, appropriately constructed?

Answering these questions, at least partially will be the focus of two followup papers.

In this paper we will address these questions in one dimension lower. That is in the context of curves on surfaces. In particular the following theorems will be proved:

Theorem 1. *Let (Σ, \mathcal{T}, g) be a triangulated Riemannian surface. Then there exists a flow defined on the set of normal homotopy classes of curves that deforms each class to a geodesic curve in that class.*

Theorem 2. Let $(\Sigma, \mathcal{T}, \mathfrak{g})$ be a hyperbolic triangulated surface isometrically embedded in some \mathbb{R}^n . Let γ be a curve on Σ that is either a shortest geodesic between two given points p and q on Σ , in a homotopy class $\text{rel}\{p, q\}$, or a shortest closed geodesic in a homotopy class in Σ . Then there exists a sequence of $p\ell$ triangulated 2-submanifolds of \mathbb{R}^n , Σ_n , each of them is a $p\ell$ -approximation of Σ , and a sequence of $p\ell$ -curves γ_n , each of which is normal with respect to Σ_n so that γ_n converges to γ .

Theorem 1 gives rise to a curve shortening flow. Such flows are of major importance both theoretically and applicatively speaking, (see [4–8]). For instance, such flows are extensively used to identify malignant tissues in medical images and many other image segmentation problems.

In Section 2 we will give an essential preliminaries. This will include basic definitions and notations. In particular, normal curves will be defined in this section. Section 3 is devoted to the proof of Theorem 1. In Section 4 we will describe some experimental implementations of the curve shortening flow described in Section 3. These results will include some synthetic examples as well as examples in which this flow was tested on real data. In Section 5 Theorem 8 will be proved. In Section 6 we regard normal curves from the view point of the dual graph. Last, in Appendix A we will give a brief introduction to the theory of normal surfaces.

2. Preliminaries

2.1. Simplicial Complexes

Definition 1. A collection K of simplices is called a simplicial complex if

1. If $\tau < \sigma$, then $\tau \in K$.
2. Let $\sigma_1, \sigma_2 \in K$ and let $\tau = \sigma_1 \cap \sigma_2$. Then $\tau < \sigma_1, \tau < \sigma_2$.
3. K is locally finite.

$|K| = \bigcup_{\sigma \in K} \sigma$ is called the underlying polyhedron (or polytope) of K .

Remark 1. A triangulated manifold is regarded as a cell complex, and the complex structure is given by the triangulation.

Definition 2. A complex K' is called a subdivision of K iff

1. $K' \subset K$;
2. if $\tau \in K'$, then there exists $\sigma \in K$ such that $\tau \subseteq \sigma$.

If K' is a subdivision of K we denote it by $K' \triangleleft K$.

Let K be a simplicial complex and let $L \subset K$. If L is a simplicial complex, then it is called a subcomplex of K .

Definition 3. Let $a \in |K|$. Then

$$\text{St}(a, K) = \bigcup_{\substack{a \in \sigma \\ \sigma \in K}} \sigma$$

is called the star of $a \in K$.

If $S \subset K$, then we define: $\text{St}(S, K) = \bigcup_{a \in S} \text{St}(a, K)$.

Definition 4. Let $f : K \rightarrow \mathbb{R}^n$ be a C^r embedding of an m -complex K . Let $\delta : K \rightarrow \mathbb{R}_+^*$ be a positive continuous function. Then a map $g : |K| \rightarrow \mathbb{R}^n$ is called a δ -approximation to f if:

- There exists a subdivision K' of K such that $g \in C^r(K', \mathbb{R}^n)$
- $d_{\text{eucl}}(f(x), g(x)) < \delta(x)$, for any $x \in |K|$
- $d_{\text{eucl}}(df_a(x), dg_a(x)) \leq \delta(a) \cdot d_{\text{eucl}}(x, a)$, for any $a \in |K|$ and for all $x \in \overline{\text{St}}(a, K')$.

Definition 5. Let $(\mathcal{M}, \mathcal{T})$ be a C^r triangulated m -manifold embedded in \mathbb{R}^n . Let $s \in \mathcal{T}$ be a simplex. The vertex map $V_s : s \rightarrow \mathbb{R}^n$, is determined by $V_s(v) = v$, where v is a vertex of s . The linear extension of V_s is called the secant map induced by \mathcal{T} , and it is denoted by L_s .

2.2. Fat Triangulations

Definition 6.

- A triangle in \mathbb{R}^2 is called φ -fat iff all its angles are larger than a prescribed value $\varphi > 0$.
- A k -simplex $\tau \subset \mathbb{R}^n$, $2 \leq k \leq n$, is φ -fat if all its facets are φ -fat for some $\varphi > 0$.
- A triangulation $\mathcal{T} = \{\sigma_i\}_{i \in \mathbf{I}}$ is fat if all its simplices are φ -fat for some $\varphi > 0$.

The importance of fat triangulations is stressed by the following example.

Schwartz Lantern

The Schwartz lantern is given by the following:

Example 1. Let \mathcal{G} be the standard cylinder $\mathbb{S}^1 \times I$ of height 1 and radius 1 in \mathbb{R}^3 . Let \mathcal{G}_n be the “lantern” that is given as the pl triangulated surface so that the vertices of its triangulation are placed on \mathcal{G} . Suppose the set of vertices of \mathcal{G}_n consists of n^4 points so that n points are evenly placed along the circle $\mathbb{S}^1 \times \{1\}$ and n^3 points are placed along the vertical direction as indicated in the following Figure 1.

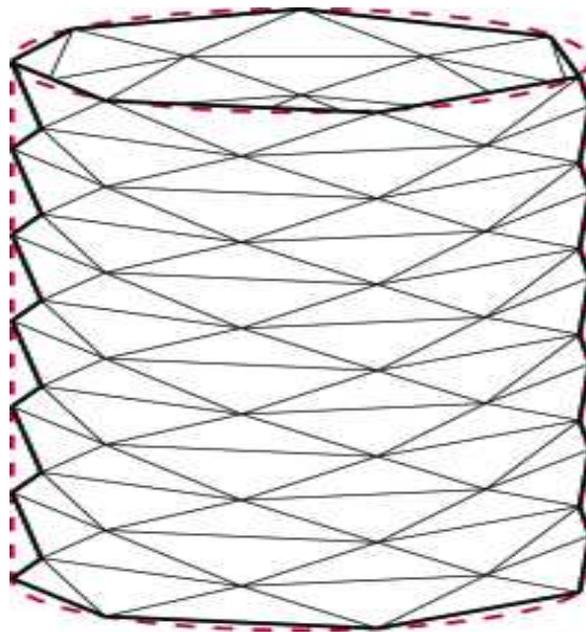


Figure 1. Schwartz Lantern.

Simple computation shows that each of the pl triangles of \mathcal{G}_n has base length $\geq \frac{1}{2n}$ and height $\geq \frac{1}{4n} \arctan(\frac{\pi}{2n})$. Following this, the area of the lantern thus satisfies:

$$\mathcal{A}(\mathcal{G}_n) \geq n^4 \frac{1}{2n} \frac{1}{4n} \arctan\left(\frac{\pi}{2n}\right). \tag{1}$$

As $n \rightarrow \infty$, $\arctan(\frac{\pi}{2n}) \rightarrow \frac{\pi}{2n} > \frac{1}{2n}$, hence the following holds:

$$\mathcal{A}(\mathcal{G}_n) \geq \frac{n}{8}, \tag{2}$$

so $\mathcal{A}(\mathcal{G}_n)$ blows up to ∞ as n goes to ∞ . On the other hand, it can be shown that if the n^4 points are places so that the triangles of \mathcal{G}_n are fat then $\mathcal{A}(\mathcal{G}_n) \rightarrow \mathcal{A}(\mathcal{G})$.

2.2.1. Existence Results

Existence of fat triangulations of Riemannian manifolds is guaranteed by the studies given in [9–13]. These are summarized below.

Theorem 3. Every compact C^2 Riemannian manifold admits a fat triangulation [10].

Theorem 4. Every open (unbounded) C^∞ Riemannian manifold admits a fat triangulation [11].

Theorem 5. Let M^n be an n -dimensional C^1 Riemannian manifold with boundary, having a finite number of compact boundary components [13]. Then, any fat triangulation of ∂M^n can be extended to a fat triangulation of M^n .

The motivation for having fat triangulations for manifolds in terms of $p\ell$ -approximations is stressed by the following theorem.

Theorem 6. Let $(\mathcal{M}, \mathcal{T})$ be a C^r triangulated m -manifold embedded in \mathbb{R}^n [14]. Then, for $\delta > 0$, there exist $\epsilon, \varphi_0 > 0$, such that, for any $\tau < \mathcal{T}$, fulfilling the following conditions:

- $\text{diam}(\tau) < \epsilon$
- τ is φ_0 -fat

then, the secant map L_τ is a δ -approximation of $f|_\tau$.

Definition 7. Let \mathcal{M} be a Riemannian manifold. The Hausdorff metric is a metric defined on the set of compact subsets of \mathcal{M} by setting: $d_H(A, B) = \max_{a \in A} \{d_{\mathcal{M}}(a, B)\} + \max_{b \in B} \{d_{\mathcal{M}}(b, A)\}$, for any two compact subsets A and B , of \mathcal{M} , see [15].

2.3. Normal Curves

This section reviews essential background on normal curves. Most of definition and facts, even if some of them are new with respect to the existing literature, are merely obvious modifications of what is considered for surfaces, (See Appendix A). Along this subsection (Σ, \mathcal{T}, g) denotes a closed, smooth orientable surface Σ , endowed with a triangulation \mathcal{T} , and a Euclidean or hyperbolic metric g . We assume \mathcal{T} is a geodesic triangulation i.e., every edge of it is a geodesic arc between two vertices.

Definition 8. A normal curve is a curve so that each intersection arc of the curve with a 2-face is an arc that has its end points on two different edges of the 2-face. See Figure 2 below.

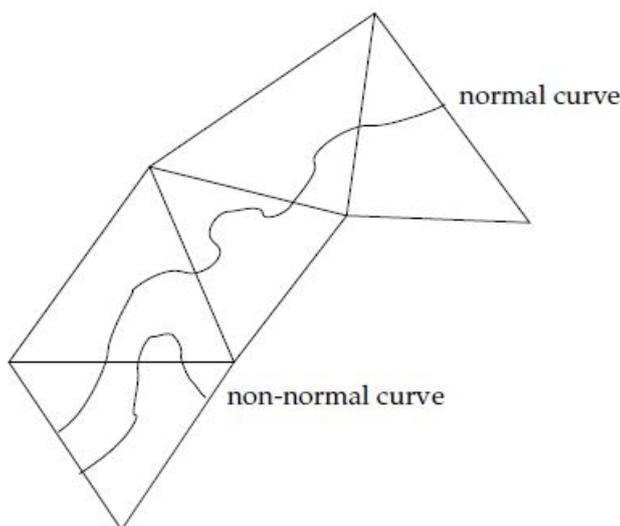


Figure 2. Normal and not normal pieces.

2.4. Geometric Measures

2.4.1. Weight and Length

Let (Σ, \mathcal{T}, g) be a triangulated Riemannian surface and let α be a rectifiable curve embedded in Σ and transverse to \mathcal{T} . Define the weight of α $w(\alpha)$, to be the total number of intersection points of the curve with $\mathcal{T}^{(1)}$.

Definition 9.

- A curve embedded in Σ and transverse to \mathcal{T} will be called normally short iff its weight is minimal with respect to its homotopy class.
- A closed normal curve that is not null homotopic (a.k.a essential), is called normally shortest if its weight is minimal with respect to all homotopy classes of essential curves in Σ .

Unless mentioned otherwise, when dealing with curves with boundary, normally short/shortest will always be considered with respect to homotopy rel- ∂ -points.

Let α be an embedded curve in (Σ, \mathcal{T}) transverse to \mathcal{T} . A local isotopy of α is an isotopy which slides a non normal piece of α across the edge intersected by this piece. Fix a given triangulation. Then the following facts hold:

- A normally short curve must be normal. This holds since if there exists a non-normal piece in some 2-face then it can be isotoped through the edge it intersects twice, thus reducing weight by at least two in contradiction to the minimality.
- For any close curve there exists a finite sequence of local isotopies which deforms the curve either to a normal curve or a curve which is completely included in just one triangle. In particular, if the curve is not null-homotopic, it is isotopic to a normal curve. This is true since the weight of a curve is an integer valued function which is bounded from below by zero, hence any curve is isotopic to some curve of minimal weight in its isotopy class which is either normal or has zero weight in which case it is entirely contained in a single 2-face.

2.4.2. Curvature

Definition 10. If Γ is a normal curve then its curvature \mathfrak{K} , is defined as follows. At a point inside a 2-face it is given by the geodesic curvature at that point, and at a point which belongs to $\Gamma \cap \mathcal{T}^{(1)}$ it is given by

$\mathfrak{K}(x) = \cos(\theta_1) + \cos(\theta_2)$, where θ_1, θ_2 are the angles between the vectors t_1, t_2 , tangent to Γ , at that point, and the vector t_e , which is tangent to the edge at this point. See Figure 3.

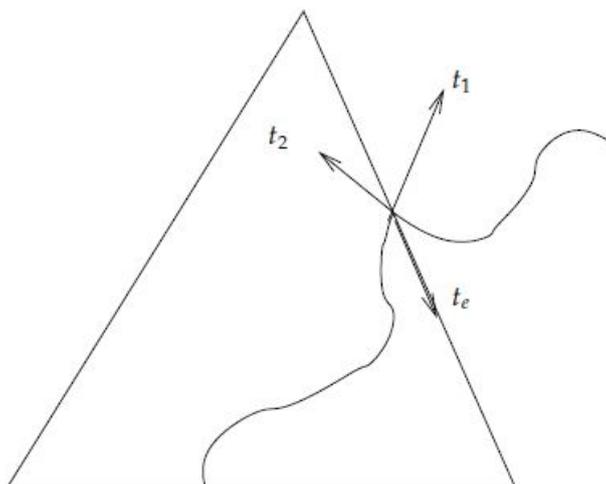


Figure 3. Curvature at intersection with the 1-skeleton.

A normal curve is straight, if it has zero curvature at each intersection point with the edges of \mathcal{T} .

Lemma 1. Let α be a piecewise smooth normal curve on a smooth triangulated surface (Σ, \mathcal{T}, g) . Suppose $x \in \alpha \cap \mathcal{T}^{(1)}$ is an intersection point of the curve on the interior of some edge e , so that α is straight at x in the sense defined above. Assume, in addition, that the restriction of α to the two faces meeting along e are geodesic curves in these two triangles. Then the geodesic curvature of α at x exists and equals zero:

$$\kappa_g(\alpha, x) = 0$$

Proof. Since the surface is smooth there exists well defined tangent plane for Σ at x . Let $T(\cdot)$ be the tangent vector field defined along α . Let $\bar{t}_1, \bar{t}_2, \bar{t}_e$ the tangent vectors of α and the tangent vector of e , all at the point x . Assuming that α is a geodesic in both triangles meeting along e we have that $\nabla_T T = \nabla_{\alpha'} \alpha' = 0$, as $\alpha' \rightarrow \bar{t}_1$, and the same holds while $\alpha' \rightarrow \bar{t}_2$. In addition, since $\bar{t}_1, -\bar{t}_2$ make an angle of π at x , T is also well defined at x and we have that the covariant derivative of T w.r.t itself is given by

$$\nabla_T T = \lim_{\epsilon \rightarrow 0} \frac{T(x - \epsilon) - T(x + \epsilon)}{\epsilon} = \frac{\bar{t}_1 - \bar{t}_2}{\epsilon} = 0$$

□

Remark 2. The curvature measure defined is a straight forward modification of the curvature measure defined in [2], for normal surfaces in 3-manifolds (see Appendix A), however Lemma 1 does not hold for surfaces in 3-manifolds. The difference between curves on surfaces and surfaces in 3-manifolds is basically the same as the difference between calculus in one variable to calculus in multiple variables.

Remark 3. Following Cheeger and Ebin ([16]—Chapter 1), one has that if Γ is a rectifiable curve (i.e., of finite total length), which minimizes length with respect to small perturbations then it has to be straight at all intersection points with $\mathcal{T}^{(1)}$ and each segment in $\Gamma \cap \mathcal{T}^{(2)}$, which connects two consecutive points in $\Gamma \cap \mathcal{T}^{(1)}$ is a geodesic. Since each edge of \mathcal{T} is a geodesic segment, and since the surface Σ is hyperbolic, in particular there is a unique geodesic that connects every pair of points on Σ , it follows that if Γ is a curve which minimizes length then every segment of Γ either coincides with an edge of \mathcal{T} , or is a normal arc, or a union of such arcs.

3. Shortening Through Straightening

In this section we will present a new method for shortening curves that is defined on normal curves. We will show that this method gives rise to a flow under which a normal curve converges to a geodesic.

Let C be a rectifiable curve on a triangulated Riemannian surface (Σ, \mathcal{T}, g) . Before describing the following straightening procedure it is necessary to extend the definition of a straight curve so that the curve is allowed to go through a vertex. Suppose Γ goes through a vertex V of \mathcal{T} , then Γ is straight at V if the two angles θ_1 and θ_2 , formed by Γ at V , as shown in Figure 4, are equal.

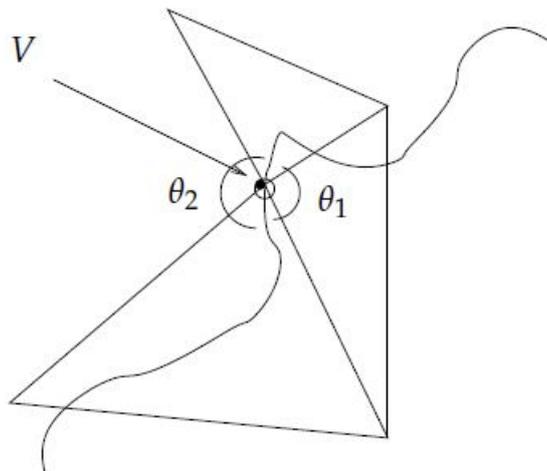


Figure 4. Straight curve at a vertex.

Consider the following procedure.

- Normalize C with respect to \mathcal{T} . For the shortening procedure the definition of normal curves is altered a bit. A normal curve is allowed to go through a vertex of \mathcal{T} .
- Take a least-weight normal curve \hat{C} , isotopic to C .
- Straighten \hat{C} at all intersections with the edges of \mathcal{T} , except at vertices. If \hat{C} intersects a vertex then this point does not change.
- Subdivide \mathcal{T} , using the median subdivision, to obtain a new triangulation \mathcal{T}_1 .
- Go to (1) while C is replaced by \hat{C} .

Remark 4. Step 3 is done by moving each internal point of a edge x , which belongs to $C \cap \mathcal{T}^1$ according to,

$$\frac{\partial x}{\partial t} = \cos(\theta_1) + \cos(\theta_2), \tag{3}$$

where θ_1, θ_2 are as before. In particular case at which x is at an edge of \mathcal{T} angles θ_1 and θ_2 are measured with respect to some chosen direction vector on the tangent plane to Σ at x . Using a partition of unity we can smoothly extend this process inside some small collar neighborhood of \mathcal{T}^1 , see Figure 5.

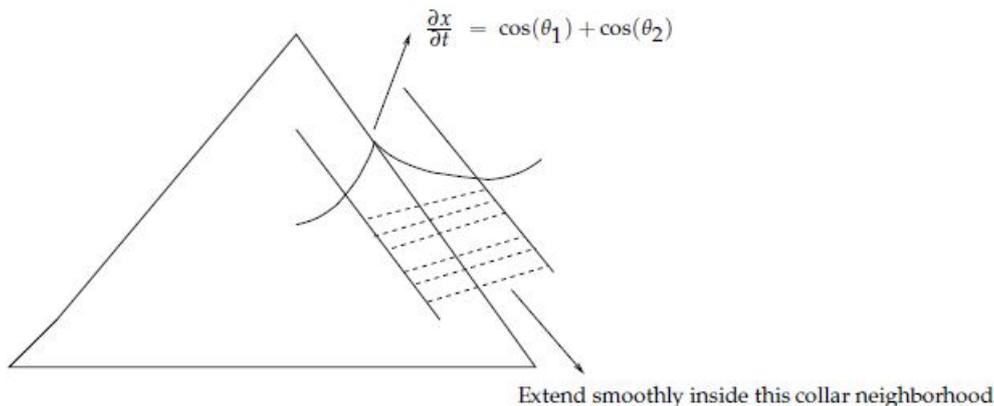


Figure 5. Shortening by straightening.

Theorem 7. *The obtained sequence of straight curves converges to a smooth geodesic on Σ .*

Proof. Let \mathcal{T}_n denote the n -th subdivision where \mathcal{T}_0 is the initial triangulation. Let λ_n be the mesh of \mathcal{T}_n . If x is an intersection point of \widehat{C}_n with some edge of \mathcal{T}_n then the curvature at x is zero, for \widehat{C}_n is assumed to be straight. Otherwise, x is a point on the interior of one of the normal pieces of \widehat{C}_n inside some triangle τ . First, observe that at each iteration while normalizing is applied (Step (1) above), each obtained normal arc is contained in the “parent” triangle of the triangle it pathes through, since it is a piece of a normal arc at the previous iteration, see Figure 6.

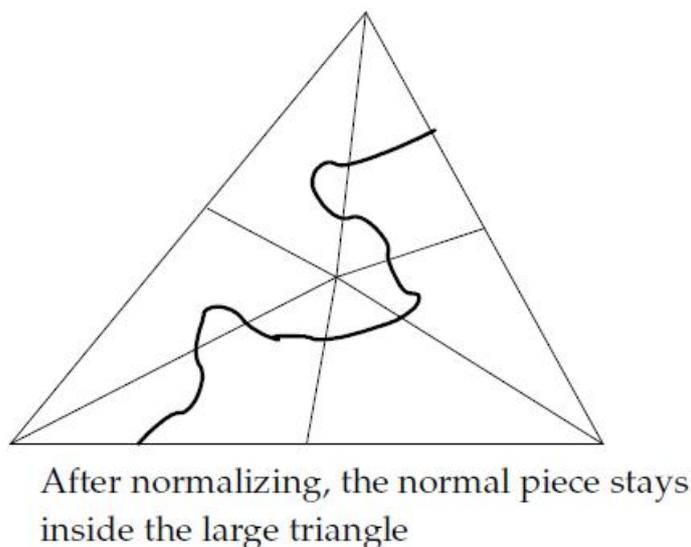


Figure 6. A normalized piece.

Hence, since vertices are becoming dense in the surface, the obtained sequence of normal curves is a Cauchy sequence, so it converges to a limit curve C_∞ in the topological completion of the space of normal curves on Σ .

Claim 3.1. *The limit curve C_∞ is a geodesic.*

Proof. Let $\epsilon > 0$ be some positive number. Let τ be some triangle of \mathcal{T} . Let $\widehat{C}_n|_\tau = \widehat{C}_n \cap \tau$ be a normal piece inside τ , and let $\{y_1, y_2\}$ be the end points of that normal piece (Figure 7).

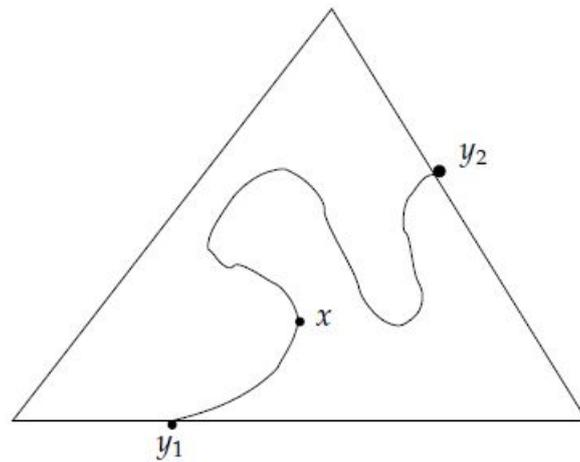


Figure 7. The curvature inside a triangle.

According to Proposition 1.2 of [17] since $\widehat{C}_n|_\tau$ is rectifiable it has a Lipschitz parametrization, hence there exists a constant M so that,

$$d_\Sigma(y_1, y_2) \leq L(\widehat{C}_n|_\tau) \leq M \cdot d_\Sigma(y_1, y_2) \leq M \cdot \lambda_n \tag{4}$$

In addition, since $\widehat{C}_n|_\tau$ is a normal piece there is some $\delta = \delta(\epsilon)$ so that $\widehat{C}_n|_\tau$ is in the δ -neighborhood of $\overline{y_1 y_2}$ in Σ . Hence,

$$\min\{L(\widehat{C}_n|_{y_1 x}), L(\widehat{C}_n|_{x y_2})\} \leq \delta. \tag{5}$$

From Lemma 1 the curvature as defined is continuous, so there exists such $\delta(\epsilon)$ satisfying,

$$\min\{|\mathfrak{K}(\widehat{C}_n, x) - \mathfrak{K}(\widehat{C}_n, y_1)|, |\mathfrak{K}(\widehat{C}_n, x) - \mathfrak{K}(\widehat{C}_n, y_2)|\} \leq \epsilon,$$

therefore

$$|\mathfrak{K}(\widehat{C}_n, x)| \leq \delta.$$

Hence, if C_∞ is the limit curve of the sequence C_n , then at every point $x \in C_\infty$ the geodesic curvature of C_∞ is zero, so C_∞ is a geodesic. \square

This completes the proof of Theorem 7. \square

Remark 5. Normality of curves, is crucial and best illustrated by the lemma above for if the curve was not normal then we could not assume a bound on the arc length of its normal pieces, relative to the distance between y_1 and y_2 above, even when the division parameter λ gets very small.

4. Experimental Results

In this subsection some initial experimental results will be presented. This is the place to thank the students Chen Harel, Amir Gilad, Dima Serf and Gleb Lander from the Technion Electrical Engineering Department for their devoted work while programming the implementation of the curve shortening algorithm described above, in the course of their undergraduate studies. As an initial stage the dual graph of \mathcal{T} is built and a shortest path between the two triangles containing these points is found using the well known Dijkstra algorithm, see [18]. This shortest path in the dual graph corresponds to the minimal number of triangles one has to travel through in order to get from one point to the other. The obtained set of triangles is called the triangle path between the two points. A triangle path is illustrated in Figure 8 below, on a surface modelling a hand. In a second stage an initial guess for the shortest geodesic is chosen and then evolved according to the straightening algorithm above.

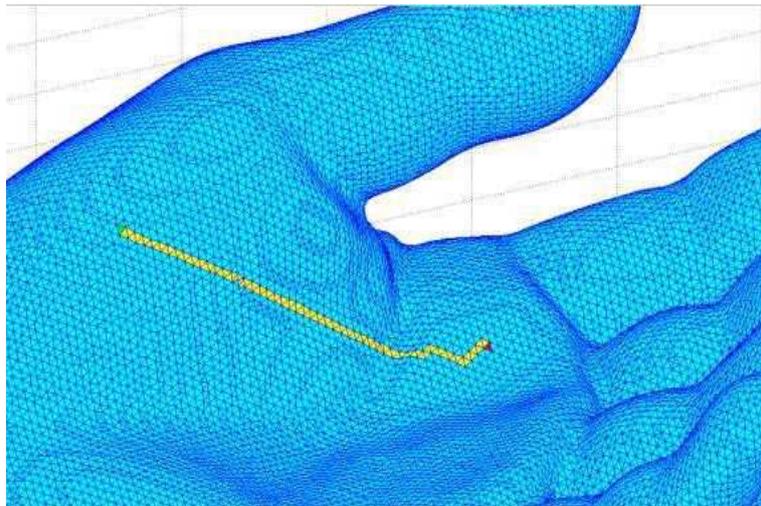


Figure 8. Triangle path.

4.1. Synthetic Surfaces

The following two examples demonstrate the algorithm as tested on synthetic surfaces.

Example 2. In Example 2 the algorithm is tested on a 2-sphere. Figure 9 shows the two antipodal points along with the obtained path are shown. The Path coincides with the half circle painted in red (the dark half circle following the bright one shown in the figure).

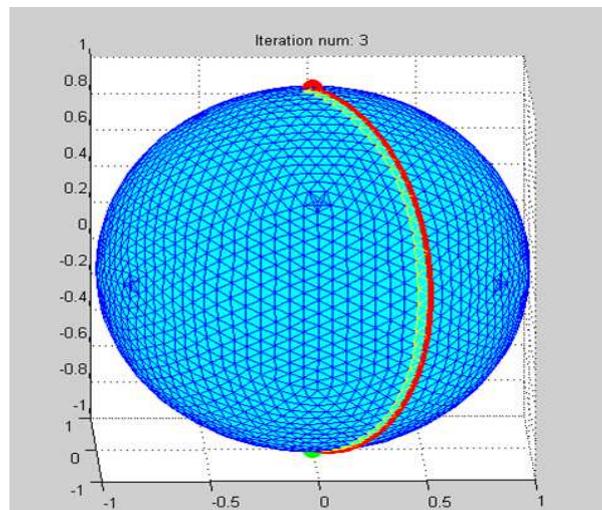


Figure 9. Sphere.

Example 3. In Example 3 the algorithm is tested on a knotted torus. Figure 10 shows the obtained path.

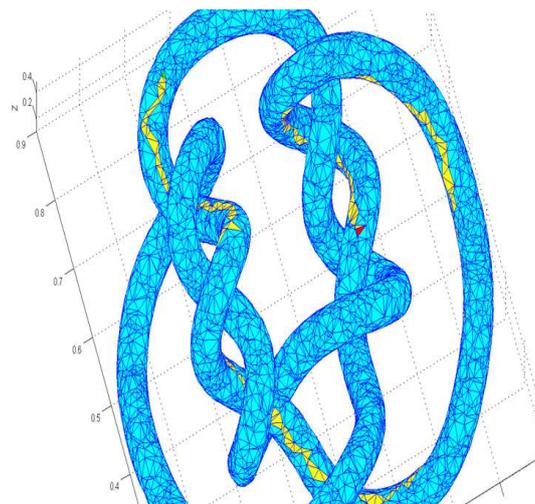


Figure 10. Knotted torus.

4.2. “Real” Data and Weighted Triangulation

As the Dijkstra algorithm for finding shortest path in a graph is suitable also to weighted graphs, where in this case a positive set of weights is assigned to the set of edges of the graph, and each weight reflects the cost of going through the specific edge. The algorithm above can be applied on weighted triangulations as well. A weighted triangulation is a triangulation endowed with a set of positive weights, assigned on the set of triangles, so that each weight reflects the cost it takes to path through the specific triangle when traveling from one chosen point to the other. Given this, the triangle path will be found using the weighted Dijkstra algorithm on the dual graph to the triangulation.

Example 4. In Example 4 the algorithm is tested on a triangulated surface presenting a patch of the surface of the Earth. A chosen area in the Nevada Desert is represented as a triangulated surface. Figure 11 presents the overall area from which the examples are taken. Figure 12 shows an example of the extracted triangulated surface representing a chosen patch in the overall area.

The weight on each triangle is proportional to its height. Figure 13 shows the obtained path between a chosen pair of points.

The Following Figure 14 shows the outcome of the algorithm in the presence of an obstacle. The chosen points are located on two sides of a lake, and the obtained route indeed avoids the lake.

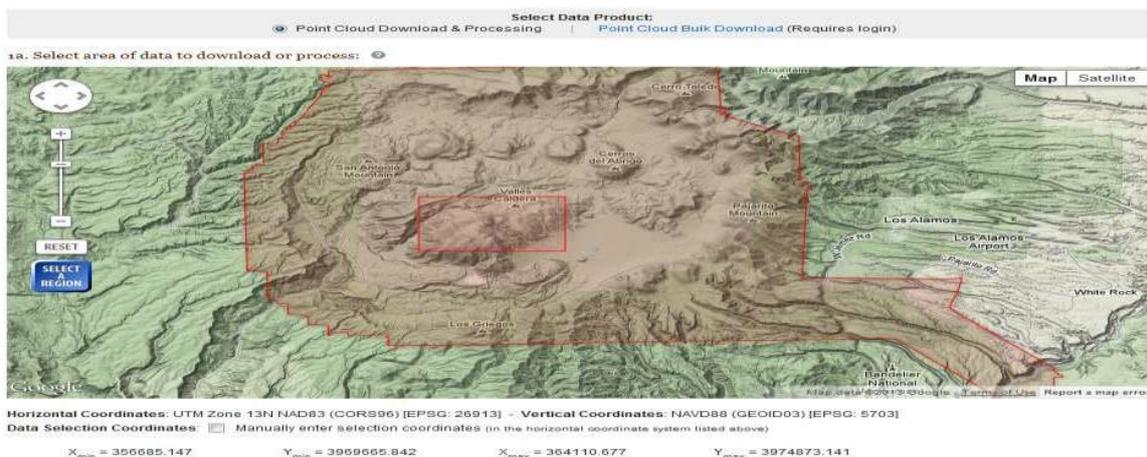


Figure 11. overall area.

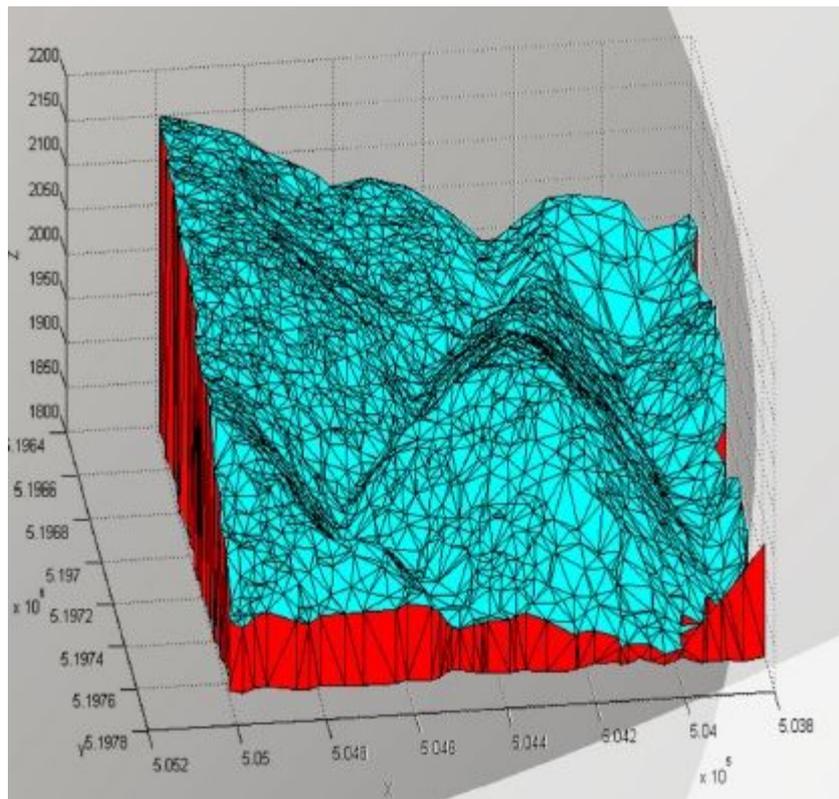
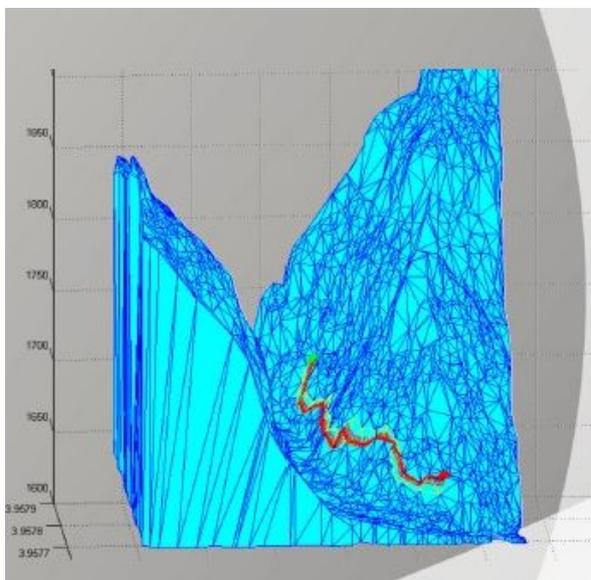


Figure 12. Triangulated surface of the Earth.



(a)



(b)

Figure 13. Earth route.

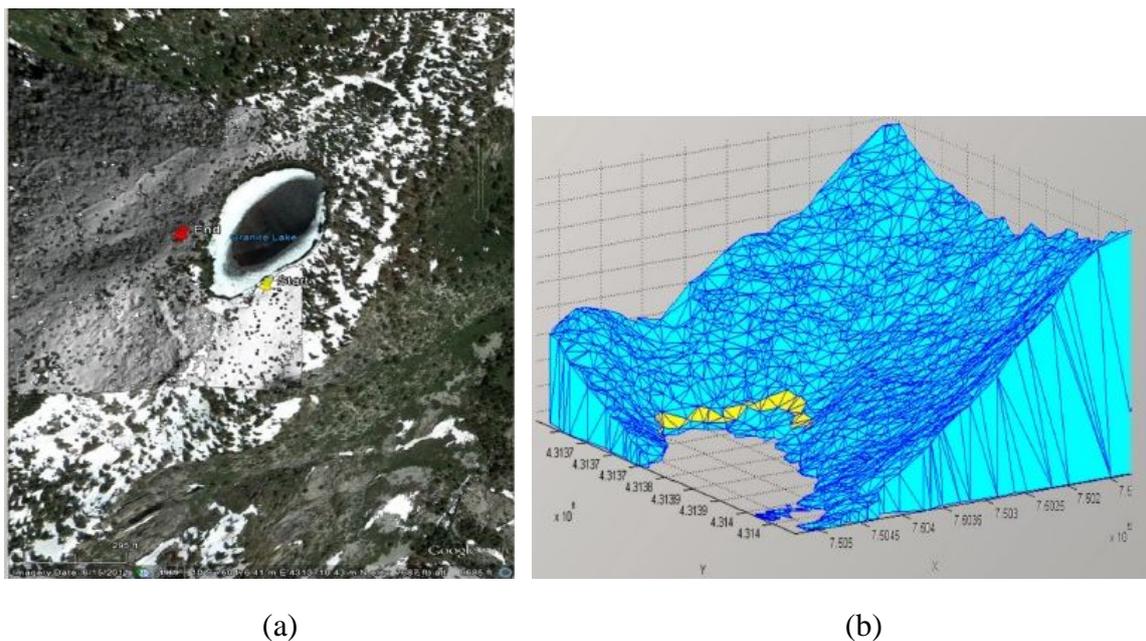


Figure 14. Walk around a lake.

5. pl Approximations of Geodesics

Theorem 8. Let Σ be a hyperbolic surface isometrically embedded in some \mathbb{R}^n . Let γ be a curve on Σ that is either a shortest geodesic between two given points p and q on Σ or, a shortest closed geodesic in Σ . Then there exists a sequence of pl 2-submanifolds of \mathbb{R}^n , Σ_n , each of them is a pl -approximation of Σ , and a sequence of pl -curves γ_n , each of which is normal with respect to Σ_n so that γ_n converges to γ .

Proof. The pl -approximation is given by the secant map as in Definition 5. Recall that according to Remark 3 every shortest path is a collection of normal pieces and pieces that coincide with edges of \mathcal{T} . Each normal piece of γ can be projected to a linear normal piece on a planar triangle of the secant approximation exactly in the same way as was done for surfaces. Every piece that coincides with an edge is projected onto the corresponding edge of the secant approximation of Σ . This projection yields the sequence of curves that approximates γ . \square

6. Reflections in the Dual Graph

Let γ be a normal curve on the triangulated surface (Σ, \mathcal{T}) . Suppose $\partial\gamma = \{p, q\}$ are two fixed points on Σ interior in some two 2-cells, τ_p, τ_q , respectively. It is possible that $p = q$. Then γ naturally defines a weighted path γ^* in the dual graph \mathcal{T}^* of \mathcal{T} , in the following way.

- Each vertex of the dual path γ^* corresponds to a triangle intersected by γ .
- Each edge of γ^* is assigned to an edge of \mathcal{T} at which γ crosses from one 2-cell to an adjacent one.
- Each edge of γ^* is weighted according to the number of times γ crosses through the corresponding edge of \mathcal{T} .
- The length of γ^* is defined to be the sum of weights of all its edges.

Remark 6. It is essential at this context to consider curves that are normally shortest, since while proving some of the following lemmas isotopy classes may change.

Lemma 2. Let γ be a normally shortest curve that is either a simple closed curve or a curve that connects two given distinct points, p and q . Then, γ has at most one normal piece in every triangle of \mathcal{T} .

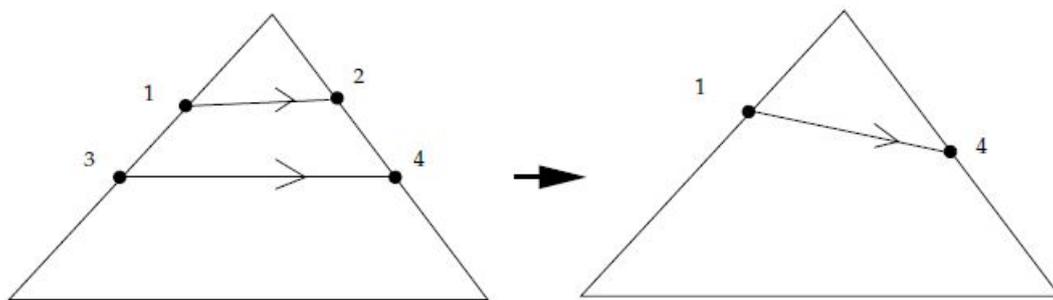
Proof. Suppose conversely, that there is some triangle τ the intersection of which with γ is made of more than one normal piece. For simplicity suppose for the moment that it is made of two normal pieces. The general case will easily follow. Endowing the curve γ with an orientation, gives rise to four different types of such intersection patterns. In any of the following figures, the numbering of the points on the edges of τ is according to their order with respect to the chosen orientation.

- Two parallel intersections: γ passes through τ twice between the same two edges of τ with the same orientation.
- Anti parallel intersection: γ as in (1) yet, with opposite orientations.
- Non parallel: Only one pair of points is on the same edge.

In any of the four cases the point 1 is connected to 4. A distinction will now be made between simple close curves and curves with boundary.

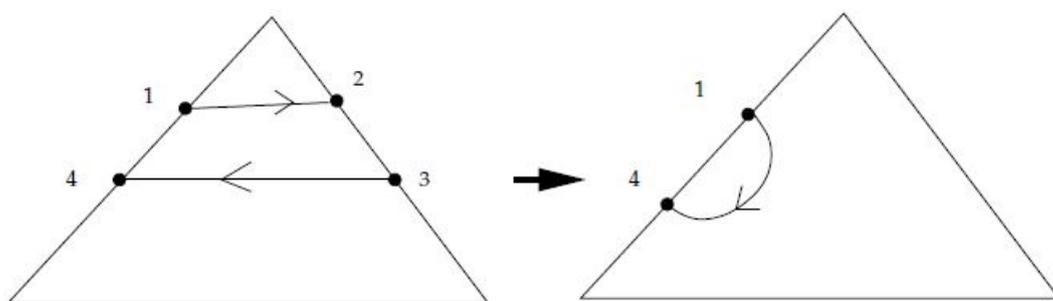
If $\partial\gamma \neq \emptyset$, then by doing the above replacement another curve is obtained that also connects p to q which may or may not stay normal with respect to the triangulation yet, its weight is strictly smaller than the weight of γ . In case it is not normal it is normalized while further reducing weight. All in all, a normal curve is obtained, which connects the two ∂ -points and has a lesser weight than γ in contradiction to the fact that γ is normally shortest. Refer to Figures 15 and 16 for two typical examples.

If γ is a simple close curve then in addition to connecting point 1 to 4, then the point 3 is connected to 2. This way two simple close curves γ' and γ'' are obtained, that either intersect each other at some point inside τ or they are disjoint in τ , as seen in Figures 17 and 18, showing the typical possibilities in this case. This of course depends on the way γ passes through τ .



Resolving parallel normal pieces. The part of γ from the points 2 and 3 is deleted and 1 is directly connected to 4. We end up with a normal curve having lesser weight.

Figure 15. Resolving parallel components.



Resolving anti-parallel normal pieces. The part of γ from the points 2 and 3 is deleted and 1 is directly connected to 4. We obtain a non-normal curve which is normalized further decreasing its weight. Finally, we end up with a normal curve having smaller weight.

Figure 16. Resolving non parallel components.

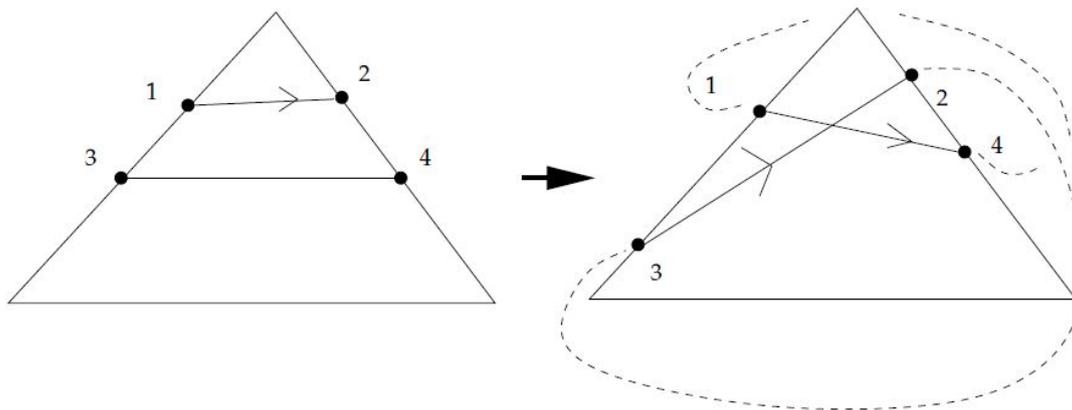


Figure 17. Intersecting closed curves.

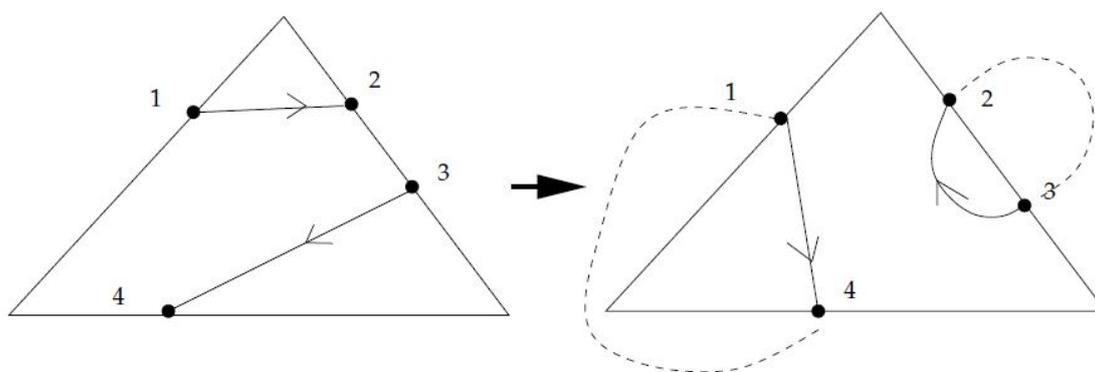


Figure 18. Disjoint closed curves.

Suppose $\gamma' \cap \gamma'' \neq \emptyset$. In this case both of them are normal with respect to \mathcal{T} , and it is necessary to show that at least one of them is essential. If say, γ' is null homotopic in Σ it must separate a disk in Σ but then by the Jordan curve theorem it must intersect with γ'' at an additional point as well. This intersection point has to be outside of τ but, the parts of both γ' and γ'' outside of the triangle τ coincide with the pieces of the initial curve γ , therefore γ has a self intersection point and this contradicts the assumption that it is simple.

Suppose now that γ' and γ'' are disjoint from each other, and suppose both of them are null homotopic in Σ . In this case the initial curve γ can be reproduced as a connected sum of γ' and γ'' inside the triangle τ , thus making γ null homotopic in contradiction to its being essential. Suppose γ' is essential. Still, as indicated in the Figure 18, it may happen that γ' is not normal. In this case it is further normalized until a normal simple closed curve is obtained. This normal curve has strictly less weight than the initial curve γ , again forming a contradiction.

The same argument will go also for the general case of more than two normal pieces inside a triangle thus completing the proof. \square

Lemma 3. A path in the dual graph for which all weights along its edges are equal to 1 is realizable as a normal curve with respect to \mathcal{T} .

Proof. The proof is obvious by assigning a normal piece to every pair of consecutive edges along the path. \square

Theorem 9. A normal curve is normally shortest if and only if its corresponding dual path is a shortest path according to the length function defined in Section 6.

Proof. From Lemmas 2 and 3 there is a correspondence between unitary normal curves, i.e., curves that have at most one normal piece inside each triangle, and unitary pathes in the dual graph, i.e., paths for which the weight along each edge is one, so that, the weight of each normal curve, is equal to the length of its dual path and vice versa. The proof thus follows. \square

Corollary 1. *There is an algorithm which finds in finite time a normally shortest curve on a triangulated surface.*

Proof. For a start find a least-weight dual curve in the dual graph using one of the numerous existing algorithms, such as the Dijkstra algorithm. Afterwards we use Theorem 9 to find its realization as a least-weight normal curve. A rough estimation on the complexity of this algorithm shows that it is in the order of the complexity of finding the least weighted curve in the dual graph times the complexity of building the dual graph. Building the dual graph is proportional to the number of triangles in the original triangulation while the complexity of say, Dijkstra algorithm is in the magnitude of n^2 , where n is the number of vertices in the dual graph which is equal to the number of triangles in the original graph. All in all a rough estimation of the complexity of the algorithm is $O(\#triangles)^3$. \square

Conflicts of Interest: The author declares no conflict of interest.

Appendix A. Normal Surfaces

In this section we will give some preliminary definitions and notations from normal surface theory that will be used later in this work.

Normal surfaces were first introduced in the study of 3-manifolds by H. Kneser [19], in the early 1930's. Kneser used normal surfaces for his proof of the existence of prime decomposition of 3-manifolds. Normal surfaces were further used by W. Haken in the 1960's [1], while giving Haken the possibility to develop a combinatorial viewpoint on topological problems such as the recognition of the unknot and other problems related to 3-manifolds, in the aim of producing computable algorithms. Nowadays, normal surfaces are extensively used and studied in overall aspects of 3-manifold topology and geometry, yielding more and more results and algorithms. The novice reader should look into [1,20]. Among other things, there is an algorithm that uses normal surfaces in order to decide in finite time whether or not a given 3-manifold is Haken, see [21].

Throughout this section $(\mathcal{M}, \mathcal{G}, \mathcal{T})$ will denote a 2-dimensional surface, \mathcal{G} , embedded in 3-manifold \mathcal{M} , which admits a triangulation, \mathcal{T} .

Definition A1. *An isotopy $H : \mathcal{M} \times I \rightarrow \mathcal{M}$ is a normal isotopy with respect to a given triangulation \mathcal{T} , if $H_x(\sigma) = \sigma$ for every $x \in I$ and for all simplices $\sigma \in \mathcal{T}$. We further say that two submanifolds of \mathcal{M} are normally isotopic if there exists a normal isotopy which carries one to the other.*

Definition A2. *Let $\mathfrak{t} \in \mathcal{T}$ denote a tetrahedron of a given triangulation and let \mathfrak{f} denote a face of \mathfrak{t} . A normal arc in \mathfrak{f} , is an arc properly embedded in \mathfrak{f} , such that its end points are on two different edges of \mathfrak{f} . A normal curve in $\mathcal{T}^{(2)}$, is a curve properly embedded in $\mathcal{T}^{(2)}$, transverse to the edges of \mathcal{T} , so that its intersection with each 2-face consists of collection of normal arcs. A normal disk is a disk \mathcal{D} , properly embedded in a tetrahedron \mathfrak{t} such that $\partial\mathcal{D} = \mathcal{D} \cap \partial\mathfrak{t}$, is a simple closed normal curve in $\partial\mathfrak{t}$. An elementary disk is a normal disk so that its boundary is either a triangle that separates one vertex on $\partial\mathfrak{t}$, from all three other vertices $\partial\mathfrak{t}$, or a quadrilateral (quad for short) that partitions the vertices of the tetrahedron into two separated pairs.*

Elementary disks are illustrated in Figure A1, and they come in seven different types corresponding to seven possibilities to partition the set of vertices of a tetrahedron into subsets according to the definition. In the sequel we will sometimes refer to the boundary of an elementary disk as a 3-gon or 4-gon according to its being a triangle or a quad, and for a general normal disk we may refer to its boundary as an n -gon.

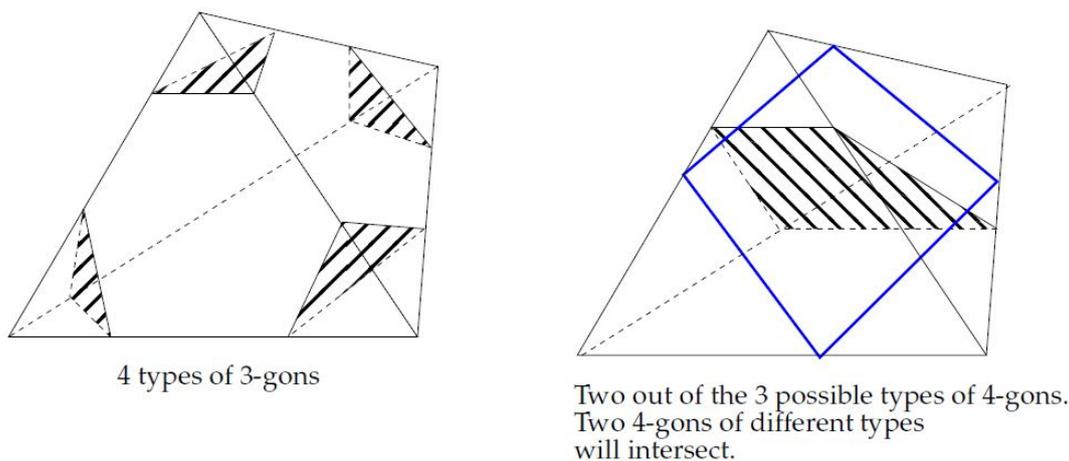


Figure A1. Elementary disks.

Definition A3. A surface \mathcal{G} properly embedded in \mathcal{M} is normal with respect to \mathcal{T} if it is transverse to the skeleta of \mathcal{T} , and if its intersections with any of the tetrahedra of \mathcal{T} is composed of a (possibly empty) collection of disjoint elementary disks.

Note that since the surface is embedded, we can have only one type of 4-gon at each tetrahedron.

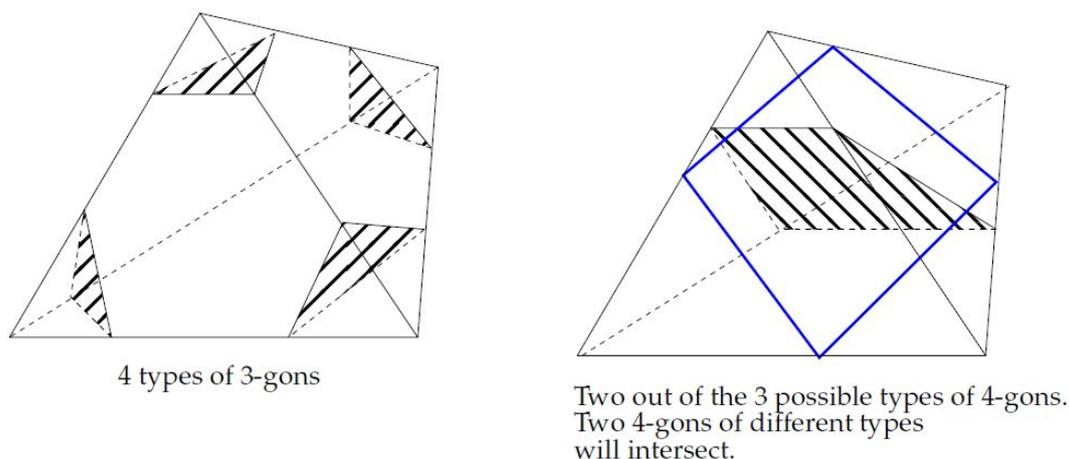


Figure A2. Normal disks.

Appendix B. Possible Normal Disk Types

The length of a normal curve is the number of normal arcs consisting it. Recall that the boundary of a normal disk is a simple closed normal curve on the boundary of a tetrahedron that forms a polygon. A thorough analysis of the possible length such a polygon can have, was done by Thompson and Stocking in [20,22] respectively. It is summarized as follows.

Lemma A1. If \mathcal{D} is a normal disk then $\partial\mathcal{D}$ is an n -gon where n is either 3, 4 or an even number ≥ 8 [20,22].

Figure A3 shows an example of a normal disk whose boundary is of length eight.

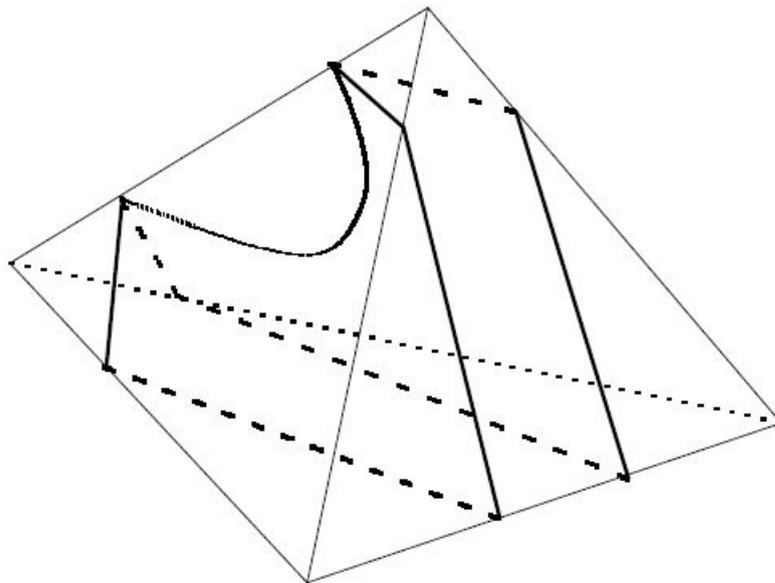


Figure A3. Octagon normal disk.

Appendix C. Weight and Area

Definition A4. Let \mathcal{G} be a surface properly embedded in $(\mathcal{M}, \mathcal{T})$ and transverse to the skeleta of \mathcal{T} , (\mathcal{G} needs not be normal).

- The weight $w(\mathcal{G})$, of \mathcal{G} is the total number of intersections of the surface with the 1-skeleton $\mathcal{T}^{(1)}$.
- A metric on each 2-face of \mathcal{T} is induced by identifying the face with an ideal triangle in the hyperbolic plane \mathbb{H}^2 . As a result we can measure the length of each arc in $\mathcal{G} \cap \mathcal{T}^2$. The length $\ell(\mathcal{G})$, of \mathcal{G} is defined as the total sum of lengths of all arcs in $\mathcal{G} \cap \mathcal{F}^{(2)}$.
- The area of \mathcal{G} is defined as the ordered pair $(w(\mathcal{G}), \ell(\mathcal{G}))$, lexicographically ordered.

Definition A5.

- A surface \mathcal{F} is called least weight iff its weight is minimal amongst all surfaces in its isotopy class.
- A normal surface is called a *pl*-minimal surface iff its length is a stationary point with respect to small variations in its normal isotopy class.
- A surface is of *pl*-least area surface iff its area is minimal in its isotopy class.

Definition A5 above is equivalent to the following. If $F : \mathcal{G} \times (-\delta \ \delta) \mapsto \mathcal{M}$, is a small variation of \mathcal{G} , and $\mathcal{G}_s = F(\mathcal{G} \times \{s\})$, and suppose $\mathcal{G} = \mathcal{G}_0$. Then, \mathcal{G} is normally minimal if,

$$\frac{d}{ds} \ell(s)|_{s=0} = 0,$$

where, $\ell(s) = \ell(\mathcal{G}_s)$.

Notice that each two normally isotopic surfaces have the same weight. In addition, one can show that a least weight surface must be normal. Using this observation, the generality of normal surfaces is stressed by the following theorem,

Theorem A1. ([1,2]) Let \mathcal{G} be a surface properly embedded in a 3-manifold \mathcal{M} . Then, for any triangulation \mathcal{T} , of \mathcal{M} , there is a finite sequence of compressions, ∂ -compressions and local isotopies, that deforms \mathcal{G} into a (possibly empty) set of surfaces that are normal, w.r.t. \mathcal{T} , and another (maybe empty) set of surfaces, each of which is entirely contained in a single tetrahedron of \mathcal{T} . In particular, if \mathcal{M} is irreducible, and \mathcal{G} is incompressible and ∂ -incompressible, then it is isotopic to a normal surface.

Remark A1. The natural candidate for a surface to be isotoped to, is a least weight surface in an isotopy class. From the proof of Theorem A1 one immediately obtains that a least weight surface is normal. Furthermore, Theorem A1 reflects nicely the analogy between normal surfaces and minimal surfaces. In [23–25] Meeks and Yau make a thorough study of minimal and least area surfaces in 3-manifolds and using partial differential equations techniques they show among other things that an incompressible surface is isotopic to a least area surface. Hence, in a way Theorem A1 can be regarded as a combinatorial counterpart of this result.

Appendix D. Curvature of Normal Surfaces

Let \mathcal{G} be a normal surface in the triangulated 3-manifold $(\mathcal{M}, \mathcal{T})$. In [2] a piecewise smooth mean curvature vector field, H , is defined on the intersection of a normal surface in the following way.

Definition A6. Suppose α is some arc of $\mathcal{G} \cap \mathcal{T}^{(2)}$ and x is a point on α .

- If x is a point of lying in the interior of a 2-simplex, then define

$$\bar{H}(x) = \nabla_{\bar{T}}(\bar{T}(x)) ,$$

where \bar{T} here stands for the tangent vector of α at x and ∇ denotes covariant derivative. Hence, $|\bar{H}|$ at an internal point is the geodesic curvature of α considered as an arc embedded in a triangular face of the triangulation (assume α is arc-length parameterized).

- If $x \in \alpha \cap \mathcal{T}^{(1)}$ is an end point of α , let $\{\alpha_1, \alpha_2, \dots, \alpha_k\}$ be all arcs of $\mathcal{G} \cap \mathcal{T}^{(2)}$ emanating from x , and define

$$H(x) = \sum_{i=1}^k \langle \bar{t}_i(x), \bar{V} \rangle \bar{V} ,$$

where \bar{t}_i is the tangent vector of α_i at x and \bar{V} is a vector tangent to the edge e of T where x is included. $|\bar{H}|$ here is the sum of all co-sines of the angles between the t_i s and \bar{V} .

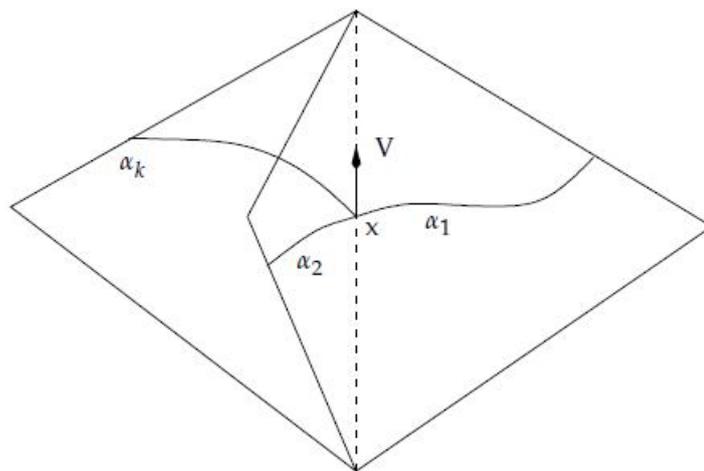


Figure A4. Curvature at an intersection point with an edge.

Using first and second variations, it is shown in [2] that \mathcal{G} is a PL minimal surface iff $\bar{H}(\mathcal{G}) \equiv 0$.

Appendix E. Mean Curvature Flow

As mentioned earlier classical results about the existence of least area surfaces such as [3] are based on techniques of partial differential equations. In particular, they are based on the analysis of

the following mean curvature flow equation. Let \mathcal{G} be a smooth surface embedded in a Riemannian 3-manifold $(\mathcal{M}, \mathfrak{g})$. Then the mean curvature flow is defined by the evolution equation, see [26]:

$$\begin{aligned} \frac{\partial \mathcal{G}}{\partial t} &= H\hat{N}; 0 \leq t \leq T < \infty, \\ \mathcal{G}(0, *) &= \mathcal{G}. \end{aligned} \tag{A1}$$

That is, every point of \mathcal{G} moves in the direction of the normal vector at that point, and at velocity that is proportional to the mean curvature at that point. In contrast, based on the definition of $p\ell$ -minimal and $p\ell$ -least area surfaces, as well as on the definition of curvature for normal surfaces, Jaco and Rubinstaein deduced, in [2], a $p\ell$ version of the mean curvature flow. By the analysing it they where able to obtain the following existence theorem for $p\ell$ -minimal surfaces: A vertex linking sphere in a triangulated 3-manifold $(\mathcal{M}, \mathcal{T})$ is a normal 2-sphere that is the boundary of a small 3-ball around a vertex of \mathcal{T} .

Theorem A2. ([2]) *In the normal homotopy class of any normal surface \mathcal{G} , that is not a vertex linking sphere, there exists a $p\ell$ -minimal surface.*

Remark A2. *In the theorem above Jaco-Rubinstein also claim uniqueness of the surface; however, since the definition of $p\ell$ -minimal as well as $p\ell$ -least area surfaces does not account for the interior of the normal disks, the uniqueness argument in [2] holds only with respect to the boundary of any normal disks while the interior of the disk may vary. In [27] the definition of a $p\ell$ -minimal surface is altered by putting an additional constraint that the interior of each normal disk in $\mathcal{G} \cap \mathcal{T}^{(3)}$, is a minimal surface in the usual differential geometric sense.*

With this altered definition of $p\ell$ -minimality it is shown,

Theorem A3 ([27]). *There is exactly one $p\ell$ -minimal surface in a normal homotopy class of any normal surface \mathcal{G} that is not a vertex linking sphere.*

Appendix F. $p\ell$ -Minimal Surfaces and Approximation of Least Area Surfaces

The above existence theorem gives hope that the answer to Question II, that was raised in the introduction, is affirmative. That is, it is possible to represent a least area surface as a limit surface of a sequence of $p\ell$ -minimal surfaces. A seemingly natural approach to prove such an affirmative answer would be the following: Show it is possible to build a sequence of triangulations of the manifold \mathcal{M} , of decreasing mesh, so that the given least area surface is normal with respect to each of these triangulations. Then, based on the above theorem, take the $p\ell$ -minimal representative of the least area surface, inside its normal homotopy class, with respect to each of the triangulations. This way one obtains a sequence of $p\ell$ -minimal surfaces that should converge to the least area surface. However, as the mesh of the triangulations decreases it may very well happen that the given least area surface will become arbitrarily close to vertices of triangulations in the sequence. This pose a major problem. In the proof of Theorem A2 above it is shown that a surface which is $p\ell$ -minimal with respect to some triangulation cannot get close to the vertices of that triangulation. Therefore, whenever the least area surface is close to a vertex it will remain far away from all $p\ell$ -minimal surfaces. Since any natural way to define convergence of a sequence of surfaces to some limit surface will evidently require that the surfaces in the sequence will become close, in some natural manner, to the limit surface, the constrain presented here prevents us from obtaining an affirmative answer to Question II. In two consecutive papers Question I and II will be addressed in the case of normal surfaces.

References

1. Haken, W. Theorie der normalflächen. *Acta Math.* **1961**, *105*, 245–375.
2. Jaco, W.; Rubinstein, J.H. PL-Minimal surfaces in 3-manifolds. *J. Differ. Geom.* **1988**, *27*, 493–524.

3. Meeks, W.H., III; Simon, L.; Yau, S.T. Embedded minimal surface, exotic spheres and manifolds with positive ricci curvature. *Ann. Math.* **1982**, *116*, 621–659.
4. Caselles, V.; Kimmel, R.; Sapiro, G. Geodesic Active Countours. *Int. J. Comput. Vis.* **1997**, *22*, 61–79.
5. Grayson, M. Shortening embedded curves. *Ann. Math.* **1989**, *129*, 71–111.
6. Grayson, M. The heat equation shrinks embedded curves to round points. *J. Differ. Geom.* **1987**, *26*, 285–314.
7. Hass, J.; Scott, P. Shortening Curves on Surfaces. *Topology* **1994**, *33*, 25–43.
8. Kimmel, R.; Malladi, R.; Sochen, N. Images as embedded maps and minimal surfaces. *Int. J. Comput. Vis.* **2000**, *39*, 111–129.
9. Breslin, W. Thick triangulations of hyperbolic n -manifolds. *Pacific J. Math.* **2009**, *241*, 215–225.
10. Cairns, S.S. A simple triangulation method for smooth manifolds. *Bull. Am. Math. Soc.* **1961**, *67*, 389–390.
11. Peltonen, K. *On the Existence of Quasiregular Mappings*; Annales Academiae Scientiarum Fennicae. Series A. I, Mathematica, Dissertationes; Suomalainen Tiedeakatemia: Helsinki, Finland, 1992.
12. Saucan, E. The existence of quasimeromorphic mappings in dimension 3. *Conform. Geom. Dyn.* **2006**, *10*, 21–40.
13. Saucan, E. Note on a theorem of Munkres. *Mediterr. J. Math.* **2005**, *2*, 215–229.
14. Munkres, J. *Elementary differential topology*; Princeton University Press: Princeton, NJ, USA, 1966.
15. Morvan, J.-M.; Thibet, B. Approximation of the normal vector field and the area of a smooth surface. *Disc. Comp. Geom.* **2004**, *32*, 383–400.
16. Cheeger, J.; Ebin, D. *Comparison Theorems in Riemannian Geometry*; North-Holland Mathematical Library; American Mathematical Society: Providence, RI, USA, 1975; p. 9.
17. Sullivan, J.M. Curves of finite total curvature. *Disc. Diff. Geom.* **2008**, *38*, 137–161.
18. Dijkstra, E.W. A note on two problems in connexion with graphs. *Numer. Math.* **1959**, *1*, 269–271.
19. Kneser, H. Geschlossene flachen in drei-dimesnionalen Mannigfaltigkeiten. *Jahresbericht der Deutschen Mathematiker-Vereinigung* **1929**, *38*, 248–259.
20. Thompson, A. Thin position and the recognition of S^3 . *Math. Res. Lett.* **1994**, *1*, 613–630.
21. Jaco, W.; Oertel, U. An Algorithm to decide if a 3-Manifold is a Haken Manifold. *Topology* **1984**, *23*, 195–209.
22. Stocking, M. Almost normal surfaces in 3-manifolds. *Trans. Am. Math. Soc.* **2003**, *352*, 171–207.
23. Meeks, W.H., III; Yau, S.T. Topology of three dimensional manifolds and the embedding problem in minimal surface theory. *Ann. Math.* **1980**, *112*, 441–484.
24. Meeks, W.H., III; Yau, S.T. The Classical Plateau problem and the topology of three dimensional manifolds. *Topology* **1982**, *21*, 409–442.
25. Meeks, W.H., III; Yau, S.T. The Existence of embedded minimal surfaces and the problem of uniqueness. *Math. Z.* **1982**, *179*, 151–168.
26. Huisken, G. Flow by mean curvaature of convex surfaces into spheres. *J. Differ. Geom.* **1984**, *20*, 237–266.
27. Ni, Y. Uniqueness of PL minimal surfaces. *Acta Math. Sin.* **2007**, *23*, 961–964.

