Homotopic Curve Shortening and the Affine Curve-Shortening Flow

Sergey Avvakumov
Institute of Science and Technology Austria (IST Austria),
Am Campus 1, 3400 Klosterneuburg, Austria
sergey.avvakumov@ist.ac.at

Gabriel Nivasch[1]
Ariel University, Ariel, Israel
gabrieln@ariel.ac.il

Abstract

We define and study a discrete process that generalizes the convex-layer decomposition of a planar point set. Our process, which we call *homotopic curve shortening* (HCS), starts with a closed curve (which might self-intersect) in the presence of a set $P \subset \mathbb{R}^2$ of point obstacles, and evolves in discrete steps, where each step consists of (1) taking shortcuts around the obstacles, and (2) reducing the curve to its shortest homotopic equivalent.

We find experimentally that, if the initial curve is held fixed and P is chosen to be either a very fine regular grid or a uniformly random point set, then HCS behaves at the limit like the affine curve-shortening flow (ACSF). This connection between HCS and ACSF generalizes the link between "grid peeling" and the ACSF observed by Eppstein et al. (2017), which applied only to convex curves, and which was studied only for regular grids.

We prove that HCS satisfies some properties analogous to those of ACSF: HCS is invariant under affine transformations, preserves convexity, and does not increase the total absolute curvature. Furthermore, the number of self-intersections of a curve, or intersections between two curves (appropriately defined), does not increase. Finally, if the initial curve is simple, then the number of inflection points (appropriately defined) does not increase.

2012 ACM Subject Classification Theory of computation → Computational geometry; Mathematics of computing → Geometric topology

Keywords and phrases affine curve-shortening flow, shortest homotopic path, integer grid, convex-layer decomposition

Digital Object Identifier 10.4230/LIPIcs.SoCG.2020.12

Related Version A full version of this paper is available at `https://arxiv.org/abs/1909.00263`.

Supplementary Material Our code is available at `https://github.com/savvakumov/`.

Funding *Sergey Avvakumov*: Supported by the Austrian Science Fund (FWF), Project P31312-N35.

Acknowledgements Thanks to Arseniy Akopyan, Imre Bárány, Jeff Erickson, Radoslav Fulek, Jeremy Schiff, Arkadiy Skopenkov, and Peter Synak for useful discussions. Thanks also to the referees for their useful comments.

[1] Corresponding author

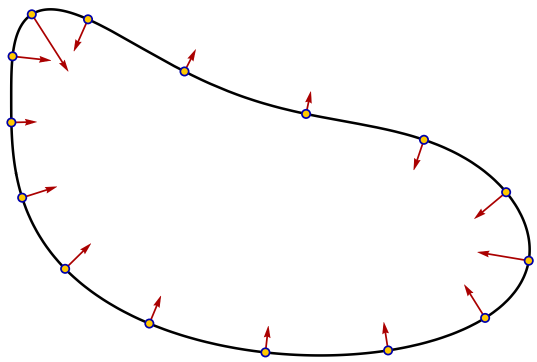

Figure 1 Affine curve-shortening flow. The arrows indicate the instantaneous velocity of different points along the curve at the shown time moment.

1 Introduction

Let $\mathbb{S}^1$ be the unit circle. In this paper we call a piecewise-smooth function $\gamma : [0,1] \to \mathbb{R}^2$ a *path*, and a piecewise-smooth function $\gamma : \mathbb{S}^1 \to \mathbb{R}^2$ a *closed curve*, or simply a *curve*. If γ is injective then the curve or path is said to be *simple*. We say that two paths or curves γ, δ are ε*-close* to each other if their Fréchet distance is at most ε, i.e. if they can be re-parametrized such that for every t, the Euclidean distance between the points $\gamma(t)$, $\delta(t)$ is at most ε.

1.1 Shortest Homotopic Curves

Let P be a finite set of points in the plane, which we regard as obstacles. Two curves γ, δ that avoid P are said to be *homotopic* if there exists a way to continuously transform γ into δ while avoiding P at all times. And two paths γ, δ that avoid P (except possibly at the endpoints) and satisfy $\gamma(0) = \delta(0)$, $\gamma(1) = \delta(1)$ are said to be *homotopic* if there exists a way to continuously transform γ into δ, without moving their endpoints, while avoiding P at all times (except possibly at the endpoints). We extend these definitions to the case where γ avoids obstacles but δ does not, by requiring the continuous transformation of γ into δ to avoid obstacles at all times except possibly at the last moment.

Then, for every curve (resp. path) γ in the presence of obstacles there exists a unique shortest curve (resp. path) δ that is homotopic to γ. The problem of computing the shortest path or curve homotopic to a given piecewise-linear path or curve, under the presence of polygonal or point obstacles, has been studied extensively. A simple and efficient algorithm for this task is the so-called "funnel algorithm" [12, 26, 27]. See also [7, 9, 18].

1.2 The Affine Curve-Shortening Flow

In the *affine curve-shortening flow*, a smooth curve $\gamma \subset \mathbb{R}^2$ varies with time in the following way. At each moment in time, each point of γ moves perpendicularly to the curve, towards its local center of curvature, with instantaneous velocity $r^{-1/3}$, where r is that point's radius of curvature at that time. See Figure 1.

The ACSF was first studied by Alvarez et al. [3] and Sapiro and Tannenbaum [28]. It differs from the more usual *curve-shortening flow* (CSF) [10, 14], in which each point is given instantaneous velocity r^{-1}. Unlike the CSF, the ACSF is invariant under affine transformations: Applying an affine transformation to a curve, and then performing the

ACSF, gives the same results (after rescaling the time parameter appropriately) as performing the ACSF and then applying the affine transformation to the shortened curves. Moreover, if the affine transformation preserves area, then the time scale is unaffected.

The ACSF was originally applied in computer vision, as a way of smoothing object boundaries [10] and of computing shape descriptors that are insensitive to the distortions caused by changes of viewpoint.

Properties of the CSF and ACSF for Simple Curves. Under either the CSF or the ACSF, a simple curve remains simple, and its length decreases strictly with time ([14], [28], resp.). Furthermore, a pair of disjoint curves, run simultaneously, remain disjoint at all times ([29], [5], resp.). More generally, the number of intersections between two curves never increases ([4], [5], resp.). The total absolute curvature[2] of a curve decreases strictly with time and tends to 2π ([21, 22], [5], resp.). The number of inflection points of a simple curve does not increase with time ([4], [5], resp.).

Under the CSF, a simple curve eventually becomes convex and then converges to a circle as it collapses to a point [21, 22]. Correspondingly, under the ACSF, a simple curve becomes convex and then converges to an ellipse as it collapses to a point [5].

Self-Intersecting Curves. When the initial curve is not simple, a self-intersection might collapse and form a cusp with infinite curvature. For the CSF, it has been shown that, as long as the initial curve satisfies some natural conditions, it is possible with some care to continue the flow past the singularity [2, 4]. Angenent [4] generalized these results to a wide range of flows, but unfortunately the ACSF is not included in this range [5]. Hence, no rigorous results have been obtained for self-intersecting curves under the ACSF. Still, ACSF computer simulations can be run on curves that have self-intersections or singularities with little difficulty.

1.3 Relation to Grid Peeling

Let P be a finite set of points in the plane. The *convex-layer decomposition* (also called the *onion decomposition*) of P is the partition of P into sets $P_1, P_2, P_3, \ldots$ obtained as follows: Let $Q_0 = P$. Then, for each $i \geq 1$ for which $Q_{i-1} \neq \emptyset$, let P_i be the set of vertices of the convex hull of Q_{i-1}, and let $Q_i = Q_{i-1} \setminus P_i$. In other words, we repeatedly remove from P the set of vertices of its convex hull. See [6, 13, 16, 17].

Eppstein et al. [19], following Har-Peled and Lidický [24], studied *grid peeling*, which is the convex-layer decomposition of subsets of the integer grid $\mathbb{Z}^2$. Eppstein et al. found an experimental connection between ACSF for convex curves and grid peeling. Specifically, let γ be a fixed convex curve. Let n be large, let $(\mathbb{Z}/n)^2$ be the uniform grid with spacing $1/n$, and let $P_n(\gamma)$ be the set of points of $(\mathbb{Z}/n)^2$ that are contained in the region bounded by γ. Then, as $n \to \infty$, the convex-layer decomposition of $P_n(\gamma)$ seems experimentally to converge to the ACSF evolution of γ, after the time scale is adjusted appropriately. They formulated this connection precisely in the form of a conjecture. They also raised the question whether there is a way to generalize the grid peeling process so as to approximate ACSF for non-convex curves as well.

Dalal [16] studied the convex-layer decomposition of point sets chosen uniformly and independently at random from a fixed convex domain, in the plane as well as in $\mathbb{R}^d$.

[2] Let $\gamma : [0,1] \to \mathbb{R}^2$ be a smooth closed curve, and let $\alpha : [0,1] \to \mathbb{S}^1$ be continuous such that $\alpha(s)$ is tangent to $\gamma(s)$ for all $s \in [0,1]$. Then the *total absolute curvature* of γ is the total distance traversed by $\alpha(s)$ in $\mathbb{S}^1$ as s goes from 0 to 1. If γ is convex then its total absolute curvature is exactly 2π; otherwise, it is larger than 2π.

1.4 Our Contribution

In this paper we describe a generalization of the convex-layer decomposition to non-convex, and even non-simple, curves. We call our process *homotopic curve shortening*, or HCS. Under HCS, an initial curve evolves in discrete steps in the presence of point obstacles. We find that, if the obstacles form a uniform grid, then HCS shares the same experimental connection to ACSF that grid peeling does. Hence, HCS is the desired generalization sought by Eppstein et al. [19]. We also find that the same experimental connection between ACSF and HCS (and in particular, between ACSF and the convex-layer decomposition) holds when the obstacles are distributed uniformly at random, with the sole difference being in the constant of proportionality.

Although the experimental connection between HCS and ACSF seems hard to prove, we do prove that HCS satisfies some simple properties analogous to those of ACSF: HCS is invariant under affine transformations, preserves convexity, and does not increase the total absolute curvature. Furthermore, the number of self-intersections of a curve, or intersections between two curves (appropriately defined), does not increase. Finally, if the initial curve is simple, then the number of inflection points (appropriately defined) does not increase.

Organization of This Paper. In Section 2 we describe homotopic curve shortening (HCS), our generalization of the convex-layer decomposition. In Section 3 we present our conjectured connection between ACSF and HCS, as well as experimental evidence supporting this connection. In Section 4 we state our theoretical results, to the effect that HCS satisfies some properties analogous to those of ACSF. In Section 5 we sketch the proofs the results stated in Section 4. Missing details can be found in the full version in the arXiv.

2 Homotopic Curve Shortening

Let P be a finite set of obstacle points. A *P-curve* (resp. *P-path*) is a curve (resp. path) that is composed of straight-line segments, where each segment starts and ends at obstacle points.

Homotopic curve shortening (HCS) is a discrete process that starts with an initial P-curve γ_0 (which might self-intersect), and at each step, the current P-curve γ_n is turned into a new P-curve $\gamma_{n+1} = \mathrm{HCS}_P(\gamma_n)$.

The definition of $\gamma' = \mathrm{HCS}_P(\gamma)$ for a given P-curve γ is as follows. Let $(p_0, \ldots, p_{m-1})$ be the circular list of obstacle points visited by γ. Call p_i *nailed* if γ goes straight through p_i, i.e. if $\angle p_{i-1}p_ip_{i+1} = \pi$.[3] Let $(q_0, \ldots, q_{k-1})$ be the circular list of nailed vertices of γ. Suppose first that $k \geq 1$. Then γ' is obtained through the following three substeps:

1. *Splitting.* We split γ into k P-paths $\delta_0, \ldots, \delta_{k-1}$ at the nailed vertices, where each δ_i goes from q_i to q_{i+1}.
2. *Shortcutting.* For each non-endpoint vertex p_i of each δ_i, we make the curve avoid p_i by taking a small shortcut. Specifically, let $\varepsilon > 0$ be sufficiently small, and let C_{p_i} be a circle of radius ε centered at p_i. Let e_i be the segment $p_{i-1}p_i$ of δ_i. Let $x_i = e_i \cap C_{p_i}$ and $y_i = e_{i+1} \cap C_{p_i}$. Then we make the path go straight from x_i to y_i instead of through p_i. Call the resulting path ρ_i, and let ρ be the curve obtained by concatenating all the paths ρ_i.
3. *Shortening.* Each ρ_i in ρ is replaced by the shortest P-path homotopic to it. The resulting curve is γ'.

[3] All indices in circular sequences are modulo the length of the sequence.

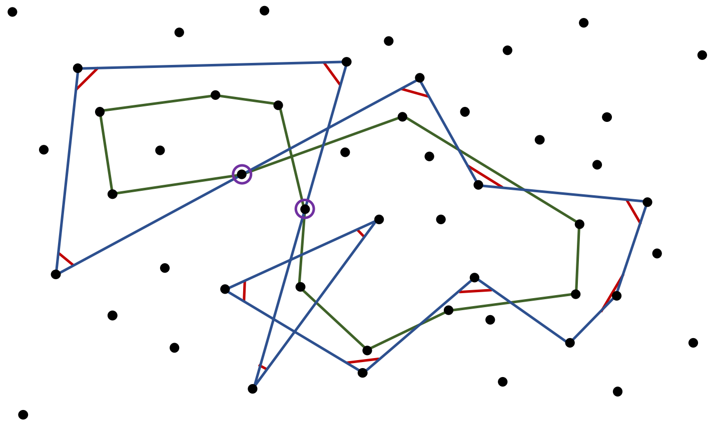

Figure 2 Computation of a single step of homotopic curve shortening: Given a P-curve γ (blue), we first identify its nailed vertices (purple). In this case, the two nailed vertices split γ into two paths δ_0, δ_1. In each δ_i we take a small shortcut around each intermediate vertex (red). Then we replace each δ_i by the shortest path homotopic to it, obtaining the new P-curve $\gamma' = \mathrm{HCS}_P(\gamma)$ (green).

If γ has no nailed vertices ($k = 0$) then γ' is obtained by performing the shortcutting and shortening steps on the single closed curve γ. Figure 2 illustrates one HCS step on a sample curve.

The process terminates when the curve collapses to a point. This will certainly happen after a finite number of steps, since at each step the curve gets strictly shorter, and there is a finite number of distinct P-curves of at most a certain length.

HCS for Convex Curves. If the initial curve γ_0 is the boundary of the convex hull of P, then the HCS evolution of γ_0 is equivalent to the convex-layer decomposition of P. Namely, for every $i \geq 0$, the curve γ_i is the boundary of a convex polygon, and the set of vertices of this polygon equals the $(i+1)$-st convex layer of P. See Section 4 below.

3 Experimental Connection Between ACSF and HCS

Our experiments show that HCS, using $P = (\mathbb{Z}/n)^2$ as the obstacle set, approximates ACSF at the limit as $n \to \infty$, just as grid peeling approximates ACSF for convex curves. The connection between the two processes is formalized in the following conjecture, which generalizes Conjecture 1 of [19].

▶ **Conjecture 1.** *There exists a constant $c_{\mathrm{g}} \approx 1.6$ such that the following is true: Let δ be a piecewise-smooth initial curve. Fix a time $t > 0$, and let $\delta' = \delta(t)$ under ACSF. For a fixed n, let γ_0 be the shortest curve homotopic to δ under obstacle set $P_n = (\mathbb{Z}/n)^2$. Let $m = c_{\mathrm{g}} t n^{4/3}$, and let $\gamma_m = \mathrm{HCS}_P^{(m)}(\gamma_0)$ be the result of m iterations of HCS starting with γ_0. Then, as $n \to \infty$, the Fréchet distance between γ_m and δ' tends to 0.*

Furthermore, we find that the connection between ACSF and HCS also holds if the uniform grid $(\mathbb{Z}/n)^2$ is replaced by a random point set, though with a different constant of time proportionality.

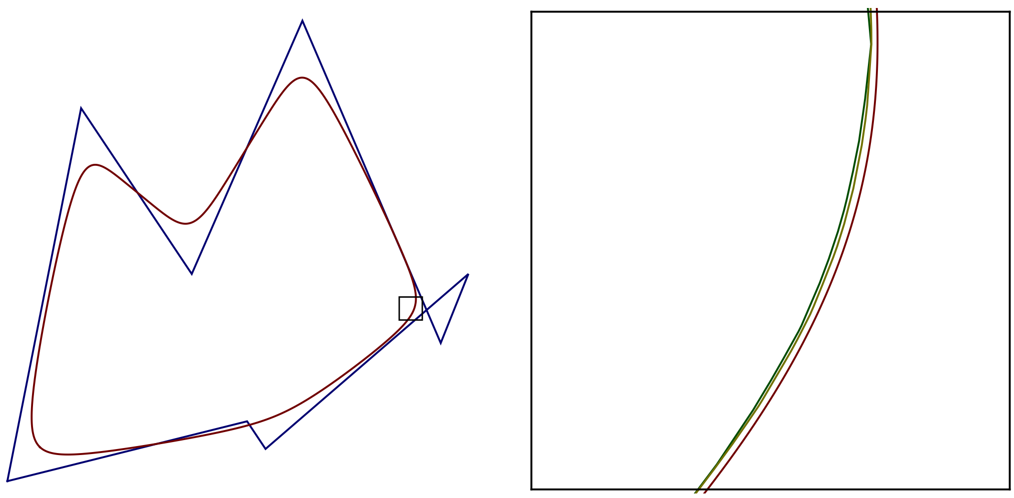

Figure 3 Left: Initial curve Δ (blue) and simulated ACSF result after the curve's length reduced to 70% of its original length (red). Right: Comparison between ACSF approximation (red), HCS with $n = 10^7$ uniform-grid obstacles (green), and HCS with $n = 10^7$ random obstacles (yellow) on a small portion of the curve.

▶ **Conjecture 2.** *There exists a constant* $c_r \approx 1.3$ *such that the following is true: Let* δ *be a piecewise-smooth initial curve, contained in a convex region* R *of area* A*. Fix a time* $t > 0$*, and let* $\delta' = \delta(t)$ *under ACSF. For a fixed* n*, let* P *be a set of* An^2 *obstacle points chosen uniformly and independently at random from* R*. Let* γ_0 *be the shortest curve homotopic to* δ *under obstacle set* P*. Let* $m = c_r t n^{4/3}$*, and let* $\gamma_m = \mathrm{HCS}_P^{(m)}(\gamma_0)$ *be the result of* m *iterations of HCS starting with* γ_0*. Then, as* $n \to \infty$*, the Fréchet distance between* γ_m *and* δ' *is almost surely smaller than* ε*, for some* $\varepsilon = \varepsilon(n)$ *that tends to* 0 *with* n*.*

3.1 Experiments

We tested Conjectures 1 and 2 on a variety of test curves. We found that for all our test curves, the result of HCS does seem to converge to the result of ACSF as $n \to \infty$, both for grid and for random obstacle sets.

We illustrate our experiments on the piecewise-liner curve Δ having vertices $(0, 0)$, $(0.16, 0.81)$, $(0.4, 0.45)$, $(0.64, 1)$, $(0.94, 0.3)$, $(1, 0.45)$, $(0.56, 0.07)$, $(0.52, 0.13)$. We approximated ACSF using an approach similar to the one in [19]. We ran our ACSF simulation on Δ until we obtained a curve Δ' whose length equals 70% of the original length of Δ. See Figure 3 (left). This happened at $t^* \approx 0.0266$. By this time, the self-intersection and an inflection point of the curve have disappeared.

Then we introduced in the unit square $[0, 1]^2 \supset \Delta$ a set P of n obstacle points, where P is either a uniform grid (i.e. a $\sqrt{n} \times \sqrt{n}$ grid) G_n, or a random set R_n. For each case, we initially snapped each vertex of Δ to its closest point in P, obtaining a P-curve, and then we ran HCS until the length of the curve shrank to 70% of its original length, obtaining a new curve $\Delta'' = \Delta''(P)$. We did this for several values of n. For each case, we computed $h(\Delta', \Delta'')$, where $h(\gamma_1, \gamma_2)$ for piecewise-linear curves γ_1, γ_2 is defined as the maximum distance between a vertex of one curve and the closest point on the other curve. (For "nice" curves as ours, there is no significant difference, if at all, between this distance h and either the Hausdorff or the Fréchet distance between the two curves.)

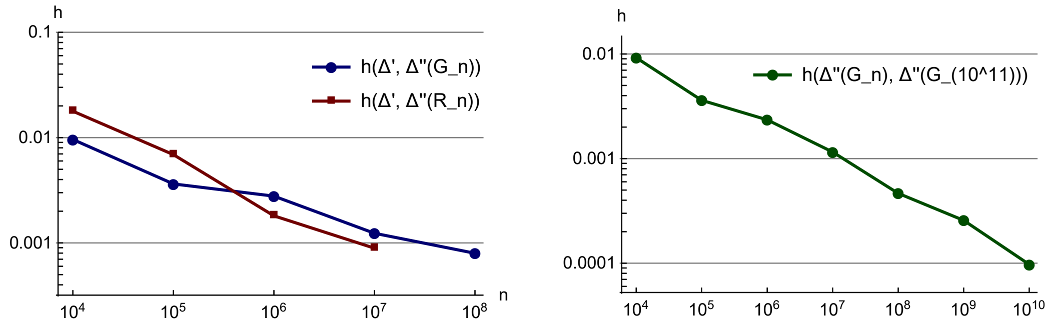

Figure 4 Left: Distance between ACSF approximation and HCS with uniform-grid obstacles (blue curve) or random obstacles (red curve, average of 5 trials), for increasing values of n, the number of obstacles. Right: Distance between HCS with uniform-grid obstacles for $n = 10^4, \dots, 10^{10}$ and with $n = 10^{11}$.

Table 1 Approximations of the constants c_g and c_r given by the experiments.

n	iterations with G_n	c_g	avg. iterations with R_n	c_r
10^4	20	1.616	15.6	1.261
10^5	93	1.619	75.2	1.309
10^6	434	1.628	351.2	1.317
10^7	2006	1.621	1628.6	1.316
10^8	9266	1.613		

For random obstacles, we conducted this experiment for $n = 10^4, 10^5, 10^6, 10^7$, taking the average of 5 samples for each value of n. Our random-obstacle program is limited by memory rather than by time, since it stores all the obstacle points in memory. For uniform-grid obstacles, we conducted this experiment also for $n = 10^8$. After this point, our ACSF approximation Δ' does not seem to be accurate enough for reliable comparisons. The results are shown in Figure 4 (left).

We also checked whether the relation between the ACSF time t^* and the number of HCS iterations m behaves as predicted by Conjectures 1 and 2. For this purpose, we computed $c = m/(t^* n^{2/3})$ for each case, and checked whether c is roughly constant. The results are shown in Table 1.

As we can see, Conjectures 1 and 2 are well supported by the experiments.

Finally, we measured the rate of convergence of the uniform-grid HCS to its limit shape as $n \to \infty$. To this end, we computed $h(\Delta''(G_n), \Delta''(G_m))$ for $n \in \{10^4, 10^5, \dots, 10^{10}\}$ and $m = 10^{11}$. See Figure 4 (right). As we can see, increasing n by a factor of 10 has the effect of multiplying the distance by roughly a factor of 0.47.

See the full version in the arXiv for some implementation details of our ACSF and HCS simulations.

4 Properties of Homotopic Curve Shortening

We now prove that HCS satisfies some properties analogous to those of ACSF.

▶ **Theorem 3.** *HCS is invariant under affine transformations. Namely, if P is a set of obstacle points, γ is a P-curve, and T is a non-degenerate affine transformation, then* $T(\mathrm{HCS}_P(\gamma)) = \mathrm{HCS}_{T(P)}(T(\gamma))$.

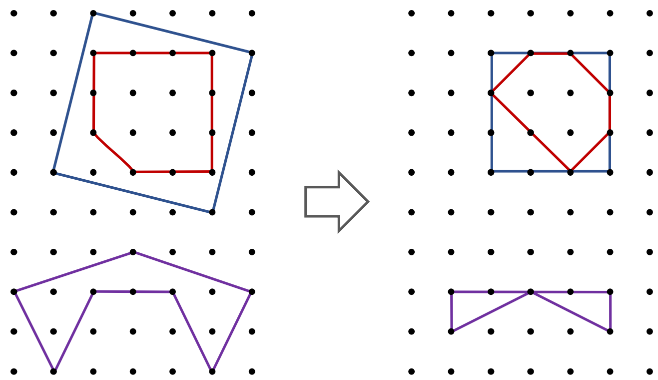

Figure 5 HCS might cause disjoint curves to intersect, or a simple curve to self-intersect.

In particular, if T is a grid-preserving affine transformation, meaning that T maps $(\mathbb{Z}/n)^2$ injectively to itself, then the HCS evolution using $P = (\mathbb{Z}/n)^2$ (as in Conjecture 1) is unaffected by T. Hence, HCS on uniform-grid obstacles is invariant under a certain subset of the area-preserving affine transformations, just as in grid peeling [19].

Also, if T is an area-preserving affine transformation, then the probability distribution of random sets P in the convex region R of Conjecture 2 stays unaffected after applying T to R.

▶ **Theorem 4.** *Let γ be a simple P-curve, and let $\gamma' = \mathrm{HCS}_P(\gamma)$. If γ is the boundary of a convex polygon, then so is γ'. Hence, under HCS, once a curve becomes the boundary of a convex polygon, it stays that way.*

The *total absolute curvature* of a piecewise-linear curve γ with vertices $(p_0, \dots, p_{m-1})$ is the sum of the exterior angles $\sum_{i=0}^{m-1}(\pi - |\angle p_{i-1}p_ip_{i+1}|)$. It equals 2π if γ is the boundary of a convex polygon, and it is larger than 2π otherwise.

▶ **Theorem 5.** *Let γ be a P-curve, and let $\gamma' = \mathrm{HCS}_P(\gamma)$. Let α, α' be the total absolute curvature of γ, γ', respectively. Then $\alpha \geq \alpha'$. Hence, under HCS, the total absolute curvature of a curve never increases.*

If γ, δ are disjoint P-curves, then $\mathrm{HCS}_P(\gamma), \mathrm{HCS}_P(\delta)$ are not necessarily disjoint. Similarly, if γ is a simple P-curve, then $\mathrm{HCS}_P(\gamma)$ is not necessarily simple. See Figure 5.

Curves γ, δ are called *disjoinable* if they can be made into disjoint curves by peforming on them an arbitrarily small perturbation. Similarly, a curve γ is called *self-disjoinable* if it can be turned into a simple curve by an arbitrarily small perturbation. Note that if γ is self-disjoinable then γ, γ are disjoinable, though the reverse is not necessarily true: Consider for example a curve γ that makes two complete clockwise turns around the unit circle.

Akitaya et al. [1] recently found an $O(n \log n)$-time algorithm for deciding whether a given mapping of a graph into the plane is a so-called *weak embedding*. This algorithm can decide, in particular, whether a given curve is self-disjoinable.

An intersection between two curves, or between two portions of one curve, is called *transversal*, if at the point of intersection both curves are differentiable and their normal vectors are not parallel at that point. If all intersections between curves γ_1 and γ_2 are transversal, then we say that γ_1, γ_2 are themselves *transversal*. Similarly, if all self-intersections of γ are transversal, then we say that γ is *self-transversal*. (Transversal and self-transversal curves are sometimes called *generic*, see e.g. [11].)

If γ is self-transversal, we denote by $\chi(\gamma)$ the number of self-intersections of γ.[4] If γ is not self-transversal, then we define $\chi(\gamma)$ as the minimum of $\chi(\widehat{\gamma})$ among all self-transversal curves $\widehat{\gamma}$ that are ε-close to γ, for all small enough $\varepsilon > 0$. Hence, $\chi(\gamma) = 0$ if and only if γ is self-disjoinable. We define similarly the number of intersections $\chi(\gamma_1, \gamma_2)$ between two curves. Then, γ_1 and γ_2 are disjoinable if and only if $\chi(\gamma_1, \gamma_2) = 0$. Fulek and Tóth recently proved that the problem of computing $\chi(\gamma)$ is NP-hard [20].

▶ **Theorem 6.** *Let γ be a P-curve, and let $\gamma' = \mathrm{HCS}_P(\gamma)$. Then their self-intersection numbers satisfy $\chi(\gamma') \le \chi(\gamma)$. Let δ be another P-curve, and let $\delta' = \mathrm{HCS}_P(\delta)$. Then their intersection numbers satisfy $\chi(\gamma', \delta') \le \chi(\gamma, \delta)$. In particular, if γ is self-disjoinable, so is γ', and if γ, δ are disjoinable, then so are γ', δ'. Hence, under HCS, the intersection and self-intersection numbers never increase.*

With the technique of Theorem 6 we can obtain an upper bound on the number of iterations of HCS:

▶ **Theorem 7.** *If $|P| = n$ then the HCS process starting with any P-curve ends in at most $n/2$ iterations. If $P = \{1, 2, \dots, \sqrt{n}\}^2$ then the process ends in at most $O(n^{2/3})$ iterations. If P is uniformly and independently chosen at random inside a fixed convex domain, then the expected number of iterations is $O(n^{2/3})$.*

We say that an obstacle set P is in *general position* if no three points of P lie on a line. Note that if P is in general position then there are no nailed vertices in HCS.

▶ **Theorem 8.** *Let P be an obstacle set in general position. Let γ be a simple P-curve. Then $\mathrm{HCS}_P(\gamma)$ is also simple. Let γ_1, γ_2 be disjoint P-curves. Then $\mathrm{HCS}_P(\gamma_1), \mathrm{HCS}_P(\gamma_2)$ are also disjoint. Hence, under HCS with obstacles in general position, a simple curve stays simple, and a pair of disjoint curves stay disjoint.*

Let γ be a simple piecewise-linear curve with vertices $(v_0, \dots, v_{n-1})$. Assume that the sequence of vertices is minimal, meaning no v_{i-1}, v_i, v_{i+1} lie on a straight line. An *inflection edge* of γ is an edge $v_i v_{i+1}$ such that the previous and next vertices v_{i-1}, v_{i+2} lie on opposite sides of the line through v_i, v_{i+1}. Let $\varphi(\gamma)$ be the number of inflection edges of γ. Note that $\varphi(\gamma)$ is always even, since every inflection edge lies either after a sequence of clockwise vertices and before a sequence of counterclockwise vertices, or vice versa.

If γ is not simple but self-disjoinable, then we define $\varphi(\gamma)$ as the minimum of $\varphi(\gamma')$ over all simple piecewise-linear curves γ' that are ε-close to γ, for all sufficiently small $\varepsilon > 0$. (Note that for a given γ there might exist different curves γ' with different values of $\varphi(\gamma')$. For example, if γ goes from a point p to a point q and back n times, then γ' could be a spiral with just two inflection edges, or a double zig-zag with $2n - 2$ inflection edges.)

▶ **Theorem 9.** *Let γ be self-disjoinable, and let $\gamma' = \mathrm{HCS}_P(\gamma)$. Then their inflection-edge numbers satisfy $\varphi(\gamma') \le \varphi(\gamma)$. Hence, under HCS on a self-disjoinable curve, the curve's number of inflection edges never increases.*

5 Proofs

In order to prove Theorems 3–9, we rely on two different approaches for computing shortest homotopic curves. The first approach uses a triangulation of the ambient space, while the second aproach consists of repeatedly releasing unstable vertices. We start by describing these two approaches in detail.

[4] A self-intersection in a curve $\gamma : \mathbb{S}^1 \to \mathbb{R}^2$ is a pair $s \neq t$ such that $\gamma(s) = \gamma(t)$. Hence, if γ passes k times through a certain point, that counts as $\binom{k}{2}$ self-intersections.

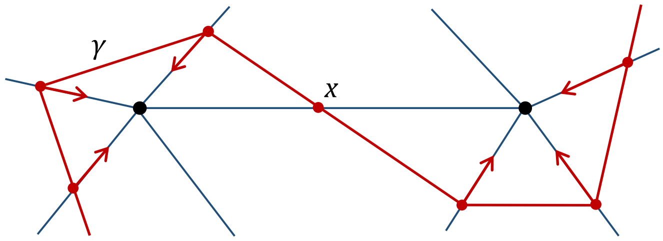

Figure 6 In the shortest curve homotopic to γ, the position of the point x is not uniquely defined.

5.1 Triangulations

Let P be a finite set of point obstacles, and let γ be a piecewise-smooth curve avoiding P. Assume without loss of generality that γ is contained in the convex hull of P (by adding points outside the convex hull of γ if necessary). Let $\mathcal{T}$ be a triangulation of the convex hull of P using the points of P as vertices.

We can assume without loss of generality that the curve γ intersects each triangle edge transversally. Let $\mathcal{E} = \mathcal{E}(\gamma)$ be the circular sequence of triangle edges intersected by γ. Then a piecewise-smooth homotopic change of γ can only have two possible types of effects on $\mathcal{E}$: Either an adjacent pair ee is inserted somewhere in the sequence, or an existing such pair is deleted. Hence, two curves γ, γ' are homotopic if and only if their corresponding edge sequences $\mathcal{E}(\gamma)$, $\mathcal{E}(\gamma')$ are *equivalent*, in the sense that they can be transformed into one another by a sequence of operations of these two types.

Call an edge sequence $\mathcal{E}(\gamma)$ *reduced* if it contains no adjacent pair ee. Then every edge sequence is equivalent to a unique reduced sequence. (Proof sketch: Supposing for a contradiction that there exist two distinct equivalent reduced sequences S_1, S_2, consider a transformation of S_1 into S_2 that uses the minimum possible number of deletions, and among those, consider one in which the first deletion is done as early as possible. Then it is easy to arrive at a contradiction.)

Hence, in order to compute the shortest curve homotopic to γ, we first compute $\mathcal{E}(\gamma)$, then we reduce this sequence by repeatedly removing adjacent pairs, obtaining a reduced sequence $\mathcal{E}'$, then we place a point $x(e)$ on each $e \in \mathcal{E}'$, and then we slide the points $x(e)$ along their edges so as to minimize the length of the curve. This last step can be done by the above-mentioned "funnel algorithm", the details of which we omit.

See the full version of this paper for a proof that there is always a unique shortest curve. Note that, even though the shortest curve is always unique, the final positions of the points $x(e)$ are not necessarily unique. This can happen if a triangulation edge is an edge of the final curve. See Figure 6.

5.2 The Vertex Release Algorithm

We now present another simple algorithm for the shortest homotopic curve problem. This algorithm is not mentioned in any previous publication that we are aware of, but it is similar in spirit to well-known algorithms, in particular to the funnel algorithm.

As a warm-up, let us first consider the case in which the obstacles are are not single points but rather polygons. Let γ be a curve that avoids all the obstacles. Call a vertex v of γ *unstable* if v does not lie on any obstacle, or if v lies on the boundary of an obstacle T, but T lies locally on the side of γ at which the angle is larger than π. If v is unstable, then the process of *releasing* v is as follows: Let u and w be the previous and next vertices

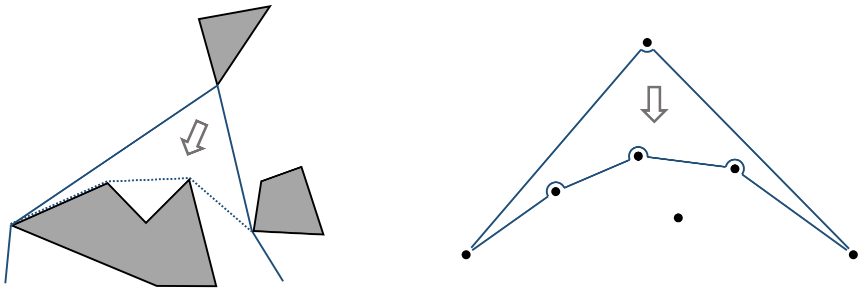

Figure 7 Releasing an unstable vertex in the presence of polygonal obstacles (left) or point obstacles (right).

of γ. Suppose first that $u \neq w$. Let Δ be the triangle uvw, and let S be the set of obstacle vertices that lie inside Δ. Let $u, z_1, \dots, z_k, w$ be the vertices of the convex hull of $S \setminus \{v\}$ in order. Then we replace v by $z_1, \dots, z_k$ in γ. The new vertices $z_1, \dots, z_k$ are necessarily stable, but u and w might change from stable to unstable or vice versa. If $u = w$ then we simply remove v and w from γ. See Figure 7 (left).

Then the algorithm consists of releasing unstable vertices one by one, in an arbitrary order, until no more unstable vertices remain.

If there are also point obstacles, then the algorithm becomes slightly more complicated. For each curve vertex v that lies on a point obstacle, we need to remember the corresponding signed angle α_v that the curve turns around the obstacle, since this angle could be larger than 2π in absolute value. The angle α_v is always congruent modulo 2π to $\angle uvw$, where u and w are the previous and next vertices. A vertex v is unstable if and only if $|\alpha_v| < \pi$.

Whenever we release an unstable vertex v preceded by u and followed by w, we proceed as described above, and we update the angles as follows (see Figure 7, right):

- If $u \neq w$ then we give to each new vertex z_i the unique appropriate angle that has the opposite sign of α_v and satisfies $\pi \leq |\alpha_{z_i}| < 2\pi$. We then update the angles α_u and α_w as follows: Denote $z_0 = u$ and $z_{k+1} = w$ (in order to handle properly the case $k = 0$). We add to α_u the angle $\angle vuz_1$, and we add to α_w the angle $\angle z_k wv$.
- If $u = w$ then we update α_u by adding to it the angle α_w.

For the proof of correctness of the vertex release algorithm, see the full version of this paper.

5.3 Proof of Theorems 3–5

Theorems 3–5 follow easily from the vertex-release algorithm.

Proof of Theorem 3. The claim follows from the fact that shortest homotopic curves and paths are invariant under affine transformations. Namely, let γ be a curve or path in the presence of obstacle points P, let δ be the shortest curve or path homotopic to γ, and let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a non-degenerate affine transformation. Then the shortest curve or path homotopic to $T(\gamma)$ in the presence of $T(P)$ is $T(\delta)$. This, in turn, follows from the fact that T does not affect whether a vertex is stable or unstable, and furthermore, if a vertex is unstable, then it does not matter whether we first release the vertex and then apply T, or do these operations in the opposite order. ◀

Theorem 4 is also trivial, since the property of being the boundary of a convex polygon is preserved by each vertex release.

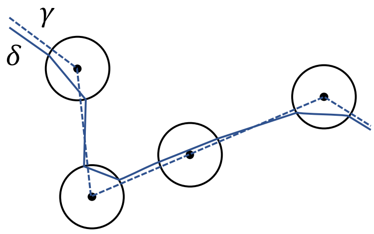

Figure 8 A P-curve γ and a corresponding type-2 curve δ.

Proof of Theorem 5. Given a curve γ with vertices $(p_0, \dots, p_{m-1})$, let $v_i \in \mathbb{S}^1$ be the unit vector parallel to $\overrightarrow{p_i p_{i+1}}$ for each i . Call a tour of $\mathbb{S}^1$ *valid* if it visits the vectors $v_0, v_1, \dots, v_{m-1}, v_0$ in this order. Then the total absolute curvature of γ equals the length of the shortest valid tour of $\mathbb{S}^1$.

Now let γ be a given P-curve, and let $\gamma' = \mathrm{HCS}_P(\gamma)$. Recall that γ' is obtained from γ by a series of vertex releases. Each vertex release replaces two adjacent vectors $v_i, v_{i+1} \in \mathbb{S}^1$ by a certain number $k \geq 1$ of vectors $w_1, \dots, w_k$ lying between them, in this order. Hence, the shortest valid tour of $\mathbb{S}^1$ for the old vector sequence goes from v_i to v_{i+1} through $w_1, \dots, w_k$, and hence this tour is also valid for the new vector sequence. ◀

5.4 Proof sketch of Theorems 6–8

The proof of Theorems 6–8 is based on the triangulation technique. Let γ be a P-curve, let $\varepsilon > 0$ be small enough, and let $\widehat{\gamma}$ be a self-transversal curve that is ε-close to γ and has the minimum possible number of self-intersections.

In order to prove Theorem 6, we proceed as follows:

1. We show that, without loss of generality, we can assume that $\widehat{\gamma}$ passes through the "correct side" of each non-nailed obstacle, as in the "shortcutting" step of HCS.
2. We modify $\widehat{\gamma}$ homotopically, by first eliminating repetitions in its edge sequence $\mathcal{E}$ and then sliding its vertices along the triangulation edges, until each vertex comes within ε of its final position as given by $\gamma' = \mathrm{HCS}_P(\gamma)$. We show that the number of self-intersections never increases in the process.

The case of two curves is similar.

In order to do the first step, we define a type of curves that are ε-close to P-curves and pass through the "correct side" of non-nailed obstacles. We call them *type-2 curves.* We also define a "snapping" operation, which transforms $\widehat{\gamma}$ into a type-2 curve without increasing its number of self-intersections.

Type-2 Curves. Let γ be a P-curve, let $(p_0, \dots, p_{k-1})$ be the circular list of obstacles visited by γ, and let $\varepsilon > 0$ be small enough. For each $p \in P$, let C_p be a circle of radius ε centered at p. For each i, let $x_i \in C_{p_i}$ be a point at distance at most ε^2 from the segment $p_{i-1}p_i$, and let $y_i \in C_{p_i}$ be a point at distance at most ε^2 from the segment $p_i p_{i+1}$. Then a type-2 curve δ corresponding to γ travels in a straight line from y_{i-1} to x_i and then in a straight line from x_i to y_i for each i. See Figure 8. We call each segment $y_{i-1}x_i$ a *long part* and each segment $x_i y_i$ a *short part.* If $\angle p_{i-1}p_i p_{i+1} \neq \pi$ and ε is chosen small enough, then p_i lies on the side of the curve at which the angle is larger than π. If $\angle p_{i-1}p_i p_{i+1} = \pi$ then the corresponding short part passes within distance ε^2 of v_i.

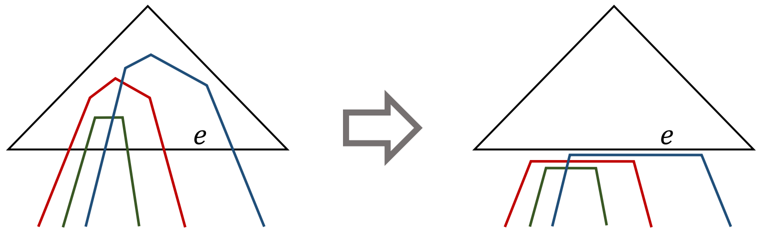

Figure 9 Reducing a curve's edge sequence without increasing its number of self-intersections. Different portions of the curve are shown in different colors.

The Snapping Operation. Let γ be a P-curve, and let $\widehat{\gamma}$ be a curve (ε^2)-close to γ. We define the type-2 curve $\mathrm{snap}(\widehat{\gamma})$ as follows. For each p_i visited by $\widehat{\gamma}$ there exists a point z_i in $\widehat{\gamma}$ that is within distance ε^2 of p_i. Let y_i be the first intersection of $\widehat{\gamma}$ with C_{p_i} that comes after z_i, and let x_i be the last intersection of $\widehat{\gamma}$ with C_{p_i} that comes before z_i. (Thus, the part of $\widehat{\gamma}$ between x_i and y_i is entirely contained in the disk bounded by C_{p_i}.) Then we let $\mathrm{snap}(\widehat{\gamma})$ be the type-2 curve that uses these points x_i, y_i for all i as vertices.

The curve $\delta = \mathrm{snap}(\widehat{\gamma})$ also is also self-transversal, and it satisfies $\chi(\delta) \le \chi(\widehat{\gamma})$. Similarly, if γ_1, γ_2 are two P-curves, and $\widehat{\gamma}_1$, $\widehat{\gamma}_2$ are transversal curves (ε^2)-close to them, respectively, such that no intersection between $\widehat{\gamma}_1$ and $\widehat{\gamma}_2$ occurs on any circle C_p, then the curves $\delta_1 = \mathrm{snap}(\widehat{\gamma}_1), \delta_2 = \mathrm{snap}(\widehat{\gamma}_2)$ are transversal and satisfy $\chi(\delta_1, \delta_2) \le \chi(\widehat{\gamma}_1, \widehat{\gamma}_2)$. See the full version of this paper.

Proof of Theorem 6. Let γ be a P-curve, let $\varepsilon > 0$ be small enough, and let $\widehat{\gamma}$ be a self-transversal curve that is ε-close to γ and has the minimum possible number of self-intersections. Fix a triangulation $\mathcal{T}$ of P. Assume without loss of generality that $\widehat{\gamma}$ does not pass through any obstacle, and that no self-intersection of γ lies on any edge of $\mathcal{T}$. Let $\eta = \mathrm{snap}(\widehat{\gamma})$. Partition η into paths $\eta_0, \dots, \eta_{k-1}$ that are ε-close to the corresponding paths $\delta_0, \dots, \delta_{k-1}$ of the HCS "splitting" step, by introducing split points as follows: For each nailed visit to an obstacle $p \in P$, we choose a split point that is within distance $O(\varepsilon)$ of p and lies on a triangle edge (where the implicit constant depends only on P).

Then we modify each η_i into a homotopic path η'_i whose edge sequence $\mathcal{E}(\eta'_i)$ is reduced. We do this without increasing the number of intersections, by repeatedly doing the following: Let e be triangulation edge such that ee appears one or more times in the sequences $\mathcal{E}(\eta'_i)$. We shortcut the corresponding paths η'_i so as to not cross e at all, instead keeping a small distance from e. We make the distance to e inversely related to the distance between the two crossing points of η'_i with e. See Figure 9.

Next, we modify each η'_i into η''_i by straightening out each part within each triangle of $\mathcal{T}$. Hence, each η''_i is determined by the position of its vertices $x(e)$ along the triangle edges e.

Finally, we slide the vertices $x(e)$ along the edges to within ε of their final positions, as given by γ'. We do this without changing the order of any pair of vertices along the same edge, unless necessary. Call the resulting paths η'''_i. Meaning, if in γ' there are several vertices along an edge that coincide, then in the paths η'''_i we place those vertices within ε of each other, conserving the order they had in η''_i. Let η'', η''' be the curves formed by concatenating the paths η''_i, η'''_i for all i, respectively. Hence, by construction, η''' is ε-close to γ'.

The number of self-intersections of η''' is not larger than that of η''. See the full version of this paper. This concludes the proof of Theorem 6 for the case of the number of self-intersections of a single curve. The case of the number of intersections of two curves is similar. ◀

Theorem 7 follows by running HCS simultaneously on the given curve γ_0 and on the boundary δ_0 of the convex hull of P. The HCS process starting with δ_0 is just the convex-layer decomposition of P, so we can apply the known bounds on the number of convex layers. Denote $\gamma_{i+1} = \mathrm{HCS}_P(\gamma_i)$ and $\delta_{i+1} = \mathrm{HCS}_P(\delta_i)$ for all i. By the proof of Theorem 6, the two curves stay disjoinable throughout the HCS process, with δ_i bounding γ_i for all i. See the full version for more details, as well as for the proof of Theorem 8.

5.5 Proof sketch of Theorem 9

The proof of Theorem 9 (regarding the number of inflection edges) is based of the vertex-release algorithm. The basic idea is that, given a self-disjoinable curve, if the vertex releases are performed in an appropriate order, then the curve stays self-disjoinable at all times. Moreover, no vertex release increases the number of inflection edges. Along the way, we develop enough machinery to re-prove Theorem 6. The proof appears in the full version of this paper.

6 Discussion

One of the reasons continuous curve-shortening flows were introduced and studied was to overcome the shortcomings of the *Birkhoff curve-shortening process* ([8], see also e.g. [15]), specifically the fact that it might cause the number of curve intersections to increase [23, 25]. As we have shown, HCS is a discrete process that overcomes this flaw without introducing analytical difficulties, at least in the plane. It would be interesting to check whether HCS can be applied on more general surfaces.

References

1. Hugo A. Akitaya, Radoslav Fulek, and Csaba D. Tóth. Recognizing weak embeddings of graphs. In *Proc. 29th Symp. on Discrete Algorithms*, pages 274–292, 2018. doi:10.1137/1.9781611975031.20.
2. Steven J. Altschuler and Matthew A. Grayson. Shortening space curves and flow through singularities. *J. Differential Geom.*, 35(2):283–298, 1992. doi:10.4310/jdg/1214448076.
3. Luis Alvarez, Frédéric Guichard, Pierre-Luis Lions, and Jean-Michel Morel. Axioms and fundamental equations of image processing. *Arch. Rational Mech. Anal.*, 123(3):199–257, 1993. doi:10.1007/BF00375127.
4. Sigurd Angenent. Parabolic equations for curves on surfaces: Part II. Intersections, blow-up and generalized solutions. *Annals of Mathematics*, 133(1):171–215, 1991. doi:10.2307/2944327.
5. Sigurd Angenent, Guillermo Sapiro, and Allen Tannenbaum. On the affine heat equation for non-convex curves. *J. Amer. Math. Soc.*, 11(3):601–634, 1998. doi:10.1090/S0894-0347-98-00262-8.
6. Vic Barnett. The ordering of multivariate data. *J. Roy. Statist. Soc. Ser. A*, 139(3):318–355, 1976. doi:10.2307/2344839.
7. Sergei Bespamyatnikh. Computing homotopic shortest paths in the plane. *Journal of Algorithms*, 49(2):284–303, 2003. doi:10.1016/S0196-6774(03)00090-7.
8. George D. Birkhoff. Dynamical systems with two degrees of freedom. *Trans. Amer. Math. Soc.*, 18:199–300, 1917.
9. Sergio Cabello, Yuanxin Liu, Andrea Mantler, and Jack Snoeyink. Testing homotopy for paths in the plane. *Discrete & Computational Geometry*, 31(1):61–81, 2004. doi:10.1007/s00454-003-2949-y.
10. Frédéric Cao. *Geometric Curve Evolution and Image Processing*, volume 1805 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2003. doi:10.1007/b10404.

11 Hsien-Chih Chang and Jeff Erickson. Untangling planar curves. *Discrete & Computational Geometry*, 58:889–920, 2017. doi:10.1007/s00454-017-9907-6.

12 Bernard Chazelle. A theorem on polygon cutting with applications. In *Proc. 23rd Annual Symposium on Foundations of Computer Science (FOCS 1982)*, pages 339–349, 1982. doi:10.1109/SFCS.1982.58.

13 Bernard Chazelle. On the convex layers of a planar set. *IEEE Trans. Inform. Theory*, 31(4):509–517, 1985. doi:10.1109/TIT.1985.1057060.

14 Kai-Seng Chou and Xi-Ping Zhu. *The Curve Shortening Problem.* Chapman & Hall/CRC, Boca Raton, FL, 2001. doi:10.1201/9781420035704.

15 Cristopher B. Croke. Area and the length of the shortest closed geodesic. *J. Differential Geometry*, 27:1–21, 1988.

16 Ketan Dalal. Counting the onion. *Random Struct. Algor.*, 24(2):155–165, 2004. doi:10.1002/rsa.10114.

17 William F. Eddy. Convex Hull Peeling. In *COMPSTAT 1982 5th Symposium held at Toulouse 1982*, pages 42–47. Physica-Verlag, 1982. doi:10.1007/978-3-642-51461-6_4.

18 Alon Efrat, Stephen G. Kobourov, and Anna Lubiw. Computing homotopic shortest paths efficiently. *Computational Geometry*, 35(3):162–172, 2006. doi:10.1016/j.comgeo.2006.03.003.

19 David Eppstein, Sariel Har-Peled, and Gabriel Nivasch. Grid peeling and the affine curve-shortening flow. *Experimental Mathematics*, page to appear, 2018. doi:10.1080/10586458.2018.1466379.

20 Radoslav Fulek and Csaba D. Tóth. Crossing minimization in perturbed drawings. In T. Biedl and A. Kerren, editors, *Proc. 26th Symp. Graph Drawing and Network Visualization*, pages 229–241. Springer, 2018. doi:10.1007/978-3-030-04414-5_16.

21 Michael Gage and Richard S. Hamilton. The heat equation shrinking convex plane curves. *J. Differential Geom.*, 23(1):69–96, 1986. doi:10.4310/jdg/1214439902.

22 Matthew A. Grayson. The heat equation shrinks embedded plane curves to round points. *J. Differential Geom.*, 26(2):285–314, 1987. doi:10.4310/jdg/1214441371.

23 Matthew A. Grayson. Shortening embedded curves. *Annals of Mathematics*, 129(1):79–111, 1989.

24 Sariel Har-Peled and Bernard Lidický. Peeling the grid. *SIAM J. Discrete Math.*, 27(2):650–655, 2013. doi:10.1137/120892660.

25 Joel Hass and Peter Scott. Shortening curves on surfaces. *Topology*, 33:25–43, 1994. doi:10.1016/0040-9383(94)90033-7.

26 John Hershberger and Jack Snoeyink. Computing minimum length paths of a given homotopy class. *Computational Geometry*, 4(2):63–97, 1994. doi:10.1016/0925-7721(94)90010-8.

27 Der-Tsai Lee and Franco P. Preparata. Euclidean shortest paths in the presence of rectilinear barriers. *Networks*, 14(3):393–410, 1984. doi:10.1002/net.3230140304.

28 Guillermo Sapiro and Allen Tannenbaum. Affine invariant scale-space. *Int. J. Comput. Vision*, 11(1):25–44, 1993. doi:10.1007/bf01420591.

29 Brian White. Evolution of curves and surfaces by mean curvature. In *Proceedings of the International Congress of Mathematicians, Vol. I (Beijing, 2002)*, pages 525–538, 2002. URL: https://www.mathunion.org/fileadmin/ICM/Proceedings/ICM2002.1/ICM2002.1.ocr.pdf.