Isoperimetric Enclosures

Greg Aloupis[*†] Luis Barba[*‡§] Jean-Lou De Carufel[‡¶] Stefan Langerman[*‖]
Diane L. Souvaine[**]

Abstract

Let S be a set of $n > 2$ points in the plane whose convex hull has perimeter t. Given a number $P \geq t$, we study the following problem: Of all curves of perimeter P that enclose S, which is the curve that encloses the maximum area? In this paper, we give a complete characterization of this curve. We show that there are cases where this curve cannot be computed exactly and provide an $O(n \log n)$-time algorithm to obtain an approximation of this curve, with arbitrary precision, having the same combinatorial structure.

1 Introduction

In geometric optimization, we are often interested in finding an object that maximizes or minimizes an objective function subject to geometric constraints. For example, given a set S of points in the plane we can ask for the minimum enclosing circle of S [9], or for the unit circle that contains the maximum number of points of S.

In this paper, we study optimization problems involving simple planar curves of fixed perimeter, in other words involving *isoperimetric curves*. In general, *isoperimetric problems* involve optimizing a given function over a family of isoperimetric curves. The classical isoperimetric problem is to maximize the area enclosed by a curve of fixed perimeter. It was known since ancient Greece that the solution to this problem is the circle. However, a rigorous proof of this statement was obtained until the late 19th century and it is known as the *Isoperimetric Theorem.* The main issue was to prove the existence of this curve as the limit of a sequence of polygons that approximate the circle [3]. An intuitive proof of the Isoperimetric Theorem can be found in [10]. A dual theorem states that among all plane figures of equal area, the circle is the one with minimum perimeter.

The closely related Dido's problem is present in greek mythology. According to the legend, Dido was fleeing from her homeland and seeking asylum in northern Africa. She was offered a piece of land as large as she could encompass with an oxhide. After cutting the hide into one long strip, she formed a curve between two coastal points, thus claiming a large area that later came to be the city of Carthage. Dido's problem is hence known as that of maximizing the area bounded by a straight line (in her case, the coast) and a curve of fixed length. As Dido figured out, the maximum area is obtained when the curve is in the shape of a half circle with endpoints on the coast line.

Some variants of isoperimetric problems have been addressed by adding geometric constraints. Given a set S of $n > 2$ points in the plane, Bose and De Carufel [5] considered the family of isoperimetric triangles enclosing S, with the additional constraint that one angle was also fixed. They provided an algorithm to find one such triangle of maximum (or minimum) area in $O(n^2)$ time and showed how to solve a dual version of the problem. That is, among all triangles sharing the same area that enclose S and have one equal angle, compute a triangle with maximum (or minimum) perimeter.

By adding geometric constraints, we extend the isoperimetric problem as follows.

[*]Département d'Informatique, Université Libre de Bruxelles, Brussels, Belgium. aloupis.greg@gmail.com, {lbarbafl,slanger}@ulb.ac.be
[†]Chargé de recherches du F.R.S.-FNRS.
[‡]School of Computer Science, Carleton University, Ottawa, Canada. jdecaruf@cg.scs.carleton.ca
[§]Boursier FRIA, FNRS.
[¶]Research supported in part by NSERC and FQRNT
[‖]Directeur de recherches du F.R.S.-FNRS.
[**]Department of Computer Science, Tufts University, Medford, MA. dls@cs.tufts.edu. NSF Grant #CCF-0830734

Problem 1 *Let S be a set of $n > 2$ points in the plane and let $P > 0$ be a given value. Among all the curves of perimeter P that enclose S, which is the curve that encloses the maximum area?*

In Section 2, we provide a full characterization of the solution to Problem 1. Among other properties, we show that the solution to this problem is a convex curve made of circular arcs of equal radius. Therefore, to obtain a full description of this solution, we need to compute their common radius. However, as described in Section 4, it is unlikely that a closed-form expression to represent this radius always exists. That is, we show that if *Schanuel's conjecture* [2, Chapter 12] is true, then there are instances of Problem 1 for which we cannot compute the radius of these arcs. However, according to Chow [6], at present, a proof of this conjecture seems to be out of reach.

In light of this result, in Section 3 we present an algorithm to compute an approximation of the solution of Problem 1. We show how to find a convex curve made of circular arcs that encloses S such that its perimeter is arbitrarily close to P, and its enclosed area is maximum among all curves that share the same perimeter.

Finally, we discuss the dual version of this problem where we ask, among all curves of some fixed area A that enclose S, what is the curve with the minimum perimeter? We show that computing the solution of this dual problem is equivalent to that of its primal formulation. Consequently, all results presented in this paper can be used to solve the dual formulation of Problem 1.

2 Fixed Perimeter, Maximum Area

We say that a closed curve is *convex* if the region it encloses is convex. The *area* of a closed curve is the area it encloses. Given a positive number P, a P*-curve* is either a simple closed curve of perimeter P or a simple open curve of length P. Given a subset R of the plane, a (P, R)*-curve* is a P-curve that encloses R. The following lemma is a simple extension of a known equivalent statement regarding area-maximizing isoperimetric curves in the absence of R.

Lemma 1 *Given a number $P > 0$ and a set of points S, any (P, S)-curve of maximum area is convex.*

Proof. Let $\mathscr{C}$ be a (P, S)-curve of maximum area. Assume that $\mathscr{C}$ is not convex for the sake of contradiction. Therefore, there exists a line ℓ passing through two points x and y on $\mathscr{C}$ such that the open segment (x, y) is contained in the exterior of $\mathscr{C}$, and $\mathscr{C}$ is completely contained in one of the closed halfplanes defined by ℓ.

Let $\gamma_{x,y}$ be the open curve along $\mathscr{C}$ that joins x with y contained in the interior of the convex hull of $\mathscr{C}$. Let $\gamma^*_{x,y}$ be the reflection of $\gamma_{x,y}$ on the line ℓ. By replacing $\gamma_{x,y}$ by $\gamma^*_{x,y}$ on $\mathscr{C}$, we obtain a curve $\mathscr{C}^*$ with perimeter P but larger area. Moreover, the region enclosed by $\mathscr{C}$ is also enclosed by $\mathscr{C}^*$. Because $\mathscr{C}$ and $\gamma^*_{x,y}$ lie on opposite sides of ℓ, $\mathscr{C}^*$ is simple and hence it is a (P, S)-curve with larger area than $\mathscr{C}$ yielding a contradiction. Therefore, $\mathscr{C}$ must enclose a convex region of the plane. □

Because any convex curve enclosing a point set also encloses its convex hull, for the rest of this paper we assume that we are given a convex n-gon Q that we want to enclose with some convex (P, Q)-curve. This assumption adds an $\Theta(n \log n)$ preprocessing step to the final algorithm.

Given two open curves $\mathscr{C}_1$ and $\mathscr{C}_2$ sharing at least one endpoint, $\mathscr{C}_1 + \mathscr{C}_2$ denotes the curve obtained by the concatenation of the two curves. The following lemma is a direct consequence of the Isoperimetric Theorem, however, for the sake of completeness we include a full proof.

Lemma 2 *Let xy be a closed segment and let P be a positive number. If $\mathscr{C}$ is an open P-curve joining x with y such that $\mathscr{C} + xy$ is a simple closed curve of maximum area, then $\mathscr{C}$ is a circular arc.*

Proof. Assume that $\mathscr{C}$ is not a circular arc for the sake of contradiction. Let D be the unique circular arc with endpoints x and y of length P (modulo reflection). Let D^o be the circular arc such that $D + D^o$ is a circle and let α be the length of D^o. By definition, $\mathscr{C}$ is the P-curve such that $\mathscr{C} + xy$ is of maximum area. Thus, $\mathscr{C} + D^o$ is a $(P + \alpha)$-curve having a larger area than $D + D^o$ as the cap formed by D^0 and xy is present on both curves. However, by the Isoperimetric Theorem proved in Chapter X of [10], we know that $D + D^0$ is the unique $(P + \alpha)$-curve of maximum area which yields a contradiction. Thus, $\mathscr{C}$ is a circular arc equal to D. □

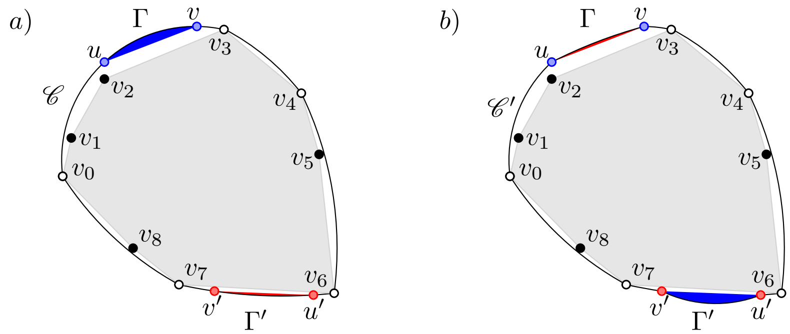

Figure 1: Illustration of the proof of Theorem 4. Polygon $Q = v_0v_1...v_8$ appears in light gray. *a*) Two arcs Γ and Γ' of $\mathscr{C}$ having different radii. Points u, v and u', v' lie on the arcs Γ and Γ', respectively, with the property that $|uv| = |u'v'|$. *b*) By swapping the arc that connects u with v along Γ with the one that connects u' with v' along Γ', we obtain a non-convex curve $\mathscr{C}'$ with the same area and perimeter.

Given a closed region R of the plane, let ∂R denote its boundary.

Lemma 3 *Let P be a positive number and let Q be a convex polygon. If $\mathscr{C}$ is the (P,Q)-curve of maximum area, then $\mathscr{C}$ is a sequence of one or more curves, each one being either a circular arc or a segment flush with ∂Q. Moreover, if $\mathscr{C}$ is not a circle, then the endpoints of these curves lie on ∂Q.*

Proof. Let z be a point on $\mathscr{C}$. Because $\mathscr{C}$ encloses Q, z lies either in the complement or on the boundary of Q. If z lies in the complement of Q, then let z^* be the closest point to z in Q and let ℓ be the perpendicular bisector of the segment zz^*. Let Π_ℓ be the open halfplane supported by ℓ that contains z. By Lemma 1, $\mathscr{C}$ is convex. Because $\mathscr{C}$ also encloses Q, ℓ intersects $\mathscr{C}$ at exactly two points x and y. Let $\gamma_z \subset \mathscr{C}$ be the curve joining x with y contained in Π_ℓ. By Lemma 2, γ_z must be a circular arc whose interior contains z. Therefore, every point of $\mathscr{C}$ is either on the interior of a circular arc if it lies in the complement of Q, or otherwise lies on the boundary of Q. □

Lemma 3 implies that no two arcs of $\mathscr{C}$, supported by different circles, can meet in the complement of Q, i.e., every arc of $\mathscr{C}$ must have its endpoints on the boundary of Q. This implies that $\mathscr{C}$ is the concatenation of circular arcs and segments flush with the boundary of Q. We proceed to prove this in the following main result of this section.

Theorem 4 *Let P be a positive number and let Q be a convex polygon. If $\mathscr{C}$ is the (P,Q)-curve of maximum area, then $\mathscr{C}$ is a convex curve that consists of a sequence of circular arcs of equal radius. Moreover, if $\mathscr{C}$ consists of at least two circular arcs, then their endpoints are vertices of Q.*

Proof. By Lemma 3 we know that $\mathscr{C}$ is a convex P-curve that consists of circular arcs, some of which can have an infinite radius. Suppose that at least two arcs of $\mathscr{C}$ have different radius (possible ∞) for the sake of contradiction. Let Γ and Γ' be two arcs of $\mathscr{C}$ with different radii; see Figure 1. Let $u, v \in \Gamma$ and $u', v' \in \Gamma'$ be points such that straight-line segments uv and $u'v'$ have equal length. Moreover, by taking this length sufficiently small, we can ensure that the segments uv and $u'v'$ do not intersect Q.

Let γ be the circular arc with endpoints u and v contained in Γ. Define γ' analogously for u', v' and Γ'. If we swap γ and γ', we obtain a (P,Q)-curve $\mathscr{C}'$ such that area$(\mathscr{C})$ = area$(\mathscr{C}')$. However, since the radii of Γ and Γ' are different, $\mathscr{C}'$ is not convex. Thus, $\mathscr{C}'$ is not optimal by Lemma 1 and hence, there is a curve with perimeter P and area larger than A, which is a contradiction to the optimality of $\mathscr{C}$. Therefore, all circular arcs of $\mathscr{C}$ must have the same radius.

By Lemma 3, we know that every arc of $\mathscr{C}$ has its two endpoints on the boundary of Q. However, if two arcs of $\mathscr{C}$ meet at a point on ∂Q that is not a vertex of Q, we would obtain a non-convex curve yielding a contradiction to Lemma 1. Therefore, if two arcs of $\mathscr{C}$ meet, they do so at a vertex of Q. □

To complete the characterization of the (P,Q)-curve of maximum area, we use the following well-known result.

Lemma 5 *(Cauchy's Arm Lemma [11, p. 110]) Let $Q = (v_0, v_1, \ldots, v_k, v_0)$ be a convex polygon where each consecutive pair of vertices is connected by an edge $e_i = v_{i-1}v_i$, and where the internal angle at vertex v_i between e_i and e_{i+1} (modulo k) is θ_i. If we remove the edge v_kv_0 from Q and increase the value of some nonempty subset of the angles θ_i while keeping the length of all remaining edges fixed and every $\theta_i \leq \pi$, then the distance between the endpoints v_0 and v_k strictly increases.*

Lemma 6 *Let $\mathscr{C} = (v_0, v_1, \ldots, v_k, v_0)$ be the (P, Q)-curve of maximum area, where v_i and v_{i+1} (modulo k) are connected by a circular arc a_i, and each a_i has the same radius. If C_i is the circle extending arc a_i, then Q is enclosed by C_i.*

Proof. If $\mathscr{C}$ is a circle, then it consists of a unique arc and the result is trivial. Therefore, assume that $\mathscr{C}$ consists of at least two circular arcs. We prove the result for C_0, however, the proof is the same for every circle extending an arc of $\mathscr{C}$. Let c_0 be the center of circle C_0. Let ℓ_0 (*resp.* ℓ_1) be the line passing through c_0 and v_0 (*resp.* v_1). Notice that ℓ_0 and ℓ_1 split the plane into four regions R_0, R_1, R_2 and R_3, labeled in clockwise order around c_0, starting with the region containing the arc a_0; see Figure 2(a). We claim that for every $0 \leq j \leq k$, the vertex v_j lies inside or on the boundary of C_0. If this claim is true, then as the curvature of C_0 and every arc along $\mathscr{C}$ is the same, $\mathscr{C}$ is contained in C_0. Moreover, because Q is enclosed by $\mathscr{C}$, we can conclude that Q is also enclosed by C_0 yielding our result. To prove that v_j lies inside C_0, we consider two cases depending on the position of c_0.

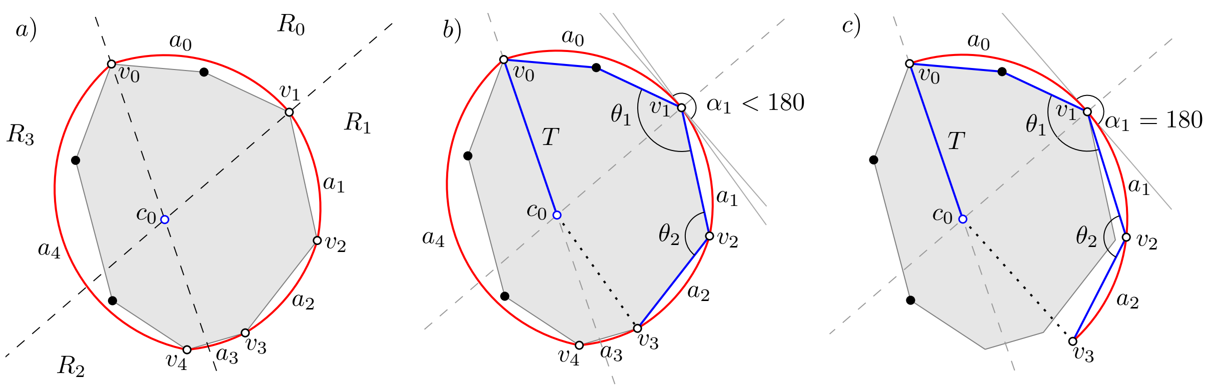

Figure 2: a) The four regions defined by the lines through v_0, c_0 and through v_1, c_0, where c_0 is the center of the circle extending the arc a_0. b) The construction to prove that v_3 is in the circle extending a_0. The path $T = (c_0, v_0, v_1, v_2, v_3)$ and the angles $\alpha_1, \alpha_2, \theta_1$ and θ_2 are depicted, all being smaller than 180 degrees. c) By opening α_1 and α_2 to 180 degrees, v_0, v_1 and v_2 become co-circular while the angles θ_1 and θ_3 increase. Points c_0 and v_3 get farther away after this deformation, which implies that v_3 was inside the circle extending a_0 before the deformation of T.

Case 1. If c_0 lies inside of Q, then assume that v_j lies in R_1. A similar proof follows if v_j lies in R_2 or R_3. Let $T = (c_0, v_0, v_1, \ldots, v_{j-1}, v_j, c_0)$ be a simple polygon enclosed by $\mathscr{C}$. Because every vertex of $\mathscr{C}$ is also a vertex of Q by Theorem 4, we know that T is a convex polygon. For $0 < i < j$, consider the two lines passing through v_i that are tangent to C_i and C_{i-1}, respectively, and let α_i be the external angle between these lines; see Figure 2(b) for an illustration. Because $\mathscr{C}$ is convex, every $\alpha_i \leq 180$ and equality holds if and only if both a_{i-1} and a_i belong to the same circle.

Remove the edge v_jc_0 from T to obtain a polygonal chain T° with endpoints c_0 and v_j. Continuously deform T° by making every angle α_i equal to 180 while keeping the length of its edges fixed, i.e., we make $v_0, v_1, \ldots, v_j$ co-circular while maintaining the distance between consecutive vertices along T°. If we assume that c_0, v_0 and v_1 remain at their original location, then every vertex v_i of T° ends up lying on the circle C_0 after this deformation. By increasing the value of each α_i to 180, the internal angle $\theta_i = \angle v_{i-1}v_iv_{i+1}$ at v_i also increases while remaining smaller than 180 degrees; see Figure 2(c). Therefore, Lemma 5 guarantees that the distance between the endpoints of T° increases after this deformation, i.e., the points c_0 and v_j get farther apart. Because every vertex of T° lies on the boundary of C_0 after the deformation, v_j was originally closer to c_0 and hence, it was enclosed by circle C_0.

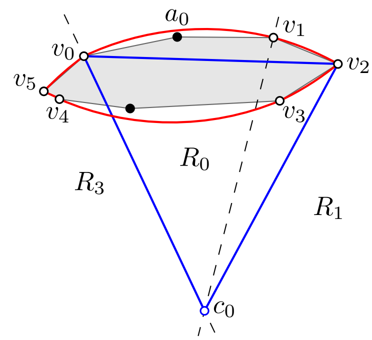

Figure 3: Proof of Case 2 of Lemma 6 where v_2 is the vertex with the largest index i such that the curve $T' = (c_0, v_0, v_1, \ldots, v_i, c_0)$ is convex.

An analogous proof follows if v_j lies in R_3. If v_j lies in R_2, let x be the intersection between the segment v_0v_1 and the line through v_j and c_0. Then, consider the convex polygon $T = (c_0, x, v_1, v_2, \ldots, v_{j-1}, v_j, c_0)$ and apply the same argumentation.

Case 2. If c_0 lies outside of Q, then no vertex of Q lies in R_2. Notice that if v_j lies in R_0, it is contained in the triangle $\triangle v_0v_1c_0$ which implies by convexity that v_j is enclosed by C_0. Assume that v_j lies in R_1 and notice that the same proof used in Case 1 holds as long as $T = (c_0, v_0, v_1, \ldots, v_j, c_0)$ defines a convex polygon. If T is not convex, let $1 \leq i \leq j$ be the largest index such that $T' = (c_0, v_0, v_1, \ldots, v_i, c_0)$ is convex. Therefore, v_i lies inside circle C_0 using the proof of Case 1. Because v_j lies in R_1, v_j lies inside the triangle $\triangle c_0, v_0, v_i$. Moreover, as c_0, v_0 and v_i are all enclosed by C_0, so is v_j by convexity; see Figure 3. An analogous proof holds if v_j lies in R_3 yielding our result. □

Let $\mathcal{B}$ be the smallest disk that contains Q. Notice that if the given perimeter P is larger than the perimeter of $\mathcal{B}$, then the (P, Q)-curve of maximum area is a circle of perimeter P that encloses Q (see Chapter X of [10]). Therefore, we assume that P is smaller than the perimeter of $\mathcal{B}$.

Given a number $r \geq radius(\mathcal{B})$, let $\mathscr{D}_r$ be the set of disks of radius r that contain Q, and let φ_r be the intersection of all disks in $\mathscr{D}_r$.

Recall that by Theorem 4, the (P, Q)-curve of maximum area, denoted by $\mathscr{C}$, consists of a sequence of circular arcs of equal radius.

Lemma 7 *Let $\mathscr{C}$ be the (P, Q)-curve of maximum area and let r be the radius of every arc along $\mathscr{C}$. If $P <$ perimeter$(\mathcal{B})$, then $\mathscr{C}$ is the boundary of φ_r.*

Proof. Because Q is contained in every $D \in \mathscr{D}_r$, Q is contained in φ_r. Moreover, every arc on the boundary of φ_r has radius r and φ_r is convex since it is the intersection of convex shapes.

Let a be an arc of $\mathscr{C}$ and let D_a be the disk of radius r extending it. Notice that $D_a \in \mathscr{D}_r$ by Lemma 6. We claim that arc a belongs to the boundary of φ_r and we prove it by contradiction. Assume that there is a point x on a such that x is not part of the boundary of φ_r. That is, there exists a disk $D \in \mathscr{D}_r$ such that x is not contained in D. Because D and D_a have the same radius, one of the endpoints of a lies in the complement of D. Therefore, D does not contain Q as both endpoints of a are vertices of Q by Theorem 4. However, Q is contained in every disk of $\mathscr{D}_r$ which is a contradiction. Consequently, every arc along $\mathscr{C}$ is contained on the boundary of φ_r proving our result. □

By Lemma 7, we need to consider only the intersection of all disks of radius r that contain Q to describe $\mathscr{C}$. However, we further simplify this description using the following result.

Proposition 1 *The intersection of every disk of radius r that contains Q is the intersection of every disk of radius r that contains Q and whose boundary passes through at least 2 vertices of Q.*

Proof. Let B be a disk containing Q. If ∂B passes through no vertex of Q, then we can continuously move B to the left (*resp.* right), until ∂B reaches a vertex of Q, to obtain a disk B^- (*resp.* B^+). Since $Q \subset B^+ \cap B^- \subset B$, we can ignore all disks whose boundary contains no vertex of Q when describing φ_r.

If the boundary of B passes through a single vertex v of Q, we can rotate B clockwise and counter-clockwise around v until ∂B reaches a second vertex of Q. In this way, we obtain two disks B^- and B^+ whose intersection is a lune contained in B. Because $Q \in B^+ \cap B^-$, by ignoring B, the intersection of all disks of radius r that contain Q remains unchanged. In other words, we can consider only the disks that pass through at least two vertices of Q to describe φ_r. □

3 Computing the (P, Q)-curve of maximum area

In this section, we show how to compute the (P, Q)-curve of maximum area using the farthest-point Voronoi diagram of the vertices of Q.

Let $c_{\mathcal{B}}$ be the center of $\mathcal{B}$ and let $r_{\mathcal{B}}$ be its radius. The farthest-point Voronoi diagram of the vertices of Q, denoted by $\mathcal{V}(Q)$, can be seen as a tree with n unbounded edges [7]. We can think of the leaves of this tree as points at infinity in the direction of these unbounded edges. Because $c_{\mathcal{B}}$ lies either on a vertex of $\mathcal{V}(Q)$ or on one of its edges, we can assume $\mathcal{V}(Q)$ to be rooted at $c_{\mathcal{B}}$ (if $c_{\mathcal{B}}$ is not a vertex, insert it by splitting the edge where it belongs). Given a point x in the plane, let $\rho(x)$ be the radius of the minimum enclosing circle of Q with center on x. The following is a well known property of the farthest point Voronoi diagram.

Lemma 8 *The map ρ is monotonically increasing along the path joining $c_{\mathcal{B}}$ with any leaf of $\mathcal{V}(Q)$.*

Given a number $r \geq r_{\mathcal{B}}$, let $X_r = \{x \in \mathbb{R}^2 : \rho(x) = r \text{ and } x \text{ is a point on } \mathcal{V}(Q)\}$.

Lemma 9 *Let $\mathscr{C}$ be the (P, Q)-curve of maximum area and let r be the radius of every arc along $\mathscr{C}$. If $P < perimeter(\mathcal{B})$, then $\mathscr{C}$ is the boundary of the intersection of every disk of radius r centered at a point of X_r.*

Proof. Let D be a disk of radius r that contains Q such that ∂D passes through two vertices v and v' of Q. Since $P < perimeter(\mathcal{B})$, by Lemma 7 and Proposition 1, it suffices to prove that D has its center in X_r. Let c be the center of D and notice that any disk of radius $r' < r$ centered at c does not contain v and v' and hence, it does not contain Q. Moreover, as D contains Q, we conclude that $\rho(c) = r$. Because ∂D passes through v and v', c is equidistant from these vertices. Furthermore, as Q is contained in D, there cannot be a vertex of Q that is farther from c than v (or v'). Therefore, c lies on the boundary of the farthest-point Voronoi cells of both v and v'. That is, c must lie on an edge of $\mathcal{V}(Q)$ which implies that $c \in X_r$ yielding our result. □

By Lemma 6, the circle extending every arc on $\mathscr{C}$ contains Q. Therefore, the radius of every arc along $\mathscr{C}$ must be greater than the radius of $\mathcal{B}$. Recall that for any value $r \geq r_{\mathcal{B}}$, φ_r is the intersection of all disks of radius r that contain Q.

Lemma 10 *Given a radius $r \geq r_{\mathcal{B}}$, φ_r and its perimeter can be obtained in $O(n)$ time after computing the farthest-point Voronoi diagram of the vertices of Q.*

Proof. Because $r_{\mathcal{B}} < r$, Lemma 8 implies that on every path joining the root with a leaf of $\mathcal{V}(Q)$, there is a point x such that $\rho(x) = r$.

By Lemma 8, we can scan every edge of $\mathcal{V}(Q)$ to find those edges that contain a point of X_r. Because each of these edges represents the bisector of two vertices of Q, we can determine the position of all the points of X_r in $O(n)$ time. Furthermore, once X_r is computed, we can reconstruct their cyclic order along the boundary of φ_r by performing a depth first search in the tree $\mathcal{V}(Q)$ in linear time. Thus, φ_r and its perimeter can be computed in $O(n)$ time as all circular arcs have the same curvature. □

The following result states that the solution to the isoperimetric problem cannot always be computed exactly. Therefore, we resort to approximation algorithms. The proof of the following theorem is deferred to Section 4 for ease of readability.

Theorem 11 *There exists a convex polygon Q and a value P such that the (P,Q)-curve of maximum area cannot be computed exactly, i.e., the radius of each arc along this curve cannot be represented as a closed-form expression.*

We finish this section by providing an algorithm to compute an approximation of the (P,Q)-curve of maximum area with fixed but arbitrary precision. Notice that if we are not given a convex polygon but a set of n points, we need to compute its convex hull in $\Theta(n \log n)$ time as a preprocessing step.

Theorem 12 *Let Q be a convex polygon and let $\mathcal{B}$ be the minimum disk containing Q. Given a value $P > 0$, it holds that: (1) If $P \geq perimeter(\mathcal{B})$, then the (P,Q)-curve of maximal area can be computed in $O(n)$ time. (2) If $P < perimeter(\mathcal{B})$, then we can compute an approximation of the (P,Q)-curve of maximal area with arbitrary but fixed precision in $O(n \log n)$ time. Moreover, this approximation has the same combinatorial structure as the optimal solution.*

Proof. Compute the smallest disk $\mathcal{B}$ containing Q in $O(n)$ time [9, 12]. Two cases arise:
(1) If $P \geq perimeter(\mathcal{B})$, let c be the center of $\mathcal{B}$ and let $\mathscr{C}$ be the circle of perimeter P centered on c.
(2) If $P < perimeter(\mathcal{B})$, then start by computing the farthest-point Voronoi diagram $\mathcal{V}(Q)$ of the vertices of Q in $O(n)$ time [1]. Let r be the radius of every arc in $\mathscr{C}$. Because $\mathscr{C}$ is equal to the boundary of φ_r by Lemma 9, we can use Lemma 10 to approximate r. First, consider every vertex v of $\mathcal{V}(Q)$ and assume that $\rho(v)$ was stored during the computation of $\mathcal{V}(Q)$.

Sort the vertices of $\mathcal{V}(Q)$ by their value under ρ in $O(n \log n)$ time. Then, we can approximate the radius of the arcs of $\mathscr{C}$ using a binary search for r in the set $\{\rho(v) : v \text{ is a vertex of } \mathcal{V}(Q)\}$. That is, for a given r' in this set, compute $\varphi_{r'}$ and its perimeter in $O(n)$ time using Lemma 10. Then, compare it with the value of P: If this perimeter is larger than P, then $r > r'$; otherwise, $r \leq r'$. In this way, we will find two vertices u and v of $\mathcal{V}(Q)$ such that $\rho(u) < r < \rho(v)$. At this point, we know that for any value $\rho(u) < r' < \rho(v)$, the curve $\varphi_{r'}$ has the same combinatorial structure, i.e., the vertices of $\varphi_{r'}$ and their order along its boundary are the same. As $\rho(u) < r < \rho(v)$, we have a combinatorial description of $\mathscr{C}$ where only the radius r of the arcs along $\mathscr{C}$ has not yet been determined.

By Theorem 11, there are instances where we cannot compute r exactly. However, we can use a binary search in the interval $[\rho(u), \rho(v)]$ to approximate it. That is, for a given constant $\varepsilon > 0$, we can compute an approximation r^* of r such that $|r^* - r| < \varepsilon$ in $O(n \log \frac{1}{\varepsilon})$ time. Therefore, we get a curve that approximates $\mathscr{C}$ with arbitrary precision and that has the same combinatorial structure. ☐

4 Hardness of the computation of exact solutions

In this section, we prove that Problem 1 cannot always be solved exactly. By an *exact solution*, we mean a solution that can be represented by a *closed-form expression*. The problem with closed-form expressions is that there is no consensus in the literature on how they should be defined [4]. According to Borwein and Crandall [4], closed-form expressions can be considered to be "a topic that intrinsically has no 'right' answer." For different reasons that we explain in the next paragraphs, we adopt the definition of Chow [6], qualified by Borwein and Crandall [4] as "the smallest plausible class of closed forms." We first make a little detour to discuss polynomial equations. Let

$$p_d(x) = 0 \tag{1}$$

be a polynomial equation of degree d. If $d \leq 4$, there exist general formulas to solve (1). These general formulas are finite (closed-form) expressions that involve $+$, $-$, $\times$, $\div$ and $\sqrt[k]{\ }$ for any integer $k \geq 2$. In this case, we say that (1) is *solvable by radicals*. If $d > 4$, some polynomial equations are solvable by radicals. For instance, $x^5 - 10x^4 + 35x^3 - 50x^2 + 24x = x(x-1)(x-2)(x-3)(x-4)$. However, some polynomial equations, such as $x^5 - x + 1 = 0$ are not solvable, as can be proven using Galois theory.

If the operations available at unit cost in the model of computation are $+$, $-$, $\times$, $\div$ and $\sqrt[k]{\ }$ for any integer $k \geq 2$, then the solutions to (1) cannot always be computed exactly when $d > 4$. For some problems, the degree of any polynomial equation involved in the solution is bounded by some constant μ. One way of getting around the unsolvability issue is to assume that any polynomial equation of degree less than μ can be solved in $O(1)$ time in the model of computation. However, for some problems, the degree of the polynomial equations involved can be shown to be unbounded in general. In such cases, the solution cannot be computed exactly and we usually turn to approximation algorithms.

Let $\mathscr{C}$ be the (P, Q)-curve of maximum area. Assume that we are given the sequence $v_0, \ldots, v_k$ of vertices of Q that connect consecutive edges along $\mathscr{C}$. That is, $\mathscr{C} = (v_0, v_1, \ldots, v_k, v_0)$ and each v_i is connected to v_{i+1} (modulo k) via a circular arc a_i. By Theorem 4, we know that every arc along $\mathscr{C}$ has the same radius r. Notice that to give an exact description of $\mathscr{C}$, it suffices to compute the value of r.

Let c_i denote the center of the circle extending a_i and notice that c_i lies on the bisector of v_i and v_{i+1}. If d_i denotes half the distance between v_i and v_{i+1}, then the angle $\alpha_i = \angle v_i c_i v_{i+1}$ is equal to $2 \cdot \arcsin(d_i/r)$. Therefore, the perimeter of arc a_i is given by the equation $2r \cdot \arcsin(d_i/r)$. Because the perimeter of $\mathscr{C}$ is P, to find the value of r it suffices to solve the following equation:

$$\sum_{i=0}^{k} 2r \cdot \arcsin(d_i/r) = P \tag{2}$$

This equation is not polynomial and at first glance, it does not seem possible to convert it into a polynomial via an appropriate change of variables. We will show that in general, there is no closed-form expression (in Chow's sense [6]) to express the solutions to (2).

Following Chow [6], we say that a number can be written in *closed-form* if it is an exponential-logarithmic number.

Definition 2 (Exponential-Logarithmic Numbers, [6]) *Let $\mathbb{E}$ be the set such that $\mathbb{Q} \subset \mathbb{E}$ and for all $x, y \in \mathbb{E}$ with $y \neq 0$, it holds that:*

1. $x + y,\ x - y,\ xy,\ x/y \in \mathbb{E}$,
2. $e^x \in \mathbb{E}$,
3. $\log(y) \in \mathbb{E}$, *where* $\log$ *is the branch of the natural logarithm function such that* $-\pi < Im(\log(y)) \leq \pi$ *for all* y.

Notice that $e \in \mathbb{E}$, $i \in \mathbb{E}$ and $\pi \in \mathbb{E}$, since $e = e^{e^0}$, $i = e^{\log(-1)/2}$ and $\pi = -i \log(-1)$. Therefore, $2\pi i \in \mathbb{E}$ and we have access to all branches of the natural logarithm. We can also compute x^y for any $x, y \in \mathbb{E}$ with $x \neq 0$, since $x^y = e^{y \log(x)}$. Consequently, we can compute the k-th root of any number in $\mathbb{E}$ (for any integer $k \geq 2$). This implies that we can compute the solutions to any polynomial equation with rational coefficients that is solvable by radicals. Finally, we have access to all trigonometric functions, since for instance,

$$\sin(x) = \frac{e^{ix} - e^{-ix}}{2i},\ \ \arcsin(x) = -i \log\left(ix + \sqrt{1 - x^2}\right),\ \ \sinh(x) = \frac{e^x - e^{-x}}{2}.$$

If we suppose that the operators available at unit cost are $+$, $-$, $\times$, $\div$, $\exp(\cdot)$ and $\log(\cdot)$, we can compute any so-called elementary functions on $\mathbb{E}$ in $O(1)$ time. Most problems from computational geometry can be solved using numbers from $\mathbb{E}$ and the operators defined on $\mathbb{E}$.

Asking whether a polynomial equation can be solved exactly corresponds to asking whether it can be solved by radicals. Asking whether a transcendental equation like $x + e^x = 0$, $\cos(x) = x$ or (2) can be solved exactly corresponds (in Chow's sense [6]) to asking whether its solutions belong to $\mathbb{E}$. Standard results in this matter are of the form: if *Schanuel's conjecture* is true, then the solution to $x + e^x = 0$ does not belong to $\mathbb{E}$ [6]. Here is the statement of Schanuel's conjecture [2, Chapter 12].

Conjecture 3 (Schanuel) *If α_1, α_2,..., α_n are complex numbers linearly independent over $\mathbb{Q}$, then the transcendence degree of the field $\mathbb{Q}\left(\alpha_1, e^{\alpha_1}, \alpha_2, e^{\alpha_2}, ..., \alpha_n, e^{\alpha_n}\right)$ over $\mathbb{Q}$ is at least n.*

According to Chow [6], at present, a proof of Schanuel's conjecture seems to be out of reach.

We now take an instance of Problem 1 and show that its solution does not belong to $\mathbb{E}$, provided that Schanuel's conjecture is true. Let T be an equilateral triangle such that each side has length 2; see Figure 4. Suppose we want to enclose T with a closed maximum-area curve with perimeter 7. To compute the optimal solution, we can solve $6r \cdot \arcsin(1/r) = 7$ for $r \in \mathbb{R}$. Equivalently, we solve

$$\sin(r') = \frac{6}{7}r' \ , \tag{3}$$

where $r' = \frac{7}{6r}$. Let r' be the solution[1] to (3) such that $r' > 0$. Is $r' \in \mathbb{E}$? Using a similar approach to Chow's for the proof of Theorem 2 in [6], we can prove the following lemma.

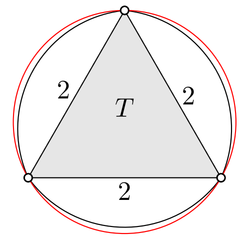

Figure 4: The red curve is the circumcircle of T. The black curve is the solution to Problem 1 when given a perimeter $P = 7$.

Lemma 13 *If Schanuel's conjecture is true, then $r' \notin \mathbb{E}$.*

As mentioned earlier, in the case of polynomial equations, one can avoid solvability questions by adding an appropriate operator to the model of computation. However, this is reasonable provided that the polynomial equations are of bounded degree. Looking at (3), one could argue that it suffices to add an appropriate operator to the model of computation. We could add an operator that solves equations of the form $\sin(x) = \lambda x$ in $O(1)$ time, where λ can be any real number. However, we claim that in general, the complexity of (2) is unbounded. As for the case of polynomial equations, it means that we need to turn to approximate solutions. To prove our claim we use the following identity.

$$\arcsin(\alpha) + \arcsin(\beta) = \arcsin\left(\alpha\sqrt{1-\beta^2} + \beta\sqrt{1-\alpha^2}\right) \tag{4}$$

Using (4) and the change of variables $r' = \frac{P}{2r}$, if all the d_i's are different, (2) becomes

$$\sin(r') = \mathcal{R}(r') \ ,$$

where $\mathcal{R}(r')$ satisfies the following: (a) It contains at least one sequence of $k-1$ nested square roots. (b) Each square root in this sequence is summed to a polynomial in r' of degree at least 2. (c) The most inner radicand is a polynomial in r' of degree at least 2. Therefore, the complexity of (2) is unbounded and r' cannot be represented as a closed-form expression. In other words, if $P = 7$, then most likely the (P, T)-curve of maximum area cannot be computed exactly. We obtain the following result.

Theorem 11 *There exists a convex polygon Q and a value P such that the (P, Q)-curve of maximum area cannot be computed exactly, i.e., the radius of each arc along this curve cannot be represented as a closed-form expression.*

We conclude this section with the following remark. Chow proved [6, Corollary 1] that if Schanuel's conjecture is true, then the polynomial equations with rational coefficients that can be solved within $\mathbb{E}$ are precisely those that are solvable by radicals. This is one more reason to define a closed-form number as in Definition 2.

5 Duality and open problems

We have presented a characterization of the (P, Q)-curve of maximum area. It is worth noting that there is a dual formulation of this problem. Given a number $A > area(Q)$, and a convex polygon Q, an $[A, Q]$-curve is a simple closed curve enclosing both Q and a region of area A. The dual formulation of Problem 1 is stated as follows.

Problem 2 *Given a value A and a convex polygon Q, find the $[A, Q]$-curve of minimum perimeter.*

The following observation will help us prove the relation between these two formulations.

Observation 4 *Let $\mathscr{C}$ be the (P, Q)-curve of maximum area. Then, the area of $\mathscr{C}$ increases monotonically as P increases.*

Lemma 14 *Given a convex polygon Q and an area $A > area(Q)$, let $\mathscr{C}^*$ be an $[A, Q]$-curve of minimum perimeter. If P is the perimeter of $\mathscr{C}^*$ then $\mathscr{C}^*$ is a (P, Q)-curve of maximum area.*

[1]We have $r' \approx 0.94683$, from which $r \approx 1.23219$.

Proof. If A is greater than the area of the minimum enclosing disk $\mathcal{B}$ of Q, then the result is equivalent to the classic isoperimetric theorem. Therefore, assume that $area(Q) \leq A \leq area(\mathcal{B})$. Let φ_r be the intersection of all disks of radius r that contain Q. Let r_A be the radius such that φ_{r_A} has area A. This value always exists because the area of φ_r attains every value in $[area(Q), area(\mathcal{B})]$ as the radius r increases continuously.

Let $\mathscr{C}$ be the boundary of φ_{r_A} and let P be the length of its perimeter. By Lemma 7, $\mathscr{C}$ is the (P, Q)-curve of maximum area. Therefore, every curve enclosing Q of perimeter smaller than P has area smaller than A by Observation 4. Consequently, the $[A, Q]$-curve of minimum perimeter must have perimeter at least P. Moreover, as $\mathscr{C}$ has area A by construction, we get that $\mathscr{C}$ is the $[A, Q]$-curve of minimum perimeter, i.e., $\mathscr{C} = \mathscr{C}^*$. □

We conclude by stating open problems that are closely related to the problem presented in this paper.

Problem 3 *Given a set of points $S \in \mathbb{R}^2$ and a value $P > 0$, find the (P, S)-curve of minimum area.*

While Problem 3 has a similar formulation to Problem 1, it belongs to a different class of problems as stated in the following result.

Theorem 15 *Problem 3 is NP-hard.*

Proof. A Steiner tree of S is a geometric tree whose vertex set contains S. Let $\mathcal{T}$ be a minimum Euclidean Steiner tree of S, i.e., a minimum cost Steiner tree where the cost of an edge is its Euclidean length. Notice that if P is equal to twice the length of $\mathcal{T}$, then the solution to Problem 3 is a closed curve of area zero obtained by going around $\mathcal{T}$. Because computing the minimum Steiner tree is NP-hard [8], computing the solution to Problem 3 is also NP-hard. □

By extending the formulation of isoperimetric problems to $\mathbb{R}^3$, we obtain the following problem statement that, as far as we know, has not yet been studied.

Problem 4 *Let $S \in \mathbb{R}^3$ be a set of points and let $A > 0$. Among all surfaces of area A, what is the closed surface of (maximum) minimum volume that encloses S?*

References

[1] A. Aggarwal, L. Guibas, J. Saxe, and P. Shor. A linear time algorithm for computing the voronoi diagram of a convex polygon. In *Proceedings of STOC*, pages 39–45, New York, NY, USA, 1987. ACM.

[2] A. Baker. *Transcendental number theory.* Cambridge University Press, 1990.

[3] V. Blåsjö. The isoperimetric problem. *The American Mathematical Monthly*, 112(6):526–566, 2005.

[4] J. M. Borwein and R. E. Crandall. Closed forms: what they are and why we care. *Notices Amer. Math. Soc.*, 60(1):50–65, 2013.

[5] P. Bose and J.-L. Carufel. Isoperimetric triangular enclosures with a fixed angle. *Journal of Geometry*, 104(2):229–255, 2013.

[6] T. Y. Chow. What is a closed-form number? *The American mathematical monthly*, 106(5):440–448, 1999.

[7] M. de Berg, M. van Kreveld, M. Overmars, and O. Schwarzkopf. *Computational Geometry: Algorithms and Applications.* Springer-Verlag, third edition, 2008.

[8] M. R. Garey, R. L. Graham, and D. S. Johnson. The complexity of computing steiner minimal trees. *SIAM journal on applied mathematics*, 32(4):835–859, 1977.

[9] N. Megiddo. Linear-time algorithms for linear programming in $\mathbb{R}^3$ and related problems. *SIAM J. Comput.*, 12(4):759–776, 1983.

[10] G. Polya. *Mathematics and plausible reasoning, Volume 1: Induction and analogy in mathematics.* Princeton University Press, 1990.

[11] D. A. Singer. *Geometry: plane and fancy.* Springer, 1998.

[12] E. Welzl. Smallest enclosing disks (balls and ellipsoids). In *Results and New Trends in Computer Science*, pages 359–370. Springer-Verlag, 1991.