Computing the center of area of a convex polygon

Peter Braß*
Institut für Informatik
Freie Universität Berlin
D-14195 Berlin, Germany
brass@inf.fu-berlin.de

Laura Heinrich-Litan†
Institut für Informatik
Freie Universität Berlin
D-14195 Berlin, Germany
litan@inf.fu-berlin.de

Abstract

The center of area of a planar convex set C is the point p for which the minimum area cut off from C by any halfplane containing p is maximized. Properties of this point were studied already long ago in classical geometry, but it is quite nontrivial to really determine this point for a given convex set. We give an $O(n^2(\log n)^3\alpha(n))$-algorithm to compute the center of area of a convex n-gon, which improves a previous $O(n^6(\log n)^2)$-algorithm of Diaz and O'Rourke.

1 Introduction

The 'center of area' of a convex set X is the point p^* for which the minimum over all lines through p^* of the area cut off from X is maximal. So p^* is the point where the cut-off area function

$$f(p) = \min\Big\{ \text{area}(X \cap h) \mid h \text{ is a halfplane bounded by a line through } p \Big\}$$

is maximal. This concept was studied already long ago [8, 5, 7]. It is a classical result of convex geometry that any line (hyperplane) through the centroid of X will always cut off at least $\frac{1}{3}$ of the total area of X (generally $\frac{1}{d+1}$ of the volume), which is worst case optimal, as can be seen by a simplex. But there are sets X for which there is a point p different from the centroid c such that $f(p) > f(c)$. On the other hand, the cut-off area can never be very big, it is always at most half the total area (since the complement of each halfplane through p is also a halfplane through p), and it is exactly half the area for each line through p if and only if the set X is centrally symmetric, with center p. For this reason the ratio of the maximum of the cut-off area function to the total area of X was proposed as measure of symmetry: $\frac{1}{2}$ for centrally symmetric sets, $\frac{1}{3}$ for the triangles, and something in between for all other convex sets. This is 'Winternitz's measure of symmetry' [4].

It is, however, quite nontrivial to actually determine that center of area point for a given convex set X. Unlike the centroid, there is no obvious way to determine it, although the classical literature already gave some balance conditions it has to satisfy. In a series of papers [1, 2, 3] Diaz and O'Rourke gave algorithms for this problem and some related chord-center problems. They obtained an $O(n^6(\log n)^2)$ algorithm to compute the center of area of a convex n-gon, which is impracticable, not only because of its complexity, but also because it uses the construction of lower envelopes of families of functions on two-dimensional domains. They also gave a 'numerical' approximate

*Supported by Deutsche Forschungsgemeinschaft (DFG), grant BR 1465/5-2.
†Supported by the IST Programme of the EU as a Shared-cost RTD (FET Open) Project under Contract No IST-2000-26473 (ECG - Effective Computational Geometry for Curves and Surfaces).

algorithm which determines the center of area with K bit precision, if the n vertices are given in G bit precision, in $O(GK(n+K))$ time.

It is the aim of this paper to give an exact algorithm with a reasonable time bound, and that does not need unimplementable algorithmic tools.

Theorem 1.1 *The center of area of a convex n-gon can be determined in $O\left(n^2(\log n)^3\alpha(n)\right)$ time.*

Here $\alpha(n)$ is the inverse Ackermann function.

2 Geometric properties of the area function

The cut-off area function $f(p)$ is defined over the convex polygon P. The *upper level sets* of the cut-off area function $f(p)$ are defined as $\Gamma_\delta = \{p \mid f(p) \geq \delta\}$. The following properties of Γ_δ are known ([3]):

1. The boundary of Γ_δ, the *level curve* $\{p \mid f(p) = \delta\}$ is a simple, connected curve.
2. If $\delta_2 > \delta_1$ then Γ_{δ_2} is strictly contained in Γ_{δ_1}.
3. If $f(p) = \delta$ then each line $L^*(p)$ which minimizes the cut-off area for p supports Γ_δ.
4. Γ_δ is convex ([7]).

For the directed line $L(p)$ through the point p we define $C(L(p))$ as the chord formed by the intersection of the polygon with $L(p)$. For a point p a directed line $L^*(p)$ which minimizes the area cut off from the polygon by the closed halfplane bounded by, and to the left of, $L^*(p)$ is called a *min-line* and the corresponding chord, a *min-chord*. It is known that if a line $L(p)$ is a min-line, then p is the midpoint of the chord $C(L(p))$. So if $L(p)$ is a min-line through p, and it cuts the boundary in the edges e_i and e_j, then p belongs to the set $R_{i,j}$ of midpoints of chords connecting e_i and e_j. This set $R_{i,j}$ is a parallelogram (Figure 2). And, unless e_i and e_j are parallel sides of the polygon (in which case $R_{i,j}$ degenerates to a line-segment), for each point $p \in R_{i,j}$ there is a unique directed line L which intersects first e_i then e_j, and has p as the midpoint of the chord $C(L)$. This line cuts off some area of the polygon P, which is the minimum area that can be cut off from P by a line $L(p)$ which intersects the sides e_i and e_j. This cut-off area is some function defined on $R_{i,j}$ and the minimum cut-off area $f(p)$, taken over all lines $L(p)$, is the minimum over those partial functions (depending on e_i and e_j) for which $p \in R_{i,j}$. But for each edge pair (e_i, e_j), $i \neq j$, this function $f(p; i, j)$ on $R_{i,j}$ can explicitly be determined; they are quadratic functions in the two coordinates of the point p, where the coefficients are constants determined by the edges e_i, e_j of the polygon, as already used in [3]. This fact is illustrated by the example presented in Figure 1. Consider the directed line $L(p)$ whose corresponding chord has midpoint p and endpoints on the edges e_i and e_j. The area cut off by the closed halfplane bounded by, and to the left of the directed line $L(p)$ is given by

$$f(p; i, j) = 2 \cdot s_1(p) \cdot s_2(p) \cdot \sin(\beta) - \text{area}(\triangle svw) + R(v, w)$$

where $s_1(p)$ and $s_2(p)$ are the side lengths of the shaded parallelogram and are linear functions in $p = (x, y)$. $R(v, w)$ is the area of the shaded piece of the polygon bounded by the segment vw and β is the angle formed by the edges e_i and e_j. Thus, $f(p; i, j)$ is a quadratic function in two variables, which are the two coordinates of the point $p = (x, y)$.

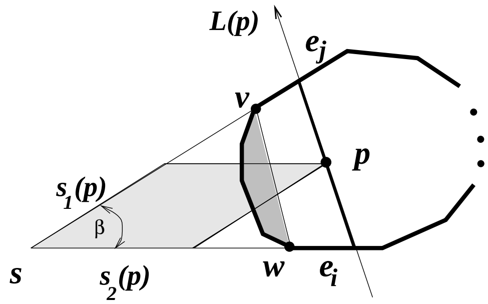

Figure 1: Computing the area function $f_l(p)$

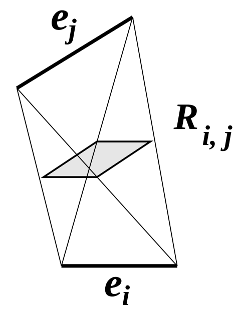

Figure 2: Domain of a function f_l

Thus, we can determine a list of these area functions $f_l, l = 1, \ldots, n(n-1)$ in $O(n^2)$ time, together with their domains $\mathrm{Dom}(f_l)$, where the domain of the partial function f_l is the parallelogram $R_{i,j}$ generated by the corresponding edge pair e_i, e_j and which is contained in the convex hull of this generating edge pair (see Figure 2). The value of $f(p)$ is then determined as the minimum among those functions f_l with $p \in \mathrm{Dom}(f_l)$:

$$f(p) = \min\{f_l(p) \mid p \in \mathrm{Dom}(f_l)\}$$

It should be noted that one reason for the difficulty of the problem is that we cannot just drop the condition $p \in \mathrm{Dom}(f_l)$; although we can evaluate the quadratic functions f_l anywhere in the plane, outside their region $\mathrm{Dom}(f_l)$ they will not correspond to some area cut off by a halfplane. On the boundary of P, e.g., most of these functions will give a negative value, which cannot be a valid area. Figure 3 shows an example of such a minimization diagram, where each region is labelled by the edge pair which generates the minimal quadratic function.

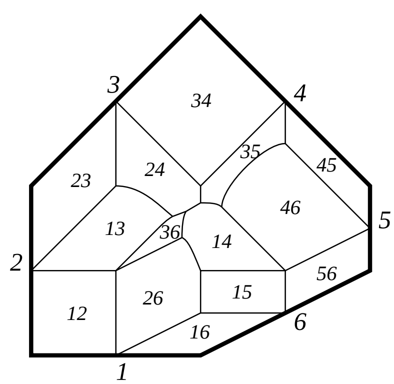

Figure 3: Minimization diagram

Diaz and O'Rourke obtained in [3] an $O(n^6(\log n)^2)$ algorithm for the problem, by determining all $O(n^4)$ regions formed by the intersection of the domains $\mathrm{Dom}(f_l)$ and by constructing the lower envelope of the functions f_l over each of these regions of intersection.

3 Outline of the algorithm

Our algorithm proceeds in two phases:

- In the first we maintain a successively smaller search area which contains the maximum, the center of area, and a superset of all functions f_l whose domain intersects the search region. This will finally result in a search region which is intersected only by the domains of $O(n)$ functions.
- In the second we will find the maximum over the remaining search region of the minimum of the remaining partial functions.

In the following we will denote the partial functions whose domain possibly intersects the current search region as *relevant* functions. The validity of this approach depends on our ability to maintain and shrink the search region which is guaranteed to contain the maximum. In the beginning the search region is the entire polygon. Then the reduction is based on:

Lemma 3.1 *Let l be a line that intersects the current search region. Given a list of the r still relevant functions, it is possible to decide on which side of l the maximum lies in time $O(\lambda_3(r)\log r) = O(r(\log r)\alpha(r))$.*

This orientation information reduces not only the search region, but also the set of relevant functions. The domain of a function is always contained in the convex hull of the generating edge pair; so if both generators lie on the wrong side of the separating line l, the domain of that partial function cannot intersect the new search region.

By performing a binary search on such lines, it is simple to show:

Lemma 3.2 *Given a vertex v_i, a list of the r still relevant functions and the current search region, we can find in time $O(r(\log r)(\log n)\alpha(r) + n)$ two consecutive vertices $v_j v_{j+1}$ such that the maximum lies between $v_i v_j$ and $v_i v_{j+1}$.*

We call this region the *cone* $C(v_i)$ of vertex v_i (see Figure 4). The search region is then the intersection of such cones. By proper choice of the sequence of the vertices v_i from which we construct the cones, we can reduce the number of relevant functions sufficiently fast:

Lemma 3.3 *We can construct the cones for the vertices v_i in such a sequence that after constructing the i-th cone for $i = 1, \dots, n$ there are only $O(\frac{n^2}{i})$ relevant functions left, which can be also listed in $O(\frac{n^2}{i})$ time.*

The first phase of the algorithm consists now of the n repetitions of the cone construction of Lemma 3.2, for the sequence of vertices given by Lemma 3.3, this takes a total time:

$$\sum_{i=1}^{n} O\left(\frac{n^2}{i}\left(\log\frac{n^2}{i}\right)(\log n)\,\alpha\left(\frac{n^2}{i}\right)\right) = O\left(n^2(\log n)^3\alpha(n)\right) \tag{1}$$

So after constructing the cones for all vertices in the proper sequence we have a search region, which is the intersection of n such cones, thus at most a $2n$-gon, and only a set of $O(n)$ still relevant functions left. It remains to find the point p^* (the center of area) at which the maximum of the lower envelope (pointwise minimum) of these partial function is reached, which is the second phase of the algorithm.

It is known that there are at least three relevant functions f_i that give that same value $f(p^*) = \min\{f_i(p^*) \mid i{:}\, p^* \in \mathrm{Dom}(f_i)\}$ at that point; they correspond to three distinct lines through p^* that all cut off the same minimal area of the polygon [4]. So, the center of area p^* is a vertex of the lower envelope of the $O(n)$ still relevant partial functions, and we could find it by just constructing that lower envelope (which is expensive and not really implementable).

There is, however, a faster way. We first restrict the search region further, to make all still relevant partial functions into total functions on that restricted search region. For this we look at each still relevant function f_i; the domain of f_i is bounded by 4 lines. We test for each line whether it intersects the current search region (in $O(n)$ time), and if it does, perform one orientation test by Lemma 3.1. These at most $4 = O(1)$ lines cut the current search region into at most $O(1)$ cells, and by the orientation test we know which of these cells contains the center of area p^*. We restrict the search region to that cell (adding at most 4 further sides), and, if the cell is not a part of $\mathrm{Dom}(f_i)$, remove f_i from the relevant functions. This further reduction takes a total of $O(n^2(\log n)\alpha(n))$ time, and results in a search region which is still an $O(n)$-gon, and a set of $O(n)$ still relevant functions, and for each still relevant function the search region is contained in its domain.

We now want to decide for each of the still relevant functions whether it is possible for them to be among the at least three functions for which equality holds in the center of area: $f_i(p^*) = f(p^*)$. We do this by a process of pairwise comparison, in which we can always exclude one of a tested pair of functions, unless we have already found the global maximum. This does not further decrease the set of relevant functions; we are now constructing a new subset of equality-candidate functions, which starts out as the set of all relevant functions and decreases by the pairwise comparison process.

If f_i, f_j are two still relevant functions, then the set $\{p \mid f_i(p) = f_j(p)\}$ defines a quadratic curve in the plane, and on the one side of it $f_i > f_j$, and on the other side $f_i < f_j$, so if we know on which side of the curve the point p^* (center of area) is, either f_i or f_j will be excluded from the set of equality-candidate functions.

Lemma 3.4 *Given a search region which is an $O(n)$-gon, a list of $O(n)$ relevant functions which are defined on the search region, and a quadratic curve γ, we can decide in time $o\left(n(\log n)^2\right)$ on which side of γ the point p^* lies, or whether it lies on γ.*

So by at most $O(n)$ such pairwise comparison tests, we can reduce the set to three equality-candidate functions, or find the center of area p^* already on the way. This takes $o\left(n^2(\log n)^2\right)$, so the total complexity of our algorithm is $O\left(n^2(\log n)^3\alpha(n)\right)$, as claimed in the theorem.

4 Details of the algorithm

Proof of Lemma 3.1: Each of the r relevant partial functions is a quadratic function in two variables; so its restriction to the line l is a quadratic function of one variable, and we can compute its coefficients in $O(1)$ time per function just by inserting a fixed linear parametrization $l(t)$ of l into the definition of the two-dimensional function. Also the domain of each of these partial functions is some convex two-dimensional domain, a parallelogram, so the intersection of the domain with the line l is a segment, which we can find in $O(1)$ time. So the restriction of each relevant function to the line l, in the parametrization $l(t)$, is a quadratic function on an interval, and we can determine these r one-dimensional partial functions in time $O(r)$.

Since they are quadratic functions, any two of them intersect in at most two points, so by [6] p.137 it is possible to construct the lower envelope (pointwise minimum) in $O(r\log(r)\alpha(r))$ time. We can find the maximum of this lower envelope function by going once through all the $\lambda_4(r) = O\left(r \cdot 2^{\alpha(r)}\right)$ intervals.

This maximum is either unique, or reached on an interval, since the upper level sets $\{p \mid f(p) \geq \lambda\}$ of the two-dimensional area function f are known to be convex sets. Let $p_{\max} \in l$ be a point in which that maximum of f over l is reached. Since f cannot be constant over an open set, the point $p_{\max} \in l$ is a boundary point of the (convex) level curve $\{p \mid f(p) = f(p_{\max})\}$, and the line l is a tangent to that level curve. We only have to decide on which side of l that level curve lies, since its convex hull (the upper level set $\{p \mid f(p) \geq f(p_{\max})\}$) contains the global maximum.

For this we determine (going once through the list of relevant functions) the set $I^{=} = \{i \mid p_{\max} \in \mathrm{Dom}(f_i) \text{ and } f_i(p_{\max}) = f(p_{\max})\}$ of functions touching the lower envelope in $p_{\max}$. For a sufficiently small $\varepsilon > 0$, f is in an ε-environment of $p_{\max}$ just the minimum of these functions. For each of these functions we determine the quadratic curve $\{p \mid f_i(p) = f(p_{\max})\}$ (which contains $p_{\max}$), and compute the tangent t_i to that curve in $p_{\max}$. By the gradient of f_i we can orient that tangent to a half-plane h_i^- so that for all v there is an ϵ such that if $p_{\max} + \epsilon v$ is in the interior of h_i^- then $f_i(p_{\max} + \epsilon v) < f(p_{\max}) < f_i(p_{\max} - \epsilon v)$ (so on the h_i^--side of t_i the function f_i locally decreases, and on the other side locally increases). Now, the line l must be contained in the union $\bigcup_{i \in I^{=}} h_i^-$ of these halfspaces, since otherwise the function $f = \min_{i \in I^{=}} f_i$ (locally, in an ε-environment of $p_{\max}$) would increase in one direction along the line.

But if the line l is contained in a union of closed halfplanes, then at most one halfplane bounded by l still contains points not covered by that union of closed halfplanes. But the level set $\{p \mid f(p) \geq f(p_{\max})\}$ is not contained in the interior of any of these halfplanes. If the halfplanes together cover the whole plane, then we have already found the global maximum; otherwise the maximum is on that side of l where there are still uncovered points. We can find that side in $O(r)$ time by computing for each h_i^- the angular interval around $p_{\max}$ covered by that halfplane; the union of such angular intervals is again an interval (since each interval has length π), so it is easy to compute the union of all these intervals and check whether there are still uncovered directions.

□

Proof of Lemma 3.2: The current search region $\mathcal{S}$ is the intersection of cones, thus at most a $2n$-gon. The tangent rays to $\mathcal{S}$ starting in v_i touches $\mathcal{S}$ either in a vertex or an edge and can be determined in time $O(n)$. We consider the region $\mathcal{T}$ which contains the maximum and is bounded by the two rays. The edges of the polygon intersecting $\mathcal{T}$ and the vertices of the polygon contained in $\mathcal{T}$ can be found in $O(n)$ time. For all these vertices $w \in \mathcal{T}$ the line $v_i w$ intersects the current search region $\mathcal{S}$, thus Lemma 3.1 can be used. If there are at least two vertices with $w \in \mathcal{T}$ we use binary search to find in time $O(r(\log r)(\log n)\alpha(r))$ two consecutive vertices $v_j v_{j+1}$ among them such that the maximum lies between $v_i v_j$ and $v_i v_{j+1}$. Otherwise there is either only one edge intersecting $\mathcal{T}$ which is then the one $e(v_i) = (v_j, v_{j+1})$ we looked for, or there are two edges intersecting $\mathcal{T}$ and the relevant one can be found by one use of Lemma 3.1.

□

Proof of Lemma 3.3: For a vertex v the cone $C(v)$ is the triangle $v w_i w_{i+1}$ which contains the center of area and $e(v) := w_i w_{i+1}$ is the edge corresponding to the cone $C(v)$ (see Figure 4).

The first cone $C(v_1)$, where v_1 is some vertex of the polygon, dissects the polygon in two polygonal chains L_1 and L_2, such that L_1 and L_2 cover the boundary of P and have exactly one edge, $e(v_1)$, in common. W.l.o.g. $|L_1| \geq |L_2|$, where the length $|L|$ of a polygonal chain L is the number of its edges. Clearly, all functions which are now relevant correspond to an edge-pair $(e, e') \in L_1 \times L_2$ since parallelograms, the domains of the functions f_l which contain the center of area have to intersect the cone $C(v)$. We call the pair (L_1, L_2) a relevant *chain-pair*. The number of relevant functions after the construction of $C(v_1)$ is bounded by $2 \cdot |L_1| \cdot |L_2|$ and since $xy \leq \frac{(x+y)^2}{4}$, $\forall\, x, y \in \mathbb{R}$ we get $2 \cdot |L_1| \cdot |L_2| \leq \frac{(n+1)^2}{2}$. Now, another cone $C(w)$ has either the apex $w \in L_1$ and the edge $e(w) \in L_2$

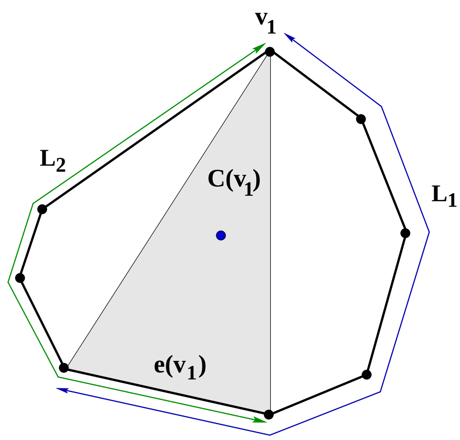

Figure 4: Cone $C(v_1)$ and chain-pair (L_1, L_2)

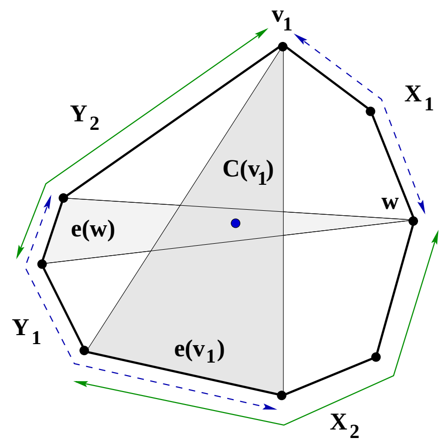

Figure 5: Dissection of chain-pair (L_1, L_2)

or the apex $w \in L_2$ and the edge $e(w) \in L_1$, with other words the cone $C(w)$ is contained in the region between L_1 and L_2. This fact is based on the observation that $C(w)$ contains the center of area as well, thus has to intersect the cone $C(v)$. The cone $C(w)$ dissects the chain-pair (L_1, L_2) in two chain-pairs (X_1, Y_1) and (X_2, Y_2) as illustrated in Figure 5. After i iterations i cones have been determined and there is a set $\mathcal{C}_i$ of i relevant left chain-pairs. In the $(i+1)$-st iteration a relevant chain-pair (X, Y) from $\mathcal{C}_i$ will be split into and replaced by two new relevant chain-pairs (X_1, Y_1) and (X_2, Y_2), such that X_1 and X_2 form a partition of X, and Y_1 and Y_2 cover exactly Y and have exactly one edge in common, which is the edge of the cone computed in the $(i+1)$-st iteration (see Figure 6).

The order $v_1, v_2, \ldots, v_n$ in which the nodes are processed for cone-computation is chosen as follows:

first part : all vertices of the chain L_1 are chosen iteratively such that after the 2^k-th iteration there are 2^k (almost) equal chains which partition L_1 and the next 2^k chosen vertices are iteratively halving those 2^k chains. A chain of length x is halved such that the two partitioning smaller chains have lengths $\lceil \frac{x}{2} \rceil$ and $\lfloor \frac{x}{2} \rfloor$, respectively.

second part : all vertices w of the chain L_2 are chosen in some order and the cones $C(w)$ are determined.

Analysis of the first part

Let $T(i)$ be the number of relevant functions left after the ith iteration. It equals twice the sum of the sizes of all relevant chain-pairs of $\mathcal{C}_i$:

$$T(i) \;=\; 2 \cdot \sum_{(L,L') \in \mathcal{C}_i} |L| \cdot |L'|$$

Now consider the iteration $i > 1$ and let k be such that $k = \max\{l \mid 2^l < i\}$. It is easy to prove by induction that after each iteration i all cones which haven't been computed yet are contained in the region of exactly one chain-pair of $\mathcal{C}_i$. Let (X, Y) with $|X| = x$ and $|Y| = y$ be the chain-pair

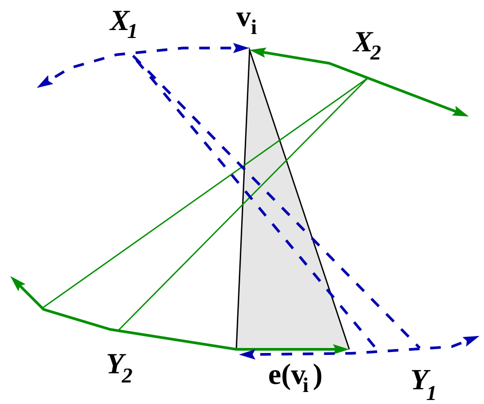

Figure 6: Cone dissection of a relevant chain-pair

which is dissected in the ith iteration. The vertex $v_i \in L$ is chosen such that it divides the chain X in two chains X_1 and X_2 of lengths $x_1 = \lceil \frac{x}{2} \rceil$ and $x_2 = \lfloor \frac{x}{2} \rfloor$, respectively. The cone $C(v_i)$ is determined and it divides the chain Y into two chains Y_1 and Y_2 such that $Y_1 \cap Y_2 = \{e(v_i)\}$ as shown in Figure 6. Obviously, all other cones which are contained in the region of (X, Y) either are contained in the region of (X_1, Y_1) or are contained in the region of (X_2, Y_2), since they contain the center of area and therefore have to intersect the cone $C(v_i)$. Now,

$$T(i) = \sum_{\substack{(L,L') \in \mathcal{C}_{i-1} \\ (L,L') \neq (X,Y)}} \left(2 \cdot |L| \cdot |L'| \right) + 2 \cdot |X_1| \cdot |Y_1| + 2 \cdot |X_2| \cdot |Y_2|$$

Using $y = y_1 + y_2 - 1$ it can be easily verified that :

$$\begin{aligned} |X_1| \cdot |Y_1| + |X_2| \cdot |Y_2| &= \left\lceil \frac{x}{2} \right\rceil \cdot (y_1 - 1) + \left\lfloor \frac{x}{2} \right\rfloor \cdot (y_2 - 1) + x \\ &\leq \frac{x+1}{2} \cdot (y-1) + x \leq \frac{|X| \cdot |Y|}{2} + \frac{|X| + |Y| - 1}{2} \end{aligned}$$

This implies

$$\begin{aligned} T\left(2^{k+1}\right) &\leq \sum_{(L,L') \in \mathcal{C}_{2^k}} \left(|L| \cdot |L'| + (|L| + |L'| - 1) \right) \\ &\leq \frac{T(2^k)}{2} + \sum_{(L,L') \in \mathcal{C}_{2^k}} (|L| + |L'| - 1) \end{aligned}$$

Since all chains L with $(L, L') \in \mathcal{C}_{2^k}$ partition chain L_1 we have $\sum_{(L,L') \in \mathcal{C}_{2^k}} |L| = |L_1|$. The chains L' with $(L, L') \in \mathcal{C}_{2^k}$ have pairwise at most an edge in common and cover exactly the chain L_2. It is easy to see that

$$\sum_{(L,L') \in \mathcal{C}_{2^k}} |L'| \leq |L_2| + 2^k - 1$$

Since $|\mathcal{C}_{2^k}| = 2^k$ and $|L_1| + |L_2| = n + 1$ we get

$$T\left(2^{k+1}\right) \;\leq\; \frac{T(2^k)}{2} + (|L_1| + |L_2|) \;=\; \frac{T(2^k)}{2} + n + 1 \;\leq\; \frac{T(1)}{2^k} + (n+1) \cdot \left(1 + \frac{1}{2} + \dots \frac{1}{2^{k-1}}\right)$$
$$\leq\; \frac{(n+1)^2}{2 \cdot 2^k} + 2n + 2$$

since $T(1) \leq \frac{(n+1)^2}{2}$. Because of $T(i) \leq T(2^k)$ we get

$$T(i) \;\leq\; \frac{(n+1)^2}{2 \cdot 2^k} + 2n + 2 \;\leq\; \frac{(n+1)^2}{i} + 2n + 2\;, \quad i = 1, \dots, |L_1|$$

Analysis of the second part

After the cones of the first part have been constructed each relevant chain-pair is of the form (e, L) where e is an edge of L_1. Thus, all cones $C(v_j)$ with $v_j \in L_2$ which are to be determined in the second phase have the corresponding edge $e(v_j) = e$ such that $v_j \in L$ and (e, L) is a relevant chain-pair left after the end of the first part. The number of relevant functions remains $T\left(|L_1|\right) = O(n)$:

$$T(i) \;=\; T(|L_1|) \;\leq\; \frac{(n+1)^2}{|L_1|} + 2n + 2 \;\leq\; \frac{(n+1)^2}{\frac{n}{2}} + 2n + 2 \;=\; O(n)\;, \quad i = |L_1| + 1, \dots, n$$

□

Proof of Lemma 3.4: The technique is the same as in Lemma 3.1: we choose a simple parametrization of that curve, determine the restrictions of all relevant functions to this curve in terms of this parametrization, construct the lower envelope of this family of one-dimensional functions of constant degree in $o\left(n(\log n)^2\right)$ time [6], choose any global maximum of this lower envelope, and find on which side of this curve the function f locally increases in this point.

There is, however, one additional difficulty by the fact that our quadratic curve, unlike the straight line in Lemma 3.1, may enter and leave the search region $O(n)$-gon up to $O(n)$ times, and we do not have any information on the behavior of the functions outside the search region, so there might be more relevant functions out there. We avoid this by increasing the family of relevant functions by one function for each edge of the search region, which is defined as 0 on the halfplane bounded by that edge not containing the final search region, and undefined else. For this extended set of (still $O(n)$) functions we have that the lower envelope on the search region agrees with the original lower envelope, and is 0 outside. This is smaller than all originally occurring function values, so the maximum, and all nontrivial level sets, do not change. Also the original functions can be extended to functions on the whole plane without changing the lower envelope, so they are defined on the whole quadratic curve γ. And the additional half-plane functions, restricted to γ, each contribute only one interval on which the restricted function is constant (0), and we can join overlapping intervals in $O(n \log n)$ time. Thus, we have indeed reduced it to a family of $O(n)$ functions, any two of them intersecting at most four times, as needed for the construction of the lower envelope. If the maximum of the lower envelope is 0, the curve γ does not intersect the search region, so we have to decide only for one arbitrary point of the search region, on which side of γ it lies; otherwise we can choose an arbitrary maximal point and decide for it, on which side of it the function f locally increases, as in Lemma 3.1.

□

5 Final Remarks

We gave in this paper an $O(n^2(\log n)^3\alpha(n))$-algorithm to compute the center of area of a convex n-gon, a problem, for which previously only an $O(n^6(\log n)^2)$-algorithm by Diaz and O'Rourke [3] was known. Our algorithm is based on some simple geometric properties of the cut-off area function: the convexity of its level sets, and its being the pointwise minimum of a set of $O(n^2)$ partial functions which can easily be determined explicitly, and whose domain of definition have a simple structure. We do not use complicated algorithmic techniques, the most demanding step being the construction of one-dimensional lower envelopes along one-dimensional cuts; so the algorithm should be well-implementable. We did not make an attempt to get the best exponent of the logarithmic factor, probably a slight further gain would be possible by a better balancing of the first and second phase; but with any algorithm that starts by constructing the partial functions we cannot do better than $\Omega(n^2)$.

References

[1] M. Diaz, J. O'Rourke. Computing the center of area of a polygon. In *Proc. of First International Workshop on Algorithms and Data Structures*, Springer LNCS 382 (1989) 171–182.

[2] M. Diaz, J. O'Rourke: Chord centers for convex polygons. In *Contemporary Mathematics* 119 (1991) (*Proc. AMS Spec. Sess. Vision Geometry* 1989) 29–43.

[3] M. Diaz and J. O'Rourke. Algorithms for computing the center of area of a convex polygon. In *The Visual Computer* 10 (1994), 432–442.

[4] B. Grünbaum. Measures of symmetry for convex sets. In *Convexity* (V. Klee, ed.), *AMS Proc. Symp. Pure Mathematics* 7 (1963) 223–270.

[5] V.E. Hoggatt. Forbidden area. In *American Mathematical Monthly* 69 (1962) 98–104.

[6] M. Sharir, P.K. Agarwal. Davenport-Schinzel sequences and their geometric applications. Cambridge Univ. Press 1995.

[7] D.S. Singmaster, G. Singmaster. Forbidden regions are convex. In *American Mathematical Monthly* 74 (1967) 184–186.

[8] W. Süss. Über Eibereiche mit Mittelpunkt. Mathematisch-Physikalische Semesterberichte 1 (1950) 273–287.