Cutting a Cake without Destroying the Toppings
— Draft —

Erel Segal-Halevi, Bar-Ilan University

On a two-dimensional cake, there are k pairwise-disjoint toppings. It is required to cut the cake to $k + h$ pieces, such that each topping is entirely contained in a unique piece, and the total number of extra pieces ("holes", h) is minimized. The minimal number of holes naturally depends on the geometric constraints of the cake and its pieces. I present tight bounds on the number of holes when the pieces must be connected, convex or rectangular.

1. INTRODUCTION

Let C ("cake") be a compact subset of the Euclidean plane. Let $C_1, \dots, C_k$ be pairwise-disjoint subsets of C ("toppings"), such that $C \supseteq C_1 \cup \cdots C_k$.

Our goal is to create a partition of C, $C = C'_1 \cup \cdots \cup C_{k+h}$, for some $h \geq 0$, such that:

$$\forall i \in 1 \dots k : \qquad\qquad C_i \subseteq C'_i$$

i.e, each topping is contained in a unique piece.

Subject to that requirement, we would like to minimize h. This h represents the number of "holes" - the number of pieces in the final partition with no topping. What is the smallest h that can be guaranteed, in the worst case, as a function of k?

Without further restrictions on the pieces, the answer is trivial: it is always possible to get $h = 0$ by setting e.g. $C'_1 = C_1 \cup C \setminus (C_2 \cup \cdots \cup C_k)$ (attach all "holes" arbitrarily to one of the toppings). The question becomes more interesting when there are geometric constraints on the pieces. A geometric constraint is represented by a family S of subsets of the plane. It is assumed that for each $i \in \{1, \dots, k\}$, $C_i \in S$. It is required that for each $i \in \{1, \dots, k+h\}$, $C'_i \in S$. Several such constraints are considered here:

(1) S is the family of path-connected sets, and the cake C is path-connected.
(2) S is the family of convex sets, and C is convex.
(3) S is the family of axes-parallel rectangles, and C is an axis-parallel rectangle.
(4) S is the family of axes-parallel rectangles, and C is a rectilinear polygon.

A-priori, one could think that a sophisticated computational-geometric algorithm is needed to find a partition with minimal h. However, I found out that the following simple algorithm provides near-optimal results:

> For each i, expand C_i arbitrarily to the largest extent possible, while keeping it contained in C and an element of S and disjoint from the other C_i-s.

I.e, it does not matter how exactly the toppings are expanded, to what direction and in what order, as long as they are expanded to the maximum possible amount. Formally, say that a sequence $C'_1, \dots, C'_k$, is *maximal*, if for each $i \in 1, \dots, k$, there is no $C^*_i \supsetneq C'_i$ such that $C^*_i \subseteq C$ and $C^*_i \in S$ and $C'_1, \dots, C'_{i-1}, C^*_i, C'_{i+1}, \dots, C'_k$ are pairwise-disjoint. The simple algorithm stated above can also be stated as:

(a) Find an arbitrary maximal set $C'_1, \dots, C'_k$ such that for all i: $C'_i \subseteq C$ and $C'_i \in S$ and $C'_i \supseteq C_i$.
(b) Partition the reminder, $C \setminus (C'_1 \cup \cdots \cup C'_k)$, to a minimal number of elements of S; let that number be h and the parts $C'_{k+1}, \dots, C'_{k+h}$.
(c) Return $C'_1, \dots, C'_{k+h}$.

Below, I analyze the performance of this algorithm for the four geometric scenarios presented above. During the analysis, for simplicity, I assume that the set $C_1, \ldots, C_k$ is already maximal. The analysis amounts to proving bounds on the number of holes when the set of toppings is maximal.

The results are summarized in the table below. The expressions are upper and lower bounds on the worst-case number of holes h, as a function on the number of toppings k and the number of reflex vertexes in the rectilinear polygon, R.

Cake C	Pieces S	Lower bound	Upper bound
Path-connected	Path-connected	0	0
Convex	Convex	$2k - O(\sqrt{k})$	$2k - 4$
Axes-parallel rectangle	Axes-parallel rectangles	$k - O(\sqrt{k})$	$k - 2$
Rectilinear polygon	Axes-parallel rectangles	$k + R - O(\sqrt{k})$	$k + R - 2$ (???)

2. PATH-CONNECTED CAKE AND PATH-CONNECTED PIECES

In this section, S is the family of path-connected sets and the cake C is also path-connected.

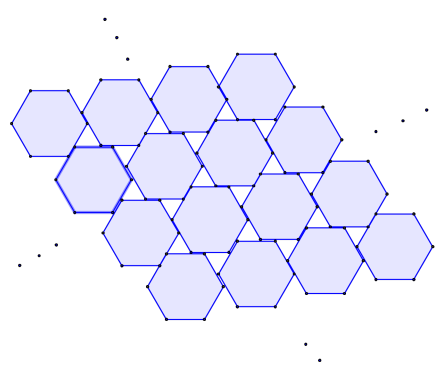
It is apparently obvious that when the set of toppings is maximal, $h = 0$ (there are no holes). Indeed, suppose there is a hole $H \subseteq C$. Since C is path-connected, H can be connected by a path to any of the C_i-s. In particular, there is some topping C_i such that a path from H to C_i through C does not pass through any other topping C_j. The union of C_i and the path and H is path-connected, contained in C and disjoint from the rest of the toppings, in contradiction to maximality.

3. CONVEX CAKE AND CONVEX PIECES

In this section, S is the family of convex sets and the cake C is also convex. Akopyan [2016] proved that the worst-case number of holes when the set of toppings is maximal is $2k - o(k)$.

3.1. Lower bound

Consider the following tiling (imagine that it fills the entire plane):



This tiling is apparently maximal - no hexagon can be expanded while remaining convex and disjoint from the other hexagons. Each hole touches three hexagons and each hexagon touches six holes. Hence, the number of holes is asymptotically twice the number of hexagons.

If the tiling is finite, then the number of holes is slightly smaller, since some holes near the boundary of C can be discarded or joined with nearby toppings. But (assuming e.g. that C is square), the number of holes near the boundary is only $O(\sqrt{k})$. Hence, in the worst case we may have $h \geq 2k - O(\sqrt{k})$ holes.

3.2. Upper bound

The proof of the upper bound is based on the following lemma, which is proved in Claim 2 of Pinchasi [2015]:

LEMMA 3.1. *When the set of toppings is maximal, every hole must be convex.*

From here, Akopyan's proof proceeds as follows.

Each side of a hole is formed by one of the C_i-s. Lets call two sets C_i and C_j *neighbors*, if they form two adjacent sides of some hole. Note that:

- To each hole correspond at least 3 neighbor-pairs - since each hole has at least 3 sides.
- To each neighbor-pair C_i, C_j correspond at most two holes - since all such holes must have a side co-linear with the segment in which the boundaries of C_i and C_j intersect.

Therefore, the number of holes is at most 2/3 the number of neighbor-pairs.

The "neighbor" relation defines a planar graph which we denote by G. Denote by V the set of vertexes of G - which are the toppings $C_1, \ldots, C_k$. Denote by E the set of edges of G - which are the neighbor-pairs defined above. Denote by F the set of faces of G. By Euler's formula: $|V| - |E| + |F| = 2$. Each face is at least a triangle so it is adjacent to at least three edges. Each edge is adjacent to at most two faces. Hence, $|F| \leq 2|E|/3$. Substituting into Euler's formula gives $|V| - |E|/3 \geq 2$ so $|E| \leq 3|V| - 6$. Hence, the number of neighbor-pairs is at most $3k - 6$ and we have at most $h \leq 2k - 4$ holes.

4. RECTANGULAR CAKE AND RECTANGULAR PIECES

In this section, S is the family of axes-parallel rectangles and the cake C is also an axes-parallel rectangle. I adapt the proofs of Pinchasi [2015] and Akopyan [2016] to prove that the worst-case number of holes when the set of toppings is maximal is $k - o(k)$.

4.1. Lower bound

Consider the following tiling, and imagine that it fills the entire 2-dimensional plane:

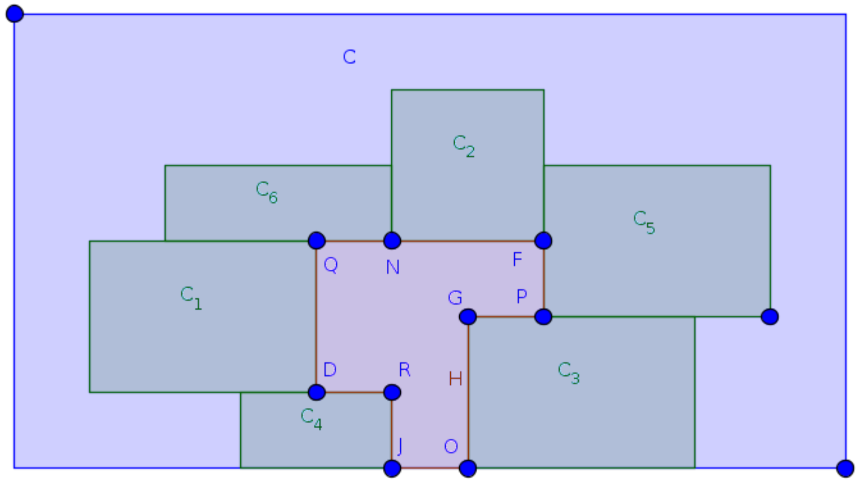
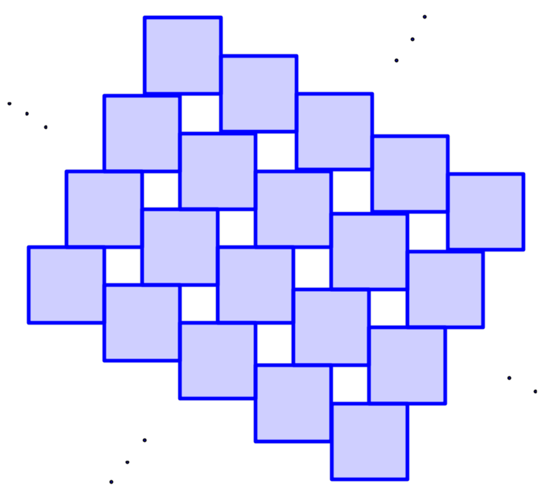


Fig. 1. A rectangular cake C, rectangular toppings C_j, and a hole H. Note the configuration is not maximal.



This set of squares is apparently maximal. Each hole touches four squares and each square touches four holes. Hence, the number of squares and holes is asymptotically the same.

If the tiling is finite, then the number of holes is slightly smaller, since some holes near the boundary of C can be discarded or joined with nearby rectangles. But, the number of holes near the boundary is only $\approx O(\sqrt{k})$. Hence, in the worst case there may be as many as $h \geq k - O(\sqrt{k})$ holes.

4.2. Upper bound part A: All holes are axes-parallel rectangles

Let H be a hole. The boundary of H is made of line-segments that are parts of the boundaries of the C_i, and possibly some line-segments that are parts of the boundary of C. The examples below refer to Figure 1.

Definition 4.1. A *vertex of* H is a point on H's boundary that is a vertex of C_i or of C or a point of contact between two adjacent C_i's.

Note that this includes the vertexes of H as a polygon, but may also include ”flat vertexes” (vertexes with angle $180°$, such as the vertex N).

Definition 4.2. An *edge of* H is a line-segment between two vertexes of H.

By our definition of the vertexes of H, it is clear that each edge of H is contained either in an edge of some topping C_i or in an edge of the cake C.

Definition 4.3. For each edge e, let $Con(e)$ be the (unique) rectangle one of whose edges contains e. This may be either some C_i or C.

Definition 4.4. Given a vertex v and an edge e of a hole H, we say that v *is exposed in* e if v is also a vertex of $Con(e)$.

For example, the vertex D is exposed in the edge DQ but not in the edge DR.

LEMMA 4.5. *If e is an edge of a hole H and $Con(e) \neq C$ (i.e, e is not part of the outer cake boundary), then e cannot have both of its vertexes exposed.*

PROOF. If $Con(e) \neq C$ then $Con(e)$ must be one of the rectangles C_i. This means that e is equal to an edge of C_i. But then, the rectangle C_i can be stretched over e and into the hole H, contradicting the maximality condition. □

For example, the edge DQ has both its vertexes exposed, and indeed the rectangle C_1 can be stretched rightwards into the hole and the configuration is not maximal.

LEMMA 4.6. *If v is a vertex of a hole H that is adjacent to two edges e, e', then v must be exposed in at least one of e, e'.*

PROOF. If $Con(e) = C$, then obviously v is exposed in e'; similarly, if $Con(e') = C$ then v is exposed in e (see e.g. vertexes J and O).

Otherwise, suppose by contradiction that v is not exposed in e nor in e'. Then v is both internal to an edge of $Con(e)$ and internal to an edge of $Con(e')$. But this implies that $Con(e)$ and $Con(e')$ overlap, contradicting the pairwise-disjointness condition. □

LEMMA 4.7. *No vertex of a hole H lies on the boundary of C.*

PROOF. By contradiction, let v be a vertex on the boundary of CS (e.g. the vertex J).

Starting from v, follow the edge that is not part of the boundary of C (e.g. JR). Continue to traverse the edges and vertexes of H in the same direction until you reach a vertex v' that lies on the boundary of C (v' may be the same as v or a different vertex, e.g. the vertex O).

In this traversal, the first vertex is exposed in the edge after it (e.g. J is exposed in JR) and the last vertex is exposed in the edge before it (e.g. O is exposed in GO). By Lemma 4.6, each edge in the path must have at least one of its vertexes exposed. Hence, by the pigeonhole principle, at least one edge in the path must have two of its vertexes exposed - a contradiction to Lemma 4.5. □

Definition 4.8. A vertex of H is called *reflex* if the internal angle adjacent to it (the angle through the interior of H) is larger than $180°$.

Since all rectangles are parallel to the axes, a reflex vertex (e.g. the vertex G) must have an internal angle of $270°$.

LEMMA 4.9. *No vertex of a hole H is reflex.*

PROOF. By contradiction, let v be a reflex vertex of H (e.g. the vertex G).

Starting from v, traverse the edges and vertexes of H until you reach v again (e.g. from G through GP to P through PF to F ... to O through OG to G). In this traversal, the first vertex v is exposed in the edge after it (e.g. G is exposed in GP) and the last vertex (which in this case is also v) is exposed in the edge before it (e.g. G is exposed in GO). By Lemma 4.6, each edge in the path must have at least one of its vertexes exposed. Hence, by the pigeonhole principle, at least one edge in the path must have two of its vertexes exposed - a contradiction to Lemma 4.5. □

COROLLARY 4.10. *Every hole H must be a rectangle and must be strictly in the interior of C (not adjacent to its boundary).*

4.3. Upper bound part B: Counting the holes

By Corollary 4.10, each hole must be a rectangle internal to the cake C. Therefore, the sides of each hole are formed by the rectangular toppings $C_1, \ldots, C_k$. Let's call two such toppings *neighbors* if they form adjacent sides of the same hole.

Each hole has 4 sides, so to each hole correspond (at least) 4 neighbor-pairs.

To each neighbor-pair C_i, C_j correspond at most two holes - since all such holes must have a side co-linear with the segment in which the boundaries of C_i and C_j intersect.

Therefore, the number of holes is at most 1/2 the number of neighbor-pairs.

The "neighbor" relation defines a planar graph which we denote by G. Denote by V the set of vertexes of G - which are the rectangles $C_1, \ldots, C_k$. Denote by E the set of edges of G - which are the neighbor-pairs defined above. Denote by F the set of faces of G. By Euler's formula: $|V|-|E|+|F| = 2$. Since all angles are multiples of $90°$, each face is adjacent to at least four edges. Each edge is adjacent to at most two faces. Hence, $|F| \leq |E|/2$. Substituting into Euler's formula gives $|V| - |E|/2 \geq 2$ so $|E| \leq 2|V| - 4$. Hence, the number of neighbor-pairs is at most $2k - 4$ and we have at most $h \leq k - 2$ holes.

5. RECTILINEAR-POLYGONAL CAKE AND RECTANGULAR PIECES

In this section, S is still the family of axes-parallel rectangles, but now the cake C can be an arbitrary axes-parallel rectilinear polygon. The "complexity" of a rectilinear polygon is characterized by the number of its *reflex vertexes* - vertexes with internal angle $270°$. It is known that a rectilinear polygon with R reflex vertexes can always be partitioned to at most $R + 1$ rectangles [Eppstein 2009; Keil 2000]. Our goal here is to bound h - the number of *rectangular* holes (rectangular pieces with no toppings) - as a function of k (the number of toppings) and R.

The bounds and the proofs are similar to those of Section 4 so I only repeat the parts that are different.

5.1. Lower bound

Consider the tiling of Subsection 4.1. Assume that the tiling is contained in a very large square, and the square is connected in a narrow pipe to a rectilinear polygon with R reflex vertexes. Then, we have $k - O(\sqrt{k})$ rectangular holes in the square. The connected rectilinear polygon is counted, not as one hole but as many holes, since the holes must be rectangular. When this rectilinear polygon is partitioned to rectangles, the partition may have as many as $R + 1$ rectangles. All in all, we may have as many as $h \geq k + R - O(\sqrt{k})$ holes.

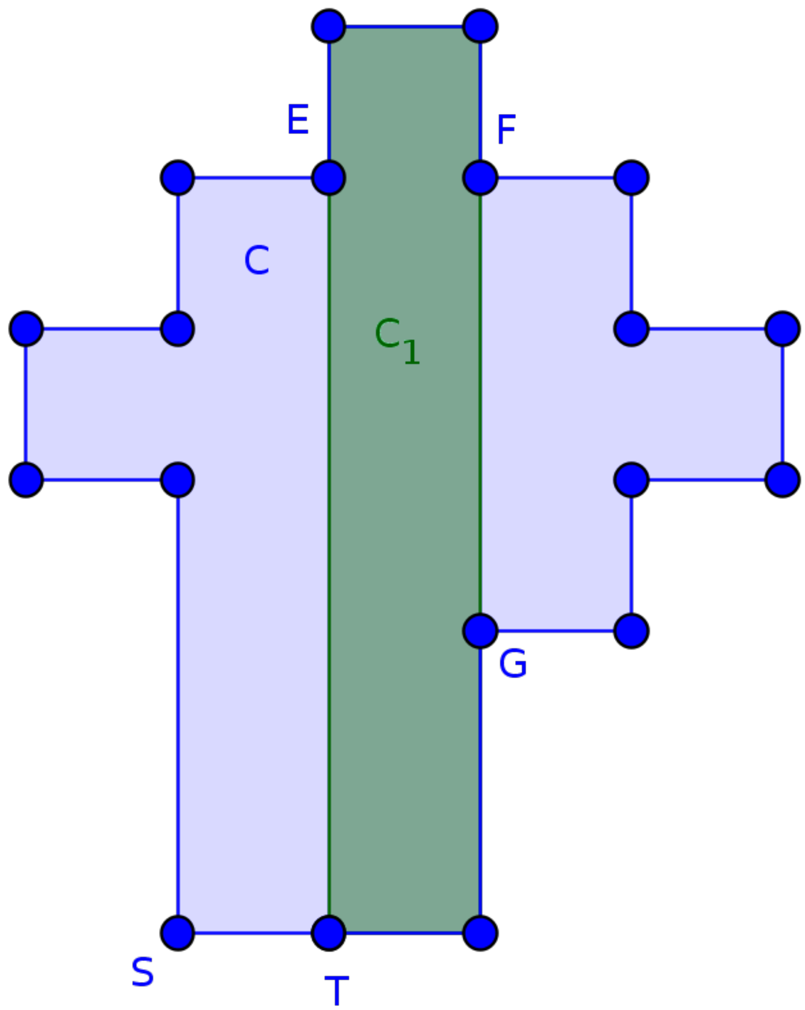


Fig. 2. A rectangular cake C, rectangular topping C_1, and two rectilinear holes. Note the configuration is maximal.

5.2. Upper bound

The main change between the current section and Subsection 4.2 is that Lemma 4.7 is no longer true, i.e, even in a maximal configuration, there may be holes adjacent to the boundary of C. See Figure 2, where there $k = 1$ topping and two holes.

However, most lemmas are still true for *internal holes* - holes that are not adjacent to C's boundary, as I explain below.

Lemma 4.5 is true as-is: *If e is an edge of a hole H and $Con(e) \neq C$ (i.e, e is not part of the outer cake boundary), then e cannot have both of its vertexes exposed.*

Lemma 4.6 becomes: *if v is a vertex of a hole H that is adjacent to two edges e, e', **and** v **is internal**, then v must be exposed in at least one of e, e'.* Because if v is internal, both $Con(e)$ and $Con(e')$ are toppings.

Instead of Lemma 4.7, we have (???): *every vertex of a hole H that is on the boundary of C, must be either an exposed vertex or a reflex vertex of C.* In Figure 2, vertex T is an exposed vertex in edge ET while vertexes E, F and G are reflex vertexes of C.

Moreover, this implies that, in every hole H, there is at least one vertex that is a reflex vertex of C, and it is a convex (non-reflex) vertex of H. (???)

Lemma 4.9 becomes: *No vertex of an **internal** hole H is reflex.*

Corollary 4.10 becomes: *Every **internal** hole H must be a rectangle.*

The upper bound in Subsection 4.3, namely $k - 2$, is a valid upper bound on the number of **internal** holes. To bound the total number of holes, we must also partition the "boundary holes" to rectangles. As mentioned above, in every boundary-hole H, there is at least one vertex that is a reflex vertex of C, and it is a convex (non-reflex) vertex of H. (???) Hence, the total number of reflex vertexes in all boundary-holes is at most R minus the number of boundary-holes. Hence, the boundary-holes can be partitioned to at most R rectangles. All in all, we have at most: $h \leq k + R - 2$ holes.

ACKNOWLEDGMENTS

I am grateful to Arseniy Akopyan, whose answer to my question in Math Overflow [Akopyan 2016] lead me to writing this paper.

REFERENCES

Arseniy Akopyan. 2016. Partitioning a convex object without harming existing convex subsets. Math Overflow. (29 Feb. 2016). http://mathoverflow.net/q/232322/34461 URL:http://mathoverflow.net/a/232445/34461 (version: 2016-03-14).

David Eppstein. 2009. Graph-Theoretic Solutions to Computational Geometry Problems. (26 Aug. 2009). DOI:http://dx.doi.org/10.1007/978-3-642-11409-0_1

J. Mark Keil. 2000. Polygon Decomposition. In *Handbook of Computational Geometry*. University of Saskatchewan Saskatoon, Sask., Canada, 491–518. DOI:http://dx.doi.org/10.1016/B978-044482537-7/50012-7

Rom Pinchasi. 2015. On the perimeter of k pairwise disjoint convex bodies contained in a convex set in the plane. *Combinatorica* (23 July 2015), 1–27. DOI:http://dx.doi.org/10.1007/s00493-015-3217-5