

We begin by showing that the process of converting an arbitrary polygonal subdivision into a trapezoidal decomposition increases its size by at most a constant factor. We derive the exact expansion factor in the next claim.

Lemma: Given set S of n line segments in the plane, the resulting trapezoidal map $\mathcal{T}(S)$ has at most $6n + 4$ vertices and $3n + 1$ trapezoids.

Proof: Each segment generates six vertices in the map, two for the segment's endpoints, two from the upward vertical rays, and two from the downward vertical rays (see Fig. 2(a)). In addition, there are four vertices for the bounding rectangle, resulting in a total of $6n + 4$ vertices.

Let us charge each trapezoid to the vertex along its left edge. Each segment is charged by three trapezoids, since the left endpoint of each segment is charged twice, one from the trapezoid above-right and one from the trapezoid below-right, and the right endpoint is charged once by the trapezoid to its right (see Fig. 2(b)). This yields a total of $3n$ trapezoids. There is one more trapezoid, namely the leftmost one, for a total of $3n + 1$.

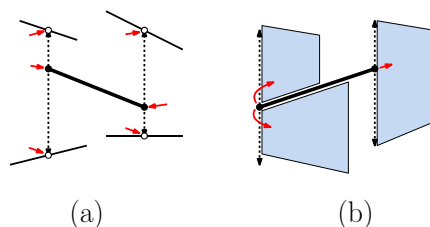


Fig. 2: (a) The six vertices associated with each segment and (b) the three left-bounding trapezoids associated with each segment.

An important fact to observe about each trapezoid is that its existence is *determined* by exactly four entities from the original subdivision: a segment on top, a segment on the bottom, a bounding vertex on the left, and a bounding vertex on the right. The bounding vertices may be endpoints of the upper or lower segments, or they may belong to completely different segments. This simple observation will play an important role later in the analysis.

Construction: We could construct the trapezoidal map by a straightforward application of plane sweep. (By now, this should be an easy exercise for you. You might think about how you would do it.) Instead, we will build the trapezoidal map by a different approach, namely a *randomized incremental algorithm*.¹

The incremental algorithm starts with the initial bounding rectangle (that is, one trapezoid) and then we add the segments of the polygonal subdivision one by one in random order. As each segment is added, we update the trapezoidal map. Let S_i denote the subset consisting of the first i (randomly permuted) segments, and let \mathcal{T}_i denote the resulting trapezoidal map.

¹Historically, the randomized incremental algorithm that we will discuss arose as a method for solving a more general problem, namely computing the intersection of a collection of line segments. Given n line segments that have I intersections, this algorithm runs in $O(I + n \log n)$ time, which is superior to the plane sweep algorithm we discussed earlier. The original algorithm is due to Ketan Mulmuley.

To perform this update, we need to know which trapezoid of the current map contains the left endpoint of the newly added segment. We will address this question later when we discuss point location. We then trace the line segment from left to right, by “walking” it through the existing trapezoidal map (see Fig. 3). Along the way, we discover which existing trapezoids it intersects. We go back to these trapezoids and “fix them up”. There are two things that are involved in fixing process.

- The left and right endpoints of the new segment need to have bullets fired from them.
- One of the earlier created walls might hit the new line segment. When this happens the wall is trimmed back. (We store which vertex shot the bullet path for this wall, so we know which side of the wall to trim.)

The process is illustrated in Fig. 3, where we insert a new segment (red) into the trapezoidal map from Fig. 1.

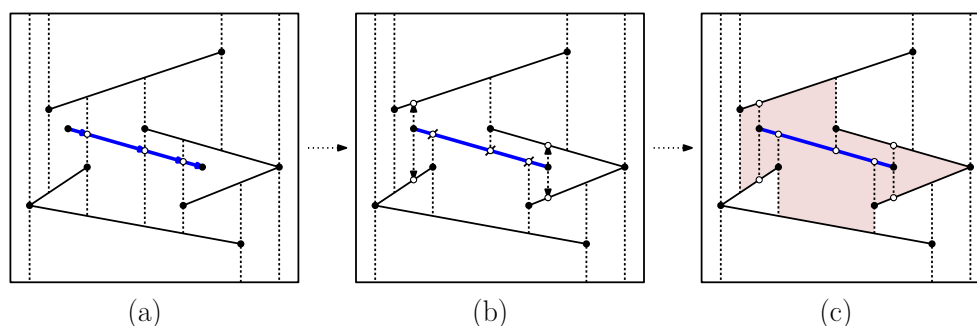


Fig. 3: Inserting a segment into the trapezoidal map: (a) Locate the left endpoint and trace the segment through trapezoids, (b) shoot bullet paths from endpoints and trim walls that have been crossed, (c) four original trapezoids have been replaced by seven new trapezoids (shaded).

Observe that the structure of the trapezoidal decomposition does *not* depend on the order in which the segments are added. (This fact will be exploited later in the running time analysis, and it is one of the reasons that trimming back the walls is so important.) The following is also important to the analysis.

Lemma: Ignoring the time spent to locate the left endpoint of an segment, the time that it takes to insert the i th segment and update the trapezoidal map is $O(k_i)$, where k_i denotes the number of newly created trapezoids.

Proof: Consider the insertion of the i th segment, and let w_i denote the number of existing walls that this segment intersects. We need to shoot four bullets (two from each endpoint) and then trim each of the w_i walls, for a total of $w_i + 4$ operations that need to be performed. If the new segment did not cross any of the walls, then we would get exactly four new trapezoids. For each of the w_i walls we cross, we add one more to the number of newly created trapezoids, for a total of $w_i + 4$. Thus, letting $k_i = w_i + 4$ be the number of trapezoids created, the number of update operations is exactly k_i . Each of these operations can be performed in $O(1)$ time given any reasonable representation

of the trapezoidal map as a planar subdivision, for example, a doubly connected edge list (DCEL).

Analysis: We will analyze the expected time to build the trapezoidal map, assuming that segments are inserted in random order. (Note that we make no assumptions about the spatial distribution of the segments, other than the fact they do not intersect.) Clearly, the running time depends on how many walls are trimmed with each intersection. In the worst case, each newly added segment could result in $\Omega(n)$ walls being trimmed, and this would imply an $\Omega(n^2)$ running time. We will show, however, that the expected running time is much smaller, in fact, we will show the rather remarkable fact that, each time we insert a new segment, the expected number of wall trimmings is just $O(1)$. (This is quite surprising at first. If many of the segments are long, it might seem that every insertion would cut through $O(n)$ trapezoids. What saves us is that, although a long segment might cut through many trapezoids, it shields later segments from cutting through many trapezoids.) As was the case in our earlier lecture on linear programming, we will make use of a backwards analysis to establish this result.

There are two things that we need to do when each segment is inserted. First, we need to determine which cell of the current trapezoidal map contains its left endpoint. We will not discuss this issue today, but in our next lecture, we will show that the expected time needed for this operation is $O(n \log n)$. Second, we need to trim the walls that are intersected by the new segment. The remainder of this lecture will focus on this aspect of the running time.

From the previous claim, we know that it suffices to count the number of new trapezoids created with each insertion. The main result that drives the analysis is presented in the next lemma.

Lemma: Consider the randomized incremental construction of a trapezoidal map, and let k_i denote the number of new trapezoids created when the i th segment is added. Then $E[k_i] = O(1)$, where the expectation is taken over all possible permutations of the segments as the insertion orders.

Proof: The analysis will be based on a backwards analysis. Recall that such an analysis involves analyzing the expected value assuming that the last insertion was random.

Let \mathcal{T}_i denote the trapezoidal map resulting after the insertion of the i th segment. Because we are averaging over all permutations, among the i segments that are present in \mathcal{T}_i , each one has an equal probability $1/i$ of being the last one to have been added. For each of the segments s we want to count the number of trapezoids that would have been created, had s been the last segment to be added.

We say that a trapezoid Δ of the existing map *depends* on a segment s , if s would have caused Δ to be created had s been the *last* segment to be inserted. (For example, in Fig. 4(a), the shaded trapezoids depend on s , and none of the others do.) We want to count the number of trapezoids that depend on each segment, and then compute the average over all segments. If we let $\delta(\Delta, s) = 1$ if segment Δ depends on s , and 0

otherwise, then the expected value is

$$\begin{aligned} E[k_i] &= \sum_{s \in S_i} (\text{prob. that } s \text{ is last}) (\text{num. of trapezoids that depend on } s) \\ &= \sum_{s \in S_i} \frac{1}{i} \left(\sum_{\Delta \in \mathcal{T}_i} \delta(\Delta, s) \right) = \frac{1}{i} \sum_{s \in S_i} \sum_{\Delta \in \mathcal{T}_i} \delta(\Delta, s). \end{aligned}$$

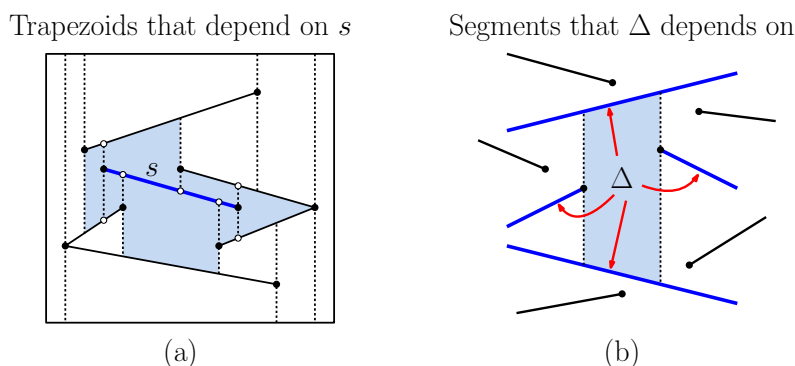


Fig. 4: Trapezoid-segment dependencies.

Some segments might have resulted in the creation of lots of trapezoids and others would have resulted in very few. How can we analyze such an unruly quantity? The trick is, rather than counting the number of trapezoids that depend on each segment, we swap the order of the two summations and instead count the number segments upon which each trapezoid depends. In other words we can express the above quantity as:

$$E[k_i] = \frac{1}{i} \sum_{\Delta \in \mathcal{T}_i} \sum_{s \in S_i} \delta(\Delta, s).$$

This quantity is much easier to analyze. In particular, each trapezoid is bounded by at most four sides. (The reason it is “at most” is that degenerate trapezoids are possible which may have fewer sides.) The top and bottom sides are each determined by a segment of S_i , and clearly if either of these was the last to be added, then this trapezoid would have come into existence as a result. The left and right sides are each determined by an endpoint of a segment in S_i , and clearly if either of these was the last to be added, then this trapezoid would have come into existence.²

In summary, each of the decomposition trapezoid is dependent on at most four segments, which implies that $\sum_{s \in S_i} \delta(\Delta, s) \leq 4$. Since \mathcal{T}_i consists of at most $3i + 1$ trapezoids we have

$$E[k_i] \leq \frac{1}{i} \sum_{\Delta \in \mathcal{T}_i} 4 = \frac{4}{i} |\mathcal{T}_i| \leq \frac{4}{i} (3i + 1) = O(1).$$

²There is a bit of a subtlety here. What if multiple segments share the endpoint? Note that the trapezoid is only dependent on the first such segment to be added, since this is the segment that caused the vertex to come into existence. Also note that the same segment that forms the top or bottom side might also provide the left or right endpoint. These considerations only decrease the number of segments on which a trapezoid depends.

We know that the total number of trapezoids in the end is at most $3n + 1 = O(n)$. Since the expected number of new trapezoids created with each insertion is $O(1)$, it follows that the total number of trapezoids that are created (and perhaps destroyed) throughout the entire process is $O(n)$. This fact is important in bounding the total time needed for the randomized incremental algorithm.

The only question that we have not considered in the construction is how to locate the trapezoid that contains left endpoint of each newly added segment. We will consider this question, and the more general question of how to do point location in our next lecture.