CMSC 754: Lecture 2
Convex Hulls in the Plane

Reading: Some of the material of this lecture is covered in Chapter 1 in the "4M's" book (by de Berg, Cheong, van Kreveld, and Overmars). The divide-and-conquer algorithm is given in Joseph O'Rourke's, "Computational Geometry in C."

Convex Hulls: In this lecture we will consider a fundamental structure in computational geometry, called the *convex hull*. We will give a more formal definition later, but, given a set P of points in the plane, the convex hull of P, denoted $\text{conv}(P)$, can be defined intuitively by surrounding a collection of points with a rubber band and then letting the rubber band "snap" tightly around the points (see Fig. 1).

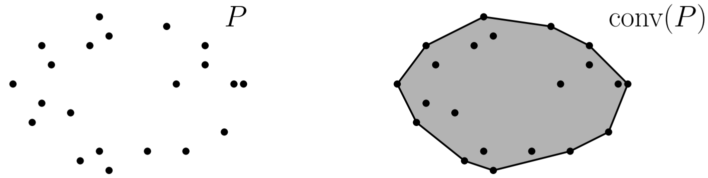

Fig. 1: A point set and its convex hull.

The (planar) *convex hull problem* is, given a discrete set of n points P in the plane, output a representation of P's convex hull. The convex hull is a closed convex polygon, the simplest representation is a counterclockwise enumeration of the vertices of the convex hull. In higher dimensions, the convex hull will be a convex polytope. We will discuss the representation of polytopes in future lectures, but in 3-dimensional space, the representation would consist of a vertices, edges, and faces that constitute the boundary of the polytope.

There are a number of reasons that the convex hull of a point set is an important geometric structure. One is that it is one of the simplest shape approximations for a set of points. (Other examples include minimum area enclosing rectangles, circles, and ellipses.) It can also be used for approximating more complex shapes. For example, the convex hull of a polygon in the plane or polyhedron in 3-space is the convex hull of its vertices.

Also many algorithms compute the convex hull as an initial stage in their execution or to filter out irrelevant points. For example, the *diameter* of a point set is the maximum distance between any two points of the set. It can be shown that the pair of points determining the diameter are both vertices of the convex hull. Also observe that minimum enclosing convex shapes (such as the minimum area rectangle, circle, and ellipse) depend only on the points of the convex hull.

Basic Definitions: Before getting to the discussion of the various convex-hull algorithms, let's begin with a few standard definitions, which will be useful throughout the semester. For any $d \geq 1$, let $\mathbb{R}^d$ denote real d-dimensional space, that is, the set of d-dimensional vectors over the real numbers. We refer to elements of $\mathbb{R}^d$ either as "points" or "vectors", depending on what we intend them to represent (a location in space or a displacement, respectively). We refer to real numbers as *scalars*.

A point/vector $p \in \mathbb{R}^d$ is expressed as a d-vector $(p_1, \ldots, p_d)$, where $p_i \in \mathbb{R}$. Following standard terminology from linear algebra, given two vectors $u, v \in \mathbb{R}^d$ and a scalar $\alpha \in \mathbb{R}$, let "$u + v$" be the vector-valued sum, αv be scalar-vector product, and let "$u \cdot v$" denote the standard scalar-valued dot-product.

Affine and convex combinations: Given two points $p = (p_x, p_y)$ and $q = (q_x, q_y)$ in $\mathbb{R}^d$, we can express any point on the (infinite) line $\overleftrightarrow{pq}$ as a linear combination of their coordinates, where the coefficient sum to 1:

$$(1-\alpha)p + \alpha q \;=\; ((1-\alpha)p_x + \alpha q_x, (1-\alpha)p_y + \alpha q_y).$$

This is called an *affine combination* of p and q (see Fig. 2(a)).

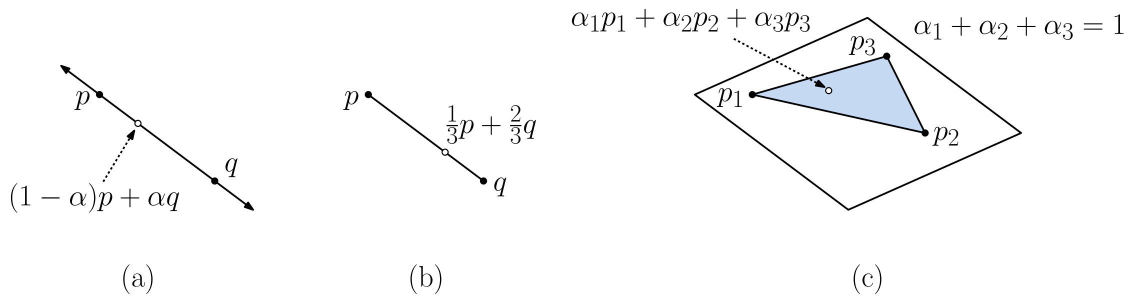

Fig. 2: Affine and convex combinations.

By adding the additional constraint that $0 \leq \alpha \leq 1$, the set of points generated lie on the *line segment* $\overline{pq}$ (see Fig. 2(b)). This is called a *convex combination*. Notice that this can be viewed as taking a weighted average of p and q. As α approaches 1, the point lies closer to p and when α approaches zero, the point lies closer to q.

It is easy to extend both types of combinations to more than two points. For example, given k points $\{p_1, \ldots, p_k\}$ an affine combination of these points is the linear combination

$$\sum_{i=1}^{k} \alpha_i p_i, \qquad \text{such that } \alpha_1 + \cdots + \alpha_k = 1.$$

When $0 \leq \alpha_i \leq 1$ for all i, the result is called a *convex combination*.

The set of all affine combinations of three (non-collinear) points generates a plane, and generally, the resulting set is called the *affine span* or *affine closure* of the points.

Hyperplanes/Halfspaces: Given a nonzero vector $v \in \mathbb{R}^d$ and a scalar $\alpha \in \mathbb{R}$, the set of points $\{p \mid p \cdot v = \alpha\}$ is a $(d-1)$-dimensional affine subspace, more often called a *hyperplane*. In the special cases $\mathbb{R}^2$ and $\mathbb{R}^3$, this defines a *line* and *plane*, respectively. If v is a unit vector and $\alpha \geq 0$, this hyperplane is orthogonal to v and lies at distance α from the origin (see Fig. 3(a)). If we change the equality to an inequality, we obtain a *halfspace* consisting of points lying on one side of the hyperplane, for example $\{p \mid p \cdot v \leq \alpha\}$.

Concepts from Topology: There are a number of useful concepts that arise from topology: open and closed sets, interior, exterior, and boundary, connectivity, etc. We will use these without definitions, since for the simple objects that we will be working with,

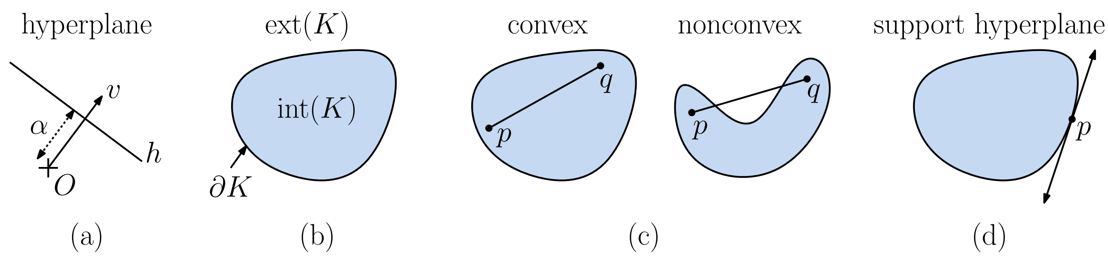

Fig. 3: Basic concepts.

these concepts are easily understood at an intuitive level.[1] We will use $\mathrm{int}(K)$, $\mathrm{ext}(K)$, and ∂K to denote the interior, exterior, and boundary of a set K, respectively (see Fig. 3(b))).

Convexity: A set $K \subseteq \mathbb{R}^d$ is *convex* if given any points $p, q \in K$, the line segment $\overline{pq}$ is entirely contained within K (see Fig. 3(c)). Otherwise, it is called *nonconvex*. This is equivalent to saying that K is "closed" under convex combinations. Examples of convex sets in the plane include circular disks (the set of points contained within a circle), the set of points lying within any regular n-sided polygon, lines (infinite), line segments (finite), rays, and halfspaces.

Support line/hyperplane: An important property of any convex set K in $\mathbb{R}^d$ is that at every point p on the boundary of K, there exists a (not necessarily unique) hyperplane h that passes through p such that K lies entirely in one of the closed halfspaces defined by h (see Fig. 3(d)). This is called a *support hyperplane* for K. (Again, in $\mathbb{R}^2$ and $\mathbb{R}^3$ we would call this a *support line* or *support plane*.) Observe that there may generally be multiple support hyperplanes passing through a given boundary point of K (e.g., when the point is a vertex of the convex hull).

Convex Hulls: Formally, the convex hull of a point set P is defined to be smallest convex set containing P. It can be characterized in two provably equivalent ways (one additive and one subtractive). First, it is equal to the set of all convex combinations of points in P and second, it is equal to the intersection of all halfspaces that contain P.

When computing convex hulls, we will usually take P to be a finite set of points. In such a case, $\mathrm{conv}(P)$ will be a convex polygon. Generally P could be an infinite set of points. For example, we could talk about the convex hull of a collection of circles. The boundary of such a shape would consist of a combination of circular arcs and straight line segments.

General Position: As in many of our algorithms, it will simplify the presentation to avoid lots of special cases by assuming that the points are in *general position*. This effectively means that "degenerate" geometric configurations (e.g., two points sharing the same x or y coordinate, three points being collinear, four points being cocircular, etc.) do not arise in the input.

To motivate this concept, suppose you design an algorithm that sorts the input points according to their x-coordinates, and then it visits them in sorted order. In the rare occasion where two points share the same x-coordinate, which point is visited first? While we could

[1] See, for example https://en.wikipedia.org/wiki/Boundary_(topology) for definitions.

take the effort to spell this out, it will be simpler to just add an assumption that this never happens, that is, there are no duplicate x-coordinates in the input set. This assumption is not unreasonable, since any random (infinitesimal) perturbation of the input coordinates will satisfy this condition with extremely high probability. This is an example of a general-position assumption.

General position assumptions can be tricky to define formally, since the definition of "degeneracy" depends on the algorithm itself. In our previous example, if our algorithm had instead sorted points in angular order about the origin, we might wish to rule our duplicate angles.

In summary the assumption of *general position* is your license to ignore special configurations, which could be easily eliminated by a random perturbation of the data set.

Graham's Scan: We will begin with a presentation of a simple $O(n \log n)$ algorithm for the convex hull problem. It is a simple variation of a famous algorithm for convex hulls, called *Graham's Scan*, which dates back to the early 1970's (and named for its inventor Ronald Graham). The algorithm is loosely based on a common approach for building geometric structures called *incremental construction*. In such a algorithm object (points here) are added one at a time, and the structure (convex hull here) is updated with each new insertion.

Let us assume that the input consists of a set P of n points $\{p_1, \ldots, p_n\}$, where $p_i = (x_i, y_i)$. While we will not do so, it is easy to prove that the convex hull of a finite set of points is a convex polygon, whose vertices are a subset of P. A natural representation of such a polygon is a cyclic (for example, counterclockwise) listing of the vertices of this polygon.

An important issue with incremental algorithms is the order of insertion. If we were to add points in some arbitrary order, we would need some method of testing whether the newly added point is inside the existing hull. Instead, we will insert points in increasing order of x-coordinates. This guarantees that each newly added point is outside the current hull. (Note that Graham's original algorithm sorted points in a different way. It found the lowest point in the data set and then sorted points cyclically around this point. Sorting by x-coordinate seems to be a bit easier to implement, however.)

Since we are working from left to right, it would be convenient if the convex hull vertices were themselves ordered from left to right. To do this, we can break the boundary of the convex hull into two *chains*, an *upper chain* consisting of the vertices along the upper part of the hull and a *lower chain* consisting of the vertices along the lower part of the hull. Both chains will start with the leftmost point of P and will end with the rightmost point of P. We will make the general-position assumption that no two vertices have the same x-coordinates, so these two points are unique (see Fig. 4(a)).

It suffices to show how to compute the upper hull, since the lower hull is symmetrical. (Just flip the picture upside down.) Once the two hulls have been computed, we can simply concatenate them with the reversal of the other to form the final hull.

Observe that a point $p \in P$ lies on the upper hull if and only if there is a support line passing through p such that all the points of P lie on or below this line. Our algorithm will be based on the following lemma, which characterizes the upper hull of P. This is a simple consequence of the convexity. The first part says that the line passing through each edge of the hull is a support line, and the second part says that as we walk from right to left along the upper hull, we make successive left-hand turns (see Fig. 4(b)).

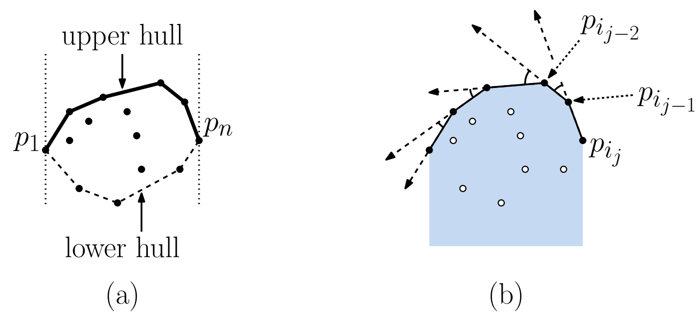

Fig. 4: (a) Upper and lower hulls and (b) the left-hand turn property of points on the upper hull.

Lemma 1: Let $\langle p_{i_1}, \ldots, p_{i_m} \rangle$ denote the vertices of the upper hull of P, sorted from left to right. Then for $1 \leq j \leq m$, (1) all the points of P lie on or below the line $\overline{p_{i_j} p_{i_{j-1}}}$ joining consecutive vertices and (2) each consecutive triple $\langle p_{i_j} p_{i_{j-1}} p_{i_{j-2}} \rangle$ forms a left-hand turn.

Let $\langle p_1, \ldots, p_n \rangle$ denote the sequence of points sorted by increasing order of x-coordinates. For i ranging from 1 to n, let $P_i = \langle p_1, \ldots, p_i \rangle$. We will store the vertices of the upper hull of P_i on a stack S, where the top-to-bottom order of the stack corresponds to the right-to-left order of the vertices on the upper hull. Let $S[t]$ denote the stack's top. Observe that as we read the stack elements from top to bottom (that is, from right to left) consecutive triples of points of the upper hull form a (strict) left-hand turn (see Fig. 4(b)). As we push new points on the stack, we will enforce this property by popping points off of the stack that violate it.

Turning and orientations: Before proceeding with the presentation of the algorithm, we should first make a short digression to discuss the question of how to determine whether three points form a "left-hand turn." This can be done by a powerful primitive operation, called an *orientation test*, which is fundamental to many algorithms in computational geometry.

Given an ordered triple of points $\langle p, q, r \rangle$ in the plane, we say that they have *positive orientation* if they define a counterclockwise oriented triangle (see Fig. 5(a)), *negative orientation* if they define a clockwise oriented triangle (see Fig. 5(b)), and *zero orientation* if they are collinear, which includes as well the case where two or more of the points are identical (see Fig. 5(c)). Note that orientation depends on the order in which the points are given.

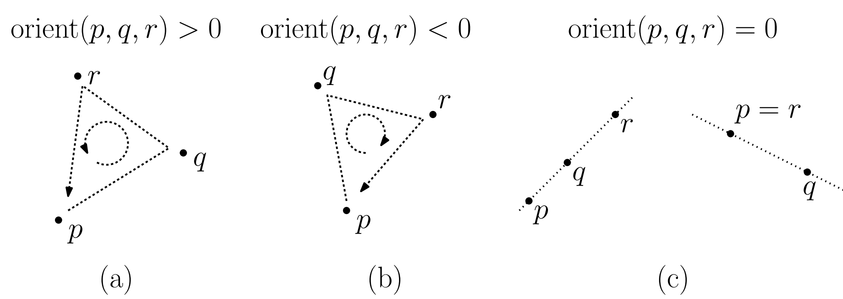

Fig. 5: Orientations of the ordered triple (p, q, r).

Orientation is formally defined as the sign of the determinant of the points given in homogeneous coordinates, that is, by prepending a 1 to each coordinate. For example, in the plane, we define

$$\text{orient}(p,q,r) = \text{sign}\left(\det\begin{pmatrix} 1 & p_x & p_y \\ 1 & q_x & q_y \\ 1 & r_x & r_y \end{pmatrix}\right).$$

Observe that in the 1-dimensional case, $\text{orient}(p,q)$ is just $q-p$. Hence it is positive if $p < q$, zero if $p = q$, and negative if $p > q$. Thus orientation generalizes the familiar 1-dimensional binary relations $<, =, >$.

Also, observe that the sign of the orientation of an ordered triple is unchanged if the points are translated, rotated, or scaled (by a positive scale factor). A reflection transformation (e.g., $f(x,y) = (-x,y)$) reverses the sign of the orientation. In general, applying any affine transformation to the point alters the *sign* of the orientation according to the *sign* of the determinant of the matrix used in the transformation. (By the way, the notion of orientation can be generalized to $d+1$ points in d-dimensional space, and is related to the notion of *chirality* in Chemistry and Physics. For example, in 3-space the orientation is positive if the point sequence defines a right-handed screw.)

Given a sequence of three points p, q, r, we say that the sequence $\langle p,q,r\rangle$ makes a (strict) *left-hand turn* if $\text{orient}(p,q,r) > 0$.

Graham's Scan Details: We can now present the full algorithm. Let us consider just the case of the upper hull, and let's see what happens when we process the insertion of the ith point, p_i (see Fig. 6(a)). First observe that p_i is on the upper hull of P_i (since it is the rightmost point seen so far). Let p_j be its predecessor on the upper hull of P_i. We know from Lemma 1 that all the points of P_i lie on or below the line $\overline{p_i p_j}$. Let p_{j-1} be the point immediately preceding p_j on the upper hull. We also know from this lemma that $\langle p_i p_j p_{j-1}\rangle$ forms a left-hand turn. Clearly then, if any triple $\langle p_i, S[t], S[t-1]\rangle$ does *not* form a left-hand turn (that is, $\text{orient}(p_i, S[t], S[t-1]) \leq 0$), we may infer that $S[t]$ is *not* on the upper hull, and hence it is safe to delete it by popping it off the stack. We repeat this until we find a left-turning triple (see Fig. 6(b)) or hitting the bottom of the stack. Once this happens, we push p_i on top of the stack, making it the rightmost vertex on the upper hull (see Fig. 6(c)). The algorithm is presented in the code block below.

Graham's Scan

(1) Sort the points according to increasing order of their x-coordinates, denoted $\langle p_1, p_2, \ldots, p_n\rangle$.

(2) push p_1 and then p_2 onto S.

(3) for $i \leftarrow 3, \ldots, n$ do:

   (a) while ($|S| \geq 2$ and $\text{orient}(p_i, S[t], S[t-1]) \leq 0$) pop S.

   (b) push p_i onto S.

Correctness: We will not give a formal proof of the correctness of Graham's Scan, but here is the gist of it. The correctness follows from the following invariant. Let $U(i)$ denote the sequence of vertices forming the upper hull of $\{p_1, \ldots, p_i\}$.

Lemma: Following step i, the stack contains (from bottom to top order) $U(i)$.

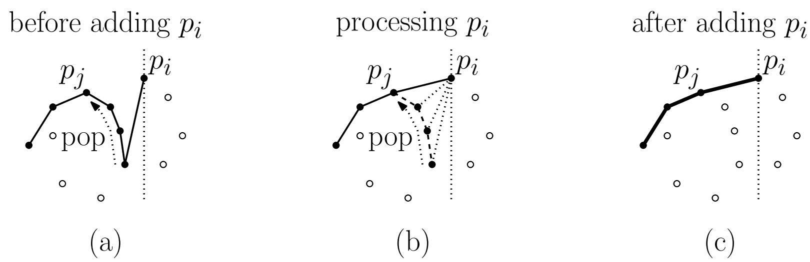

Fig. 6: Graham's scan.

The proof of this lemma naturally involves induction. The basis of induction is trivial, since clearly $U(2)$ just consists of p_1 and p_2. We assume inductively that $U(i-1)$ was correctly computed after stage $i-1$. As we argued before, p_i must be the final vertex of $U(i)$ (since it has the largest x-coordinate). Let p_j denote the vertex that immediately precedes p_i on $U(i)$. We need to argue (1) p_j is on $U(i-1)$ (and hence it is in the stack at the start of stage i), (2) all the vertices following p_j on $U(i-1)$ will be popped from the stack, and (3) the popping process will end when it reaches p_j. We will leave the details as an exercise.

Running-time analysis: We will show that Graham's algorithm runs in $O(n \log n)$ time. Clearly, it takes this much time for the initial sorting of the points. After this, we will show that $O(n)$ time suffices for the rest of the computation.

Let d_i denote the number of points that are popped (deleted) on processing p_i. Because each orientation test takes $O(1)$ time, the amount of time spent processing p_i is $O(d_i + 1)$. (The extra $+1$ is for the last point tested, which is not deleted.) Thus, the total running time is proportional to

$$\sum_{i=1}^{n}(d_i + 1) \;=\; n + \sum_{i=1}^{n} d_i.$$

To bound $\sum_i d_i$, observe that each of the n points is pushed onto the stack once. Once a point is deleted it can never be deleted again. Since each of n points can be deleted at most once, $\sum_i d_i \leq n$. Thus after sorting, the total running time is $O(n)$. Since this is true for the lower hull as well, the total time is $O(2n) = O(n)$.

Convex Hull by Divide-and-Conquer: As with sorting, there are many different approaches to solving the convex hull problem for a planar point set P. Next, we will consider another $O(n \log n)$ algorithm, which is based on divide-and-conquer. It can be viewed as a generalization of the well-known MergeSort sorting algorithm (see any standard algorithms text). Here is an outline of the algorithm. As with Graham's scan, we will focus just on computing the upper hull, and the lower hull will be computed symmetrically.

The algorithm begins by sorting the points by their x-coordinate, in $O(n \log n)$ time. In splits the point set in half at its median x-coordinate, computes the upper hulls of the left and right sets recursively, and then merges the two upper hulls into a single upper hull. This latter process involves computing a line, called the *upper tangent*, that is a line of support for both hulls. The remainder of the algorithm is shown in the code section below.

Divide-and-Conquer (Upper) Convex Hull

(1) If $|P| \leq 3$, then compute the upper hull by brute force in $O(1)$ time and return.

(2) Otherwise, partition the point set P into two sets P' and P'' of roughly equal sizes by a vertical line.

(3) Recursively compute upper convex hulls of P' and P'', denoted H' and H'', respectively (see Fig. 7(a)).

(4) Compute the upper tangent $\ell = \overline{p'p''}$ (see Fig. 7(b)).

(5) Merge the two hulls into a single upper hull by discarding all the vertices of H' to the right of p' and the vertices of H'' to the left of p'' (see Fig. 7(c)).

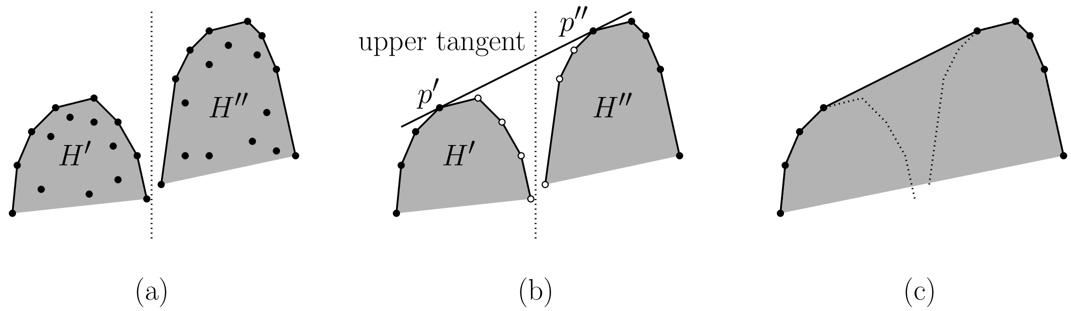

Fig. 7: Divide and conquer (upper) convex hull algorithm.

Computing the upper tangent: The only nontrival step is that of computing the common tangent line between the two upper hulls. Our algorithm will exploit the fact that the two hulls are separated by a vertical line. The algorithm operates by a simple "walking procedure." We initialize p' to be the rightmost point of H' and p'' to be the leftmost point of H'' (see Fig. 8(a)). We will walk p' backwards along H' and walk p'' forwards along H'' until we hit the vertices that define the tangent line. As in Graham's scan, it is possible to determine just how far to walk simply by applying orientation tests. In particular, let q' be the point immediately preceding p' on H', and let q'' be the point immediately following p'' on H''. Observe that if $\text{orient}(p', p'', q'') \geq 0$, then we can advance p'' to the next point along H'' (see Fig. 8(a)). Symmetrically, if $\text{orient}(p'', p', q') \leq 0$, then we can advance p' to its predecessor along H' (see Fig. 8(b)). When neither of these conditions applies, that is, $\text{orient}(p', p'', q'') < 0$ and $\text{orient}(p'', p', q') > 0$, we have arrived at the desired points of mutual tangency (see Fig. 8(c)).

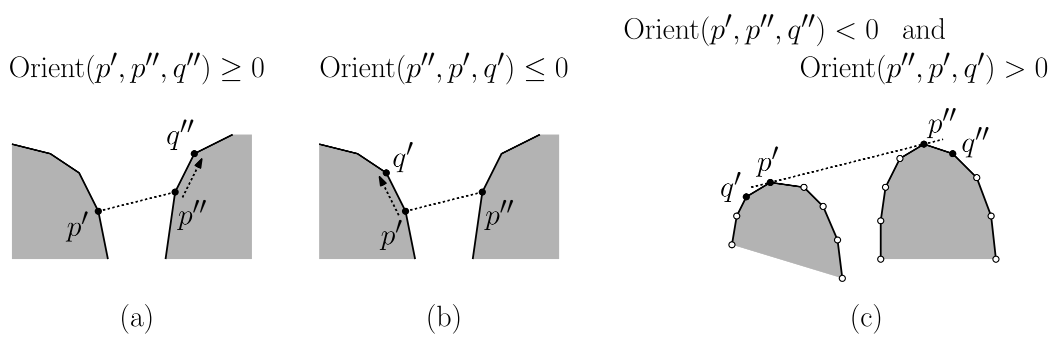

Fig. 8: Computing the upper tangent.

There is one rather messy detail in implementing this algorithm. This arises if either q' or q'' does not exist because we have arrived at the leftmost vertex of H' or the rightmost vertex of H''. We can avoid having to check for these conditions by creating two *sentinel points*. We create a new leftmost vertex for H' that lies infinitely below its original leftmost vertex, and we create a new rightmost vertex for H'' that lies infinitely below its original rightmost vertex. The tangency computation will never arrive at these points, and so we do not need to add a special test for the case when q' and q'' do not exist. The algorithm is presented in the following code block.

Computing the Upper Tangent

UpperTangent(H', H'') :

(1) Let p' be the rightmost point of H', and let q' be its predecessor.
(2) Let p'' be the leftmost point of H'', and let q'' be its successor.
(3) Repeat the following until orient$(p', p'', q'') < 0$ and orient$(p'', p', q') > 0$:
  (a) while (orient$(p', p'', q'') \geq 0$) advance p'' and q'' to their successors on H''.
  (b) while (orient$(p'', p', q') \leq 0$) advance p' and q' to their predecessors on H'.
(4) return (p', p'').

Running-time analysis: The asymptotic running time of the algorithm can be expressed by a recurrence. Given an input of size n, consider the time needed to perform all the parts of the procedure, ignoring the recursive calls. This includes the time to partition the point set, compute the upper tangent line, and return the final result. Clearly, each of these can be performed in $O(n)$ time, assuming any standard list representation of the hull vertices. Thus, ignoring constant factors, we can describe the running time by the following recurrence:

$$T(n) = \begin{cases} 1 & \text{if } n \leq 3 \\ n + 2T(n/2) & \text{otherwise.} \end{cases}$$

This is the same recurrence that arises in Mergesort. It is easy to show that it solves to $T(n) \in O(n \log n)$ (see any standard algorithms text).