CMSC 754: Lecture 6
Halfplane Intersection and Point-Line Duality

Reading: Chapter 4 in the 4M's, with some elements from Sections 8.2 and 11.4.

Halfplane Intersection: Today we begin studying another fundamental topic in geometric computing and convexity. Recall that any line in the plane splits the plane into two regions, one lying on either side of the line. Each such region is called a *halfplane*. We say that a halfplane is either *closed* or *open* depending, respectively, on whether or not it contains the line. Unless otherwise stated, we will assume that halfplanes are closed.

In the *halfplane intersection problem*, we are given a set of n closed halfplanes $H = \{h_1, \ldots, h_n\}$. We may assume that each halfplane is represented by a linear inequality, for example, $h_i = \{(x, y) \mid a_i x + b_i y \leq c_i\}$. The objective is to compute their intersection. It is easy to see that the intersection of halfspaces is a convex polygon (see Fig. 1(a)), but this polygon may be unbounded (see Fig. 1(b)) or even empty (see Fig. 1(c)).

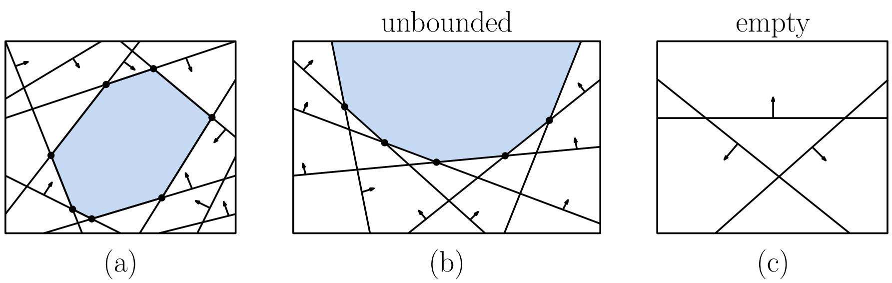

Fig. 1: Halfplane intersection.

Clearly, the number of sides of the resulting polygon is at most n, but may be smaller since some halfspaces may not contribute to the final shape. A natural choice for the output would be the (possibly empty) sequence of indices of the halfplanes that contribute to the final boundary, listed in cyclic order around the boundary. Note that we can easily compute the vertices as the points where adjacent bounding lines intersect.

Divide-and-Conquer Algorithm: Our first approach is based on applying divide and conquer.

(1) If $n = 1$, then just return this halfplane as the answer.

(2) Otherwise, partition H into subsets H_1 and H_2, each of size roughly $n/2$.

(3) Compute the intersections $K_1 = \bigcap_{h \in H_1} h$ and $K_2 = \bigcap_{h \in H_2} h$ recursively.

(4) If either either K_1 or K_2 is empty, return the empty set. Otherwise, compute the intersection of the convex polygons K_1 and K_2 (by the procedure described below).

Let's consider how to carry out step (4), where we intersect two convex polygons, K_1 and K_2 (see Fig. 2(a)). Note that these are somewhat special convex polygons because they may be empty or unbounded.

We can compute the intersection by a left-to-right plane sweep in $O(n)$ time (see Fig. 2(b)). We begin by breaking the boundaries of the convex polygons into their upper and lower chains. (This can be done in $O(n)$ time.) By convexity, the sweep line intersects the boundary of each convex polygon K_i in at most two points, one for the upper chain and one for the lower chain. Hence, the sweep-line status contains at most four points. This implies that updates to the sweep-line status can be performed in $O(1)$ time. Also, we need keep track of a constant number of events at any time, namely the right endpoints of the current segments in the sweep-line status, and the intersections between consecutive pairs of segments. Thus, each step of the plane-sweep process can be performed in $O(1)$ time.

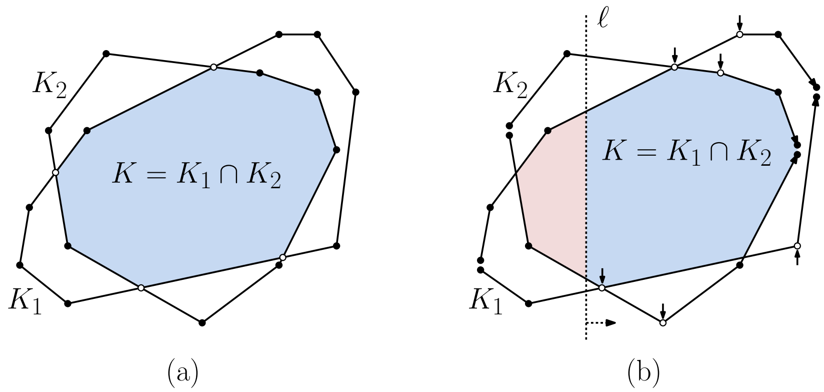

Fig. 2: Intersecting two convex polygons by plane sweep.

The total number of events is equal to the total number vertices, which is n, and the total number of intersection points. It is an easy exercise (which we leave to you) to prove that two convex polygons with a total of n sides can intersect at most $O(n)$ times. Thus, the overall running time is $O(n)$.

Given that we can compute the intersection of two polygons of size n in $O(n)$ time, we obtain the following recurrence (up to constant factors) for the overall running time of the divide-and-conquer algorithm.

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

It follows by standard results on recurrences (consult the *Master Theorem* in CLRS) that $T(n)$ is $O(n \log n)$.

Upper and Lower Envelopes of Lines: Let's next consider a variant of the halfplane intersection problem. Given any set of nonvertical lines $L = \{\ell_1, \ell_2, \ldots, \ell_n\}$ in the plane. Each line defines two natural halfplanes, and upper and lower halfplane. The intersection of all the lower halfplanes is called the *lower envelope* of L and the *upper envelope* is defined analogously (see Fig. 3). Let's assume that each line ℓ_i is given explicitly as $y = a_i x - b_i$.

The lower envelope problem is a restriction of the halfplane intersection problem, but it an interesting restriction. Notice that any halfplane intersection problem that does not involve any vertical lines can be rephrased as the intersection of two envelopes, a lower envelope defined by the lower halfplanes and an upper envelope defined by the upward halfplanes.

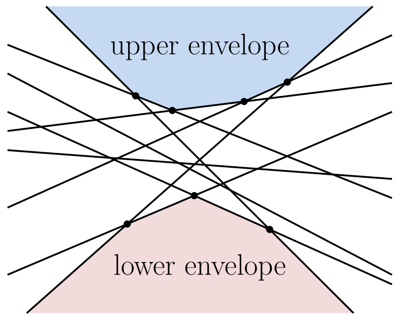

Fig. 3: Lower and upper envelopes.

We will see that solving the lower envelope problem is very similar (in fact, essentially the same as) solving the upper convex hull problem. Indeed, they are so similar that exactly the same algorithm will solve both problems, without changing even a single character of code! All that changes is the way in which you interpret the inputs and the outputs.

Motivating Duality: Consider any line $\ell : y = cx - d$ in the plane. Just as with point, it takes two real numbers (c and d) to define any line in $\mathbb{R}^2$. Thus, we can identify the lines in $\mathbb{R}^2$ with the set of points (c, d) in an alternative or "dual" space. For example, the line $\ell : y = 2x + 1$ corresponds to the point $(2, -1)$ in dual space, which we denote by ℓ^*. Conversely, each point $p = (a, b)$ is associated with a dual line, $p^* : y = ax - b$, which we denote by p^*.

This insight would not be of much use unless we could say something about how geometric relationships in one space relate to the other. The connection between the two involves incidences between points and line.

Primal Relation	**Dual Relation**
Two (nonparallel) lines meet in a point	Two points join to form a line
A point lies on a line	A line passes through a point
A point lies above/below a line	A line passes above/below a point
Three points may be collinear	Three lines may pass through the same point

We'll show that these relationships are preserved by duality. For example, consider the two lines $\ell_1 : y = 2x + 1$ and the line $\ell_2 : y = -\frac{x}{2} + 6$ (see Fig. 4(a)). These two lines intersect at the point $p = (2, 5)$. The duals of these two lines are $\ell_1^* = (2, -1)$ and $\ell_2^* = \left(-\frac{1}{2}, -6\right)$. The line in the (a, b) dual plane passing through these two points is easily verified to be $b = 2a - 5$. Observe that this is exactly the dual of the point p (see Fig. 4(b)). (As an exercise, prove this for two general lines.)

Dual Transformation: Let us explore this dual transformation more formally. Duality (or more specifically *point-line duality*) is a transformation that maps points in the plane to lines and lines to point. Given point $p = (a, b)$ and line $\ell : y = cx - d$, we define[1]

$$p^* : y = ax - b \qquad \text{and} \qquad \ell^* = (c, d).$$

[1]Duality can be generalized to higher dimensions as well. In $\mathbb{R}^d$, let us identify the y axis with the d-th coordinate vector, so that an arbitrary point can be written as $p = (x_1, \ldots, x_{d-1}, y)$ and a $(d-1)$-dimensional hyperplane can be written as $h : y = \sum_{i=1}^{d-1} a_i x_i - b$. The dual of this hyperplane is $h^* = (a_1, \ldots, a_{d-1}, b)$ and the dual of the point p is

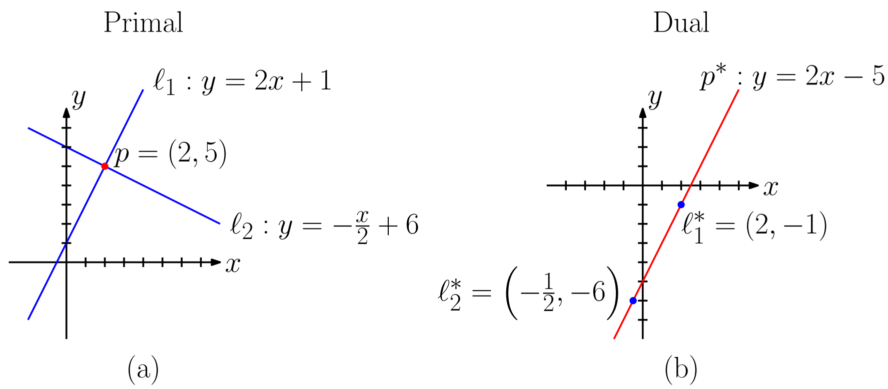

Fig. 4: The primal and dual planes.

Properties of Point-Line Duality: Duality has a number of interesting properties, each of which is easy to verify by substituting the definition and a little algebra.

Self Inverse: $p^{**} = p$, $\ell^{**} = \ell$ (This is trivial.)

Incidence Preserving: Point p is on line ℓ if and only if point ℓ^* is on line p^*. (Proof: The first assertion is equivalent to $b = ac - d$ and the second is equivalent to $d = ca - b$. Clearly, one is satisfied if and only if the other is.)

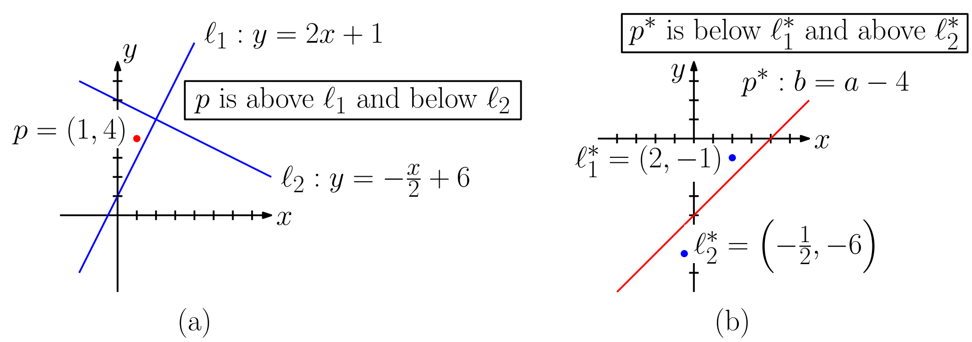

Fig. 5: Order reversal.

Order Reversal: Point p is above/below line ℓ if and only if line p^* is below/above point ℓ^*, respectively (see Fig. 5). (Proof: The first assertion is equivalent to $b > ac - d$ and the second to $d > ca - b$. Clearly, one is satisfied if and only if the other is.)

Intersection Preserving: Lines ℓ_1 and ℓ_2 intersect at point p if and only if the dual line p^* passes through points ℓ_1^* and ℓ_2^*. (Follows directly from incidence preservation.)

Collinearity/Coincidence: Three points are collinear in the primal plane if and only if their dual lines intersect in a common point. (Follows directly from incidence preservation.)

$p^* : b = \sum_{i=1}^{d-1} x_i a_i - y$. All the properties defined for point-line relationships generalize naturally to point-hyperplane relationships, where notions of above and below are based on the assumption that the y (or b) axis is "vertical."

Convex Hulls and Envelopes: Let us return now to the question of the relationship between convex hulls and the lower/upper envelopes of a collection of lines in the plane. The following lemma demonstrates the, under the duality transformation, the convex hull problem is dually equivalent to the problem of computing lower and upper envelopes.

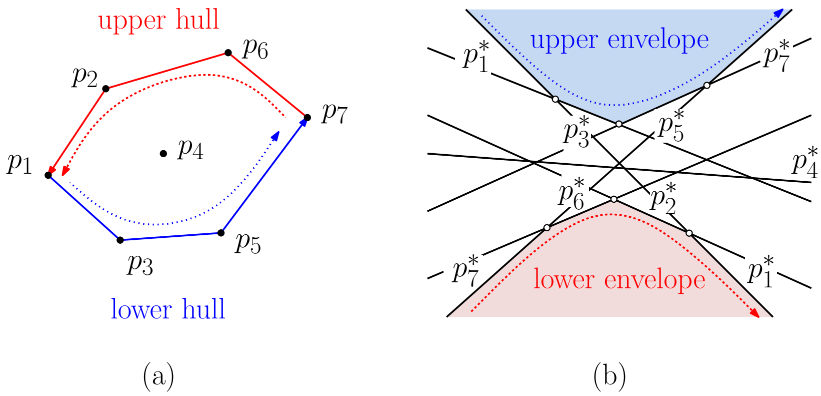

Fig. 6: Equivalence of hulls and envelopes.

Lemma: Let P be a set of points in the plane. The counterclockwise order of the points along the upper (lower) convex hull of P (see Fig. 6(a)), is equal to the left-to-right order of the sequence of lines on the lower (upper) envelope of the dual P^* (see Fig. 6(b)).

Proof: We will prove the result just for the upper hull and lower envelope, since the other case is symmetrical. For simplicity, let us assume that no three points are collinear.

Consider a pair of points p_i and p_j that are consecutive vertices on the upper convex hull. This is equivalent to saying that all the other points of P lie beneath the line ℓ_{ij} that passes through both of these points.

Consider the dual lines p_i^* and p_j^*. By the incidence preserving property, the dual point ℓ_{ij}^* is the intersection point of these two lines. (By general position, we may assume that the two points have different x-coordinates, and hence the lines have different slopes. Therefore, they are not parallel, and the intersection point exists.)

By the order reversing property, all the dual lines of P^* pass above point ℓ_{ij}^*. This is equivalent to saying the ℓ_{ij}^* lies on the lower envelope of P^*.

To see how the order of points along the hulls are represented along the lower envelope, observe that as we move counterclockwise along the upper hull (from right to left), the slopes of the edges increase monotonically. Since the slope of a line in the primal plane is the a-coordinate of the dual point, it follows that as we move counterclockwise along the upper hull, we visit the lower envelope from left to right.

One rather cryptic feature of this proof is that, although the upper and lower hulls appear to be connected, the upper and lower envelopes of a set of lines appears to consist of two disconnected sets. To make sense of this, we should interpret the primal and dual planes from the perspective of *projective geometry*, and think of the rightmost line of the lower envelope as "wrapping around" to the leftmost line of the upper envelope, and vice versa. The places where the two envelopes wraps around correspond to the vertical lines (having infinite slope) passing through the left and right endpoints of the hull. (As an exercise, can you see which is which?)

Primal/Dual Equivalencies: There are a number of computational problems that are defined in terms of affine properties of point and line sets. These can be expressed either in primal or in dual form. In many instances, it is easier to visualize the solution in the dual form. We will discuss many of these later in the semester. For each of the following, can you determine what the dual equivalent is?

- Given a set of points P, find the narrowest slab (that is, a pair of parallel lines) that contains P. Define the width of the slab to be the vertical distance between its bounding lines (see Fig. 7(a)).

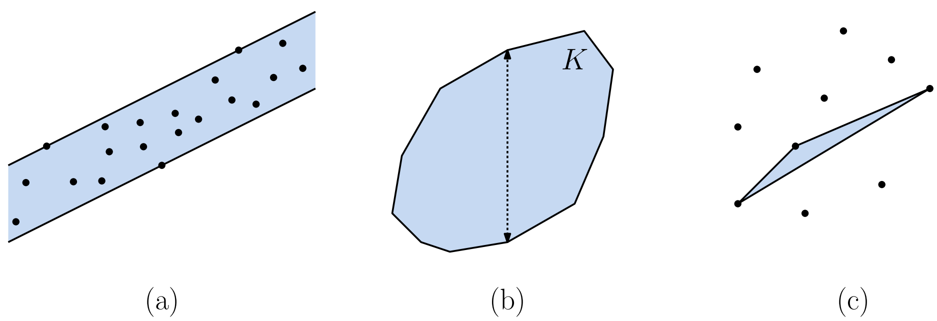

Fig. 7: Equivalence of hulls and envelopes.

- Given a convex polygon K, find the longest vertical line segment with one endpoint on K's upper hull and one on its lower hull (see Fig. 7(b)).
- Given a set of points P, find the triangle of smallest area determined by any three points of P (see Fig. 7(c)). (If three points are collinear, then they define a degenerate triangle of area 0.)

Polar Dual: (Optional) The dual transform we have described (mapping the point $p = (a, b)$ to the line $y = ax - b$) is just one example of the general concept of a dual transformation. A very popular alternative, called *polar transformation*, maps the vector $v = (a, b)$ to the line $v^\circ : ax + by = 1$ and vice versa (see Fig. 8(a)).

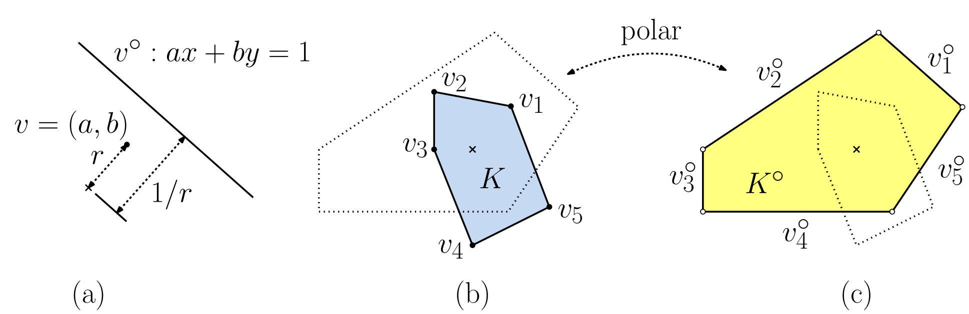

Fig. 8: Polar transformation.

This has the following geometric interpretation. The vector v is transformed to the line that is perpendicular to this vector and whose distance from the origin is the reciprocal of the length of the vector. That is, if $r = \sqrt{a^2 + b^2}$ then the line p° is perpendicular to v and its distance from the origin (along this vector) is $1/r$ (see Fig. 8(a)).

This transformation is often applied to convex bodies. Suppose that K is a convex polygon that contains the origin. Let $\{v_1, \ldots, v_n\}$ denote the vertices of the polygon, given in counterclockwise order. Then K° is the convex polygon whose edges lie on the polar lines $\{v_1^\circ, \ldots, v_n^\circ\}$, also in counterclockwise order. The operation is an involution in the sense that $(K^\circ)^\circ = K$.

There is an interesting reciprocal relationship between K and K°. As the vertices of K get closer to the origin, the edges of K° get further from the origin, by the reciprocal. Thus, one would expect that the volumes $\mathrm{vol}(K)$ and $\mathrm{vol}(K^\circ)$ vary inversely—as one gets larger the other gets smaller. Indeed, it is famous result from the theory of convex bodies that for any bounded convex body K that contains the origin in any dimension d, the product $\mathrm{vol}(K) \cdot \mathrm{vol}(K^\circ)$, called the *Mahler volume*, is both upper and lower bounded by a constant. A major open problem in mathematics involves determining which shapes achieve the extreme Mahler volume values.

Higher Dimensions: (Optional) In d-dimensional space the corresponding notion is a *halfspace*, which is the set of points lying to one side of a $(d-1)$-dimensional hyperplane. The intersection of halfspaces is a *convex polytope*. The resulting polytope will have at most n facets (at most one per halfspace).

The number of vertices, and more generally, the *combinatorial complexity* (total number of vertices, edges, and faces of all dimensions) can vary significantly. Assuming general position, it is reasonable to expect that the combinatorial complexity should be at least $\Omega(n)$, but it can be much higher. A famous result, called *McMullen's Upper-Bound Theorem* states that a polytope with n facets in dimension d can have up to $O(n^{\lfloor d/2 \rfloor})$ vertices. (In dimensions 2 and 3, this is linear in the number of halfspaces, but even in dimension 4 the number of vertices can grow to $O(n^2)$.) Obtaining such a high number of vertices takes some care, but the bound is tight in the worst case. There is a famous class of polytopes, called the *cyclic polytopes*, that achieve this bound. Symmetrically, the convex hull of n points in dimension d defines a convex polytope that can have $O(n^{\lfloor d/2 \rfloor})$ facets, and this bound is also tight.

Representing Lines and Hyperplanes: (Optional) While we will usually treat geometric objects rather abstractly, it may be useful to explore a bit regarding how lines, halfspaces, and their higher dimensional counterparts are represented. These topics would be covered in a more complete course on projective geometry or convexity.

Explicit Representation: If we think of a line as a linear function of the variable x, we can express any (nonvertical) line by the equation $y = ax + b$, where a is the slope and b is the y-intercept.

In dimension d, we can think of the dth coordinate as being special, and we will make the convention of referring to the d-th coordinate axis as pointing vertically upwards. We can express any "nonvertical" $(d-1)$-dimensional hyperplane by the set of points

$(x_1, \ldots, x_d)$, where $x_d = \sum_{i=1}^{d-1} a_i x_i + b$, thus x_d is expressed "explicitly" as a linear function of the first $d-1$ coordinates.

The associated halfspaces arise replacing "=" with an inequality, e.g., the *upper halfplane* is the set (x, y) such that $y \geq ax + b$, and the *lower halfplane* is defined analogously.

Implicit Representation: The above representation has the shortcoming that it cannot represent vertical objects. A more general approach (which works for both hyperplanes and curved surfaces) is to express the object implicitly as the zero-set of some function of the coordinates. In the case of a line in the plane, we can represent the line as the set of points (x, y) that satisfy the linear function $f(x, y) = 0$, where $f(x, y) = ax + by + c$, for scalars a, b, and c. The corresponding halfplanes are just the sets of points such that $f(x, y) \geq 0$ and $f(x, y) \leq 0$.

This has the advantage that it can represent any line in the Euclidean plane, but the representation is not unique. For example, the line described by $5x - 3y = 2$ is the same as the line described by $10x - 6y = 4$, or any scalar multiple thereof. We could apply some normalization to overcome this, for example, by requiring that $a^2 + b^2 = 1$. If this normalization is performed, then the line has the convenient geometric interpretation that it is orthogonal to the unit vector (a, b) and is at distance c from the origin.

Parametric Representation: One shortcoming of the explict and implicit representations is that it is not particularly easy to define lower dimensional objects. For example, suppose that you wanted to represent the points lying on a line segment in $\mathbb{R}^3$. A *parametric representation* defines a geometric object as the union of points, where an additional numeric parameter is used to specify these points.

To make this more concrete, suppose that we wanted to represent a line segment between two points p and q (in any dimension). We could do this by creating a real-valued parameter t, and defining the segment to be the set of convex combinations

$$\overline{pq} = \left\{(1-t)p + tq \;\middle|\; t \in [0, 1]\right\}.$$

Another use is if we are given a point p and a directional vector u, we can represent the points on the ray in the direction of u as the set of points $\mathrm{ray}(p, u) = \{p + tu \mid t \geq 0\}$.

Parametric forms are handy for representing curves. For example, we can describe a helix winding around the x-axis in $\mathbb{R}^3$ as $\{(t, \cos t, \sin t) \mid t \in \mathbb{R}\}$. We can also represent surfaces of any dimension by adding additional parameters. For example, by thinking of s and t as proxies for latitude and longitude, respectively, we can describe any point on a unit sphere in $\mathbb{R}^3$ as

$$\left\{(\sin(s)\cos(t), \sin(s)\sin(t), \cos(s)) \;\middle|\; s \in [0, \pi], t \in [0, 2\pi]\right\}.$$

Homogeneous Coordinates: A popular variant to the implicit representation is to use *homogeneous coordinates* to represent points. This involves adding an additional coordinate. For example, we could represent a point (x, y) with the triple $(x, y, 1)$. This additional coordinate is usually set to 1, but there are other uses, depending on the context.

For example, in *projective geometry*, we use $(x, y, 1)$ to represent a "regular" point, and we represent a points "at infinity" using the notation $(x, y, 0)$. This refers to the point that is infinitely far away in the direction of the vector (x, y).

Another example is in *affine geometry*, where we distinguish between points and vectors as separate geometric entities. (Points specify locations and vectors specify direction and magnitude.) We use $(x, y, 1)$ to represent the point with coordinates (x, y) and we use $(x, y, 0)$ to represent the vector from the origin to this point.

One advantage of homogeneous coordinates is that it simplifies dual representations. For example, we can describe any line in the plane as homogeneous 3-vector $\ell = (a, b, c)$. A point $p = (x, y, 1)$ lies on this line if and only if $\ell \cdot p = ax + by + c = 0$. (The name "homogeneous" derives from the fact that by eliminating the scalar term, we have a homogeneous equation.)

In dimension d, a point is represented by a $(d + 1)$-vector $p = (x_1, \ldots, x_d, 1)$ and a hyperplane by a $(d + 1)$-vector $h = (a_1, \ldots, a_{d+1})$, and a point lies on the associated hyperplane if $h \cdot p = \sum_{i=1}^{d+1} a_i x_i = 0$.