CS348a: Computer Graphics Handout #9
Geometric Modeling Original Handout #9
Stanford University Tuesday, 3 November 1992

Original Lecture #2: 6 October 1992
Topics: Coordinates and Transformations
Scribe: Steve Farris

1 Affine and Projective Coordinate Notation

Recall that we have chosen to denote the point with Cartesian coordinates (X,Y) in affine coordinates as $(1;X,Y)$. Also, we denote the line with implicit equation $a+bX+cY=0$ as the triple of coefficients $[a;b,c]$. For example, the line $Y=mX+r$ is denoted by $[r;m,-1]$, or equivalently $[-r;-m,1]$. In three dimensions, points are denoted by $(1;X,Y,Z)$ and planes are denoted by $[a;b,c,d]$. This transfer of dimension is natural for all dimensions.

The geometry of the line — that is, of one dimension — is a little strange, in that hyperplanes are points. For example, the hyperplane with coefficients $[3;7]$, which represents all solutions of the equation $3+7X=0$, is the same as the point $X=-3/7$, which we write in coordinates as $(1;-\frac{3}{7})$.

2 Transformations

2.1 Affine Transformations of a line

Suppose

$$F(X) := a+bX$$

and

$$G(X) := c+dX,$$

and that we want to compose these functions. One way to write the two possible compositions is:

$$G(F(X)) = c+d(a+bX) = (c+da)+bdX$$

and

$$F(G(X)) = a+b(c+d)X = (a+bc)+bdX.$$

Note that it makes a difference which function we apply first, F or G.

When writing the compositions above, we used *prefix* notation, that is, we wrote the function on the left and its argument on the right. There is also *postfix* notation, where the function is on the right and the argument is on the left. Using F as the function and x as the argument, people who use the prefix convention usually write $F(x)$, while people who use postfix write either xF or x^F. In this course we will adopt the prefix notation for expressing function application. Hence, when we compose two functions F and G to get the composition $F \circ G$, we mean to apply first G and then F:

$$(F \circ G)(X) = F(G(X)).$$

Affine transformations will be denoted by $F(X) := a + bX$, or in the matrix transformation form as

$$\begin{pmatrix} 1 & 0 \\ a & b \end{pmatrix} \begin{pmatrix} 1 \\ X \end{pmatrix} = \begin{pmatrix} 1 \\ a + bX \end{pmatrix}.$$

Note that in postfix notation this matrix transformation would be written

$$\begin{pmatrix} 1 & X \end{pmatrix} \begin{pmatrix} 1 & a \\ 0 & b \end{pmatrix} = \begin{pmatrix} 1 & a + bX \end{pmatrix}.$$

2.2 Affine Transformations of a plane

The general form of an affine transformation of the plane has the form

$$\begin{pmatrix} 1 & 0 & 0 \\ a & b & c \\ d & e & f \end{pmatrix} \begin{pmatrix} 1 \\ X \\ Y \end{pmatrix} = \begin{pmatrix} 1 \\ a + bX + cY \\ d + eX + fY \end{pmatrix}.$$

Here are some examples of affine transformations of the plane:

1. Translation

$$\begin{pmatrix} 1 & 0 & 0 \\ t_X & 1 & 0 \\ t_Y & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ X \\ Y \end{pmatrix} = \begin{pmatrix} 1 \\ X + t_X \\ Y + t_Y \end{pmatrix}.$$

2. Rotation through an angle θ has the matrix form

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}.$$

   Note that the second column of this matrix gives the coordinates of the vector that results from rotating $(0;1,0)$ — the unit vector in the X direction – through the angle θ, as shown in Figure 1. In a similar way, the third column gives the coordinates of the vector that results from rotating $(0;0,1)$, the unit vector in the Y direction.

3. Scale transformation

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & S_X & 0 \\ 0 & 0 & S_Y \end{pmatrix}$$

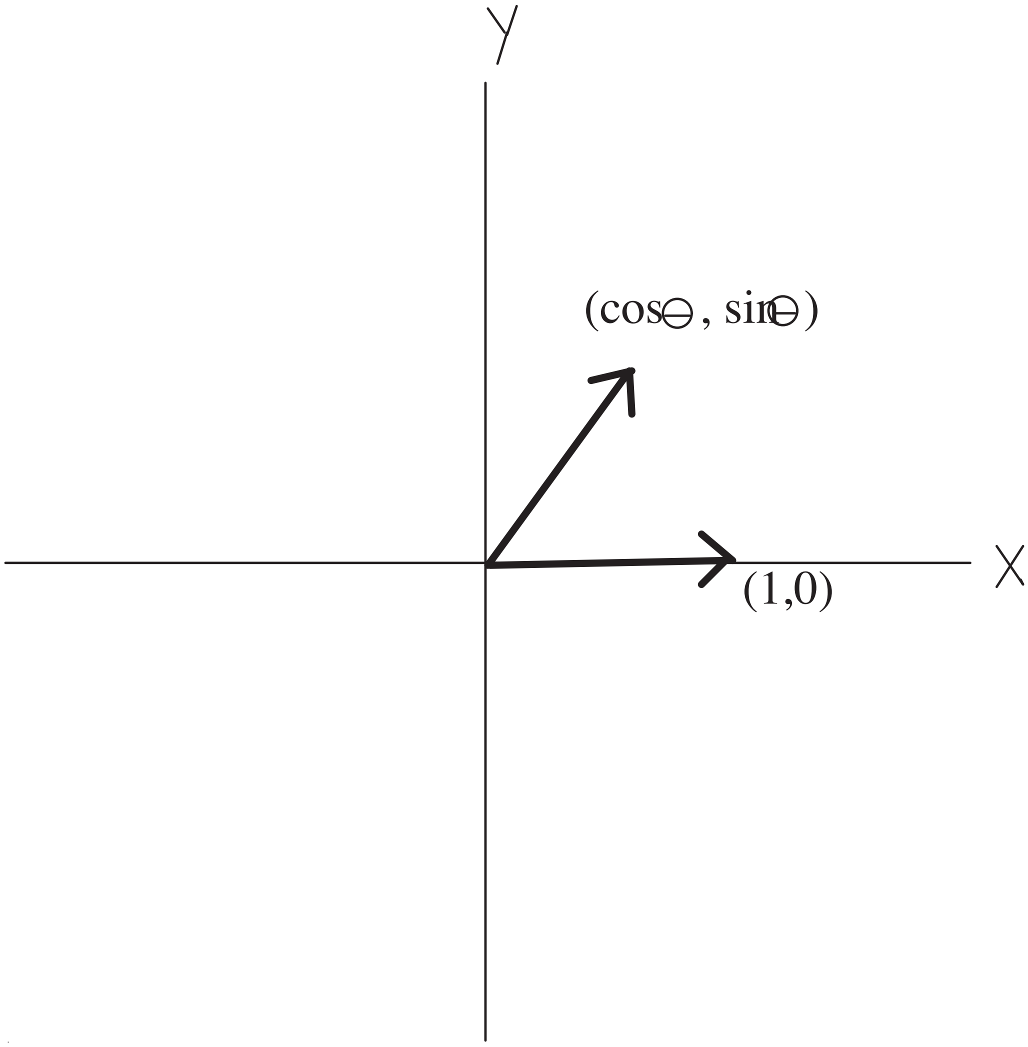

Figure 1: The rotation of the unit vector along the x-axis

What happens to lines under affine transformations? We represent a point as a column of coordinates, while we represent a line as a row of coefficients. Let $M\colon U \to V$ be an affine transformation, so M takes the point $(1;X,Y)$ in U to the point

$$M\begin{pmatrix}1\\X\\Y\end{pmatrix}$$

in V. What does M do to some line — say, the line $[p;q,r]$ — in U? Well, the line $[p;q,r]$ contains precisely those points $(1;X,Y)$ in U for which the dot product $[p;q,r]\cdot(1;X,Y)$ is

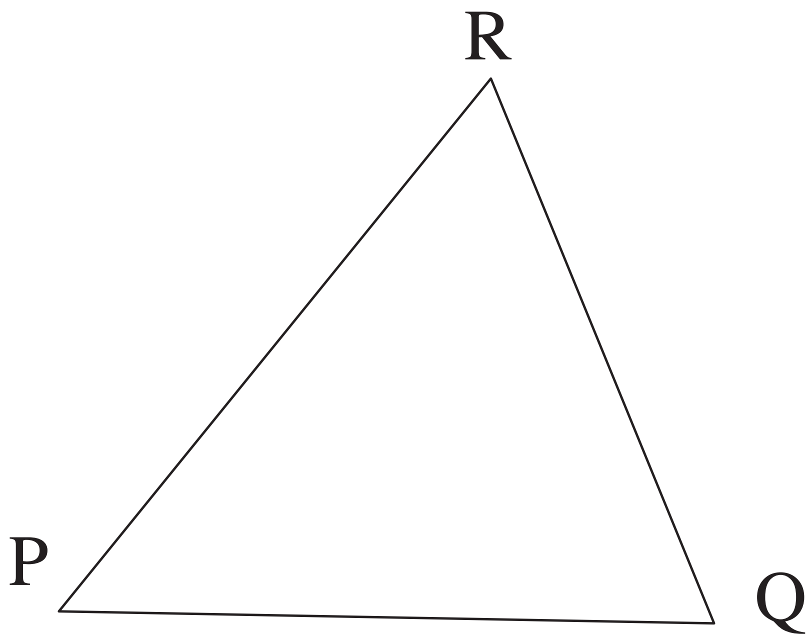

Figure 2: Barycentric Basis Points

zero. So the image line $[p';q',r']$ in V should have the property that

$$\begin{pmatrix} p' & q' & r' \end{pmatrix} M \begin{pmatrix} 1 \\ X \\ Y \end{pmatrix} = 0$$

precisely when

$$\begin{pmatrix} p & q & r \end{pmatrix} \begin{pmatrix} 1 \\ X \\ Y \end{pmatrix} = 0.$$

The way to achieve this is to set the image line $[p';q',r']$ to be the matrix product

$$\begin{pmatrix} p' & q' & r' \end{pmatrix} := \begin{pmatrix} p & q & r \end{pmatrix} M^{-1}.$$

So, if multiplying on the left by M takes points in U to points in V, then multiplying on the right by M^{-1} takes lines in U to lines in V.

3 Cartesian and Barycentric Coordinates

Sometimes it is convenient to coordinatize points with respect to given, non-collinear points. For example, see Figure 2.

Suppose we coordinatize points with respect to P, Q, and R of Figure 2. Then $\lambda P + \mu Q + \nu R$ is a point if $\lambda + \mu + \nu = 1$. For example, $\frac{1}{2}P + \frac{1}{2}Q + 0R$ is the midpoint of PQ, while $\frac{1}{3}P +$

$\frac{1}{3}Q+\frac{1}{3}R$ is the centroid of the triangle PQR and $Q-P$ is the vector from P to Q. Representing points as triples (λ,μ,ν) with $\lambda+\mu+\nu=1$ in this way is called *barycentric coordinates*; by contract, our previous coordinate triples $(1;X,Y)$ are called *Cartesian coordinates*. The barycentric coordinates (λ,μ,ν) of a vector have $\lambda+\mu+\nu=0$.

Affine transformations from one space to another always correspond to 3-by-3 matrices, but the constraints on the components of the matrix for the resulting transformation to be affine depend upon the type of coordinate system we are using. In Cartesian coordinates (with the weight coordinate written first), the first row of the matrix must be $(1,0,0)$. In barycentric coordinates, the transformation matrix

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}\begin{pmatrix} \lambda \\ \mu \\ \nu \end{pmatrix} = \begin{pmatrix} \lambda' \\ \mu' \\ \nu' \end{pmatrix}$$

must satisfy

$$\begin{aligned} a+d+g &= 1 \\ b+e+h &= 1 \\ c+h+i &= 1 \end{aligned}$$

in order to represent an affine transformation from one barycentric coordinate system to another.

When using barycentric coordinates, lines are again represented as triples of homogeneous coefficients, say $[\alpha,\beta,\gamma]$, under the rule that the line $[\alpha,\beta,\gamma]$ contains the point (λ,μ,ν) just when

$$[\alpha,\beta,\gamma]\begin{pmatrix} \lambda \\ \mu \\ \nu \end{pmatrix} = \alpha\lambda+\beta\mu+\gamma\nu = 0.$$

As a special case, note that

$$[\alpha,\alpha,\alpha]\begin{pmatrix} \lambda \\ \mu \\ \nu \end{pmatrix} = \alpha \neq 0,$$

since $\lambda+\mu+\nu=1$. Thus, the line whose three barycentric coefficients are equal contains no (affine) points. Recall that the Cartesian coefficients of this same line are $[1;0,0]$, corresponding to the implicit equation $1+0X+0Y=0$. In affine geometry, this line is not allowed. (In projective geometry, as we will see later, this line is allowed, and is called the 'line at infinity'.)

4 Affine to Projective Geometry

4.1 Site Space

When doing affine geometry in the plane, using Cartesian coordinates, triples of coordinates $(W;X,Y)$ with $W=1$ are points, while triples with $W=0$ are vectors. We will use the non-

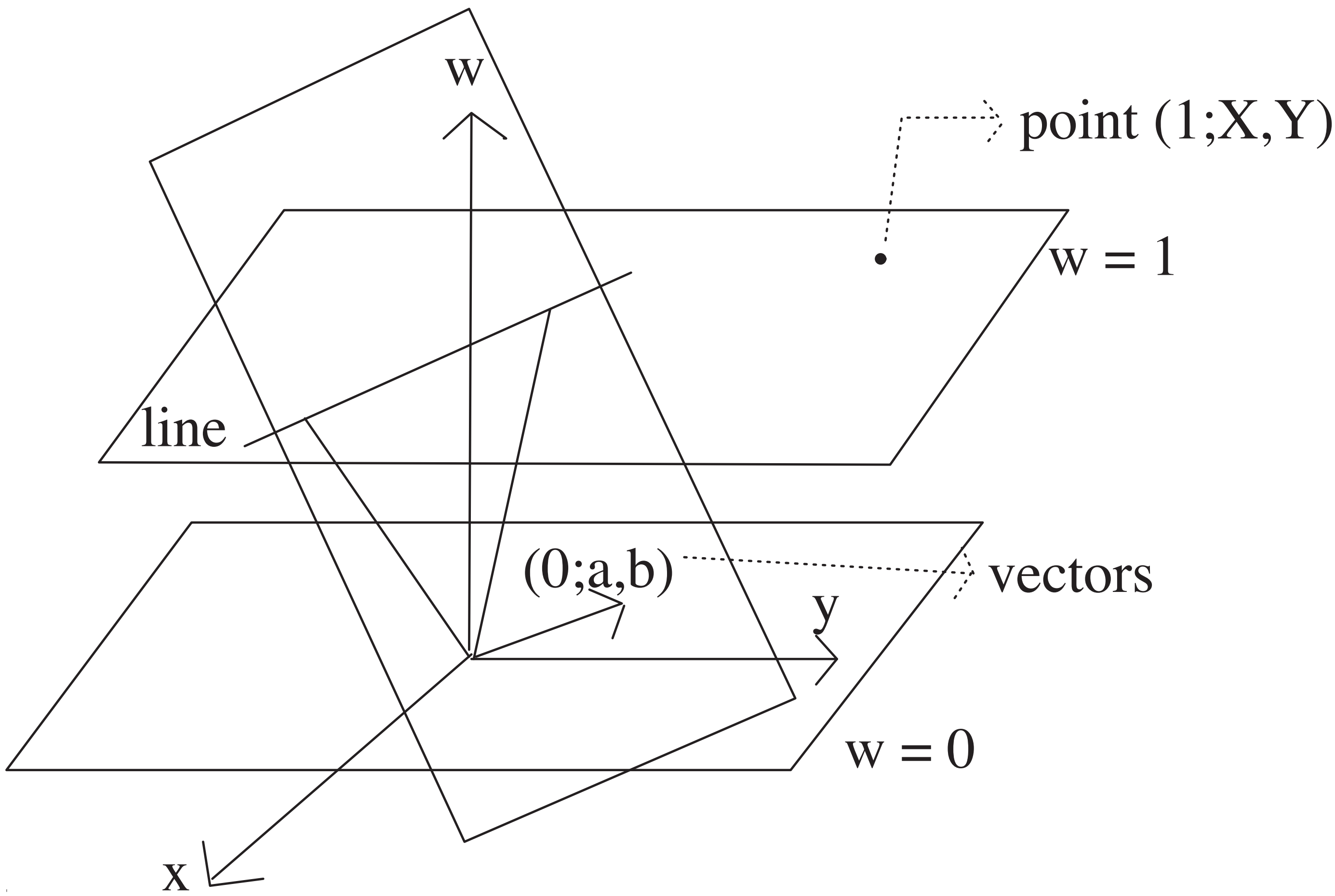

Figure 3: Site Space

standard word *site* to refer to any triple of coordinates $(W;X,Y)$, with no constraint on the weight coordinate W.

All such triples form a 3-dimensional space, which we call *site space*. In this space, all affine vectors lie in the plane $W=0$, which passes through the origin, while all affine points lie in the parallel plane $W=1$.

The line with homogeneous coefficients $[a;b,c]$ contains precisely those points $(1;X,Y)$ for which the dot product $(a,b,c)\cdot(1,X,Y)$ is zero. Thinking about site space, we can view the triple $[a;b,c]$ as giving a vector in site space. There is a unique plane through the origin orthogonal to the vector (a,b,c), and that plane cuts the plane $W=1$ of all points precisely in those points that line on the line. Thus, a line in affine geometry corresponds to a plane through the origin of site space. Figure 3 is a picture of site space.

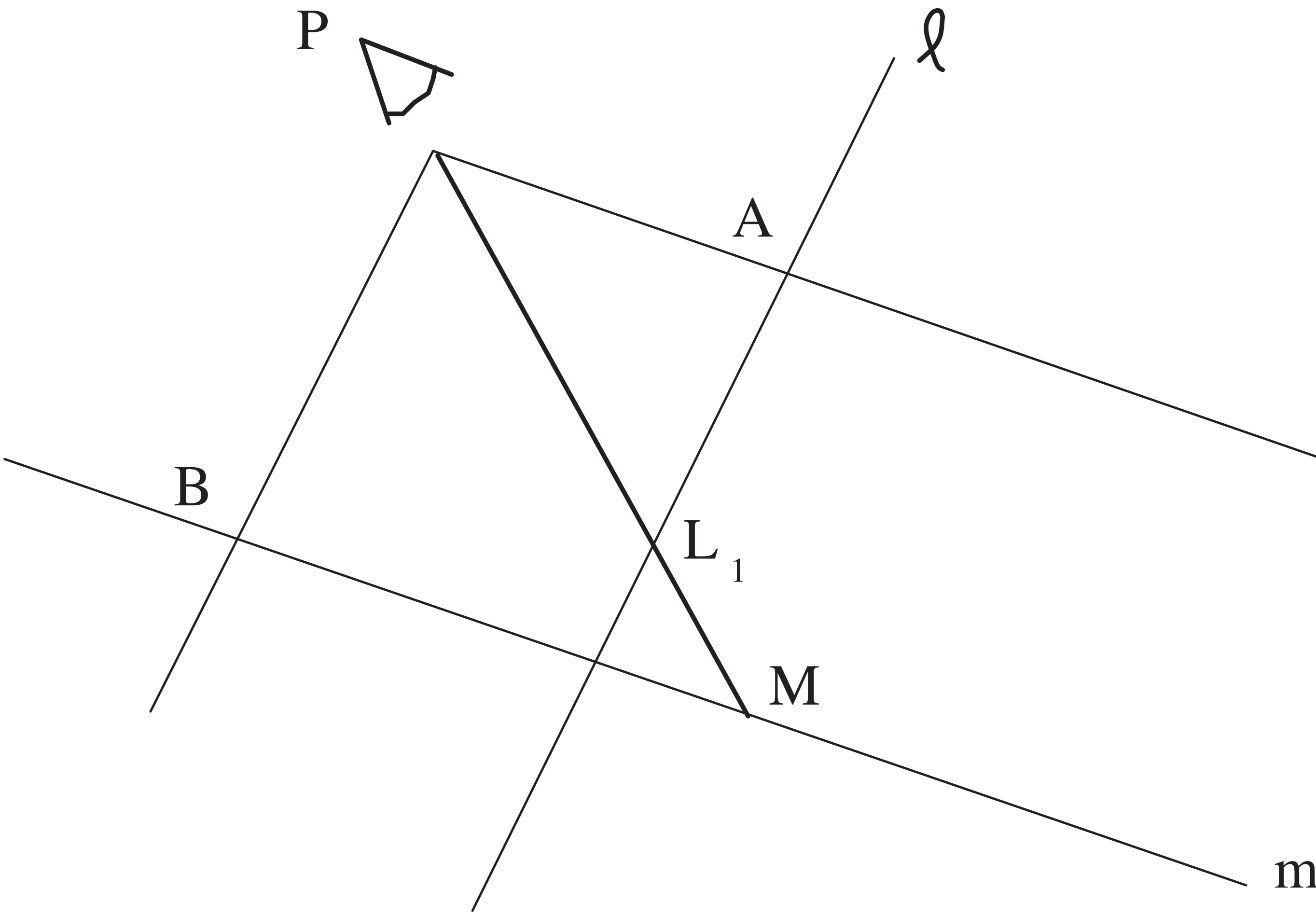

Figure 4: Perspectivity

4.2 Perspective Correspondence

Figure 4 below depicts a perspectivity between two lines ℓ and *m*, the building block of projective geometry.

The point *P* is called the *center* of perspectivity, and we say line ℓ is perspectively mapped onto *m* under this perspectivity. The image of the point *L* in ℓ is the point *M* in *m*, since those two points lie on a common line through *P*. But what points correspond to *A* and *B* under this perspectivity? The perspectivity *P* gives a one-to-one correspondence between the two lines ℓ and *m*, except that the point *A* on ℓ doesn't go anywhere on *m* and the point *B* on *m* doesn't come from anywhere on ℓ. To resolve this we add one *point at infinity* to our plane for each set of parallel lines — that is, for each slope. Having done this, the perspectivity centered at *P* can take *A* to the point at infinity on *m*; it can also take the point at infinity on ℓ to the point *B* on *m*.

Thinking about site space, we define a *projective point* to be a line through the origin of site space, and we define a *projective line* to be a plane through the origin of site space. Lines through the origin of site space that aren't horizontal are finite, affine points; lines through the origin of site space that are horizontal are points at infinity. The horizontal plane through the origin of site space is the *line at infinity*, the line whose points are precisely all of the points at infinity.

4.3 Coordinatization

The homogeneous coefficients of the line at infinity are [1;0,0]. All other lines have the same homogeneous coefficients in projective geometry that they had in affine geometry.

Points in projective geometry have *homogeneous coordinates*. If $(1;X,Y)$ is a point in affine geometry, then $[1;X,Y]$ is one set of homogeneous coordinates for the same point in projective geometry. But $[2;2X,2Y]$ is another set of homogeneous coordinates for the same point, as is $[-1;-X,-Y]$. In general, we can scale all three coordinates by any nonzero factor.

Let $[w;x,y]$ be one set of homogeneous coordinates for a point in the projective plane. If w is nonzero, we can scale the coordinates by dividing by w, getting $[1;x/w,y/w]$ as another valid set of homogeneous coordinates. Thus, points $[w;x,y]$ with $w \neq 0$ are affine points — that is, are not points at infinity — and we can determine their Cartesian coordinates X and Y by the formulas $X = x/w$ and $Y = y/w$.

Points at infinity have triples of homogeneous coordinates $[w;x,y]$ in which $w = 0$ and the ratio y/x gives the slope; in particular, the point $[0;x,y]$ is the point at infinity common to all lines with the slope y/x. For example, $[0;1,1]$ is the point at infinity for lines at slope 1 — that is, at 45 degrees — while $[0;1,0]$ is the point at infinity in the horizontal direction and $[0;0,1]$ is the point at infinity in the vertical direction.

4.4 Duality

In any theorem of projective geometry, if we systematically replace all occurrences of 'point' by 'line' and simultaneously replace all occurrences of 'line' by 'point', the resulting statment is also a theorem. For example, any two distinct points determine a unique line; but also any two distinct lines determine a unique point (possibly a point at infinity). This is called the *principle of duality*.

4.5 Sides of Lines

In Figure 5(a) we see a line in the affine plane, with two points P and Q. We say P and Q are on opposite sides of the line if the segment PQ intersects the line.

This definition of 'opposite sides' doesn't make sense in the projective plane, since we can get from any point P to any other point Q without intersecting the line. Indeed, there are two segments from P to Q, as shown in Figure 5(b), one of which intersects the line and the other of which contains the point at infinity on the line PQ. So drawing a single line in the projective plane is not enough to divide the projective plane into two pieces.

But drawing two lines in the projective plane is enough to divide the projective plane into two pieces, as shown in Figure 5(c). The two white sectors form one piece, while the two black sectors form the second piece. Even going off to infinity, we can't get from white to black or vice-versa without crossing one of the two lines.

4.6 Two-Sided Projective Geometry

There are some times in graphics when it would be helpful if a single line did divide the projective plane into two pieces. We can make this be the case by switching from the standard, *one-sided* projective plane to the *two-sided projective plane*.

The two-sided projective plane has twice as many points as the standard projective plane. Half of those points form the *top* of the two-sided projective plane, while the other half form the *bottom*. When we move from one point to another through a point at infinity, we switch from top to bottom (or bottom to top). This allows us to define the two sides of a line once again, as shown at the right in Figure 6. The side containing P on the top consists of all points on the top that are above and to the left of the line, along with all points on the bottom that are below and to the right of the line. That side contains the point Q on the bottom, but not the point Q on top. The other side contains the points on the bottom that are above and to the left and the points on the top that are below and to the right. So one side contains P on the top and Q on the bottom, while the other contains P of the bottom and Q on top. And there are two sides.

We denote points on the bottom with hollow circles, and on the top as solid points. Lines on the bottom are dashed, while lines on the top are solid.

Algebraically, points $[w;x,y]$ on top are those whose weight coordinate w is positive, while points on the bottom have w negative. Note that, in two-sided projective geometry, we are not allowed to rescale homogeneous coordinates by negative factors. In one-sided projective geometry, rescaling by any nonzero factor is legal; but in two-sided projective geometry, only positive rescaling factors are allowed. In two-sided geometry, the two points $[w;x,y]$ and $[-w;-x,-y]$ are sitting right on top of each other, but one is on top and the other is on the bottom.

In a similar way, a line $[a;b,c]$ in the two-sided plane is distinguished from the line $[-a;-b,-c]$. Lines in the two-sided plane should be thought of as oriented, either one way or the other. If $[a;b,c]$ is the line PQ oriented, say, from P to Q, then the line $[-a;-b,-c]$ is the same line, but oriented in the opposite direction, from Q to P.

If we want, we can think of points on the top and points on the bottom as being oriented as well: a point on the top is oriented counterclockwise, while a point on the bottom is oriented clockwise, as shown in Figure 7.

Given an oriented line $[a;b,c]$ and an oriented point $[w;x,y]$ in the two-sided plane, we can test whether the two orientations agree or disagree by testing the sign of the dot product:

$$[a;b,c]\begin{bmatrix} w \\ x \\ y \end{bmatrix} = aw+bx+cy = \begin{cases} > \;0 & \text{the orientations agree} \\ = \;0 & \text{the point lies on the line} \\ < \;0 & \text{the orientations disagree.} \end{cases}$$

Fixing the line, the points whose orientations agree with that line are the ones to its left; the points whose orientations disagree lie to the line's right.

For example, the line $[1;0,0]$ in the two-sided plane is the line at infinity, oriented in a counterclockwise fashion. So that line has all points on the top of the plane to its left and all points on the bottom of the plane to its right. The line $[-1;0,0]$ is also the line at infinity, but oriented in a clockwise fashion.

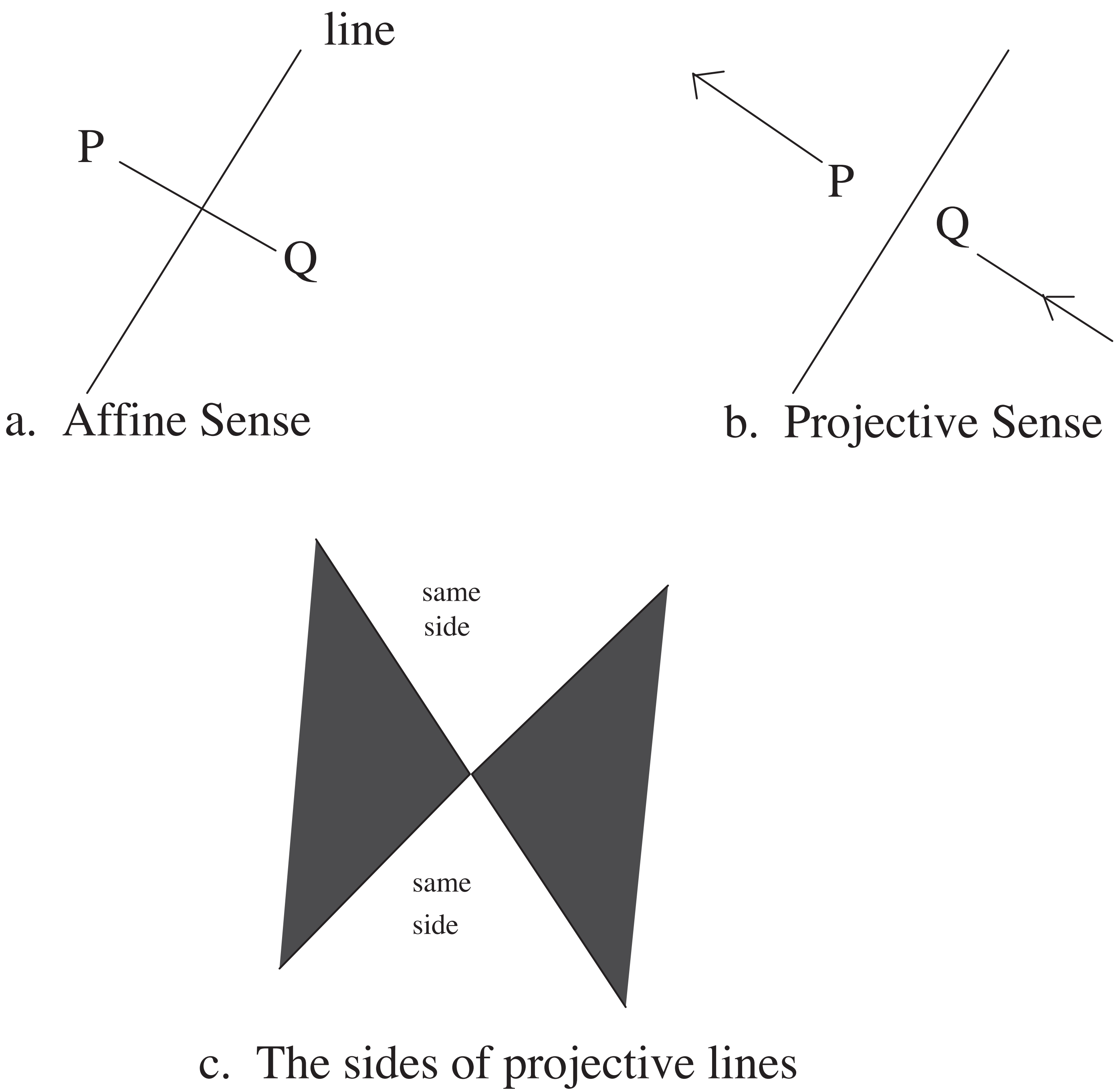

Figure 5: Sides of Various Lines

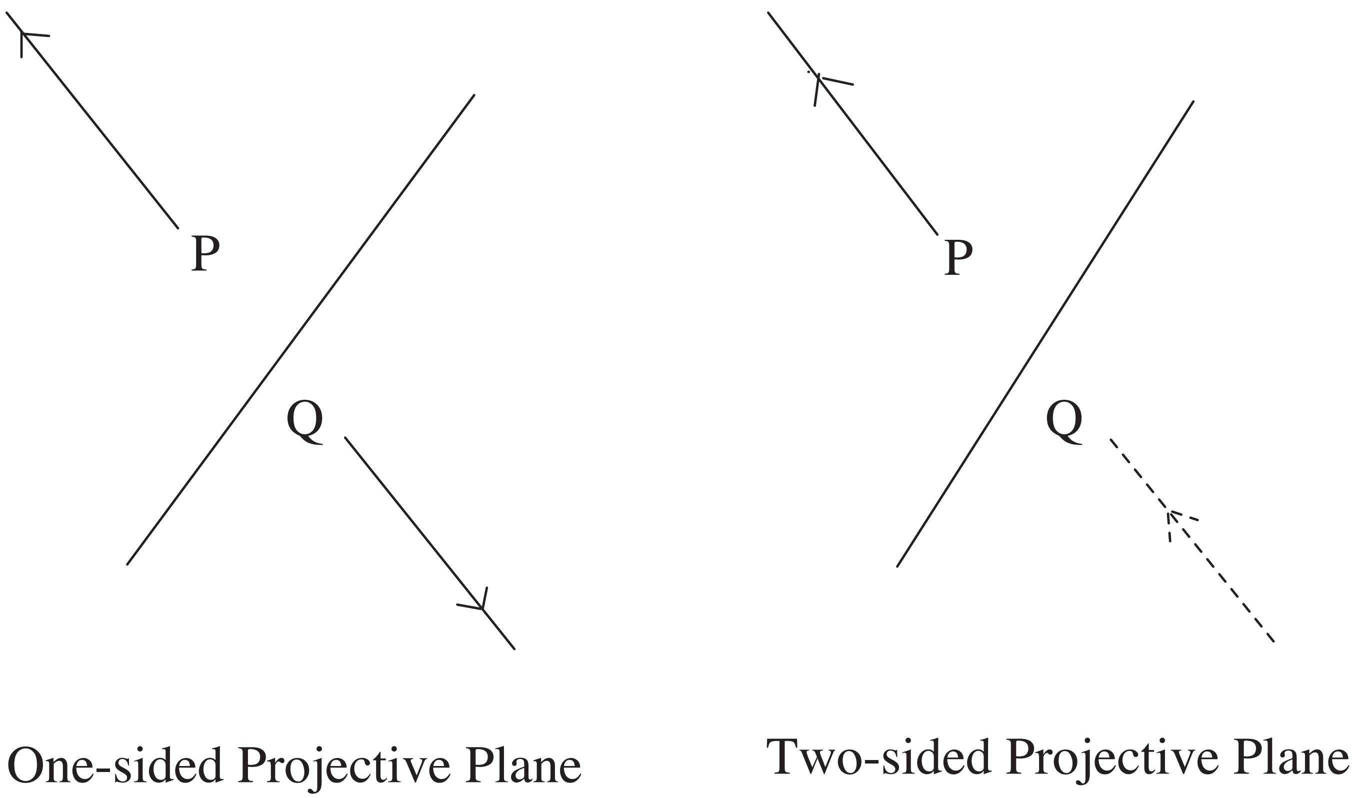

Figure 6: Points and Lines in Two-Sided Geometry

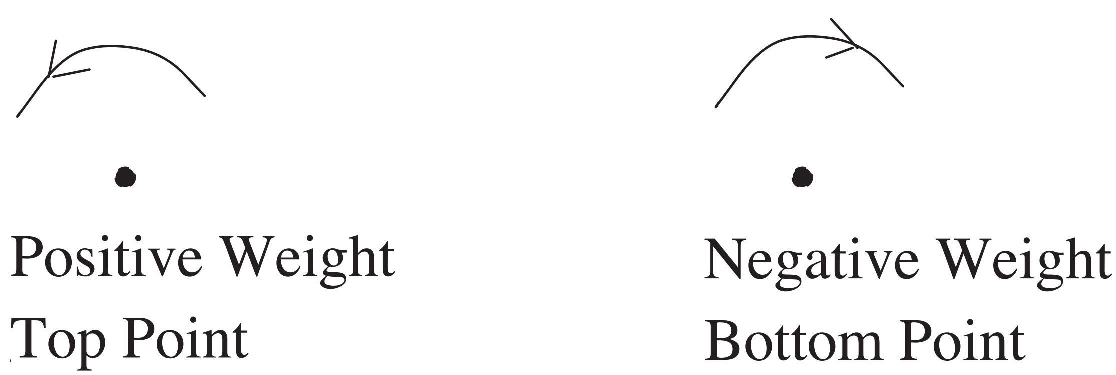

Figure 7: Oriented points in the two-sided plane