Basic Geometry for COMP 4560

This document summarizes a number of issues associated with the mathematical description and manipulation of points, lines, planes and similar basic geometric objects along with some discussion of related programming issues.

Coordinate Systems

Coordinate systems are tools for keeping track of locations in a plane (for two-dimensional work) and in space (for three-dimensional work). These are set up when, where, and in a manner convenient to the user.

Figure 1

Traditionally in mathematical work in a plane, the coordinate **axes** are represented by two straight lines intersecting at right angles as shown in Figure 1. The horizontal line is called the **horizontal axis**, or quite commonly, the **x-axis**, and the vertical line is called the **vertical axis**, or quite commonly, the **y-axis**. (The terms 'x-axis' and 'y-axis' would be changed if you were working with variables with different names, of course.) Each axis has a marked numerical scale of distances, starting in principle at $-\infty$ at the extreme left for the x-axis or the extreme bottom for the y-axis, and increasing to, in principle $+\infty$ at the extreme right for the x-axis and the extreme top for the y-axis. Arrowheads are normally drawn on the right end of the horizontal axis and the top of the vertical axis to indicate the direction of increase in these scales of distance values.

The point where the two axes cross is called the **origin**, and occurs where x = 0 and y = 0.

Generally, the scales along these axes are uniform in principle (equal physical distances along the axis correspond to equal numerical increments along the scale), and the scales along both the horizontal and vertical axes are considered to be identical in principle, even if in practice they may not be drawn that way physically.

Once such a coordinate system is set up, the location of any point in the plane is specified by stating its **coordinates** as an ordered pair of numbers, (*horizontal coordinate*, *vertical coordinate*), relative to the scales on the axes. For example, the point **P** in Figure 1 is said to have coordinates (3, 2), since it is above the horizontal scale value of 3 and on the horizontal level corresponding to the vertical scale value of 2. In the same way, point **Q** in Figure 1 has coordinates (-3, 1) and point **R** has coordinates (2, -3). A negative horizontal coordinate means that the point is to the left of the origin, and a negative vertical coordinate means that the point is below the origin.

Thus, a point and its location in a plane is represented mathematically by an ordered pair of numbers specifying its location relative to some predefined coordinate system. More complicated geometric figures (straight lines, curves, etc.) are specified by providing a formula or procedure for determining the coordinates of the sequence of points from which they are formed.

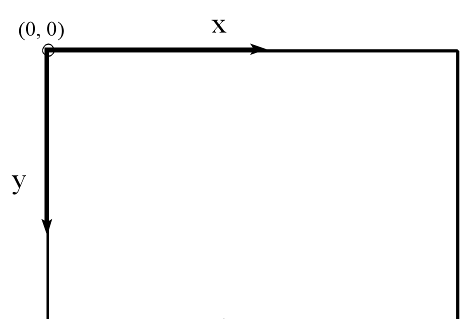

Figure 2

As mentioned, Figure 1 is the usual format of coordinate systems used for two-dimensional work in mathematics textbooks and the like. However, given the general principles described above, it is possible to define coordinate systems in other useful ways. For instance, as a result of hardware design decisions in the early days of microcomputers, many video systems assume a coordinate system on screen as shown in Figure 2. That is, the origin is at the top left corner of the screen, x increases to the right, and y increases downwards. The so-called **video mode** in use will determine the largest physical coordinates appearing on screen in both the x- and the y-directions. (For example, in standard high resolution VGA mode on an IBM-type microcomputer, the value of y ranges from 0 at the top of the screen to 479 at the bottom, and the value of x ranges from 0 at the left edge of the screen to 639 at the right edge. Various "super VGA" video modes with coordinate ranges several times larger are commonly in use now as well.) Locations with coordinates outside of these ranges still exist -- they just don't show up on the screen.

You'll see when we discuss transformations (also see the document on transformations) that we can easily redefine the screen coordinate system to some more convenient form via software. In addition, operating systems such as MS Windows often provide a variety of automatic mapping modes, which amount to alternative coordinate systems.

In three-dimensions, things are a little more complicated. First of all, the video screen remains a two-dimensional surface, and all sorts of problems arise when we try to accurately represent a three-dimensional object with a two-dimensional image. Nevertheless, you can think of a three-dimensional coordinate system as arising by adding a third coordinate axis (often called the **z-axis**) perpendicular to the **x, y-plane** of a two-dimensional coordinate system so that the three axes intersect at the coordinates x = 0, y = 0, z = 0. Figure 3 attempts to show this in a sort of oblique view -- you need to imagine the x- and y-axes as being in the plane of the paper, and the z-axis coming out of the plane towards you.

Figure 3

Locations of points in three-dimensions now are specified by ordered triples of numbers, (x, y, z), giving the location of the point relative to the distance scales measured out along the three mutually perpendicular coordinate axes. For example, the point with coordinates (2, 3, 5) is located by, say, starting at the origin, moving two units rightwards (in the same direction as the x-axis), then turning upwards and moving 3 units (in the same direction as the y-axis), and finally moving 5 units in the same direction as the z-axis.

We've referred to the three-dimensional coordinate axes in Figure 3 as a **right-handed coordinate system**. This terminology comes from the fact that if you hold your right hand so that the fingers curl from x to y, then the thumb points in the positive direction of z.

In Figure 4 we show just the positive coordinate parts of a three-dimensional coordinate system in which the x- and y-axes are in the plane of the page and the z-axis is to be viewed as going away from the viewer into the page. The difference between the coordinate systems in Figures 3 and 4 is that the z-axis has opposite orientations in the two. Figure 4 depicts a so-called **left-handed coordinate system**, because here, if you curl the fingers of your left hand from x to y, the thumb of your left hand points in the positive direction of the z-axis. With a bit of work, you should be able to convince yourself that the two coordinate systems depicted in Figures 3 and 4 really are distinct arrangements of the three axes -- you cannot take one of them and by simply re-orienting it in space, make it identical to the other.

Figure 4

A common model when creating video images is to work with a three-dimensional coordinate system in which the xy-plane is the plane of the video screen, and the video screen is simulating a "viewport" through which the viewer "looks" at three-dimensional objects on the other side of the viewport. Computer graphics practitioners often prefer to use a left-handed coordinate system, because then the z-coordinates of the objects are positive numbers. This is a choice of convenience or style, only. Any geometrical operation you can do within a left-handed coordinate system has a readily devised equivalent for a right-handed coordinate system. Switching from one type of coordinate system to another (not recommended unless absolutely necessary!) usually just introduces minus signs here and there in formulas. It doesn't really matter whether you opt to use a right-handed or left-handed coordinate system in any particular application, but once you make the choice, you need to ensure that all calculations are done and all results are interpreted consistent with that choice.

(Technically, you can't have a distinction similar to handedness in two dimensions, but what appears to be a two-dimensional situation can sometimes really be a three-dimensional situation in disguise, and the distinction of handedness can rear its ugly head, and -- one might say -- kick the unwary programmer in the head. One such situation is in modelling rotations in two dimensions. This is really a three dimensional situation because it involves at least implicitly introducing an axis of rotation which is perpendicular to the two-dimensional plane in which the image exists. One effect of this is that formulas for a rotation counterclockwise will be different for the coordinate systems in Figures 1 and 2, since they generalize to three-dimensional coordinate systems of opposite handedness when the same axis of rotation through the origin is affixed to each.)

<u>Points</u>

The location of points is always specified by stating coordinates relative to some coordinate system. Locations on the planar video screen are specified by a pair of coordinates, often denoted (x, y). Locations in three-dimensional space require three coordinates: (x, y, z).

It is often convenient to append a somewhat fictitious coordinate, h, to obtain so-called **homogeneous coordinates**:

(x, y, h)	in a plane (two dimensions)
(x, y, z, h)	in three dimensions

By convention, the "real world" corresponds to points with h = 1. Before plotting a point for which $h \neq 1$, you must usually restore homogeneity by dividing all coordinates by the value of h:

$$(x, \quad y, \quad z, \quad h) \to \left(\frac{x}{h}, \quad \frac{y}{h}, \quad \frac{z}{h}, \quad 1\right) \qquad \text{(BG-1)}$$

Many programmers and program libraries define structures or records of type POINT with members x, y, z, h (in this case to handle points in three dimensions). In this course, we will generally simply represent points as rows of two-subscript arrays with the second subscript running from 2 to 4, depending on how many coordinates are required to represent the location of each point.

<u>Example</u>:
Figure 5 shows a "wireframe" rectangular box with size 5 units in the x direction, 7 units in the y direction and 12 units in the z direction. When one corner is located at the origin and the box is on the positive side of the origin with respect to each coordinate axis, the eight numbered points forming the vertices of the box have space coordinates:

Figure 5

point number:	**x**	**y**	**z**
0	0	0	0
1	5	0	0
2	5	0	12
3	0	0	12
4	0	7	0
5	5	7	0
6	5	7	12
7	0	7	12

The corresponding array of homogeneous coordinates would have the coordinates of each point forming a row of four values:

	x	*y*	*z*	*h*
point 0:	0	0	0	1
point 1:	5	0	0	1
point 2:	5	0	12	1
point 3:	0	0	12	1
point 4:	0	7	0	1
point 5:	5	7	0	1
point 6:	5	7	12	1
point 7:	0	7	12	1

◆◆◆

Distances

The distance between points $\mathbf{P}_1 = (x_1, y_1, z_1)$ and $\mathbf{P}_2 = (x_2, y_2, z_2)$ in three dimensions, is easily calculated from the coordinates of the points using Pythagoras's Theorem:

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \qquad \text{(BG-2)}$$

For two dimensions, just omit reference to *z* in this formula.

Example

The distance between $\mathbf{P}_1 = (5, 6, -3)$ and $\mathbf{P}_2 = (-4, 8, 7)$ is

$$d = \sqrt{(-4-5)^2 + (8-6)^2 + (7-(-3))^2} = \sqrt{81 + 4 + 100} = \sqrt{185} \cong 13.601$$

◆◆◆

Lines (Two Dimensions)

Any two distinct points in a plane ($\mathbf{P}_1$ and $\mathbf{P}_2$ in Figure 6) determine a straight line. The "equation of the line" is a formula or a set of formulas which we can use to compute the coordinates, (x, y), of points along the line.

In two dimensions, the most common form of the equation of a straight line is the so-called **slope-intercept** form:

$$y = mx + b \qquad \text{(BG-3a)}$$

where

$$m = \frac{y_2 - y_1}{x_2 - x_1} \qquad \text{(BG-3b)}$$

is the **slope** of the line, and

$$b = y_1 - mx_1 \qquad \text{(BG-3c)}$$

is the coordinate of the point where the line cuts the vertical axis, the so-called **vertical intercept**.

This slope-intercept form of the equation of a line is actually a refinement of what is more commonly referred to as the **symmetric form** of the equation of the line (particularly in a three dimensional context):

Figure 6.

$$\frac{y - y_1}{y_2 - y_1} = \frac{x - x_1}{x_2 - x_1} \qquad \text{(BG-4)}$$

We can rearrange this symmetric form into a form call the **general form** of the equation of the straight line, in which the right-hand side is zero. Start by moving the right-hand side expression to the left hand side:

$$\frac{y - y_1}{y_2 - y_1} - \frac{x - x_1}{x_2 - x_1} = 0$$

Now, split the fractions up according to the terms in their numerators:

$$\frac{y}{y_2 - y_1} - \frac{y_1}{y_2 - y_1} - \frac{x}{x_2 - x_1} + \frac{x_1}{x_2 - x_1} = 0$$

or

$$\left(\frac{-1}{x_2 - x_1}\right)x + \left(\frac{1}{y_2 - y_1}\right)y + \left(\frac{x_1}{x_2 - x_1} - \frac{y_1}{y_2 - y_1}\right) = 0$$

Remember that all the symbols with subscripts here will be actual numbers in any specific situation. Thus, the last equation can be written in the much simpler form:

$$ax + by + c = 0 \qquad \text{(BG-5a)}$$

with

$$a = \frac{-1}{x_2 - x_1} \qquad b = \frac{1}{y_2 - y_1} \qquad \text{and} \qquad c = \frac{x_1}{x_2 - x_1} - \frac{y_1}{y_2 - y_1} \qquad \text{(BG-5b)}$$

In practice, the form (BG-5a) is probably as easily obtained by first writing down (BG-4) with the actual values for x_1, y_1, x_2, and y_2 substituted in, and then rearranging that result into the form (BG-5a), rather than necessarily applying formulas (BG-5b) to get the equation (BG-5a). This last bit of work demonstrates that the form (BG-5a) can be achieved.

A fourth way of writing the equation of a straight line (and perhaps the most useful form of all in computer graphics) is the so-called **parametric form**:

$$\begin{aligned} x &= x_1 + t(x_2 - x_1) \\ y &= y_1 + t(y_2 - y_1) \end{aligned} \qquad \text{(BG-6)}$$

Figure 7.

In this case, the points on the line are generated by substituting different values of t (with $t = -\infty$ to $t = +\infty$ generating the entire infinitely extended line passing through the two points). In particular, $t = 0$ gives the point $\mathbf{P}_1$ and $t = 1$ gives the point $\mathbf{P}_2$. In fact, as t varies uniformly from 0 to 1, (x, y) will traverse the line segment from $\mathbf{P}_1$ to $\mathbf{P}_2$ uniformly. Thus $t = 0.5$ gives a point on the line midway between $\mathbf{P}_1$ and $\mathbf{P}_2$, $t = 0.6$ gives another point on the line 60% of the way from $\mathbf{P}_1$ to $\mathbf{P}_2$, and so on. When $t < 0$ or when $t > 1$, we get points on the line outside of the segment between $\mathbf{P}_1$ and $\mathbf{P}_2$ ($t < 0$ gives the part of the line from $\mathbf{P}_1$ in a direction away from $\mathbf{P}_2$, and $t > 1$ gives the part of the line from $\mathbf{P}_2$ in a direction away from $\mathbf{P}_1$ — see Figure 7).

In class, we will look at one immediately interesting application of the parametric form of the equation of a straight line to carry out so-called "tweening" or "morphing" operations, where one simple image is smoothly distorted step-by-step to form another image (taking advantage of the fact that replotting (x, y) for successive values of t in a uniform sequence from 0 to 1 will give a succession of points starting at $\mathbf{P}_1$ and moving smoothly to end at $\mathbf{P}_2$).

The properties of the parametric form of the equation of a straight line are also exploited in line clipping algorithms, in which, at its simplest, we need to determine the part of a line which lies inside a particular rectangular viewport. More generally, parametric representations of equations of more general shapes than straight lines provide significant advantages over the more usual equations giving y in terms of x directly. Most of the work we will do in generating curves (and hence representations of curved surfaces, etc.) will involve parametric equations.

Example:
Find the equation of the straight line through the points $\mathbf{P}_1 = (80, 85)$ and $\mathbf{P}_2 = (120, 145)$, expressed in each of the four forms just described.

Solution:
Here, $\mathbf{P}_1 = (80, 85)$ and so $x_1 = 80$, $y_1 = 85$, and $\mathbf{P}_2 = (120, 145)$ so that $x_2 = 120$, and $y_2 = 145$.

Start with the slope-intercept form:

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{145 - 85}{120 - 80} = \frac{60}{40} = 1.5$$

using formula (BG-3b), and then from formula (BG-3c) we get

$$b = y_1 - mx_1 = 85 - (1.5)(80) = -35$$

So, the desired equation is

$$y = 1.5x - 35 \qquad (*)$$

Note that having initially determined the value of m, we could solve for b by realizing that the equation,

$$y = 1.5x + b$$

must be satisfied by the coordinates of either of the two points, $\mathbf{P}_1$ or $\mathbf{P}_2$ that the line passes through. Substituting in the coordinates of the first point, $\mathbf{P}_1 = (80, 85)$, we get

$$85 = (1.5)(80) + b$$

which can be solved for b to get

$$b = 85 - (1.5)(80) = -35,$$

as before.

The symmetric form of the equation of this line can be written down directly from formula (BG-4) above:

$$\frac{y - 85}{145 - 85} = \frac{x - 80}{120 - 80}$$

or simplified,

$$\frac{y-85}{60}=\frac{x-80}{40} \qquad (**)$$

Although it is possible to get the general form of the equation of this line by substituting the coordinates of $\mathbf{P}_1$ and $\mathbf{P}_2$ into formulas (BG-5), it is easier here simply to rearrange either (*) or (**) so that the right-hand side is zero. For instance, starting with (*),

$$y = 1.5x - 35$$

we get

$$-1.5x + y + 35 = 0 \qquad (***a)$$

which is fine. Had we used formulas (BG-5), we would have gotten

$$-\frac{1}{40}x+\frac{1}{60}y+\frac{7}{12}=0 \qquad (***b)$$

Despite it's different appearance, this is actually the same equation as (***a) just above — to see this, just multiply (***b) by 60.

Finally, the parametric form of the equation of this line is obtained by simply substituting the values of the coordinates of these two points into formulas (BG-6):

$$x = 80 + t(120 - 80) \quad \text{or} \quad x = 80 + 40t$$
$$y = 85 + t(145 - 85) \quad \text{or} \quad y = 85 + 60t \qquad (****)$$

You can easily confirm that substituting t = 0 into (****) gives the coordinates of point $\mathbf{P}_1$, and that substituting t = 1 into (****) gives the coordinates of point $\mathbf{P}_2$. Further, if you were to draw an actual graph of this situation, you could easily confirm that substituting a succession of values of t between 0 and 1 into (****) will produce coordinates of a succession of points on the line between $\mathbf{P}_1$ and $\mathbf{P}_2$. (You might try this as an experiment using a spreadsheet application such as Microsoft Excel to reduce the tedium of arithmetic and graphing.)

◆◆◆

Finding the Point Where Two Lines Intersect

Non-parallel lines in two dimensions always intersect at exactly one point. This is the unique point with coordinates that satisfy the equations of both lines simultaneously.

The best procedure for determining the coordinates of this point of intersection depends on the form in which the equations of the lines are given. If the equations are in the slope intercept form,

$$y = m_1x + b_1 \quad \text{and} \quad y = m_2x + b_2$$

then, since the y-coordinate on both lines is the same at the point of intersection, we must have that

$$m_1x + b_1 = m_2x + b_2$$

at that point only. But, this equation can be solved for the x-coordinate at that point:

$$x=\frac{b_2-b_1}{m_1-m_2}$$

Then, substituting this x-coordinate into either original equation must give the value of the corresponding y-coordinate, completing the determination.

Example:

Find the coordinates of the point at which the lines

$$y = 8x - 25 \quad \text{and} \quad y = 5x + 50$$

intersect.

Solution:
At the point where the lines intersect, we must have

$$8x - 25 = 5x + 50$$

so that

$$3x = 75$$

or

$$x = 25.$$

Substituting this into the first equation then gives

$$y = (8)(25) - 25 = 200 - 25 = 175$$

You would have gotten the same result by substituting x = 25 into the second equation as well. Thus, we conclude that the point where these two lines intersect has coordinates (25, 175).

◆◆◆

Two lines with the same slope (same value of m) are parallel. It they also have the same value of b, they are obviously the same line, and so intersect, so to speak, at every point along that line. Two lines with the same value of m but different values of b are distinct parallel lines and so do not intersect. In such a case, the procedure described above fails, because in solving for the x-coordinate of the intersection point, we would need to divide by zero.

If the equations of the lines are given in symmetric form, it's probably best to rearrange them into slope-intercept form or general form before determining the coordinates of the point of intersection.

If the equations are given in general form,

$$a_1 x + b_1 y + c_1 = 0$$
$$a_2 x + b_2 y + c_2 = 0$$

the coordinates of the point of intersection are obtained by solving these two equations as a system of two simultaneous linear equations. There are many essentially equivalent ways of doing this — several are described in the accompanying document on matrices.

Example:
Find the coordinates of the point of intersection of the lines specified by

$$5x - 3y + 7 = 0 \qquad \text{and} \qquad 4x + 7y - 3 = 0$$

Solution:
The system of equations to be solved can be written

$$5x - 3y = -7$$
$$4x + 7y = 3$$

One way to solve these is to multiply the first equation by -4 and the second equation by +5 (these factors are chosen so that in the next step all x's disappear):

$$-20x + 12y = 28$$
$$20x + 35y = 15$$

Now, add these two equations together to get

$$47y = 43$$

indicating $y = 43/47 \cong 0.91489$ is the y-coordinate of the point of intersection of the two lines. We could substitute this value of y into either original equation and solve for the corresponding value of x (we'd get the same result regardless of which of the two equations we used). Instead, here, we repeat the same procedure, only now multiplying the first equation by +7 and the second equation by +3 to get

$$35x - 21y = -49$$
$$12x + 21y = 9$$

which, upon adding, gives

$$47x = -40$$

or

$$x = \frac{-40}{47} \cong -0.85106$$

Thus, the two lines intersect at the point with coordinates approximately (-0.85106, 0.91489). You can easily verify that these coordinates satisfy both of the original equations, so this point does indeed lie on both lines simultaneously.

◆◆◆

When the equations of the lines are given in parametric form (or are to be first written in parametric form), some care must be taken. We'll illustrate the process with an example.

Example:
One line passes through the points (80, 85) and (120, 145). A second line passes through the points (360, 30) and (40, 210). Find the coordinates of the point where these two lines intersect, working from the parametric form of the equations of these lines.

Solution:
The first line has parametric equations

$$x = 80 + 40t$$
$$y = 85 + 60t$$

where the parameter t specifies position along this line relative to the "starting point" (80, 85). When we write down the parametric equations of the second line, we must use a different symbol for the parameter (since t now indicates specifically position on the first line). We will use the symbol s, so that the parametric equations for the second line look like

$$x = 360 - 320s$$

$$y = 30 + 180s$$

Now, the point where these two lines intersect is the point where both lines have the same x-coordinates:

$$80 + 40t = 360 - 320s \qquad (\#)$$

<u>and</u> the same y-coordinates

$$85 + 60t = 30 + 180s \qquad (\#\#)$$

Thus, the conditions determining the location of the point of intersection form a system of two linear equations in the two unknowns s and t. Rearranging (#) and (##) to get the terms containing s and t on the left-hand side, gives

$$\begin{aligned} 320s + 40t &= 280 \\ -180s + 60t &= -55 \end{aligned}$$

Solving this system of equations in one of the usual ways gives $s = 95/132 \cong 0.71920$ and $t = 41/33 \cong 1.24242$. To get the (x, y)-coordinates of the point of intersection, just substitute this value of t into the parametric equations of the first line:

$$x = 80 + 40(41/33) = 4280/33 \cong 129.70$$

and

$$y = 85 + 60(41/33) = 1755/11 \cong 159.54.$$

You would get exactly the same results if you substituted $s = 95/132$ into the parametric equations of the second line. Thus, the two lines described intersect at the point $(x, y) \cong (129.70, 159.54)$.

◆◆◆

<u>Parallel and Perpendicular Lines</u>

As already noted, **parallel lines** are related by having the same slopes. This is easy to recognize from the slope-intercept form of their equations. For example,

$$y = 5x + 7$$

gives a line parallel to the line with equation

$$y = 5x - 43$$

On the other hand, the lines with equations

$$y = 5x + 7 \qquad \text{and} \qquad y = 12x + 7$$

are not parallel, because the first has a slope of 5 whereas the second has a slope of 12.

A little more work is required to determine whether two lines are parallel given their equations in parametric form. If we have

first line:	$x = a + bt$	$y = c + dt$	
second line:	$x = a' + b't$	$y = c' + d't$	(BG-7)

then, the two lines will be parallel only if

$$\frac{d}{b} = \frac{d'}{b'},$$

since, you can verify, these two ratios are actually the slopes of the two lines.

Two lines are perpendicular if their slopes are negative reciprocals of each other. Thus,

$$y = m_1x + b_1 \qquad \text{and} \qquad y = m_2x + b_2$$

will represent perpendicular lines if

$$m_1 = -\frac{1}{m_2} \qquad \text{or} \qquad m_1 m_2 = -1.$$

When the equations are given in the parametric form (BG-7), this condition becomes

$$\frac{d}{b} = -\frac{b'}{d'}$$

Example:
Find the equations of the lines through the point (10, 30) which are parallel to and perpendicular to the line $y = 5x + 12$.

Solution:
From the coefficient of x in the stated equation, the line given above has a slope of 5. The slope of the parallel line must be 5 as well, so its equation must have the form

$$y = 5x + b.$$

But the point (x, y) = (10, 30) must be on this line, so substituting in the values of x and y, we get

$$30 = (5)(10) + b$$

giving

$$b = 30 - (5)(10) = -20$$

Thus, the equation of the parallel line is

$$y = 5x - 20.$$

The perpendicular line must have a slope of -1/5, and so its equation must have the form

$$y = -\frac{1}{5}x + b$$

Again, since (10, 30) is a point on this line, it must be true that

$$30 = -\frac{1}{5}(10) + b$$

so that

$$b = 30 + 2 = 32.$$

Thus, the equation of the perpendicular line is

$$y = -\frac{1}{5}x + 32$$

◆◆◆

Angle Between Two Lines

The easiest approach to getting the angle between two lines is to first work out their slopes, m_1 and m_2, as shown in Figure 8. Number the lines so that line #1 makes the biggest angle counterclockwise with the x-axis (so that angle $\alpha_1 > \alpha_2$). Then, the angle θ (from line #1 to line #2, counterclockwise) is specified by

$$\tan\theta = \frac{m_1 - m_2}{1 - m_1 m_2} \qquad \text{(BG-8)}$$

Special cases are:

$m_1 = m_2$: the lines are parallel and so the angle between them is taken to be 0^o.

$m_1 m_2 = -1$: this causes division by zero, so the formula (BG-8) fails. However, this is just the condition that the two lines are perpendicular, so we can take $\theta = 90^o$ or $\pi/2$ radians in this case.

Distance From a Point to a Line

We just state the formula here. If the equation of the line is

$$Ax + By + C = 0,$$

in general form, then the distance from the point **P** = (p_x, p_y) to this line is

$$\frac{\left|A p_x + B p_y + C\right|}{\sqrt{A^2 + B^2}}$$

The vertical bars in the numerator stand for **absolute value** — meaning that if $Ap_x + Bp_y + C$ turns out to be negative, just ignore the minus sign. This distance is measured along the direction perpendicular to the given line and so is the shortest distance from the point **P** to the given line. If **P** is on the line, the result here has a value of zero since the coordinates of **P** satisfy the equation of the line, which states that the numerator is equal to zero.

Lines in Three Dimensions

As in two dimensions, a line in three dimensions is fixed by stating the locations of two points through which it passes. In this case, each point is specified by three coordinates, (x, y, z). The notion of slope isn't as useful in three dimensions, and so the two most commonly used forms for the equation of a straight line in three dimensions is the symmetric form and the parametric form.

The symmetric form of the equation of a line through points $\mathbf{P}_1 = (x_1, y_1, z_1)$ and $\mathbf{P}_2 = (x_2, y_2, z_2)$ is simply:

$$\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1} = \frac{z - z_1}{z_2 - z_1} \qquad \text{(BG-10)}$$

This is really a collection of three distinct equations obtained by equating each possible pair of these expressions.

The parametric form of the equation of this line in the triple:

$$\begin{aligned} x &= x_1 + t(x_2 - x_1) \\ y &= y_1 + t(y_2 - y_1) \\ z &= z_1 + t(z_2 - z_1) \end{aligned} \qquad \text{(BG-11)}$$

As in the two-dimensional case, t = 0 corresponds here to the point $\mathbf{P}_1$ and t = 1 corresponds to the point $\mathbf{P}_2$.

Example:
Write down the equation of the line through the points $\mathbf{P}_1 = (5, 6, -3)$ and $\mathbf{P}_2 = (-4, 8, 7)$ in both symmetric and parametric forms. Then, demonstrate how each form of equation can be used to determine the coordinates of the point where this line cuts the xy-plane (z = 0).

Solution:
The symmetric form of the equation of this line is

$$\frac{x-5}{-4-5} = \frac{y-6}{8-6} = \frac{z-(-3)}{7-(-3)}$$

or

$$\frac{x-5}{-9} = \frac{y-6}{2} = \frac{z+3}{10}$$

This can also be written as the three simple equations:

$$\frac{x-5}{-9} = \frac{y-6}{2}, \qquad \frac{x-5}{-9} = \frac{z+3}{10}, \qquad \frac{y-6}{2} = \frac{z+3}{10}$$

The parametric form of the equation of this line consists of the three formulas:

$$\begin{array}{lcl} x = 5 + t(-4 - 5) & \text{or} & x = 5 - 9t \\ y = 6 + t(8 - 6) & \text{or} & y = 6 + 2t \\ z = -3 + t(7 - (-3)) & \text{or} & z = -3 + 10t \end{array}$$

To find the coordinates of the point where this line intersects the xy-plane, we need to determine the values of x and y when z = 0. To do this starting with the symmetric form of the equation of the line, we begin with the version which involves x and z to solve for the desired x-coordinate,

$$\frac{x-5}{-9} = \frac{z+3}{10} \Rightarrow \frac{x-5}{-9} = \frac{0+3}{10} \quad \text{which gives: } x-5 = -9\left(\frac{3}{10}\right) \quad \text{or} \quad x = 2.3$$

and the version involving y and z to solve for the desired y-coordinate:

$$\frac{y-6}{2} = \frac{z+3}{10} \Rightarrow \frac{y-6}{2} = \frac{0+3}{10} \quad \text{which gives: } y-6 = 2\left(\frac{3}{10}\right) \quad \text{or} \quad y = 6.6$$

Thus, it appears that this line intersects the xy-plane at the point (x, y) = (2.3, 6.6).

To use the parametric form of the equation of this line to determine the point of intersection with the xy-plane, we first note that since this point will have z = 0, it must correspond to the value of t given by

$$0 = -3 + 10t$$

or

$$10t = 3$$

or

$$t = 3/10.$$

Substituting this value of t into the equations for x and y then gives

$$x = 5 - 9(0.3) = 5 - 2.7 = 2.3$$

and

$$y = 6 + 2(0.3) = 6 + 0.6 = 6.6.$$

This result agrees, of course, with the result obtained using the symmetric form of the equation of this line.

Note also that since $\mathbf{P}_1$ has a negative z-coordinate, and $\mathbf{P}_2$ has a positive z-coordinate, the point where this line cuts the z = 0 plane must be between $\mathbf{P}_1$ and $\mathbf{P}_2$. Since the point we found corresponds to t = 0.3, a value between 0 and 1, this observation is confirmed — the line does intersect the xy-plane at a location between $\mathbf{P}_1$ and $\mathbf{P}_2$.

◆◆◆

Direction Cosines

The denominators of the terms in the symmetric form of the equation of a straight line, or the coefficients of t in the parametric form, essentially contain the information specifying the direction of the line in space. As a result, writing

$$\frac{x-a}{\alpha} = \frac{y-b}{\beta} = \frac{z-c}{\gamma} \qquad \text{(BG-12a)}$$

or

$$\begin{aligned} x &= a + \alpha t \\ y &= b + \beta t \\ z &= c + \gamma t \end{aligned} \qquad \text{(BG-12b)}$$

we could refer to the values (α, β, γ) as "direction numbers" for the line. In fact, it is really the relative sizes of these numbers which determine the direction of the line, and so, we can scale the values by some constant factor without changing the direction of the line represented by the equations. One particularly useful choice of scale is the one in which the sum of the squares of these values is equal to 1. This is achieved by dividing all three values by the value of $\sqrt{\alpha^2+\beta^2+\gamma^2}$, (a process called normalization), to get:

$$\ell = \frac{\alpha}{\sqrt{\alpha^2+\beta^2+\gamma^2}}, \quad m = \frac{\beta}{\sqrt{\alpha^2+\beta^2+\gamma^2}}, \quad n = \frac{\gamma}{\sqrt{\alpha^2+\beta^2+\gamma^2}} \qquad \text{(BG-13)}$$

These numbers are called the **direction cosines** of the line because they are equal to the cosines of the angles between the line and the direction of the x-, y-, and z-axes, respectively. Note that

$$\ell^2 + m^2 + n^2 = 1 \qquad \text{(BG-14)}$$

as a result of the way in which these values are constructed.

The equation of a line in three dimensions can thus be written in the slightly revised forms:

$$\frac{x-a}{\ell} = \frac{y-b}{m} = \frac{z-c}{n} \qquad \text{(BG-15a)}$$

or

$$\begin{aligned} x &= a + \ell t \\ y &= b + mt \\ z &= c + nt \end{aligned} \qquad \text{(BG-15b)}$$

(though plugging the same value of t into equations (BG-12b) and (BG-15b) will generally give different points on the line, because the scale of distance has been changed.)

For lines in three dimensions, the direction cosines play somewhat the same role as does slope for lines in two dimensions (though their precise interpretations are different — the slope of a line in two dimensions is more of a "direction tangent"). For instance, two lines in three dimensions are parallel if their direction cosines are identical. Two lines with direction cosines (l_1, m_1, n_1) and (l_2, m_2, n_2) are perpendicular if

$$l_1 l_2 + m_1 m_2 + n_1 n_2 = 0 \qquad \text{(BG-16)}$$

Planes

There are two ways of "nailing down" a plane in three-dimensional space.

(i) By giving the coordinates of one point (x_0, y_0, z_0) in the plane, and the direction numbers, (A, B, C), of a line perpendicular to the plane. The equation of this plane is

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0 \qquad \text{(BG-17)}$$

which can be written more simply as

$$Ax + By + Cz + D = 0 \qquad \text{(BG-18a)}$$

where D is the constant

$$D = -Ax_0 - By_0 - Cz_0 \qquad \text{(BG-18b)}$$

(ii) By giving the coordinates of three points in the plane which are not all in a straight line. There are several ways to get the equation of the plane from this information. One approach is to use the coordinates of the three points to get direction numbers for two lines which lie within the plane, use these to compute direction numbers of lines perpendicular to the plane, and then apply procedure (i), formula (BG-17), above. Suppose the coordinates of the three given points in the plane are $\mathbf{P}_1 = (x_1, y_1, z_1)$, $\mathbf{P}_2 = (x_2, y_2, z_2)$, and $\mathbf{P}_3 = (x_3, y_3, z_3)$. Then, direction numbers of two different lines, L_1 and L_2 in the plane are given by

$$L_1: \ (x_2 - x_1, y_2 - y_1, z_2 - z_1) = (\alpha_1, \beta_1, \gamma_1), \qquad \text{say, and,}$$
$$L_2: \ (x_3 - x_1, y_3 - y_1, z_3 - z_1) = (\alpha_2, \beta_2, \gamma_2), \qquad \text{say.}$$

These are guaranteed to be direction numbers of distinct lines because of our earlier requirement that the three points not be on the same straight line (that would mean these two lines form two of the three sides of an actual triangle). It may already have occurred to you that direction numbers of a line are equivalent to components of a vector in the direction of the line. Here, we essentially now have two vectors parallel to the plane of interest, and we need to find the components of a third vector perpendicular to these two, since its components will then amount to direction numbers of a line perpendicular to the plane. You can refer to Remark 8 on page 3 of the document "Vectors" to find that the so-called cross product of two vectors gives a third vector which is perpendicular to the first two. Applying the formulas given there, we get for our line perpendicular to the plane here the direction numbers:

$$\begin{aligned} \alpha_3 &= \beta_1 \gamma_2 - \gamma_1 \beta_2 \\ \beta_3 &= \gamma_1 \alpha_2 - \alpha_1 \gamma_2 \\ \gamma_3 &= \alpha_1 \beta_2 - \beta_1 \alpha_2 \end{aligned} \qquad \text{(BG-19a)}$$

and so the equation of the plane can be written, for instance, as

$$\alpha_3 (x - x_1) + \beta_3 (y - y_1) + \gamma_3 (z - z_1) = 0 \qquad \text{(BG-19b)}$$

A second method for determining the equation of a plane from the coordinates of the three non-collinear points $\mathbf{P}_1$, $\mathbf{P}_2$ and $\mathbf{P}_3$ lying in the plane exploits the fact that the coordinates of each point must satisfy the equation (BG-18a) above, for instance. Thus, we must have

$$\begin{aligned} Ax_1 + By_1 + Cz_1 + D &= 0 \\ Ax_2 + By_2 + Cz_2 + D &= 0 \\ Ax_3 + By_3 + Cz_3 + D &= 0 \end{aligned} \qquad \text{(BG-20)}$$

Here, the x_k, y_k, and z_k will be known numbers, and the goal is to determine the coefficients A, B, C, and D. At first this looks like a system of three linear equations, but involving the four unknowns A, B, C, and D. If that were really true, we would not be able to determine these constants (since generally, to have a unique solution, you must have exactly as many independent equations as you have unknowns). We know, this problem can only be an illusion, since we know that the coordinates of three points in the plane are adequate information to determine its equation (after all, we've already seen one procedure for doing so). The trick is to note the following. First of all, divide each of these equations by D. The first one becomes

$$(A/D)x_1 + (B/D)y_1 + (C/D)z_1 = -1 \qquad \text{(BG-21)}$$

after we move the fourth term (D/D) = 1 to the right hand side. Since A, B, C, and D stand for constant numerical values for a plane, so do A/D, B/D and C/D. This demonstrates that there really are only three unknowns here: the ratios A/D, B/D and C/D. (This result is not really a surprise, because the four constants in (BG-18a) above really came from just three in (BG-17)). Thus, our problem reduces to solving a system of three equations that look like (BG-21) for the three unknowns (A/D), (B/D) and (C/D).

Now, the second "trick", if you like. One of the methods described in the "Matrices" document for solving systems of n linear equations was Cramer's rule, which states that the value of the j^{th} unknown can be expressed as the quotient of two n × n determinants. The denominator of this quotient is the determinant of the coefficients of the n unknowns in the left-hand sides of the system of equations, and so is the same in calculating the values of all n unknowns. The numerator of the quotient for the j^{th} unknown is obtained by replacing the j^{th} column in this determinant of coefficients by the column of right-hand sides of the system of equations. Doing that here, we get:

$$\frac{A}{D} = \frac{\begin{vmatrix} -1 & y_1 & z_1 \\ -1 & y_2 & z_2 \\ -1 & y_3 & z_3 \end{vmatrix}}{\begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}}, \qquad \frac{B}{D} = \frac{\begin{vmatrix} x_1 & -1 & z_1 \\ x_2 & -1 & z_2 \\ x_3 & -1 & z_3 \end{vmatrix}}{\begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}}, \qquad \frac{C}{D} = \frac{\begin{vmatrix} x_1 & y_1 & -1 \\ x_2 & y_2 & -1 \\ x_3 & y_3 & -1 \end{vmatrix}}{\begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}} \qquad \text{(BG-22)}$$

Notice that each of these quotients has the same denominator, D, on the left-hand side, and the same determinantal denominator on the right-hand side. Since these three values amount to direction numbers of a line perpendicular to the plane of interest, and since they are only required to within a constant common factor, we can achieve the desired result — an equation for the plane through the three given points, by simply identifying D with the common denominator in all three expressions above, and A, B, and C, respectively, with each of the numerators. Applying the procedure for evaluating 3 × 3 determinants as given in the document on matrices, the final result is:

$$A = \begin{vmatrix} -1 & y_1 & z_1 \\ -1 & y_2 & z_2 \\ -1 & y_3 & z_3 \end{vmatrix} = y_1(z_3 - z_2) + y_2(z_1 - z_3) + y_3(z_2 - z_1)$$

$$B = \begin{vmatrix} x_1 & -1 & z_1 \\ x_2 & -1 & z_2 \\ x_3 & -1 & z_3 \end{vmatrix} = z_1(x_3 - x_2) + z_2(x_1 - x_3) + z_3(x_2 - x_1)$$

$$C = \begin{vmatrix} x_1 & y_1 & -1 \\ x_2 & y_2 & -1 \\ x_3 & y_3 & -1 \end{vmatrix} = x_1(y_3 - y_2) + x_2(y_1 - y_3) + x_3(y_2 - y_1)$$

and

$$D = \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix} = z_1(x_2 y_3 - y_2 x_3) + z_2(y_1 x_3 - x_1 y_3) + z_3(x_1 y_2 - y_1 x_2)$$

(BG-23)

These formulas look rather formidable, however, their evaluation involves just simple arithmetic (and what's a few multiplications to a fast microcomputer anyway?). You can probably recognize certain cyclic patterns in the terms in each formula as well.

Example:
Find the equation of the plane through the point **P** = (5, -3, 7), and perpendicular to the line given by

$$\{x = 3 + 7t,\ \ y = 4 - 5t,\ \ z = 8 + 11t\}$$

Solution:
The coefficients of t in the parametric equations of the line are direction numbers for that line. So, substituting these numbers for the symbols A, B, and C in formula (BG-17), we get

$$7(x - 5) - 5(y - (-3)) + 11(z - 7) = 0$$

or

$$7x - 5y + 11z - 127 = 0$$

as the equation of the plane.

Notice that if we substitute x = 5, y = -3 and z = 7 into this equation, the left-hand side becomes zero, confirming that whatever this figure describes, the given point **P** is part of it. We haven't taken the space in these notes to attempt to explain why formula (BG-17) should give a plane, but the ideas aren't complicated and can be found in many introductory textbooks on analytic geometry or vector analysis.

◆◆◆

Example:
Find the equation of the plane determined by the points $\mathbf{P}_1$ = (5, -3, 7), $\mathbf{P}_2$ = (-2, 9, 3) and $\mathbf{P}_3$ = (3, 4, -6).

Solution:
The procedure based on formulas (BG-19a) and (BG-19b) give the following results.

$$\alpha_1 = -2 - 5 = -7 \qquad \alpha_2 = 3 - 5 = -2$$
$$\beta_1 = 9 - (-3) = 12 \qquad \beta_2 = 4 - (-3) = 7$$
$$\gamma_1 = 3 - 7 = -4 \qquad \gamma_2 = -6 - 7 = -13$$

Then

$$\alpha_3 = (12)(-13) - (-4)(7) = -128$$
$$\beta_3 = (-4)(-2) - (-7)(-13) = -83$$
$$\gamma_3 = (-7)(7) - (-2)(12) = -25$$

Thus, an equation of the plane containing these three points is

$$-128(x - 5) - 83(y - (-3)) - 25(z - 7) = 0$$

or

$$-128x - 83y - 25z + 566 = 0.$$

You can easily verify that this equation is satisfied by the coordinates of all three of the originally given points, $\mathbf{P}_1$, $\mathbf{P}_2$, and $\mathbf{P}_3$. If the many minus signs here bother you, just multiply the whole equation by -1 to get the equivalent

$$128x + 83y + 25z - 566 = 0.$$

Formulas (BG-23) give the following results:

$$A = (-3)(-6 - 3) + (9)(7 - (-6)) + (4)(3 - 7) = (-3)(-9) + (9)(13) + (4)(-4) = 128$$
$$B = (7)(3 - (-2)) + (3)(5 - 3) + (-6)(-2 - 5) = (7)(5) + (3)(2) + (-6)(-7) = 83$$
$$C = (5)(4 - 9) + (-2)(-3 - 4) + (3)(9 - (-3)) = (5)(-5) + (-2)(-7) + (3)(12) = 25$$

and

$$D = (7)[(-2)(4) - (9)(3)] + (3)[(-3)(3) - (5)(4)] + (-6)[(5)(9) - (-3)(-2)]$$
$$= (7)(-35) + (3)(-29) + (-6)(39) = -566$$

This gives exactly the same equation as obtained above:

$$128x + 83y + 25z - 566 = 0.$$

◆◆◆

We wrap up this discussion of planes in three dimensions with a few miscellaneous remarks and results.

(i) The perpendicular distance between a point $\mathbf{P}_0 = (x_0, y_0, z_0)$ and the plane with equation $Ax + By + Cz + D = 0$ is given by

$$\text{distance} = \frac{|Ax_0 + By_0 + Cz_0 + D|}{\sqrt{A^2 + B^2 + C^2}} \qquad \text{(BG-24)}$$

This formula is very similar to (BG-9) for the perpendicular distance from a point to a line in two dimensions. Notice that by virtue of the equation of the plane, if $\mathbf{P}_0$ lies in the plane, this distance works out to be zero (which is sensible!).

(ii) Two planes are parallel if they are perpendicular to the same or parallel lines. This shows up as the coefficients of x, y, and z in the two equations being simple multiples of each other. Now, you can picture in your mind that two non-parallel planes intersect along a line. It is relatively easy to get the equation of that line in parametric form. Write the equations of the two planes generically as:

$$A_1 x + B_1 y + C_1 z + D_1 = 0$$
and
$$A_2 x + B_2 y + C_2 z + D_2 = 0 \qquad (*)$$

A point $\mathbf{P}_0 = (x_0, y_0, z_0)$ on this line of intersection is given by

$$x_0 = \frac{B_1 D_2 - B_2 D_1}{A_1 B_2 - A_2 B_1}, \qquad y_0 = \frac{A_2 D_1 - A_1 D_2}{A_1 B_2 - A_2 B_1}, \qquad z_0 = 0 \qquad \text{(BG-25a)}$$

(Just substitute $z_0 = 0$ into (*), and solve the resulting system of two equations in two unknowns.) Furthermore, if u, v, and w are direction numbers of this line of intersection, we must also have that

$$A_1 u + B_1 v + C_1 w = 0$$
and
$$A_2 u + B_2 v + C_2 w = 0 \qquad (**)$$

since the normals to the plane will be perpendicular to this line of intersection (remember that the coefficients A_k, B_k, and C_k in the equations of the two planes are direction numbers of lines perpendicular to their respective planes, and so perpendicular to any line lying in that plane). Now, one solution of the system (**) is

$$u = \frac{B_1 C_2 - B_2 C_1}{A_1 B_2 - A_2 B_1}, \qquad v = \frac{C_1 A_2 - C_2 A_1}{A_1 B_2 - A_2 B_1}, \qquad w = 1 \qquad \text{(BG-25b)}$$

(Here, substitute w = 1 into equations (**) and solve the resulting system of two linear equations in the two unknowns u and v.) Thus, we have coordinates of one point on this line of intersection, and direction numbers for this line, so the parametric equations of the line are simply:

$$x = x_0 + ut$$
$$y = y_0 + vt$$
$$z = t \qquad \text{(BG-25c)}$$

Of course, the direction numbers of this line can be scaled by any desired non-zero factor. Formulas (BG-25a) and (BG-25b) are obviously essential here, and in the version given above depend in detail on the assumptions that $z_0 = 0$ and $w = 1$. In the event that the numerators, $A_1 B_2 - B_1 A_2$, turn out to be zero, it will be necessary to pick another coordinate of $\mathbf{P}_0$ to be zero, and another direction number of the line of intersection to be 1.

Since two non-parallel planes intersect along a line in three dimensions, and any third non-parallel plane will intersect this line at exactly 1 point, we conclude that any three non-parallel planes intersect at exactly one point. You can determine the coordinates of this point by solving the equations of the three planes as a system of three simultaneous linear equations.

The Midpoint of a Line Segment

As mentioned in passing earlier, one way to find the coordinates of the point $\mathbf{P}_m = (x_m , y_m , z_m)$ on the line segment joining $\mathbf{P}_1 = (x_1 , y_1 , z_1)$ and $\mathbf{P}_1 = (x_2 , y_2 , z_2)$, and midway between these two points is to write down the parametric equations of this line and then substitute t = 0.5. However, the coordinates, (x_m , y_m , z_m), of this point can also be calculated directly using the formulas

$$x_m = \frac{x_1 + x_2}{2}, \qquad y_m = \frac{y_1 + y_2}{2}, \qquad z_m = \frac{z_1 + z_2}{2} \qquad \text{(BG-26)}$$

(These formulas work for two dimensions as well — just ignore the reference to z.)

Miscellaneous Results for Lines and Planes in Three Dimensions

This is a list of sometimes-useful results reproduced here without proofs or derivations.

(i) Non-parallel lines in three dimensions don't necessarily intersect. However, it is possible to compute the perpendicular between such lines, and from this, a distance of "closest approach" for the two lines. Writing the equations of two such lines in parametric form as

$$\begin{aligned} L_1: &\quad \mathbf{P} = \mathbf{P}_1 + \mathbf{u}\cdot t \\ L_2: &\quad \mathbf{P} = \mathbf{P}_2 + \mathbf{v}\cdot t \end{aligned} \qquad \text{(BG-27)}$$

where $\mathbf{P}_1 = (x_1 , y_1 , z_1)$ is a point on L_1 and $\mathbf{u} = (u_x , u_y , u_z)$ are direction numbers for L_1, and similarly with $\mathbf{P}_2$ and $\mathbf{v}$ for line L_2, we need to distinguish two cases:

(a) $\mathbf{u} \times \mathbf{v} \neq \mathbf{0}$ (the lines are not parallel)

$$\text{distance} = \frac{\left|(\mathbf{P}_1 - \mathbf{P}_2)\bullet(\mathbf{u} \times \mathbf{v})\right|}{|\mathbf{u} \times \mathbf{v}|} \qquad \text{(BG-28a)}$$

(b) $\mathbf{u} \times \mathbf{v} = \mathbf{0}$ (the lines are parallel)

$$\text{distance} = \frac{\left|(\mathbf{P}_1 - \mathbf{P}_2)\times \mathbf{u}\right|}{|\mathbf{u}|} \qquad \text{(BG-28b)}$$

These formulas involve vector arithmetic (the difference of two vectors, the dot product and the cross product), and the calculation of the lengths of vectors (indicated by the vertical bars). Details of this arithmetic are found in the document on vectors.

(ii) Although two non-parallel lines in three dimensions may not intersect, we can define the angle between them to be equal to the angle between the corresponding parallel lines through the origin. Thus, for two lines with equations written as in (BG-27), the smaller of the two angles between them is given by

$$\varphi = \cos^{-1}\left(\frac{\mathbf{u} \bullet \mathbf{v}}{|\mathbf{u}|\cdot|\mathbf{v}|}\right) \qquad \text{(BG-29)}$$

(iii) It is particularly easy to calculate the angle between a plane and a line if the point-normal form of the equation of the plane is given. Suppose the plane is perpendicular to the direction $\mathbf{n}$. Then, the cosine of the angle between this normal and the line $\mathbf{P} = \mathbf{P}_0 + \mathbf{u}t$ is given by
$\mathbf{u}\bullet\mathbf{n}/(|\mathbf{u}|\cdot|\mathbf{n}|)$. But, from basic trigonometry, this is equal to the sine of the complementary angle (the angle between the line and the plane). Thus, the angle between the line and the plane is given by

$$\phi = \sin^{-1}\left(\frac{\mathbf{u} \bullet \mathbf{n}}{|\mathbf{u}| \cdot |\mathbf{n}|}\right) \qquad \text{(BG-30)}$$

Note that this angle is measured in a plane containing **n**.

Circles and Spheres

Circles are a two-dimensional geometric shape and spheres are the corresponding three-dimensional shape. The points on the circumference of a circle of radius r, centered at the point (x_0, y_0), satisfy the equation

$$(x - x_0)^2 + (y - y_0)^2 = r^2 \qquad \text{(BG-31a)}$$

and the points on the surface of a sphere of radius r, centered at (x_0, y_0, z_0), satisfy the equation

$$(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 = r^2 \qquad \text{(BG-32a)}$$

These equations can also be written in parametric form. For the circle,

$$\begin{aligned} x &= x_0 + r\cdot\cos\theta \\ y &= y_0 + r\cdot\sin\theta, \qquad 0 \le \theta \le 2\pi \end{aligned} \qquad \text{(BG-31b)}$$

and for the sphere, one of many possibilities is:

$$\begin{aligned} x &= x_0 + r\cdot\cos\theta\cdot\sin\phi \\ y &= y_0 + r\cdot\sin\theta\cdot\sin\phi \\ z &= z_0 + r\cdot\cos\phi, \qquad 0 \le \theta \le 2\pi,\ \ 0 \le \phi \le \pi. \end{aligned} \qquad \text{(BG-32b)}$$

Neither of formulas (BG-31) are used for constructing actual images of circles on the video screen, since they involve "expensive" floating point calculations, and there are much more efficient algorithms available for generating the pixel pattern formed by a circle. These formulas are used in algorithms concerned with operations such as ray-tracing, lighting models, and so forth, in the generation of more sophisticated images on screen.

An important property of circles and spheres is that lines (or, perhaps more correctly, rays) drawn from the center of the figure intersect the figure perpendicularly. This property is exploited, for example, when modelling reflections of light rays from such shapes.

Other Geometric Shapes

So far, we've dealt in some detail with the properties of points, lines, and planes in two and three dimensions, and very briefly with circles and spheres. These are by far the most important basic geometric shapes in computer graphics.

There are, of course, many other mathematical formulas that give rise to useful basic shapes. These include ellipses/ellipsoids, and other so-called conic sections (parabolas and hyperbolas in two dimensions, and their three-dimensional generalizations). In two dimensions, line segments can be combined to form triangles, quadrilaterals, and so on — polygons of any desired shape. It is more usual to refer to irregular planar shapes bounded by straight line segments as **polylines**. A three-dimensional shape with planar faces is called a **polyhedron**, which are rendered on a two-dimensional video screen as a collection of polylines.

An important topic in computer graphics is the generation of empirical curves and surfaces (curves and surfaces which pass through or with shape controlled by the user-specified locations of a collection of points). The detailed study of such surfaces is (at least at present) well beyond the scope of this course, but a separate document is available which describes some of the methods for generating such curves.

As already noted in connection with circles, the equations given in this document for basic geometric objects such as lines and circles are rarely used directly for the production of images of such figures on the video screen (that is, for the detailed determination of which pixels must be set in order to produce the image of a line segment or a circular arc between two endpoints). These formulas require unnecessary amounts of floating point arithmetic in many cases. A variety of sophisticated algorithms involving just integer arithmetic are described in textbooks and the technical literature of computer graphics. Because most programming environments have access to functions implementing such basic algorithms (functions with names like Lineto() and Arc()), it isn't necessary for us to look in detail at these algorithms here.

Programming Issues

A considerable amount of appropriately organized information my be required to specify a three-dimensional shape sufficiently to render its image on a video screen. In its crudest form, such a shape would consist of a collection of planar faces (each of which may have a distinct "inside" and "outside"). These planar faces are bounded by edges consisting of consecutive line segments joining pairs of points or vertices. Essentially, the program must keep track of at least the following information:

(i) the coordinates of all vertices

As mentioned before, in this course we will usually do this using a two-subscript array, in which each 4-element row holds the (x, y, z, h) coordinates of one such point. Call this array Vertex[][].

(ii) an array of edge specification information

This is most simply done as a two-subscript integer-type array where each two element row gives the sequence numbers of the two vertices forming the endpoints of the edge. Call this array Edge[][]. Thus, to render the I'th edge in the image, we'd create the image of a line from

(Vertex[k1][1], Vertex[k1][2], Vertex[k1][3])

to

(Vertex[k2][1], Vertex[k2][2], Vertex[k2][3])

where

k1 = Edge[I][1]

and

k2 = Edge[I][2]

(iii) specification of a "poly-linal" face

This is a fairly complex task, and may be done in a variety ways depending on how much information you require for maximum efficiency in rendering the eventual image of the shape. For a given face, I, we generally need to store at least the list of sequence numbers of points forming the vertices of that face (usually ordered in some specific way, such as counterclockwise when viewed from the outside), as well as a list of sequence numbers of the edges of the face. It may be convenient to have one list giving the edges associated with each face, and another list giving the faces associated with each edge, or a combination of both. The information on the vertices of a face is often used to determine whether or not a particular face is visible, and the information about the edges is necessary in rendering the actual image of visible faces.

Specific implementation details are discussed in class, as well as in some of the other documents produced for this course (see the document on Viewing, for example). Most computer graphics reference books also devote considerable space to such details.