

1. COORDINATES

In order to locate a point in space one needs directions on how to reach it from some other known point. For example, suppose you are in a city and you want get to 1001 Tenth Street, 9th floor. If your starting point is First Street and Fifth Avenue and you asked for directions you might get the following instructions: “Go 5 blocks west along Fifth Avenue to Tenth Street and then go right on Tenth street until you get to the building with number 1001”. Once there, you would get to the 9th floor by taking an elevator. In your journey you used several rulers (numbered or scaled lines) to guide you and tell you when to make certain decisions about your getting to the place you wanted to go. The first two rulers were streets and the units of measurement were blocks or addresses of buildings. The last ruler was the path of the elevator numbered by the floor indicator. Whether the streets and avenues were at right angles or not was of no importance. What was important was that you went the correct distance in the right direction. In analytic geometry, the process of assigning addresses to points is called **coordinatisation**, the addresses being a sequence of numbers, called **coordinates**, with references to different rulers which make up the **coordinate frame**. In our example above, the place we wanted to go would have coordinates (10, 1001, 9) where the first number located the street, the second was the number of the building and the third was the floor.

1.1. Line Coordinates. If one is restricted to move on a given line L , then a single number suffices to specify one's location on that line once we specify an **origin** O and a point I of unit distance from the origin. The point I is called a **unit point**. A point P which is x units away from the origin is given the coordinate x . The number x can be positive or negative, the sign determining which side of the origin we are on. The origin has coordinate 0 and the unit point has coordinate 1. In this way, each point P of the line L is given a coordinate $x = x(P)$. We obtain an ordering of the points of the line as follows: given points A, B on the line, we have $A < B$ in this ordering if and only if $x(A) < x(B)$.

Such a numbered line is called a **ruler**. The numbering on the line L is determined by the function x which assigns to each point P of L its coordinate $x = x(P)$. For this reason, we will denote the ruler by x . If A, B are two points of L , we let

$$\overline{AB} = x(B) - x(A).$$

be the **oriented length** of the line segment AB . It is positive if and only if $A < B$ in the ordering of L specified by the ruler x . The length of AB with respect to the ruler x is

$$|AB| = |\overline{AB}| = |x(B) - x(A)|,$$

where the absolute value $|c|$ is defined to be c if $c \geq 0$ and $-c$ otherwise.

If we change the origin or unit point (or both), we obtain another ruler x' . The coordinates of a point P relative to these two rulers are related by the formula

$$x(P) = bx'(P) + a.$$

The new origin O' has coordinate $x'(O') = 0$ relative to second ruler; so $a = x(O')$, the coordinate of the O' with respect to the first ruler. The new unit point I' has coordinate $x(I') = a + b$ and $b = x(I') - x(O')$ is the **scale factor** between the rulers. The coordinate of P , relative to the second ruler, is therefore

$$x'(P) = b^{-1}(x(P) - a)$$

Since

$$x(B) - x(A) = b(x'(B) - x'(A)),$$

the two rulers x and x' give the same number for the length of AB iff the scale factor $b = \pm 1$. The two rulers define the same ordering or **orientation** if $b > 0$. If $b < 0$, they define opposite orderings or orientations. Thus a line has only two orderings or orientations.

Example. If C is the temperature in the Celsius scale and F the temperature in the Fahrenheit scale, we have

$$C = \frac{5}{9}(F - 32) = \frac{5}{9}F - 17\frac{7}{9}, \quad F = \frac{9}{5}C + 32.$$

Zero degrees Celsius is the same as 32 degrees Fahrenheit and zero degrees Fahrenheit is $-17\frac{7}{9}$ degrees Celsius. The scale factor from the Fahrenheit scale to Celsius scale is $9/5$ since a change in temperature of one degree Celsius is a change in temperature of $9/5$ degrees Fahrenheit.

1.2. Exercises. 1. If P is a point on a line having coordinate t with respect to a ruler x , find the coordinate of t with respect to the ruler x' having origin the point O' with x -coordinate 12 and unit point the point A' with x -coordinate 18. What are the x' -coordinates of the origin and unit point of x ?

2. Let A, B be distinct points on line L and let C be the midpoint of the line segment AB of L . If a, b are the coordinates of A, B for some ruler, show that the coordinate of C is $(a + b)/2$. Show also that C is the unique point with $\overline{AC}/\overline{CB} = 1$.

3. Let A, B be distinct points on a line L . Show that for each real number $c \neq -1$ there is a unique point P of L with $\overline{AP}/\overline{PB} = c$.

1.3. Plane coordinates. We now consider the case where we are in a plane Π . To introduce coordinates we need to choose two lines L_1, L_2 in Π meeting in a point O and two rulers x, y on L_1, L_2 respectively with the same origin O . One can think of the lines of Π parallel to L_1 as streets and the lines parallel to L_2 as avenues. To find out what avenue or street you are on, look at its intersection with L_1 or L_2 respectively; the line coordinate of the point of intersection gives the name. For example, the line L_1 would be 0-th street and L_2 would be 0-th avenue. If P is the point in the plane Π which is at the intersection of a -th avenue and b -th street, then the coordinates of P are

$$x(P) = a, \quad y(P) = b.$$

The pair $(x(P), y(P))$ is called the **coordinate vector** of P . The line L_1 is called the x -axis and the line L_2 is called the y -axis.

The x -axis is the set of points P with $y(P) = 0$; in other words, it has the equation $y = 0$. The y -axis has the equation $x = 0$. More generally, a line parallel to the x -axis has the equation $y = a$ while a line parallel to the y -axis has equation $x = a$. If we denote by $P(a, b)$ the point P with coordinate vector (a, b) , then $P(a, b)$ is the intersection of the lines $x = a, y = b$. If I, J are respectively the unit points for the rulers x, y , then the triple (O, I, J) completely determines L_1, L_2 and the rulers x, y ; it is called a **coordinate frame**. The pair (x, y) is called a **coordinate system** for the plane Π .

If we change the scale on L_i by a factor of b_i we obtain a new coordinate frame (O, I', J') with $x(I') = b_1, y(J') = b_2$. If (x', y') is the associated coordinate system, we have

$$x(P) = b_1 x'(P), \quad y(P) = b_2 y'(P).$$

The formula for a general change of coordinates will be derived later. If L_1 and L_2 are perpendicular and we use the same scale on both coordinate axes, the coordinate system is said to be **rectangular**. While rectangular coordinates are most often used, we shall see that non-rectangular coordinates can be very useful in solving problems. Such coordinates

are also called **affine coordinates**. When we use different scales for the coordinate axes, in order that a curve we are graphing fits on the page, we in fact are using affine coordinates. We shall see that having the freedom to choose our axes to be oblique (non-rectangular) will enormously simplify the graphing of plane curves such as conics.

Problem 1.1. Let x, y be a rectangular coordinate system. Sketch the curve whose equation in the coordinate system x', y' associated to the frame

$$(O'(-3, 4), I'(-2, 3), J'(-1, 5))$$

is $x' = y'^2$. This curve is called a parabola.

1.4. **Exercises.** 1. Draw two oblique intersecting lines on a sheet of paper and construct a ruler on each line, with center the point of intersection of the two lines, by choosing a unit point on each line. Plot the points $(10, 3)$, $(3, 10)$, $(-4, 0)$, $(-4, 2)$, $(-3, -5/2)$, $(3/2, -5)$, draw the lines $x = 4$, $y = -2$ and sketch the curve $y = x^2$. This curve is called a parabola.

2. In an affine coordinate system x, y in which the axes are perpendicular but the scale on the x -axis is twice that on the y -axis, sketch the curve $x^2 + y^2 = 4$. What is the equation of this curve if you use the frame $O(0, 0)$, $I(1, 0)$, $J(0, 1/2)$? This curve is called an ellipse.

3. Draw the coordinate axes for a rectangular coordinate system and plot the points $O'(2, 3)$, $I'(3, 1)$, $J'(3, 4)$. If x', y' is the coordinate system with the frame O', I', J' , sketch the curve whose equation is $x'y' = 1$. This curve is called a hyperbola.

1.5. **Space Coordinates.** Now suppose that we are not restricted to lie on a plane. To introduce a system of coordinates, we need to choose three non-coplanar lines L_1, L_2, L_3 which meet in a point O and rulers x, y, z on L_1, L_2, L_3 respectively, each with origin O . The plane containing the rulers x and y is called the xy -plane. Any point in this plane has well-defined x and y -coordinates.

Any point P lies on some plane which is parallel to the xy -plane. The level of this plane is defined to be the coordinate of the point of intersection of this plane with the line L_3 . If P is on level c we define the z -coordinate of P to be $z(P) = c$. The x and y -coordinates $x(P)$, $y(P)$ of P are obtained as follows: Let Q be the point of intersection of the line through P , parallel to L_3 , with the xy -plane. Then $x(P) = x(Q)$ and $y(P) = y(Q)$. The triple $(x(P), y(P), z(P))$ is called the coordinate vector of P .

The xy -plane has equation $z = 0$ and any plane parallel to it has equation $z = a$. Similarly, the planes parallel to the xz -plane have equation $y = a$ while the equations parallel to the yz -plane have equation $x = a$. The x -axis is the intersection of the xy and xz -planes and has equation $y = z = 0$. The y -axis is the intersection of the yz and xy -planes and has equation $x = z = 0$ while the z -axis is the intersection of the yz and xz -planes and has equation $x = y = 0$.

If $P(a, b, c)$ denotes the point P with coordinate vector (a, b, c) , then $P(a, b, c)$ is the intersection of the planes $x = a, y = b, z = c$. If I, J, K are respectively the unit points for the rulers x, y, z , the triple (O, I, J, K) completely determines the lines L_1, L_2, L_3 and the rulers x, y, z , and is called a **coordinate frame**. The triple (x, y, z) is called a coordinate system.

If we change the scale on L_i by a factor of b_i we obtain a new coordinate frame

$$(O, I', J', K')$$

with $x(I') = b_1$, $y(J') = b_2$ and $z(K') = b_3$. If (x', y', z') is the associated coordinate system, we have

$$x(P) = b_1 x'(P), \quad y(P) = b_2 y'(P), \quad z(P) = b_3 z'(P).$$

The formula for a general change of coordinates will be derived later. If the lines L_i are mutually perpendicular and we use the same scale on each of the coordinate axes, the coordinate system is said to be **rectangular**. Otherwise, it is said to be *affine*.

1.6. Plane Drawings. Plane drawings of objects in space are actually projections of the object. For example, this is the case when we are outside during the day and we see our shadow on the ground or wall which act as screens upon which an image is projected. The light rays from the sun travel on parallel lines and light which strikes an object is prevented from continuing, resulting in a shadow at the point of the screen where the ray would have struck. The higher the sun is in the sky the shorter the shadows for objects pointing straight up. Geometrically, the transformation (function) which sends a point to its shadow on the wall or ground is called **parallel projection** and we use this to draw planar pictures of objects in space. To do this, we imagine a screen behind the object in space and try to visualize its projection on the screen by choosing suitable points on the boundary of the object and plotting the projections on the screen. The projection of a point P parallel to a given line L is the intersection of the line through the point P and parallel to L with the plane of the screen. The basic thing to remember is that lines project to lines except for lines that point directly our direction of sight; they project to points. Also parallel lines which do not point directly at us project to parallel lines and line segments appear shorter the more they are aligned with our line of sight. Useful information about the shape of the surface can be obtained by drawing the projections of the intersections of the surface with planes parallel to the coordinate planes. We will discuss this in more detail in a later section.

The things we see with our eyes are also projections on our retinas. But, because the lens of the eye focuses light, the images we see are **central projections**. The farther an object is, the smaller the projection, and lines parallel to our direction of sight appear to converge in the distance. This is the type of projection used when one draws 'in perspective'. Geometrically, a central projection, with center A , onto a plane Π not containing A is the function which sends a point P which is not on the plane Π' through A parallel to Π into the point of intersection of Π with the line through A and P . The plane Π' is called the **vanishing plane**. For simplicity, we will use only parallel projections to make our plane drawings.

1.7. Exercises. 1. Draw the coordinate axes of a coordinate system in space as seen from the point $(10, 10, 10)$. Choosing unit points on each axis, plot the points $(0, 4, 7)$, $(4, 2, 0)$, $(1, 9, -2)$, $(-2, -4, 0)$.

2. Draw the box whose vertices have the coordinates $(2, 2, 0)$, $(2, 2, 3)$, $(2, 6, 3)$, $(2, 6, 0)$, $(5, 6, 0)$, $(5, 2, 0)$, $(5, 2, 3)$, $(5, 6, 3)$. Do this for both a rectangular and a non-rectangular coordinate system.

3. Sketch the surface $x^2 + y^2 = z^2$.

2. GEOMETRICAL VECTORS

A vector is usually described as something that has direction and magnitude and is represented geometrically by a **directed line segment**, namely, a line segment AB together with one of its two orderings or orientations. The length of the line segment and its ordering represent respectively the magnitude and direction of the vector. This directed line segment is completely determined by the pair (A, B) consisting respectively of its **initial** and **terminal** points. The direction of the line segment is indicated by an arrowhead at the terminal point and is denoted by \overrightarrow{AB} . A directed line segment is called a **bound vector** because the initial point is fixed. Such vectors are used to represent forces geometrically since the point of application of a force is fixed.

If A, B are points on a line L and x is a ruler with on L with $c = x(B) - x(A)$, the bound vector \overrightarrow{AB} is equivalent to the instruction “Starting at A , go c units along L ” since this completely determines B . The sign of c gives the direction to take along L . If we remove the first part of the instruction “Starting at A ”, then the instruction becomes “Go c units along L ”. This can be viewed as a vector in which the initial point is allowed to be any point of L ; this type of vector is called a **sliding vector**. If we allow the initial point to be any point in space and modify the instruction to read “Go c units parallel to L ”, we obtain what is called a **free vector**. Geometrically, it is what is called a **translation**. A translation is a function which sends points to points and has the following property: If P is sent to P' and Q to Q' by the translation, then PP' , QQ' are parallel and PQ , $P'Q'$ are parallel. Given points A, B , there is a unique translation which sends A to B .

Vectors are denoted by letters with arrows over them, for example, \vec{v} . The free vector which sends the point A to the point B is also denoted by \overrightarrow{AB} . This will not cause any confusion since, unless otherwise indicated, a geometrical vector will be taken to be free. Every bound vector \overrightarrow{AB} uniquely determines a free vector, namely, the translation which sends A to B . Two bound vectors \overrightarrow{AB} and $\overrightarrow{A'B'}$ determine the same free vector if and only if the translation which takes A to A' also takes B to B' . Two such bound vectors are called **equivalent**.

A free vector \vec{v} can be defined by choosing a ruler x on a line L and a number b representing how far we want to move in the direction parallel to L . Thus the free vector

$$\vec{v} = \text{“Go } b \text{ units in the } x\text{-direction”}$$

is defined as follows: If P is on L , then the point Q to which P is sent is the unique point Q with $x(Q) = x(P) + b$. If P is any point, we can, by means of parallel projection, transport the ruler x to a ruler x' on the line L' parallel to L and passing through P . Then the image of P is the unique point Q of L' with $x'(Q) = x'(P) + b$. For this reason, we let

$$P + \vec{v}$$

denote the point Q resulting from the translation of P by \vec{v} .

If \vec{v} is the geometrical vector which sends P to a point b units in a given direction and c is a number, we let $c\vec{v}$ be the geometrical vector which sends P to a point cb units in the given direction. We have

$$(c + d)\vec{v} = c\vec{v} + d\vec{v}.$$

If $c = 0$, then $P + c\vec{v} = P$ for any point P ; in other words

$$0 \cdot \vec{v} = \text{“Go } 0 \text{ units in the } x\text{-direction”}.$$

This geometrical vector is called the **zero vector** and is denoted by $\vec{0}$. A vector \vec{v} is said to be parallel to the non-zero vector \vec{u} if \vec{v} is proportional to \vec{u} , i.e., $\vec{v} = c\vec{u}$ for some

number c . If $\vec{u} = \overrightarrow{AB}$, then $c\overrightarrow{AB} = \overrightarrow{AC}$ for a unique point C with A, B, C collinear and $c\overrightarrow{AB} = \overrightarrow{AC}$. This can also be used as the definition of $c\overrightarrow{AB}$ when \overrightarrow{AB} is a bound geometrical vector.

If \vec{u} and \vec{v} are two geometrical vectors, then $\vec{u} + \vec{v}$ is the geometrical vector defined by

$$P + (\vec{u} + \vec{v}) = (P + \vec{u}) + \vec{v}.$$

This is well defined since one translation followed by another is again a translation; if $\vec{u} = \overrightarrow{AB}$ and $\vec{v} = \overrightarrow{BC}$, we have

$$\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}.$$

This is the **triangle law** for addition of geometrical vectors. For example, going 3 miles east and then 3 miles north gets you to the same place as going $3\sqrt{2}$ miles northeast. In addition, if $D = A + \overrightarrow{BC}$, we have $\overrightarrow{AD} = \overrightarrow{BC}$ and $\overrightarrow{CD} = \overrightarrow{AB}$. Thus

$$\overrightarrow{AB} + \overrightarrow{AD} = \overrightarrow{AD} + \overrightarrow{AB} = \overrightarrow{AC},$$

which is the **parallelogram law** for addition of geometrical vectors. This law can be used to define the sum of two bound vectors with the same initial point. It also shows that

$$\vec{u} + \vec{v} = \vec{v} + \vec{u},$$

which is the **commutative law for addition** of geometrical vectors. If $\vec{w} = \overrightarrow{CE}$, then

$$\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$$

since both sides are equal to \overrightarrow{AE} . This is the **associative law** for addition of geometrical vectors.

If we let $-\vec{v} = (-1)\vec{v}$, we have $\vec{u} + (-\vec{u}) = \vec{0}$. Note that $-\overrightarrow{AB} = \overrightarrow{BA}$ for free vectors. If we also define

$$\vec{v} - \vec{u} = \vec{v} + (-\vec{u}),$$

we have $(\vec{v} - \vec{u}) + \vec{u} = \vec{v}$. It follows that

$$\overrightarrow{AB} = \overrightarrow{AO} + \overrightarrow{OB} = \overrightarrow{OB} - \overrightarrow{OA}$$

for any point O . This only makes sense for free vectors.

Now suppose that we are given a coordinate system with origin O and unit points I, J, K . We let

$$[P] = (x(P), y(P), z(P))$$

and call it the **coordinate vector** of P ; this type of vector is called a **numerical vector** as opposed to a geometrical vector. In the next section we will give a more general definition of the word vector so as to include numerical vectors.

The geometrical vector \overrightarrow{OP} is called the **position vector** of P . Let $\vec{i} = \overrightarrow{OI}$, $\vec{j} = \overrightarrow{OJ}$, $\vec{k} = \overrightarrow{OK}$ be the position vectors of the unit points I, J, K . By the definition of the coordinates of P , we have

$$\overrightarrow{OP} = x\vec{i} + y\vec{j} + z\vec{k} \iff [P] = (x, y, z).$$

In particular, if \vec{v} is a geometrical vector and $O + \vec{v} = A(a, b, c)$, so that $\vec{v} = \overrightarrow{OA}$, we have

$$\vec{v} = a\vec{i} + b\vec{j} + c\vec{k}.$$

The numerical vector (a, b, c) is called the **coordinate vector** of \vec{v} and is denoted by $[\vec{v}]$. The numbers a, b, c are also called **coordinates** or **components** or **direction numbers** of \vec{v} . Notice that

$$[P] = [\vec{OP}]$$

so that a point and a geometrical vector can have the same coordinate vector. If we were to, as is too often the case, identify points and vectors with their coordinate vectors then there would be a great risk of confusing these two distinct geometric objects. Also points and geometrical vectors may have different coordinates in different coordinate systems and we may want to introduce a second coordinate system to simplify a problem.

Theorem 2.1. *If $\vec{v} = a\vec{i} + b\vec{j} + c\vec{k}$ then*

$$P(x, y, z) + \vec{v} = Q(x + a, y + b, z + c).$$

Proof: We have $Q = P + \vec{v}$ and $P = O + \vec{OP}$ so that

$$Q = (O + \vec{OP}) + \vec{v} = O + (\vec{OP} + \vec{v}).$$

Now

$$\begin{aligned} \vec{OP} + \vec{v} &= (x\vec{i} + y\vec{j} + z\vec{k}) + a\vec{i} + b\vec{j} + c\vec{k} \\ &= (x\vec{i} + a\vec{i}) + (y\vec{j} + b\vec{j}) + (z\vec{k} + c\vec{k}) \\ &= (x + a)\vec{i} + (y + b)\vec{j} + (z + c)\vec{k} \end{aligned}$$

so that $[Q] = (x + a, y + b, z + c)$.

Corollary 2.1. *If $A(a_1, b_1, c_1) + \vec{v} = B(a_2, b_2, c_2)$ then*

$$\vec{v} = \vec{AB} = (a_2 - a_1)\vec{i} + (b_2 - b_1)\vec{j} + (c_2 - c_1)\vec{k}.$$

For example, if $A(1, 2, 3) + \vec{v} = B(-1, 3, -1)$, we have

$$\vec{v} = \vec{AB} = 2\vec{i} + \vec{j} - 4\vec{k}$$

$$P(x, y, z) + \vec{v} = P(x + 2, y + 1, z - 4).$$

If we define the sum of the two numerical vectors (x, y, z) and (a, b, c) to be the numerical vector

$$(x, y, z) + (a, b, c) = (x + a, y + b, z + c),$$

we have

$$[P + \vec{v}] = [P] + [\vec{v}]$$

$$[\vec{u} + \vec{v}] = [\vec{u}] + [\vec{v}].$$

If we define the product of the number t and the numerical vector (a, b, c) to be the numerical vector

$$t(a, b, c) = (ta, tb, tc),$$

we have

$$[t\vec{v}] = t[\vec{v}]$$

since $[t\vec{v}] = (ta, tb, tc)$.

As an application, let us find a formula for the midpoint Q of the line segment joining the points $A(a_1, b_1, c_1)$ and $B(a_2, b_2, c_2)$. The coordinate vector of \vec{AB} is

$$(a_2 - a_1, b_2 - b_1, c_2 - c_1).$$

Then, since $Q = A + (1/2)\overrightarrow{AB}$, we get

$$\begin{aligned}[Q] &= [A] + (1/2)[\overrightarrow{AB}] \\ &= (a_1, b_1, c_1) + (1/2)(a_2 - a_1, b_2 - b_1, c_2 - c_1) \\ &= \left(\frac{a_1 + a_2}{2}, \frac{b_1 + b_2}{2}, \frac{c_1 + c_2}{2}\right).\end{aligned}$$

This also shows that $\overrightarrow{OQ} = \frac{1}{2}(\overrightarrow{OA} + \overrightarrow{OB})$. Similarly, one can get the coordinates of the point which divides a given line segment in any given ratio (see the exercises).

2.1. Exercises. 1. If $\overrightarrow{v} = \overrightarrow{AB}$, with $[A] = (2, 4, 6)$, $[B] = (8, 10, 12)$, and $[P] = (-2, 3, 5)$, find

$$[P + \overrightarrow{v}], [P + (1/2)\overrightarrow{v}], [P + (2/3)\overrightarrow{v}], [P + t\overrightarrow{v}].$$

Graph these points.

2. If $[\overrightarrow{u}] = (1, 1, 3)$, $[\overrightarrow{v}] = (3, 4, 2)$ compute

$$[2\overrightarrow{u} + 3\overrightarrow{v}], [(-1)\overrightarrow{u} + 2\overrightarrow{v}], [a\overrightarrow{u} + b\overrightarrow{v}].$$

3. If $A(1, 2, 4)$, $B(2, -3, 7)$, $C(6, -3, 5)$ are given points, find the coordinates of the mid-points D, E of AB and AC respectively. Using this, show that \overrightarrow{DE} is parallel to \overrightarrow{BC} . Show also that

$$D + (1/3)\overrightarrow{DC} = B + (2/3)\overrightarrow{BE}.$$

4. Let $A(x_1, y_1, z_1)$, $B(x_2, y_2, z_2)$ be distinct points on the line L passing through A and B . If P is any point on L with $\overrightarrow{AP} = r\overrightarrow{AB}$, show that

$$((1-r)x_1 + rx_2, (1-r)y_1 + ry_2, (1-r)z_1 + rz_2))$$

is the coordinate vector of P .

2.2. Planar Vectors. Let Π be a plane. A vector \overrightarrow{v} is said to be parallel to Π if $P + \overrightarrow{v}$ is a point of Π for every point P of Π ; in other words, if \overrightarrow{v} sends every point of Π to another point of Π . Then \overrightarrow{v} is parallel to Π iff it is parallel to every plane Π' which is parallel to Π . If $\overrightarrow{u}, \overrightarrow{v}$ are parallel to Π and a, b are scalars then $a\overrightarrow{u} + b\overrightarrow{v}$ is again parallel to Π .

If (O, I, J) is a frame for Π and $\vec{i} = \overrightarrow{OI}$, $\vec{j} = \overrightarrow{OJ}$, every vector \overrightarrow{v} can be uniquely written in the form

$$\overrightarrow{v} = a\vec{i} + b\vec{j}.$$

If P is any point of Π and $[P] = (x(P), y(P)) = (x, y)$ is the coordinate vector of P , we have

$$[P + \overrightarrow{v}] = (x + a, y + b).$$

If we define the coordinate vector of \overrightarrow{v} to be

$$[\overrightarrow{v}] = (a, b)$$

and define $(a_1, b_1) + (a_2, b_2) = (a_1 + b_1, a_2 + b_2)$, $c(a, b) = (ca, cb)$, we have

$$[\overrightarrow{u} + \overrightarrow{v}] = [\overrightarrow{u}] + [\overrightarrow{v}], [c\overrightarrow{v}] = c[\overrightarrow{v}].$$

2.3. **Exercises.** 1. If $\vec{v} = \overrightarrow{AB}$, with $[A] = (7, 4)$, $[B] = (-5, 10)$, and $[P] = (2, -3)$, find

$$[P + \vec{v}], [P - (1/2)\vec{v}], [P + (2/3)\vec{v}], [P + t\vec{v}].$$

Graph these points.

2. If $[\vec{u}] = (4, -3)$, $[\vec{v}] = (-5, 4)$ compute

$$[2\vec{u} + 3\vec{v}], [(-1)\vec{u} + 2\vec{v}], [a\vec{u} + b\vec{v}].$$

3. If $A(x_1, y_1)$, $B(x_2, y_2)$ are given points, find the coordinates of the vector \overrightarrow{AB} and the midpoint of the line segment AB . Prove your assertions.

4. If $A(6, 4)$, $B(-3, 7)$, $C(3, 5)$ are given points, find the coordinates of the mid-points D, E of AB and AC respectively. Using this, show that \overrightarrow{DE} is parallel to \overrightarrow{BC} . Show also that

$$D + (1/3)\overrightarrow{DC} = B + (2/3)\overrightarrow{BE}.$$

5. Let $A(x_1, y_1)$, $B(x_2, y_2)$ be distinct points on the line L passing through A and B . If P is any point on L with $\overrightarrow{AP} = r\overrightarrow{AB}$, show that

$$((1-r)x_1 + rx_2, (1-r)y_1 + ry_2)$$

is the coordinate vector of P .

2.4. **Vector Spaces.** Let V be the set of geometrical vectors. If \vec{u}, \vec{v} are in V we can add them to get another vector $\vec{u} + \vec{v}$ of V and, if c is a number (also called a scalar), we can multiply c and \vec{u} to get a vector $c\vec{u}$ of V . These operations obey the usual laws of arithmetic:

- (1) $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$;
- (2) There is a vector $\vec{0}$ in V such that $\vec{0} + \vec{u} = \vec{u} + \vec{0} = \vec{u}$ for any \vec{u} in V ;
- (3) For any \vec{u} in V , there is a vector \vec{v} with $\vec{u} + \vec{v} = \vec{v} + \vec{u} = \vec{0}$;
- (4) $\vec{u} + \vec{v} = \vec{v} + \vec{u}$;
- (5) If a, b are scalars, $a(b\vec{u}) = (ab)\vec{u}$;
- (6) $1 \cdot \vec{u} = \vec{u}$;
- (7) If a, b are scalars, then $(a+b)\vec{u} = a\vec{u} + b\vec{u}$ and $a(\vec{u} + \vec{v}) = a\vec{u} + a\vec{v}$.

We have proved all these properties except for the very last one. This one can be obtained by taking coordinate vectors, using the fact that $\vec{u} = \vec{v}$ iff $[\vec{u}] = [\vec{v}]$. If $[\vec{v}] = (x_1, y_1, z_1)$, $[\vec{v}] = (x_2, y_2, z_2)$, we have

$$\begin{aligned} [a(\vec{u} + \vec{v})] &= a[\vec{u} + \vec{v}] \\ &= a([\vec{u}] + [\vec{v}]) \\ &= a((x_1, y_1, z_1) + (x_2, y_2, z_2)) \\ &= a((x_1 + x_2, y_1 + y_2, z_1 + z_2)) \\ &= (a(x_1 + x_2), a(y_1 + y_2), a(z_1 + z_2)) \\ &= (ax_1 + ax_2, ay_1 + ay_2, az_1 + az_2) \\ &= (ax_1, ay_1, az_1) + (ax_2, ay_2, az_2) \\ &= a(x_1, y_1, z_1) + a(x_2, y_2, z_2) \\ &= [a\vec{u}] + [a\vec{v}] \\ &= [a\vec{u} + a\vec{v}] \end{aligned}$$

which shows that $a(\vec{u} + \vec{v}) = a\vec{u} + a\vec{v}$. Note that we also, in the process, obtain

$$a((x_1, y_1, z_1) + (x_2, y_2, z_2)) = a(x_1, y_1, z_1) + a(x_2, y_2, z_2).$$

More generally, any set V , having operations of addition and multiplication by scalars satisfying the above seven properties, is called a **vector space**; the elements of V are called vectors. According to this definition, a vector is simply an element of some vector space. The set of scalars, together with the usual laws of addition and multiplication, satisfy these properties. So they form vector space according to this definition.

An important example of a vector space is the set \mathbb{R}^n consisting of n -tuples of scalars (x_1, x_2, \dots, x_n) with addition and multiplication by scalars defined by

$$(x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n),$$

$$c(x_1, x_2, \dots, x_n) = (cx_1, cx_2, \dots, cx_n).$$

The proof of this is left as an exercise for the reader.

Other examples of vector spaces are the set of vectors parallel to a given plane or the set of vectors parallel to a given line. These two vector spaces are subspaces of the vector space of all geometric vectors. By definition, a **subspace** of a vector space V is a non-empty subset S of V such that, for every \vec{u}, \vec{v} in S and scalars a, b the vector $a\vec{u} + b\vec{v}$ is again in S .

The first property of vector spaces is called the **associative law for addition**, the fourth is called the **commutative law for addition**, the fifth is called the **associative law for scalar multiplication** and the last property is called the **distributive law**. The distinguished vector $\vec{0}$ whose existence is asserted in the second item above is unique. Indeed, if $\vec{0}'$ satisfies $\vec{0}' + \vec{u} = \vec{u} + \vec{0}'$ for any \vec{u} in V , we have $\vec{0} = \vec{0} + \vec{0}' = \vec{0}'$. This vector is called the **zero vector** of V . If $V = \mathbb{R}^n$, the zero vector is $(0, 0, \dots, 0)$. If V is the vector space of geometric vectors, the zero vector is the translation which moves nothing.

Given a vector \vec{u} there is a vector \vec{v} with

$$\vec{u} + \vec{v} = \vec{v} + \vec{u} = \vec{0}.$$

This vector is unique since

$$\vec{u} + \vec{v} = \vec{v}' + \vec{u} = \vec{0}$$

implies that

$$\vec{v} = \vec{v} + (\vec{u} + \vec{v}') = (\vec{v} + \vec{u}) + \vec{v}' = \vec{v}'.$$

This vector is called the **additive inverse** of \vec{u} and is denoted by $-\vec{u}$. Using this, we can show that the only vector \vec{v} satisfying $\vec{v} + \vec{v} = \vec{v}$ is the zero vector. Indeed, adding $-\vec{v}$ to both sides and using the associative law, we get $\vec{v} = \vec{0}$. If 0 is the zero scalar, we have

$$0 \cdot \vec{v} = (0 + 0) \vec{v} = 0 \cdot \vec{v} + 0 \cdot \vec{v},$$

and so $0 \cdot \vec{v} = \vec{0}$ for any vector \vec{v} . If $V = \mathbb{R}^n$, we have

$$-(x_1, x_2, \dots, x_n) = (-x_1, -x_2, \dots, -x_n);$$

if V is the vector space of geometrical vectors, and $\vec{v} = \overrightarrow{AB}$, then $-\vec{v} = \overrightarrow{BA}$. For any vector space we have

$$(-a)\vec{v} = -(a\vec{v})$$

since

$$a\vec{v} + (-a)\vec{v} = (a + (-a))\vec{v} = 0 \cdot \vec{v} = \vec{0}.$$

If we define $\vec{u} - \vec{v}$ to be

$$\vec{u} + (-\vec{v}) = \vec{u} + (-1)\vec{v},$$

then $\vec{u} - \vec{v}$ is the unique vector \vec{w} such that $\vec{w} + \vec{v} = \vec{u}$.

If $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n$ are vectors in a vector space V and a_1, a_2, \dots, a_n are scalars, then the vector

$$\vec{v} = a_1 \vec{u}_1 + a_2 \vec{u}_2 + \dots + a_n \vec{u}_n$$

is called a linear combination of the vectors $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n$. If

$$\vec{w} = b_1 \vec{u}_1 + b_2 \vec{u}_2 + \dots + b_n \vec{u}_n$$

is another linear combination of these vectors, then

$$c\vec{v} + d\vec{w} = (ca_1 + db_1)\vec{u}_1 + (ca_2 + db_2)\vec{u}_2 + \dots + (ca_n + db_n)\vec{u}_n.$$

and so $c\vec{v} + d\vec{w}$ is again a linear combination of the given vectors.

A sequence of vectors $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n$ is said to be a **basis** for the vector space V if every vector \vec{v} of V can be uniquely written in the form

$$\vec{v} = x_1 \vec{u}_1 + x_2 \vec{u}_2 + \dots + x_n \vec{u}_n.$$

The numerical vector $[\vec{v}] = (x_1, x_2, \dots, x_n)$ is called the **coordinate vector** of \vec{u} with respect to the given basis. We will show later that the number n is uniquely determined by V ; it is called the **dimension** of V . For example, the position vectors $\vec{i}, \vec{j}, \vec{k}$ of the unit points of a coordinate frame with respect to the origin of that frame are a basis for the vector space of geometrical vectors. This vector space is 3-dimensional. The vector space of geometrical vectors parallel to a given plane is 2-dimensional and the vector space of geometrical vectors parallel to a given line is one-dimensional. The vector space \mathbb{R}^n is n -dimensional with basis

$$(1, 0, \dots, 0), (0, 1, \dots, 0), (0, 0, 1, \dots, 0), \dots, (0, 0, \dots, 1).$$

This basis is called the **usual basis** of \mathbb{R}^n .

2.5. Exercises. 1. Show that \mathbb{R}^n with the operations of vector addition and multiplication by scalars defined above is a vector space.

2. Show that the subset of \mathbb{R}^3 consisting of those triples (x_1, x_2, x_3) with $x_1 + 2x_2 = x_3$ is a subspace of \mathbb{R}^3 .

3. Show that the subset of \mathbb{R}^2 consisting of those vectors x, y with $x + y = 1$ is not a subspace of \mathbb{R}^2 .

2.6. Using vectors to solve geometrical problems. Let's give some examples of the use of vectors to solve geometrical problems.

Problem 2.1. *Show that the diagonals of a parallelogram bisect each other.*

Solution. Let A, B, C, D be a parallelogram with AB parallel to CD and AC parallel to BD . Let $\vec{u} = \overrightarrow{AB}$, $\vec{v} = \overrightarrow{AC}$. Then $\overrightarrow{AD} = \vec{u} + \vec{v}$ by the parallelogram law for addition and $\overrightarrow{BC} = \vec{v} - \vec{u}$ since

$$B + (\vec{v} - \vec{u}) = (A + \vec{u}) + (\vec{v} - \vec{u}) = A + \vec{v} = C.$$

If P is the midpoint of AD and Q the midpoint of BC , we have

$$P = A + \frac{1}{2}(\vec{u} + \vec{v}), \quad Q = A + (\vec{u} + \frac{1}{2}(\vec{v} - \vec{u})).$$

But

$$\begin{aligned}\vec{u} + \frac{1}{2}(\vec{v} - \vec{u}) &= \vec{u} + \frac{1}{2}\vec{v} - \frac{1}{2}\vec{u} \\ &= \frac{1}{2}(\vec{u} + \vec{v}),\end{aligned}$$

which shows that $P = Q$.

Remark. The above point P is a **center of symmetry** of the parallelogram: Any line L through P (in the plane of the parallelogram) meets the sides of the parallelogram in two points R, S with $\overrightarrow{PR} = -\overrightarrow{PS}$. The easiest way to see this is to choose a coordinate system with center P and axes parallel to the sides of the parallelogram. The equations of the sides are therefore $x = \pm a$, $y = \pm b$. The point $R(x, y)$ lies on one of the sides iff either $|x| = |a|$, $|y| \leq |b|$ or $|y| = |b|$, $|x| \leq |a|$, in which case the point $T(-x, -y)$ also lies on one of the sides. Since P, S, T lie on a line we must have $S = T$ as L meets the sides of the parallelogram at precisely two points. But then $\overrightarrow{PS} = -\overrightarrow{PR}$, which is what we wanted to show.

Problem 2.2. Show that the medians of a triangle meet in a point.

Solution. Let ABC be a triangle. Its medians are the lines joining the vertices to the midpoints of the opposite sides. Let $\vec{u} = \overrightarrow{AB}$ and $\vec{v} = \overrightarrow{AC}$. If D is the midpoint of BC and P is any point on AD , we have

$$\begin{aligned}\overrightarrow{AP} &= t\overrightarrow{AD} = t(\overrightarrow{AB} + \overrightarrow{BD}) \\ &= t(\overrightarrow{AB} + \frac{1}{2}\overrightarrow{BC}) \\ &= t(\overrightarrow{AB} + \frac{1}{2}(\overrightarrow{BA} + \overrightarrow{AC})) \\ &= t(\overrightarrow{AB} + \frac{1}{2}(-\overrightarrow{AB} + \overrightarrow{AC})) \\ &= t(\overrightarrow{AB} - \frac{1}{2}\overrightarrow{AB} + \frac{1}{2}\overrightarrow{AC}) \\ &= t(\frac{1}{2}\overrightarrow{AB} + \frac{1}{2}\overrightarrow{AC}) \\ &= \frac{t}{2}\vec{u} + \frac{t}{2}\vec{v}.\end{aligned}$$

If E is the midpoint of AB and Q is any point on CE , we have

$$\begin{aligned}\overrightarrow{AQ} &= \overrightarrow{AC} + \overrightarrow{CQ} \\ &= \overrightarrow{AC} + s\overrightarrow{CE} \\ &= \overrightarrow{AC} + s(\overrightarrow{CA} + \overrightarrow{AE}) \\ &= \overrightarrow{AC} + s(-\overrightarrow{AC} + \frac{1}{2}\overrightarrow{AB}) \\ &= (1-s)\overrightarrow{AC} + \frac{s}{2}\overrightarrow{AB} \\ &= (1-s)\vec{u} + \frac{s}{2}\vec{v}.\end{aligned}$$

We have $P = Q$ iff

$$\frac{t}{2}\vec{u} + \frac{t}{2}\vec{v} = (1-s)\vec{u} + \frac{s}{2}\vec{v}$$

for suitable scalars s, t . Since \vec{u}, \vec{v} are the position vectors of the unit points of the frame (A, B, C) , we see that $P = Q$ iff $t/2 = 1 - s$ and $t/2 = s/2$. But these equations hold iff $s = t = 2/3$. Thus the medians AD, CE meet at a point which is two thirds of the way from the vertex to the midpoint of the opposite side. Since this result applies to any pair of medians, we see that the medians meet in a point which is two thirds of the way from the vertex to the midpoint of the opposite side. In particular, the medians trisect each other.

Remark. The point P of intersection of the medians is called the **centroid** of the triangle. Since $\vec{OB} = \vec{OA} + \vec{u}$ and $\vec{OC} = \vec{OA} + \vec{v}$, we have

$$\vec{OA} + \vec{OB} + \vec{OC} = 3\vec{OA} + \vec{u} + \vec{v}$$

for any point O and hence that

$$\vec{OP} = \frac{1}{3}(\vec{OA} + \vec{OB} + \vec{OC})$$

since $\vec{OA} + \frac{1}{3}(\vec{u} + \vec{v}) = \vec{OA} + \vec{AP} = \vec{OP}$.

2.7. Exercises. 1. If P is the center of the parallelogram $ABCD$ and O is any point, show that

$$\vec{OP} = \frac{1}{4}(\vec{OA} + \vec{OB} + \vec{OC} + \vec{OD}).$$

2. Show that $A(1, 3, 6), B(2, 5, 5), C(4, 2, 8), D(5, 4, 7)$ are the vertices of a parallelogram and find the coordinates of its center.

3. If $A(x_1, y_1, z_1), B(x_2, y_2, z_2), C(x_3, y_3, z_3)$ are not collinear, show that

$$D\left(\frac{x_1 + x_2 + x_3}{3}, \frac{y_1 + y_2 + y_3}{3}\right)$$

is the centroid of the triangle ABC . Use this to find the centroid of the triangle with vertices $A(1, -2, 4), B(-2, 3, -3), C(5, 2, 1)$.

4. What is the formula for the coordinates of the centroid of a triangle with respect to a planar frame? Using this formula, find the centroid of the triangle with vertices $A(3, 4), (6, -2), C(2, 5)$.

5. Show that the lines joining each vertex of a tetrahedron to the centroid of the opposite face meet in a point P . Find the coordinates of P in terms of the coordinates of the vertices of the tetrahedron. In what ratio do these lines divide each other?

3. EQUATIONS OF LINES AND PLANES

We first derive the equations of lines in space. Let L be a line and let A, B be distinct points on L . If $\vec{v} = \vec{AB}$, a point P lies on L iff

$$P = A + t\vec{v}$$

for some scalar t , in which case t is the coordinate of P with respect to the ruler on L with origin A and unit point B . This equation is called a **affine equation** for L . If O is a fixed point and P is any point, then P is on L iff

$$\vec{OP} = \vec{OA} + t\vec{AB}$$

for some scalar t . This equation is called a **vector equation** for L . Both affine and vector equations for are independent of any coordinate system. The vector \vec{v} can be taken to be any non-zero geometric vector parallel to L .

Now suppose that we have a coordinate system with origin O . If A and B have coordinate vectors (a_1, b_1, c_1) and (a_2, b_2, c_2) then, taking coordinate vectors, the vector and affine equations for L each transform into the following equation of numerical vectors

$$(x, y, z) = (a_1, b_1, c_1) + t(\alpha, \beta, \gamma),$$

where $\alpha = a_2 - a_1$, $\beta = b_2 - b_1$, $\gamma = c_2 - c_1$, from which we get the following set of equations:

$$\begin{aligned} x &= a_1 + t\alpha \\ y &= b_1 + t\beta \\ z &= c_1 + t\gamma. \end{aligned}$$

These equations are called **parametric equations** for the line L . If we change the points A, B on L , these equations will change and so a line may have many different parametric equations for it. As we saw above, the parameter t is the coordinate of the point $P(x, y, z)$ on L with respect to the ruler with origin $A(a_1, b_1, c_1)$ and unit point $B(a_1 + \alpha, a_2 + \beta, a_3 + \gamma)$. The numerical vector (α, β, γ) is called a **direction vector** for the line L ; it is determined by L up to multiplication by a non-zero constant. A direction vector for a line can be found by taking the coordinate vector of any non-zero vector parallel to L . Two lines have proportional direction vectors iff they are parallel. Therefore, if two lines don't have proportional direction vectors, they either meet or are **skew lines** (non-coplanar lines).

If α, β, γ are non-zero, the above equations can be written in the more **symmetric form**

$$\frac{x - a_1}{\alpha} = \frac{y - b_1}{\beta} = \frac{z - c_1}{\gamma}.$$

The original equations can be recovered by setting the common value equal to t and solving for x, y, z . If $\alpha = 0$ while β, γ are non-zero, the equations become

$$x = a_1, \quad \frac{y - b_1}{\beta} = \frac{z - c_1}{\gamma},$$

and so is the equation of a line in the plane $x = a_1$. If $\alpha = \beta = 0$, the equation of the line becomes $x = a_1, y = b_1$ (z arbitrary), which is the equation of the line through $(a_1, b_1, 0)$ which is parallel to the z -axis.

Problem 3.1. Find parametric and symmetric equations for the line through the points $A(-1, 2, 3)$, $B(1, 3, 2)$. Determine whether or not $C(3, 4, 1)$ is on the line.

If L is the line through the points $A(-1, 2, 3)$, $B(1, 3, 2)$ then L has direction vector $\overrightarrow{AB} = (2, 1, -1)$ and so has parametric equations

$$\begin{aligned} x &= -1 + 2t \\ y &= 2 + t \\ z &= 3 - t. \end{aligned}$$

The symmetric form of these equations are

$$\frac{x + 1}{2} = \frac{y - 2}{1} = \frac{z - 3}{-1}.$$

The point $C(3, 4, 1)$ is not on the line since $\overrightarrow{AC} = (4, 2, 1)$ which is not a multiple of $(2, 1, -1)$.

Remark. If we choose $t = 4$ we get a point $C(7, 6, -1)$ on L . If we use B, C to get equations for L we get

$$\begin{aligned}x &= 1 + 6t \\y &= 3 + 3t \\z &= 2 - 3t\end{aligned}$$

which, in symmetric form, are

$$\frac{x-1}{6} = \frac{y-3}{3} = \frac{z-2}{-3}.$$

Notice that the direction numbers found here are proportional to the direction numbers found above, i.e., $(6, 3, -3) = 3(2, 1, -1)$.

Problem 3.2. Find (if any) the points of intersection of two lines

$$\begin{aligned}x &= -1 + 2t & x &= 2 + t \\y &= 2 + t & \text{and } y &= 1 - 2t \\z &= 3 - t, & z &= 3 + t.\end{aligned}$$

Solution. First note that, since the direction vectors $(2, 1, -1)$ and $(1, -2, 1)$ are not proportional, these two lines are not parallel; so they either don't meet or they meet in a single point. If $P(x, y, z)$ were a point of intersection of these two lines, we must have $(x, y, z) = (-1 + t, 2 + t, 3 - t)$ for some number t since P lies on the first line. But, since P also lies on the second line, we must have $(x, y, z) = (2 + s, 1 - 2s, 3 + s)$ for some number s (possibly different from the number t found above because s, t are the coordinates of P with respect to two different rulers). We therefore have $-1 + 2t = 2 + s$, $2 + t = 1 - 2s$, $3 - t = 3 + s$ which can be written

$$\begin{aligned}-s + 2t &= 3 \\2s + t &= -1 \\s + t &= 0.\end{aligned}$$

If we add the first equation to the third and add 2 times the first equation to the second, we get

$$\begin{aligned}-s + 2t &= 3 \\5t &= 5 \\3t &= 3.\end{aligned}$$

This system has the same solutions as the original system since the above process can be reversed to give the original system. Indeed, subtracting the first equation from the third and adding -2 times the first to the second gives back the original system. If we now multiply the second equation by $1/5$ and the third equation by $1/3$, we get

$$\begin{aligned}-s + 2t &= 3 \\t &= 1 \\t &= 1.\end{aligned}$$

Subtracting the second equation from the third, we get

$$\begin{aligned}-s + 2t &= 3 \\t &= 1 \\0 &= 0.\end{aligned}$$

Since $t = 1$, $s = 1$ are the only values of s, t which satisfy all three of these equations, the two lines meet in the unique point $P(1, 3, 2)$. The equation $0 = 0$ is called the **zero equation**. Since it imposes no condition on the variables, it can safely be deleted from any system.

Problem 3.3. Repeat the previous problem with the second line replaced by the line

$$\begin{aligned}x &= 2 + t \\y &= 1 - 2t \\z &= 2 + t.\end{aligned}$$

Solution. Again, the lines are not parallel since the direction vectors $(2, 1, -1)$, $(1, -2, 1)$ of the lines are not proportional. A point $P(x, y, z)$ lies on both lines iff there are scalars s, t with $-1 + 2t = 2 + s$, $2 + t = 1 - 2s$, $3 - t = 2 + s$ and we get the system

$$\begin{aligned}-s + 2t &= 3 \\2s + t &= -1 \\s + t &= 1.\end{aligned}$$

If we add the first equation to the third and add twice the first to the second, we get

$$\begin{aligned}-s + 2t &= 3 \\5t &= 5 \\3t &= 4.\end{aligned}$$

Subtracting $3/5$ times the second equation from the third, we get

$$\begin{aligned}-s + 2t &= 3 \\5t &= 5 \\0 &= 1.\end{aligned}$$

Since the equation $0 = 1$ has no solutions, the system itself has no solutions and hence the two lines do not meet. The two lines are skew since they are not parallel.

The procedure used in the last two problems for solving linear systems is called **Gauss-Jordan elimination**. The main step in this procedure is to pick an equation, called a **pivot equation**, and select a variable that appears in this equation (has a non-zero coefficient); this variable is called a **pivot variable** and we make the term containing this variable the first term in the equation. After possibly interchanging two equations, which does not change the solution set, we can assume that the pivot equation is the first equation. We then eliminate this pivot variable from all the other equations by adding a multiple of the working equation to each of the other equations. (Note that subtracting c times an equation to another is the same as adding $-c$ times that equation.) Since the original system can be recovered by reversing this procedure, we obtain an equivalent system, one with the same solution set as the original system. Multiplying an equation by a non-zero constant also yields an equivalent system. We then repeat the above procedure on the system formed by the equations other than the first. When this process stops we will have a number of pivot equations with the pivot variable appearing in no succeeding equation and we will possibly have a number of equations of the form $0 = 0$ which can be deleted or we will have an equation $0 = a$ with $a \neq 0$, in which case the system is inconsistent, i.e., has no solutions. Such a system is said to be in **echelon form**. If we now eliminate the pivot variables from the other equations, starting with the last pivot equation to minimize the calculation, and then make the coefficient of the pivots 1 by multiplying each pivot equation by a suitable

constant, we obtain a system which is in **reduced echelon form**. In this case we will have solved for the pivot variables in terms of the non-pivot variables. If the system is consistent, we can get all solutions by arbitrarily assigning values to the non-pivot variables.

Let us now find equations for planes. Let Π be a plane and let A, B, C be non-collinear points on Π . Let $\vec{u} = \overrightarrow{AB}$, $\vec{v} = \overrightarrow{AC}$. Then a point P is on Π iff

$$P = A + s\vec{u} + t\vec{v}$$

with s, t scalars, in which case, (s, t) is the coordinate vector of P with respect to the frame (A, B, C) . This equation is called an **affine equation** for Π . If O is a fixed point then P is a point of Π iff

$$\overrightarrow{OP} = \overrightarrow{OA} + s\vec{u} + t\vec{v}$$

with s, t scalars. This is a **vector equation** for Π . Both vector and affine equations for Π are independent of any coordinate system. The vectors \vec{u} , \vec{v} could be chosen to be any two non-zero vectors parallel to Π which are not scalar multiples of each other since $A, B = A + \vec{u}, C = A + \vec{v}$ would then be three non-collinear points of Π .

Now suppose we are given a coordinate system with origin O and suppose that the coordinate vectors of A, B, C are respectively

$$(a_1, b_1, c_1), (a_2, b_2, c_2), (a_3, b_3, c_3)$$

It is easy to test for the non-collinearity of A, B, C ; in fact, they are non-collinear iff $[\overrightarrow{AB}]$ is not a scalar multiple of $[\overrightarrow{AC}]$. We have

$$[\overrightarrow{AB}] = (\alpha_1, \beta_1, \gamma_1) = (a_2 - a_1, b_2 - b_1, c_2 - c_1),$$

$$[\overrightarrow{AC}] = (\alpha_2, \beta_2, \gamma_2) = (a_3 - a_1, b_3 - b_1, c_3 - c_1).$$

Taking coordinate vectors the affine and vector equations for Π each transform to

$$(x, y, z) = (a_1, b_1, c_1) + s(\alpha_1, \beta_1, \gamma_1) + t(\alpha_2, \beta_2, \gamma_2),$$

from which we obtain the equations

$$x = a_1 + s\alpha_1 + t\alpha_2$$

$$y = b_1 + s\beta_1 + t\beta_2$$

$$z = c_1 + s\gamma_1 + t\gamma_2.$$

These are **parametric equations** for Π with parameters s, t which are the coordinates of a point P of Π with respect to the plane coordinate system with origin A and unit points B, C . The numerical vectors $(\alpha_1, \beta_1, \gamma_1), (\alpha_2, \beta_2, \gamma_2)$ are called **direction vectors** for Π . A direction vector for Π can be found by taking the coordinate vector of the geometric vector determined by any pair of distinct points on a line parallel to Π .

Problem 3.4. Show that the three points

$$A(1, 0, -2), B(3, -1, 1), C(4, 1, -4).$$

are non-collinear and find parametric equations for the plane passing through them.

Solution. Since the numerical vectors $[\overrightarrow{AB}] = (2, -1, 3)$, $[\overrightarrow{AC}] = (3, 1, -2)$ are non-proportional, the points A, B, C do not lie on a line. Also, since these vectors are direction vectors for the plane passing through A, B, C , we get

$$x = 1 + 2s + 3t$$

$$y = -s + t$$

$$z = -2 + 3s - 2t,$$

as parametric equations for this plane.

If, in the above problem, we solve for s, t using the first two equations we get

$$s = (x - 3y - 1)/5, \quad t = (x + 2y - 1)/5.$$

Substituting this in the third equation, we get the equation

$$x - 13y - 5z = 11$$

whose solutions are precisely the coordinates of the points P which lie on Π . Indeed, if (x, y, z) is a solution of this equation and we set $s = (x - 3y - 1)/5$, $t = (x + 2y - 1)/5$ we get $z = -2 + 3s - 2t$. The equation $x - 13y - 5z = 0$ is called a **normal equation** for Π . We will show later that every plane has a normal equation. For now, we content ourselves to prove the following result:

Theorem 3.1. *If a, b, c, d are scalars with a, b, c not all zero, the points $P(x, y, z)$ whose coordinates satisfy the equation $ax + by + cz = d$ form a plane.*

Proof. We first suppose that $a \neq 0$. Then

$$x = \frac{d}{a} + \frac{-b}{a}y + \frac{-c}{a}z.$$

We get all solutions of $ax + by + cz = d$ by setting $y = s$, $z = t$ arbitrarily and solving for x . We thus get parametric equations

$$\begin{aligned} x &= d/a - (b/a)s - (c/a)t \\ y &= s \\ z &= t, \end{aligned}$$

which are parametric equations for the plane Π passing through the points

$$A(d/a, 0, 0), \quad B((d - b)/a, 1, 0), \quad C((d - c)/a, 0, 1).$$

If $a = 0$ and $b \neq 0$, we get the solutions of $y = d/b - (c/b)z$ by setting $x = s$, $z = t$ arbitrarily. We thus get the parametric equations

$$\begin{aligned} x &= s \\ y &= d/b - (c/b)t \\ z &= t, \end{aligned}$$

which are parametric equations for the plane Π passing through the points

$$A(0, d/b, 0), \quad B(1, d/b, 0), \quad C(0, (d - c)/b, 1).$$

If $a = b = 0$, then the solutions of $z = -(d/c)$ can be described in parametric form by

$$\begin{aligned} x &= s \\ y &= t \\ z &= d/c, \end{aligned}$$

which are parametric equations for the plane Π passing through the points $A(0, 0, d/c)$, $B(1, 0, d/c)$, $C(0, 1, d/c)$. **Q.E.D.**

Remark. We will show later that the planes $ax + by + cz = d$ and $a'x + b'y + c'z = d'$ are equal if and only if there is a constant t such that

$$(a', b', c', d') = t(a, b, c, d).$$

Problem 3.5. Find the points of intersection of the line

$$\begin{aligned}x &= 1 + 6t \\y &= 3 + 3t \\z &= 2 - 3t\end{aligned}$$

and the plane

$$\begin{aligned}x &= 1 + 2s + 3t \\y &= -s + t \\z &= -2 + 3s - 2t.\end{aligned}$$

Solution 1. A point $P(x, y, z)$ will lie on both planes if and only if we can find scalars r, s, t so that

$$\begin{aligned}x &= 1 + 6r = 1 + 2s + 3t \\y &= 3 + 3r = -s + t \\z &= 2 - 3r = -2 + 3s - 2t.\end{aligned}$$

This will happen precisely when r, s, t satisfy the following system of linear equations:

$$\begin{aligned}6r - 2s - 3t &= 0 \\3r + s - t &= -3 \\-3r - 3s + 2t &= -4.\end{aligned}$$

Since a line can either (i) meet a plane in a single point, (ii) lie in the plane or (iii) not meet the plane, this system of equations could have (i) one solution, (ii) an infinity of solutions or (iii) no solution.

To see which it is we use Gauss-Jordan elimination on this system. Choose the second equation as our pivot equation and r as the pivot variable. Interchanging the first two equations we get the system

$$\begin{aligned}3r + s - t &= -3 \\6r - 2s - 3t &= 0 \\-3r - 3s + 2t &= -4.\end{aligned}$$

Then add -2 times the first equation to the second equation and add the first equation to the third equation to get

$$\begin{aligned}3r + s - t &= -3 \\-4s - t &= 6 \\-2s + t &= -7.\end{aligned}$$

We now repeat the procedure by choosing a new pivot equation, the third, and a new pivot variable, s and interchange the second and third equations to get

$$\begin{aligned}3r + s - t &= -3 \\-2s + t &= -7 \\-4s - t &= 6.\end{aligned}$$

We then subtract -2 times the second equation from the third to eliminate s . The resulting system is

$$\begin{aligned} 3r + s - t &= -3 \\ -2s + t &= -7 \\ -3t &= 20. \end{aligned}$$

The last equation gives $t = -20/3$ and substituting this in the second equation gives $s = 1/6$. Finally, substituting $s = 1/6$, $t = -20/3$ in the first equation, we get

$$s = 1/6, r = -59/18, t = -20/3.$$

Hence the given line and plane meet in the point

$$P(1 + 6r, 3 + 3r, 2 - 3r) = (-56/3, -41/6, 71/6).$$

As a check, note that this point is the same as

$$P(1 + 2s + 3t, -s + t, -2 + 3s - 2t) = (-56/3, -41/6, 71/6).$$

Solution 2. We could also have solved this problem by using a normal equation of the plane, which in this case is $x - 13y - 5z = 11$, and substituting $x = 1 + 6r$, $y = 3 + 3r$, $z = 2 - 3r$ in this equation to get $1 + 6r - 39 - 39r - 10 + 15r = 11$ from which $r = -59/18$.

Solution 3. There is a third way of solving this problem by using the symmetric form of the given line:

$$\frac{x-1}{6} = \frac{y-3}{3} = \frac{z-2}{-3}.$$

This equation can be written as two equations $(x-1)/6 = (y-3)/3$, $(y-3)/3 = -(z-2)/3$ which, on simplifying, yield the two equations

$$\begin{aligned} x - 2y &= -5 \\ y + z &= 3. \end{aligned}$$

Since each of these equations are equations of a plane, this yields two planes which intersect in the given line. The intersection of the given line and plane is therefore the intersection of the given plane with the above two planes. In other words, the points $P(x, y, z)$ in the intersection are given by the solutions of the system

$$\begin{aligned} x - 2y &= -5, \\ y + z &= 3, \\ x - 13y - 5z &= 11. \end{aligned}$$

We leave it as an exercise for the reader to show, using Gauss-Jordan elimination, that this system has the unique solution $x = -56/3$, $y = -41/6$, $z = 71/6$.

Problem 3.6. Find the intersection of the two planes

$$\begin{aligned} x &= 1 + 2s + 3t & x &= s + t \\ y &= -s + t & y &= s - t \\ z &= -2 + 3s - 2t & z &= 1 + s. \end{aligned}$$

Solution 1. A point $P(x, y, z)$ lies on the first plane if and only if there are numbers s_1, t_1 such that $x = 1 + 2s_1 + 3t_1$, $y = -s_1 + t_1$, $z = -2 + 3s_1 - 2t_1$. This point lies on the second plane if and only if there are numbers s_2, t_2 such that $x = s_2 + t_2$, $y = s_2 - t_2$,

$z = 1 + s_2$. To find the points intersection, we have to find the solutions of the system of equations

$$\begin{aligned} 1 + 2s_1 + 3t_1 &= s_2 + t_2 \\ -s_1 + t_1 &= s_2 - t_2 \\ -2 + 3s_1 - 2t_1 &= 1 + s_2. \end{aligned}$$

This system simplifies to

$$\begin{aligned} 2s_1 + 3t_1 - s_2 - t_2 &= -1 \\ -s_1 + t_1 - s_2 + t_2 &= 0 \\ 3s_1 - 2t_1 - s_2 &= 3. \end{aligned}$$

We choose the second equation as our first pivot equation and s_1 as the pivot variable in this equation. Interchanging the first two equations, we get

$$\begin{aligned} -s_1 + t_1 - s_2 + t_2 &= 0 \\ 2s_1 + 3t_1 - s_2 - t_2 &= -1 \\ 3s_1 - 2t_1 - s_2 &= 3. \end{aligned}$$

To eliminate s_1 from the second and third equations, add twice the first equation to the second and three times the first equation to the third to get

$$\begin{aligned} -s_1 + t_1 - s_2 + t_2 &= 0 \\ 5t_1 - 3s_2 + t_2 &= -1 \\ t_1 - 4s_2 + 3t_2 &= 3. \end{aligned}$$

We now choose the third equation as our next pivot equation and the variable t_1 as the pivot variable in this equation. Interchanging the second and third equations, we get

$$\begin{aligned} -s_1 + t_1 - s_2 + t_2 &= 0 \\ t_1 - 4s_2 + 3t_2 &= 3 \\ 5t_1 - 3s_2 + t_2 &= -1. \end{aligned}$$

To eliminate t_1 from the third equations, we add -5 times the second equation to the third equation to get

$$\begin{aligned} -s_1 + 3s_2 - 2t_2 &= -3 \\ t_1 - 4s_2 + 3t_2 &= 3 \\ 17s_2 - 14t_2 &= -16. \end{aligned}$$

We now choose the third equation as our last pivot equation and s_2 as the pivot variable. If we add $-3/17$ times the third equation to the first and $4/17$ times the third equation to the first we get the system

$$\begin{aligned} -s_1 + (8/17)t_2 &= -54/17 \\ t_1 - (5/17)t_2 &= 55/17 \\ 17s_2 - 14t_2 &= -16. \end{aligned}$$

Solving for the pivot variables, we get

$$\begin{aligned} s_2 &= (14/17)t_2 - 16/17 \\ t_1 &= (5/17)t_2 + 55/17 \\ s_1 &= (8/17)t_2 + 54/17. \end{aligned}$$

Since the solutions of this system are obtained by giving t_2 any value and solving for s_1, s_2, t_1 , we see that the points of intersection of the two planes are

$$P(s_2 + t_2, s_2 - t_2, 1 + s_2) = P((31/17)t_2 - 16/17, -(3/17)t_2 - 16/17, (14/17)t_2 + 1/17),$$

with t_2 taking arbitrary values. This shows that the two planes intersect in the line

$$\begin{aligned} x &= -16/17 + (31/17)t \\ y &= -16/17 - (3/17)t \\ z &= 1/17 + (14/17)t. \end{aligned}$$

Solution 2. Another way to solve this problem is to use normal equations for these planes.

The first plane has the normal equation $x - 13y - 5z = 11$ and the second has normal equation $x + y - 2z = -2$. The intersection of these two planes is given by the solution set of the following system:

$$\begin{aligned} x - 13y - 5z &= 11 \\ x + y - 2z &= -2. \end{aligned}$$

Using Gauss-Jordan elimination, one finds that the solutions of this system are

$$x = (31/14)z - 1/14, \quad y = -(3/14)z - 13/14$$

with z arbitrary and hence that the line of intersection is

$$\begin{aligned} x &= -15/14 + (31/14)t \\ y &= -13/14 - (3/14)t \\ z &= t. \end{aligned}$$

This parametric representation is different from the one obtained above but they both describe the same line. As a check, note that the direction numbers are proportional (so the lines they describe are parallel) and that, setting $t = 1/17$ in the second parametrization, we get $x = y = -16/17, z = 1/17$ which is the point corresponding corresponding to $t = 0$ in the first parametrization; so the parallel lines described by the two parametrizations have a point in common and hence are equal. There is a third way to solve this problem but we leave this to the reader.

3.1. Exercises. 1. Determine how the following lines intersect each other. In cases they don't meet, determine if the lines are skew or parallel.

$$\begin{array}{lll} x = 1 + t & x = -6t & x = 3t \\ y = 2 - t & y = 6t & y = 3 + t \\ z = 3 + 2t, & z = -12t, & z = 1 + 3t. \end{array}$$

2. Find parametric equations for the line of intersection of the two planes

$$\begin{array}{ll} x = 1 + s + t & x = 2 + 3s - t \\ y = 2 - s - 2t & y = 1 + s \\ z = s - t, & z = 2 + s + t. \end{array}$$

3. Find the equation of the plane passing through $(1, 2, 3)$ containing the line of intersection of the two planes $x + y + z = 1, x - 2y - z = 2$.

4. Show that any plane containing the line of intersection of the two planes

$$10x - 21y + 30z = 12, \quad 42x - 121y + 30z = 40$$

is of the form $a(10x - 21y + 30z - 12) + b(42x - 121y + 30z - 40) = 0$ for suitable scalars a, b . Use this to find the equation of the plane containing the line of intersection of the given two planes and passing through the point $(1, 2, 1)$.

3.2. Equations of Lines in a Plane. Let (O, I, J) be a frame for a plane Π and let

$$\vec{i} = \overrightarrow{OI}, \quad \vec{j} = \overrightarrow{OJ}.$$

If \vec{v} is a geometric vector parallel to Π , we have

$$P(x, y) + \vec{v} = (x + a, y + b)$$

for a unique pair of scalars (a, b) . This pair is the coordinate vector of \vec{v} in the given frame and is denoted by $[\vec{v}]$. The point P has coordinate vector $[P] = (x, y)$ with respect to this frame iff

$$\overrightarrow{OP} = x\vec{i} + y\vec{j}.$$

If L is a line in Π and $A(a_1, b_1), B(a_2, b_2)$ are distinct points on L , a vector equation for L is

$$\overrightarrow{OP} = \overrightarrow{OA} + t\overrightarrow{AB}.$$

Since $[\overrightarrow{AB}] = (a_2 - a_1, b_2 - b_1)$, the equation of L in terms of coordinate vectors is

$$(x, y) = (a_1, b_1) + t(\alpha, \beta),$$

where $(\alpha, \beta) = (a_2 - a_1, b_2 - b_1)$. We thus obtain

$$\begin{aligned} x &= a_1 + t\alpha \\ y &= b_1 + t\beta, \end{aligned}$$

which are the parametric equations of the line in terms of plane coordinates. The parameter t is the coordinate of a point $P(x, y)$ on the line with respect to the ruler with origin A and unit point B . The numerical vector (α, β) is a direction vector of the line. It is determined by the line up to multiplication by a non-zero scalar. If $\alpha = 0$, this is the line $x = a_1$; if $\beta = 0$, this is the line $y = b_1$. If b_1 and b_2 are both non-zero, the equations can be written in symmetric form

$$\frac{x_1 - a_1}{\alpha} = \frac{x_2 - a_2}{\beta},$$

which can be simplified to

$$\beta x - \alpha y = \beta a_1 - \alpha a_2.$$

The general equation of a line in a plane with plane coordinates x, y is therefore

$$ax + by = c$$

with a, b not both zero. A direction vector for this line is $(-b, a)$. If $b \neq 0$, the line has the equation

$$y = mx + b$$

with $m = -a/b$ (the **slope** of the line) and b the y -coordinate of the point of intersection of this line with the y -axis (the y -intercept of the line). If $(x_1, y_1), (x_2, y_2)$ are two points on this line, we have

$$y_1 = mx_1 + b, \quad y_2 = mx_2 + b$$

and so $y_1 - y_2 = m(x_1 - x_2)$ yielding

$$m = \frac{y_1 - y_2}{x_1 - x_2}.$$

If a is the x -intercept, then $0 = ma + b$ and $b = -ma$ and the equation of the line is then $y = m(x - a)$.

Problem 3.7. Find the equations of the sides of the triangle ABC in the coordinate system having the frame (O, I, J) with $O = A, I = B, J = C$. Also, find the equations of the medians and verify that they meet in a point that trisects each median.

Solution. The line through A, B is the x -axis and has equation $y = 0$; the side through A, C is the y -axis and has equation $x = 0$. The line through $B(1, 0)$ and $C(0, 1)$ has slope -1 and has y -intercept 1 . Its equation is therefore $y = -x + 1$ or $x + y = 1$. The midpoint of AB is the point $D(1/2, 0)$ and so the slope of the median CD is -2 . The equation of the line through C, D is therefore $y = -2(x - 1)$ or $2x + y = 1$ since the x -intercept is 1 . By symmetry, the median BE , with $E(0, 1/2)$ the mid-point of AC , has equation $x + 2y = 1$. The line joining A and the midpoint $F(1/2, 1/2)$ of BC has equation $y = x$. The lines $y = x$ and $2x + y + 1 = 0$ meet in the point $P(1/3, 1/3)$ and the line $x + 3y = 1$ passes through this point. Since $\overrightarrow{AP} = \frac{2}{3}\overrightarrow{AF}$, the point P divides the line segment AF in the ratio $2 : 1$ and, since any one of the vertices could have been chosen as origin, the same result holds for the other two medians.

We now consider the problem of finding the point of intersection of two non-parallel lines $ax + by = c$, $dx + ey = f$. That the lines are not parallel is equivalent to the statement that $(a, b) \neq k(d, e)$ for any k . If $P(x, y)$ is the point of intersection of these two lines then

$$\begin{aligned} ax + by &= c \\ dx + ey &= f. \end{aligned}$$

If we multiply the first equation by e , the second by $-b$ and add the resulting two equations, we get

$$(ae - bd)x = ce - bf.$$

Similarly, adding $-d$ times the first equation to a times the second, we get

$$(ae - bd)y = af - cd.$$

If $ae - bd \neq 0$ we would then have

$$x = \frac{ce - bf}{ae - bd}, \quad y = \frac{af - cd}{ae - bd}.$$

But this is the case as the following result shows.

Theorem 3.2. If d, e are not both zero, then $(a, b) = k(d, e)$ for some k if and only if $ae - bd = 0$.

Proof. If $(a, b) = k(d, e)$ then $a = kd, b = ke$ and so $ae - be = kde - ked = 0$. Conversely, suppose that $ae - bd = 0$. If $d \neq 0$, we then have $b = e(a/d)$ so that $(a, b) = k(d, e)$ with $k = a/d$. If $e \neq 0$, we have $(a, b) = k(d, e)$ with $k = b/e$. **Q.E.D**

The above formula for the point of intersection of two non-parallel lines $ax + by = c$, $dx + ey = f$ expresses the coordinates of the point of intersection these lines as ratios of similar looking expressions. To bring out the connection, we write down the coefficients of x, y in the equations of the two lines as a vertical list containing two rows of numbers, the first row being $a \ b$ (the coefficients of x, y in the first equation) and the second being $d \ e$ (the coefficients of x, y in the second equation). We get the following array

$$A = \begin{bmatrix} a & b \\ d & e \end{bmatrix},$$

where the brackets are used to delineate the list. Such an array is called also called a **matrix**, in this case a 2×2 matrix to indicate that it has 2 rows with 2 entries in each row (or, equivalently, 2 columns with 2 entries in each column. The matrix A is called the

coefficient matrix of the system. The number $\Delta = ae - bd$ is called the *determinant* of the matrix A and is denoted by $\det(A)$ or simply $|A|$. With this notation, we have

$$\det(A) = \begin{vmatrix} a & b \\ d & e \end{vmatrix} = ae - bd.$$

If A_1 is the 2×2 matrix obtained from A by replacing the first column of A by the column of constants c, f , and A_2 is the 2×2 matrix obtained from A by replacing the second column of A by the column of constants c, f , we obtain

$$\begin{aligned} \Delta_1 &= |A_1| = \begin{vmatrix} c & b \\ f & e \end{vmatrix} = ce - bf, \\ \Delta_2 &= |A_2| = \begin{vmatrix} a & c \\ d & f \end{vmatrix} = af - bc. \end{aligned}$$

The coordinates of the point of intersection are thus given by

$$x = \Delta_1/\Delta, \quad y = \Delta_2/\Delta.$$

This formula is known as **Cramer's Rule**. The above yields the following general theorem about systems of two linear equations in two variables.

Theorem 3.3. *The system of equations*

$$\begin{aligned} ax + by &= c \\ dx + ey &= f \end{aligned}$$

has a unique solution iff the determinant of the coefficient matrix is not zero, in which case the solution is given by

$$x = \frac{\begin{vmatrix} c & b \\ f & e \end{vmatrix}}{\begin{vmatrix} a & b \\ d & e \end{vmatrix}}, \quad y = \frac{\begin{vmatrix} a & c \\ d & f \end{vmatrix}}{\begin{vmatrix} a & b \\ d & e \end{vmatrix}}.$$

3.3. Exercises. 1. Find the equation of the line passing through the points $A(12, 3)$, $B(4, 11)$.

2. Find the equation of the line passing through the points $A(111, 23)$, $B(111, 234)$.

3. Find the intersection of the two lines $23x + 64y = 12$, $12x - 23y = 10$.

4. Find the equation of the line parallel to the line $13x - 65y = 45$ and passing through $A(5, 1)$.

5. Verify that the diagonals of a parallelogram bisect each other by finding the equations of the diagonals in a coordinate system having one vertex as origin and the two adjacent vertices as unit points.

6. Use Cramer's rule to find parametric equations for the line of intersection of the planes $12x - 13y + 7z = 1$, $10x + 11y - 5z = 2$. (Hint: Bring the terms involving z to the right of the equality sign and solve for x, y .)

7. Using determinates, solve the following system of linear equations

$$\begin{aligned} 12x - 13y + 7z &= 1 \\ 10x + 11y - 5z &= 2 \\ 13x + 7y - 11z &= 3 \end{aligned}$$

8. Let $A(x_1, y_1), B(x_2, y_2)$ be distinct points and let $P \neq A$ be point on the line L joining

A, B . If $ax + by + c = 0$ is the equation of any line which is not parallel to L and passes through P , show that

$$\frac{\overline{BP}}{\overline{AP}} = \frac{ax_2 + by_2 + c}{ax_1 + by_1 + c}.$$

9. Let A, B, C be distinct points on a line L and let A', B', C' be distinct points on a line L' which does not meet L in any of the above points. If the line joining C and B' meets the line joining C' and B in a point P , the line joining B and A' meets the line joining A and B' in a point Q and the line joining A and C' meets the line joining C and A' in a point R , show that P, Q, R are collinear.

3.4. Normal Equations of Planes. In a previous section we have shown that, in a given coordinate system, the points P whose coordinates x, y, z satisfy the equation

$$ax + by + cz = d$$

form a plane. Such an equation is called a normal equation for the plane. The following result shows the the converse is true.

Theorem 3.4. *Any plane has a normal equation.*

Proof. Let $x = a + a_1s + a_2t$, $y = b + b_1s + b_2t$, $z = c + c_1s + c_2t$ be parametric equations for a plane Π . If

$$\begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} \neq 0,$$

we can, using Cramer's rule, solve the system

$$\begin{aligned} a_1s + a_2t &= x - a \\ b_1s + b_2t &= y - b \end{aligned}$$

for s, t in terms of x, y . This gives

$$\begin{aligned} s &= \frac{\begin{vmatrix} x - a & a_2 \\ y - b & b_2 \end{vmatrix}}{\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}} = \frac{b_2(x - a) - a_2(y - b)}{a_1b_2 - a_2b_1}, \\ t &= \frac{\begin{vmatrix} a_1 & x - a \\ b_1 & y - b \end{vmatrix}}{\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}} = \frac{a_1(y - b) - b_1(x - a)}{a_1b_2 - a_2b_1}. \end{aligned}$$

Substituting this in $z - c = c_1s + c_2t$ and simplifying, we get

$$(a_1b_2 - a_2b_1)(z - c) = (c_1b_2 - c_2b_1)(x - a) + (a_1c_2 - a_2c_1)(y - b),$$

from which we obtain

$$n_1(x - a) + n_2(y - b) + n_3(z - c) = 0$$

with

$$n_1 = \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix}, \quad n_2 = -\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}, \quad n_3 = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}.$$

Since the above equation is the equation of a plane containing Π , it must be a normal equation for Π .

If

$$n_1 = \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} \neq 0,$$

we can solve the system

$$\begin{aligned} b_1s + b_2t &= y - b \\ c_1s + c_2t &= z - c \end{aligned}$$

for s, t in terms of y, z to get

$$\begin{aligned} s &= \frac{\begin{vmatrix} y-b & b_2 \\ z-c & c_2 \end{vmatrix}}{\begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}} = \frac{c_2(y-b) - b_2(z-c)}{b_1c_2 - b_2c_1}, \\ t &= \begin{vmatrix} b_1 & y-b \\ c_1 & z-c \end{vmatrix} \left| \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix} \right|^{-1} = \frac{b_1(z-c) - c_1(y-b)}{b_1c_2 - b_2c_1}. \end{aligned}$$

Substituting this in $x - a = a_1s + a_2t$ and simplifying, we get

$$(b_1c_2 - b_2c_1)(x - a) = (a_1c_2 - a_2c_1)(y - b) + (a_2b_1 - a_1b_2)(z - c),$$

from which we obtain

$$n_1(x - a) + n_2(y - b) + n_3(z - c) = 0$$

with n_1, n_2, n_3 as above. If $n_2 \neq 0$, we leave it to the reader to show that, solving

$$\begin{aligned} a_1s + a_2t &= x - a \\ c_1s + c_2t &= z - c \end{aligned}$$

for s, t in terms of x, z and substituting this in $y - a = b_1s + b_2t$, we get the same equation for Π after simplifying.

The proof will be finished if we can show that one of n_1, n_2, n_3 is not zero. This is done in the following Lemma.

Lemma 3.1. *If the scalars a_1, b_1, c_1 are not all zero, then (a_2, b_2, c_2) is a scalar multiple of (a_1, b_1, c_1) iff*

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 \\ c_1 & c_2 \end{vmatrix} = \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix} = 0.$$

Proof. If $(a_2, b_2, c_2) = k(a_1, b_1, c_1)$, then each of the determinants is zero. Conversely, suppose that each of the determinants is zero. If $a_1 \neq 0$, we have $(a_2, b_2) = k_1(a_1, b_1)$ and $(a_2, c_2) = k_2(a_1, c_2)$. But then $k_1 = a_2/a_1 = k_2$ and so $(a_2, b_2, c_2) = k_1(a_1, b_1, c_1)$. If $b_1 \neq 0$ we have $(a_2, b_2) = k_2(a_1, b_1)$ and $(b_2, c_2) = k_3(b_1, c_1)$ which gives $k_2 = b_2/b_1 = k_3$. If $c_1 \neq 0$ we have $(a_2, c_2) = k_2(a_1, c_1)$ and $(b_2, c_2) = k_3(b_1, c_1)$ which gives $k_2 = c_2/c_1 = k_3$. **Q.E.D.**

Problem 3.8. *Find a normal equation for the plane $x = 1 + 2t + 3s, y = 2 - 4t + s, z = 1 - t + s$.*

Solutiuon. Using Cramer's Rule, we solve

$$\begin{aligned} 2t + 3s &= x - 1 \\ -4t + s &= y - 2 \end{aligned}$$

to get

$$t = \frac{\begin{vmatrix} x-1 & 3 \\ y-2 & 1 \end{vmatrix}}{\begin{vmatrix} 2 & 3 \\ -4 & 1 \end{vmatrix}} = \frac{frac{x-1}{-4} - \frac{3}{1}}{-4 - 3} = \frac{frac{x-1}{-4} - 3}{-7},$$

$$s = \frac{\begin{vmatrix} 2 & x-1 \\ -4 & y-2 \end{vmatrix}}{\begin{vmatrix} 2 & 3 \\ -4 & 1 \end{vmatrix}} = \frac{4x + 2y - 8}{14}.$$

Substituting this in $z = 1 - t - s$ and simplifying, we get $5x - y + 15z = 5$ as a normal equation for the given plane.

Our next result shows that a plane has, up to a scalar multiple, only one normal equation.

Theorem 3.5. *Let Π_1, Π_2 be planes with normal equations $a_1x + b_1y + c_1z = d_1$ and $a_2x + b_2y + c_2z = d_2$ respectively. Then $\Pi_1 = \Pi_2$ iff $(a_2, b_2, c_2, d_2) = k(a_1, b_1, c_1, d_1)$ for some scalar k . The two planes are parallel iff (a_1, b_1, c_1) and (a_2, b_2, c_2) are proportional.*

Proof. If $(a_2, b_2, c_2, d_2) = k(a_1, b_1, c_1, d_1)$ for some scalar k , the equation for Π_2 is a multiple of the equation for Π_1 and so $\Pi_1 = \Pi_2$. Now suppose that

$$(a_2, b_2, c_2, d_2) \neq k(a_1, b_1, c_1, d_1)$$

for any scalar k . We have to show that $\Pi_1 \neq \Pi_2$.

Suppose that $(a_2, b_2, c_2) = k(a_1, b_1, c_1)$ for some scalar k . This implies that $d_2 \neq kd_1$. If the two planes had a point $P(x, y, z)$ in common, we could multiply the first equation by k and subtract it from the second to get $d_2 - kd_1 = 0$, contradicting $d_2 \neq kd_1$. So the planes do not meet and hence are distinct and parallel.

If (a_2, b_2, c_2) is not a scalar multiple of (a_1, b_1, c_1) , then one of the determinants

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}, \begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 \\ c_1 & c_2 \end{vmatrix}, \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} = \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}$$

is not zero. If the first one is not zero then, using Cramer's Rule, we can solve the system

$$\begin{aligned} a_1x + b_1y &= d_1 - c_1z \\ a_2x + b_2y &= d_2 - c_2z \end{aligned}$$

for any z to get

$$x = \frac{\begin{vmatrix} d_1 - c_1z & b_1 \\ d_2 - c_2z & b_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}} = \frac{d_1b_2 - d_2b_1}{a_1b_2 - a_2b_1} - \frac{c_1b_2 - c_2b_1}{a_1b_2 - a_2b_1}z,$$

$$y = \frac{\begin{vmatrix} a_1 & d_1 - c_1z \\ a_2 & d_2 - c_2z \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}} = \frac{a_2d_1 - a_1d_2}{a_1b_2 - a_2b_1} - \frac{a_1c_2 - a_2c_1}{a_1b_2 - a_2b_1}z.$$

Setting $z = t$, we get the parametric equations of a line which shows that Π_1 and Π_2 meet in a line and hence that $\Pi_1 \neq \Pi_2$. We get the same result if either of the other two determinants is not zero. **Q.E.D.**

Problem 3.9. Show that the planes $2x - 5y - 3z = 5$ and $-6x + 15y + 9z = 10$ are distinct and parallel.

Solution. It suffices to show that the given planes have no point in common. If they did then, adding 3 times the equation for the first plane to the second, we would get $0 = 25$ which is impossible. Hence the two planes do not meet and so are distinct and parallel.

Problem 3.10. Find the line of intersection of the planes $2x - 3y + 5z = 2$, $6x - 9y + 7z = 4$.

Solution. Using Cramer's Rule, we solve

$$\begin{aligned} 2x + 5z &= 3y + 2 \\ 6x + 7z &= 9y + 4 \end{aligned}$$

to get $x = (12y + 3)/8$, $z = 1/4$. Setting $y = t$, we obtain

$$x = 3/8 + 3t/4, \quad y = t, \quad z = 1/4$$

as parametric equations for the line of intersection.

The numerical vector

$$(n_1, n_2, n_3) = (b_1c_2 - b_2c_1, a_2c_1 - a_1c_2, a_1b_2 - a_2b_1)$$

is called the **vector product** of the numerical vectors $\vec{u} = (a_1, b_1, c_1)$, $\vec{v} = (a_2, b_2, c_2)$ and is denoted by $\vec{u} \times \vec{v}$. If we define the **dot (or scalar) product** of \vec{u} and \vec{v} to be the scalar

$$\vec{u} \cdot \vec{v} = a_1a_2 + b_1b_2 + c_1c_2,$$

the proof of Theorem 3.4 shows that the normal equation of a plane having non-parallel direction vectors \vec{u} , \vec{v} and passing through $A(a, b, c)$ is

$$\vec{u} \times \vec{v} \cdot (x - a, y - b, z - c) = 0.$$

This also shows that, the vector \vec{w} is a linear combination of the non-collinear vectors \vec{u} and \vec{v} iff

$$\vec{u} \times \vec{v} \cdot \vec{w} = 0.$$

The scalar $\vec{u} \times \vec{v} \cdot \vec{w}$ is called the **triple scalar product** of \vec{u} , \vec{v} , \vec{w} . If vectors \vec{u} , \vec{v} , \vec{w} are coordinate vectors of three geometrical vectors, these geometrical vectors are coplanar iff $\vec{u} \times \vec{v} \cdot \vec{w} = 0$. The triple $(\vec{u}, \vec{v}, \vec{w})$ is said to be **positively oriented** if $\vec{u} \times \vec{v} \cdot \vec{w} > 0$.

Problem 3.11. Show that the points

$$A(1, 2, 3), \quad B(2, -1, -2), \quad C(2, 1, -1)$$

are not collinear and find a normal equation for the plane Π passing through them.

Solution. We first find the direction vectors $[\overrightarrow{AB}] = (1, -3, -5)$, $[\overrightarrow{AC}] = (1, -1, -4)$. Since $(1, -3, -5) \times (1, -1, -4) = (7, -1, 3) \neq 0$, the given points are not collinear and the equation of Π is

$$(7, -1, 3) \cdot (x - 1, y - 2, z - 3) = 0.$$

This simplifies to $7x - y + 3z - 14 = 0$ or $7x - y + 3z = 14$.

The vector and scalar products have the following properties:

- (1) $\vec{u} \cdot \vec{u} \geq 0$ with equality iff $\vec{u} = (0, 0, 0)$;
- (2) $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$;
- (3) $(a\vec{u} + b\vec{v}) \cdot \vec{w} = a\vec{u} \cdot \vec{w} + b\vec{v} \cdot \vec{w}$;
- (4) $\vec{u} \times \vec{u} = \vec{0}$, $\vec{u} \times \vec{v} = -\vec{v} \times \vec{u}$;
- (5) $(a\vec{u} + b\vec{v}) \times \vec{w} = a\vec{u} \times \vec{w} + b\vec{v} \times \vec{w}$;

$$(6) \quad \vec{u} \times \vec{v} \cdot \vec{w} = \vec{w} \times \vec{u} \cdot \vec{v} = \vec{v} \times \vec{w} \cdot \vec{u};$$

$$(7) \quad \text{If } \vec{v}_1 = a_1 \vec{u} + a_2 \vec{v} + a_3 \vec{w}, \vec{v}_2 = b_1 \vec{u} + b_2 \vec{v} + b_3 \vec{w}, \text{ then}$$

$$\vec{v}_1 \times \vec{v}_2 = (a_1 b_2 - b_2 a_1) \vec{u} \times \vec{v} + (a_1 b_3 - a_3 b_1) \vec{u} \times \vec{w} + (a_2 b_3 - a_3 b_2) \vec{v} \times \vec{w}.$$

The proofs of these properties are left as exercises.

There is a very pretty application of this to the theory of linear equations. Consider the system of three equations in three variables x, y, z

$$a_1 x + b_1 y + c_1 z = d_1$$

$$a_2 x + b_2 y + c_2 z = d_2$$

$$a_3 x + b_3 y + c_3 z = d_3$$

If we let $\vec{u}_1 = (a_1, a_2, a_3)$, $\vec{u}_2 = (b_1, b_2, b_3)$, $\vec{u}_3 = (c_1, c_2, c_3)$ and $\vec{v} = (d_1, d_2, d_3)$, then the above system is equivalent to the single vector equation

$$x\vec{u}_1 + y\vec{u}_2 + z\vec{u}_3 = \vec{v}.$$

If $\vec{u}_1 \times \vec{u}_2 \cdot \vec{u}_3 \neq 0$, this equation has a unique solution (x, y, z) , namely, the coordinate vector of the point $D(d_1, d_2, d_3)$ with respect to the coordinate system with origin $O(0, 0, 0)$ and unit points

$$A(a_1, a_2, a_3), B(b_1, b_2, b_3), C(c_1, c_2, c_3).$$

Taking the vector product of both sides with \vec{u}_2 , we get

$$x\vec{u}_1 \times \vec{u}_2 + z\vec{u}_3 \times \vec{u}_2 = \vec{v} \times \vec{u}_2.$$

Taking the dot product of both sides of this equation with \vec{u}_3 , we get

$$x(\vec{u}_1 \times \vec{u}_2 \cdot \vec{u}_3) = \vec{v} \times \vec{u}_2 \cdot \vec{u}_3$$

since $\vec{u}_3 \times \vec{u}_2 \cdot \vec{u}_2 = \vec{u}_2 \times \vec{u}_2 \cdot \vec{u}_3 = 0$. Similarly, taking the vector product with \vec{u}_1 and then the dot product with \vec{u}_3 , one gets

$$y(\vec{u}_2 \times \vec{u}_1 \cdot \vec{u}_3) = \vec{v} \times \vec{u}_1 \cdot \vec{u}_3,$$

and, taking the vector product with \vec{u}_1 and then the dot product with \vec{u}_2 we get

$$z(\vec{u}_3 \times \vec{u}_1 \cdot \vec{u}_2) = \vec{v} \times \vec{u}_1 \cdot \vec{u}_2.$$

Hence, if $\vec{u}_1 \times \vec{u}_2 \cdot \vec{u}_3 \neq 0$, the system has the unique solution

$$x = \frac{\vec{v} \times \vec{u}_2 \cdot \vec{u}_3}{\vec{u}_1 \times \vec{u}_2 \cdot \vec{u}_3}, y = \frac{\vec{u}_1 \times \vec{v} \cdot \vec{u}_3}{\vec{u}_1 \times \vec{u}_2 \cdot \vec{u}_3}, z = \frac{\vec{u}_1 \times \vec{u}_2 \cdot \vec{v}}{\vec{u}_1 \times \vec{u}_2 \cdot \vec{u}_3}.$$

On the other hand $\vec{u}_1 \times \vec{u}_2 \cdot \vec{u}_3 = 0$ iff either (i) $\vec{u}_1 = (0, 0, 0)$, (ii) \vec{u}_2 is a multiple of \vec{u}_1 or (iii) \vec{u}_3 is a linear combination of \vec{u}_1 and \vec{u}_2 . One of the conditions (i), (ii), (iii) hold iff we can find x, y, z not all zero with

$$x_1 \vec{u}_1 + y_1 \vec{u}_2 + z_1 \vec{u}_3 = (0, 0, 0).$$

Such a relation is called a **dependence relation** for the vectors $\vec{u}_1, \vec{u}_2, \vec{u}_3$. If we multiply this relation by t and then add it to the original vector equation, we get

$$(x + tx_1) \vec{u}_1 + (y + ty_1) \vec{u}_2 + (z + tz_1) \vec{u}_3 = \vec{v},$$

which shows that the given system of equations either has no solution or an infinite number if $\vec{u}_1 \times \vec{u}_2 \cdot \vec{u}_3 = 0$.

This leads us to define the determinant of the 3×3 matrix

$$A = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}$$

to be the number

$$\det(A) = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = (a_1, b_1, c_1) \times (a_2, b_2, c_2) \cdot (a_3, b_3, c_3).$$

If $\Delta = \det(A)$ and $\Delta_i = \det(A_i)$, where A_i is the matrix obtained from A by replacing the i -th column of A by the column of constants d_1, d_2, d_3 , the above yields **Cramer's Rule** for the above system of equations: the system has a unique solution iff $\Delta \neq 0$ in which case the solution is $x = \Delta_1/\Delta$, $y = \Delta_2/\Delta$, $z = \Delta_3/\Delta$.

The following result gives another interpretation of the determinant, namely, as the ratio of two triple scalar products.

Theorem 3.6. If $\vec{v}_1 = a_1\vec{u}_1 + a_2\vec{u}_2 + a_3\vec{u}_3$, $\vec{v}_2 = b_1\vec{u}_1 + b_2\vec{u}_2 + b_3\vec{u}_3$, $\vec{v}_3 = c_1\vec{u}_1 + c_2\vec{u}_2 + c_3\vec{u}_3$, and

$$A = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix},$$

we have $\vec{v}_1 \times \vec{v}_2 \cdot \vec{v}_3 = \det(A) \vec{u}_1 \times \vec{u}_2 \cdot \vec{u}_3$.

Proof. We have

$$\vec{v}_1 \times \vec{v}_2 = (a_1b_2 - b_2a_1)\vec{u}_1 \times \vec{u}_2 + (a_1b_3 - a_3b_1)\vec{u}_1 \times \vec{u}_3 + (a_2b_3 - a_3b_2)\vec{u}_2 \times \vec{u}_3,$$

and so $\vec{u}_1 \times \vec{u}_2 \cdot \vec{u}_3$ is equal to

$$(a_1b_3 - a_3b_1)c_3 \vec{u}_1 \times \vec{u}_2 \cdot \vec{u}_3 + (a_2b_3 - a_3b_1)c_2 \vec{u}_1 \times \vec{u}_3 \cdot \vec{u}_2 + (a_2b_3 - a_3b_2)c_1 \vec{u}_2 \times \vec{u}_3 \cdot \vec{u}_1$$

since $\vec{u} \times \vec{v} \cdot \vec{w} = 0$ if any two the vectors $\vec{u}, \vec{v}, \vec{w}$ are equal. This yields the result since $\vec{u}_2 \times \vec{u}_3 \cdot \vec{u}_1 = \vec{u}_1 \times \vec{u}_2 \cdot \vec{u}_3$ and

$$\vec{u}_1 \times \vec{u}_3 \cdot \vec{u}_2 = \vec{u}_2 \times \vec{u}_1 \cdot \vec{u}_3 = -\vec{u}_1 \times \vec{u}_2 \cdot \vec{u}_3.$$

Q.E.D.

3.5. Exercises. 1. If $\vec{u}_1 = (1, 2, -1)$, $\vec{v} = (2, 4, 3)$, $\vec{w} = (3, 4, 1)$, find

- (a): $\vec{u} \times \vec{v}$, $(\vec{u} \times \vec{v}) \cdot \vec{w}$.
- (b): $(\vec{u} \times \vec{v}) \times \vec{w}$, $\vec{u} \times (\vec{v} \times \vec{w})$,
- (c): $\vec{u} \times \vec{v} \cdot \vec{w}$.

2. Find the normal equation of the plane passing through the points $A(1, 2, 3)$, $B(2, 4, 1)$, $C(5, 1, 2)$.

3. Find the equation of the plane containing the line

$$x = 1 + 3t, y = 7 - 6t, z = 2 + 3t$$

and not meeting the line

$$x = 10 + 2t, y = 23 + 7t, z = 39 - t.$$

4. Let $a_1x + b_1y + c_1z + d_1 = 0$, $a_2x + b_2y + c_2z + d_2 = 0$ be distinct planes which meet in a line L . Show that any plane containing L has an equation of the form

$$a(a_1x + b_1y + c_1z + d_1) + b(a_2x + b_2y + c_2z + d_2) = 0$$

for suitable a, b .

5. Find the plane passing through the point $(1, 1, -1)$ and the line of intersection of the two planes $12x + 23y + 16z = 12$, $11x + 10y - 13z = 11$.

6. Show that

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_1 & b_3 \\ c_2 & c_3 \end{vmatrix} - b_1 \begin{vmatrix} a_2 & a_3 \\ c_2 & c_3 \end{vmatrix} + c_1 \begin{vmatrix} a_2 & a_3 \\ c_2 & c_3 \end{vmatrix}.$$

7. Show that $|\vec{u} \times \vec{v}|^2 + |\vec{u} \cdot \vec{v}|^2 = |\vec{u}|^2 |\vec{v}|^2$.

4. CHANGE OF PLANE COORDINATES

Using the arithmetic of vectors, we can easily describe what happens when we change our coordinate system. For simplicity, we start with plane coordinates.

Let Π be a plane, let (O, I, J) be a frame for Π with associated coordinate system x, y . If $\vec{i} = \overrightarrow{OI}$, $\vec{j} = \overrightarrow{OJ}$ are the position vectors of the unit points of the first frame and P is a point of Π , the coordinate vector of P is (x, y) iff

$$\overrightarrow{OP} = x\vec{i} + y\vec{j}.$$

The points

$$O'(a, b), I'(a_1, b_1), J'(a_2, b_2)$$

of Π form a frame for Π iff O', I', J' are not collinear or, equivalently, if $\vec{i}' = \overrightarrow{O'I'}$ is not a scalar multiple of $\vec{j}' = \overrightarrow{O'J'}$. Since

$$\vec{i}' = (a_1 - a)\vec{i} + (b_1 - b)\vec{j}, \quad \vec{j}' = (a_2 - a)\vec{i} + (b_2 - b)\vec{j},$$

this will be true precisely when

$$\begin{vmatrix} a_1 - a & a_2 - a \\ b_1 - b & b_2 - b \end{vmatrix} \neq 0.$$

Let (x', y') be the coordinate vector of P with respect to the second frame. Then

$$\overrightarrow{O'P} = x'\vec{i}' + y'\vec{j}'.$$

We now use the fact that

$$\begin{aligned} \overrightarrow{OP} &= \overrightarrow{OO'} + \overrightarrow{O'P} \\ &= \overrightarrow{OO'} + x'\vec{i}' + y'\vec{j}' \\ &= (a\vec{i} + b\vec{j}) + x'((a_1 - a)\vec{i} + (b_1 - b)\vec{j}) + y'((a_2 - a)\vec{i} + (b_2 - b)\vec{j}) \\ &= (a + (a_1 - a)x' + (a_2 - a)y')\vec{i} + (b + (b_1 - b)x' + (b_2 - b)y')\vec{j} \end{aligned}$$

and take coordinates with respect to the first frame to get

$$\begin{aligned} x &= a + (a_1 - a)x' + (a_2 - a)y', \\ y &= b + (b_1 - b)x' + (b_2 - b)y'. \end{aligned}$$

As a check, note that setting (x', y') respectively equal to $(0, 0)$, $(1, 0)$, $(0, 1)$ gives the coordinates of the points O', I', J' . To get the coordinate vector of P with respect to the second frame we solve the equations

$$\begin{aligned} (a_1 - a)x' + (a_2 - a)y' &= x - a, \\ (b_1 - b)x' + (b_2 - b)y' &= y - b. \end{aligned}$$

for x', y' in terms of x, y .

Problem 4.1. Show that the points

$$O'(1, -1), I'(2, 1), J'(3, -2)$$

form a frame and find the coordinate vector of the point $P(3, -4)$ with respect to this frame.

Solution. Since

$$\begin{vmatrix} 1 & 2 \\ 2 & -2 \end{vmatrix} = -5 \neq 0,$$

the given points form a frame. The equations giving the change of coordinates are

$$\begin{aligned} x &= 1 + x' + 2y', \\ y &= -1 + 2x' - y'. \end{aligned}$$

Solving for x', y' in terms of x, y , we get

$$\begin{aligned} x' &= x/5 + 2y/5 + 1/5, \\ y' &= 2x/5 - y/5 - 3/5. \end{aligned}$$

Setting $x = 3, y = -4$ we get $x' = -4/5, y' = 7/5$. So $(-4/5, 7/5)$ is the coordinate vector of $P(3, -4)$ with respect to the second frame.

The system of equations

$$\begin{aligned} x &= a + \alpha_1 x' + \alpha_2 y', \\ y &= b + \beta_1 x' + \beta_2 y', \end{aligned}$$

defines a change of plane coordinates iff (α_1, β_1) is not a scalar multiple of (α_2, β_2) or, equivalently, iff the matrix

$$P = \begin{bmatrix} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \end{bmatrix}$$

has a non-zero determinant. The origin of the new coordinate system is $O'(a, b)$ with $I'(a + \alpha_1, b + \beta_1)$ the unit point on the x' -axis and $J'(a + \alpha_2, b + \beta_2)$ the unit point on the y' -axis. The matrix P is called the **transition matrix** from the xy -coordinate system to the $x'y'$ -coordinate system. The $x'y'$ -coordinate system is said to be **positively oriented** relative to the xy -coordinate system if $\det(P) > 0$.

We have

$$\begin{aligned} \vec{i}' &= \alpha_1 \vec{i} + \beta_1 \vec{j}, \\ \vec{j}' &= \alpha_2 \vec{i} + \beta_2 \vec{j}. \end{aligned}$$

The matrix

$$P^t = \begin{bmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{bmatrix}$$

is the **transpose** of the matrix P . By definition, the transpose of an $m \times n$ matrix C is the $n \times m$ matrix A^t whose entry in the i -th row j -th column is the entry in the j -th row i -th column of A . Note that $\det(P) = \det(P^t)$.

Example. Since $(-1, 1)$ is not a scalar multiple of $(2, 1)$, the equations

$$\begin{aligned} x &= 1 - x' + 2y', \\ y &= 1 + x' + y', \end{aligned}$$

are the equations for the change of x, y coordinates to a coordinate system x', y' with origin $O'(1, 1)$ and unit point $I'(0, 2)$ on the x' -axis and unit point $J'(3, 2)$ on the y' -axis. The transition matrix is

$$P = \begin{bmatrix} -1 & 2 \\ 1 & 1 \end{bmatrix}$$

and its transpose is the matrix

$$P^t = \begin{bmatrix} -1 & 1 \\ 2 & 1 \end{bmatrix}.$$

We have

$$\begin{aligned} \vec{i}' &= -\vec{i} + \vec{j}, \\ \vec{j}' &= 2\vec{i} + \vec{j}. \end{aligned}$$

The parametric equations of the x' and y' -axes are respectively $x = 1 - t, y = 1 + t$, $x = 1 + 2t, y = 1 + t$. The normal form of these equations are respectively $x + y = 2$, $x - 2y = -1$.

Suppose now that we make two changes of coordinates

$$\begin{aligned} x &= \alpha_1 x' + \alpha_2 y' & x' &= \alpha'_1 x'' + \alpha'_2 y'' \\ y &= \beta_1 x' + \beta_2 y' & y' &= \beta'_1 x'' + \beta'_2 y'' \end{aligned}$$

where, for simplicity, we have assumed that the origin does not change. To find the change of coordinates from the x, y -coordinate system to the $x'' y''$ -coordinate system, we substitute the expressions for x', y' given by the second set of equations into the first set to get

$$\begin{aligned} x &= \alpha_1(\alpha'_1 x'' + \alpha'_2 y'') + \alpha_2(\beta'_1 x'' + \beta'_2 y'') \\ y &= \beta_1(\alpha'_1 x'' + \alpha'_2 y'') + \beta_2(\beta'_1 x'' + \beta'_2 y'') \end{aligned}$$

which, on simplification, becomes

$$\begin{aligned} x &= (\alpha_1 \alpha'_1 + \alpha_2 \beta'_1) x'' + (\alpha_1 \alpha'_2 + \alpha_2 \beta'_2) y'' \\ y &= (\beta_1 \alpha'_1 + \beta_2 \beta'_1) x'' + (\beta_1 \alpha'_2 + \beta_2 \beta'_2) y''. \end{aligned}$$

If

$$P = \begin{bmatrix} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \end{bmatrix} \quad P' = \begin{bmatrix} \alpha'_1 & \alpha'_2 \\ \beta'_1 & \beta'_2 \end{bmatrix},$$

are respectively the transition matrices from the xy -coordinate system to the $x' y'$ -coordinate system and from the $x' y'$ -coordinate system to the $x'' y''$ -coordinate system, then the transition matrix from the xy -coordinate system to the $x'' y''$ -coordinate system is the matrix

$$P'' = \begin{bmatrix} \alpha_1 \alpha'_1 + \alpha_2 \beta'_1 & \alpha_1 \alpha'_2 + \alpha_2 \beta'_2 \\ \beta_1 \alpha'_1 + \beta_2 \beta'_1 & \beta_1 \alpha'_2 + \beta_2 \beta'_2 \end{bmatrix}.$$

This matrix is called the **product** of P and P' and is denoted by PP' . If we define

$$\begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} a' \\ b' \end{bmatrix} = aa' + bb',$$

then the i, j -th entry of PP' is equal to the i -th row of P times the j -th column of P' . This allows one to define the product of an $m \times 2$ matrix B and a $2 \times n$ matrix C to be the $m \times n$ matrix whose i, j -th entry is the product of the i -th row of B and the j -th column of C . For example, we have

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax + by \\ cx + dy \end{bmatrix}.$$

With these definitions, the equations giving the change of coordinates can be written in the form

$$X = PX', \quad X' = P'X''$$

with

$$X = \begin{bmatrix} x \\ y \end{bmatrix}, \quad X' = \begin{bmatrix} x' \\ y' \end{bmatrix}, \quad X'' = \begin{bmatrix} x'' \\ y'' \end{bmatrix}.$$

We then have

$$X = PX' = P(P'X'') = (PP')X''$$

which yields the associative law for multiplication of 2×2 matrices. Indeed, it suffices to note that, for 2×2 matrices A, B, C , the i -column of AB is A times the i -column of B and hence that, if C_i is the i -th column of C , the i -column of $A(BC)$ is $A(BC_i)$ since the i -th column of BC is BC_i . But, by the above, $A(BC_i) = (AB)C_i$ which is the i -th column of $(AB)C$.

Example. If

$$\begin{aligned} x &= 2x' + 3y' & x' &= 5x'' - 3y'' \\ y &= 5x' - 7y' & y' &= 4x'' + 3y'', \end{aligned}$$

we have

$$\begin{aligned} x &= 22x' + 3y' \\ y &= -3x' - 36y'. \end{aligned}$$

since

$$\begin{bmatrix} 2 & 3 \\ 5 & -7 \end{bmatrix} \begin{bmatrix} 5 & -3 \\ 4 & 3 \end{bmatrix} = \begin{bmatrix} (2)(5) + (3)(4) & (2)(-3) + (3)(3) \\ (5)(5) + (-7)(4) & (5)(-3) + (-7)(3) \end{bmatrix} = \begin{bmatrix} 22 & 3 \\ -3 & -36 \end{bmatrix}.$$

To get the transition matrix from the $x'y'$ -coordinate system to the xy -coordinate system, we solve the equations

$$\begin{aligned} \alpha_1 x' + \alpha_2 y' &= x \\ \beta_1 x' + \beta_2 y' &= y. \end{aligned}$$

for x', y' , using Cramer's Rule, to get

$$\begin{aligned} x' &= (\beta_2/\Delta)x - (\alpha_2/\Delta)y, \\ y' &= (-\beta_1/\Delta)x + (\alpha_1/\Delta)y, \end{aligned}$$

where $\Delta = \det(P)$. Hence $X' = QX$ with

$$Q = \begin{bmatrix} \beta_2/\Delta & -\alpha_2/\Delta \\ -\beta_1/\Delta & \alpha_1/\Delta \end{bmatrix} = \frac{1}{\Delta} \begin{bmatrix} \beta_2 & -\alpha_2 \\ -\beta_1 & \alpha_1 \end{bmatrix}.$$

Since $X = PX' = (PQ)X$, $X' = QX = (QP)X'$, we must have $PQ = QP = I$ where

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

The 2×2 matrix I is called the **identity matrix**. The matrix Q is called the inverse of P and is denoted by P^{-1} . The reader should verify directly that $PP^{-1} = P^{-1}P = I$, $\det(P^{-1}) = (\det P)^{-1}$ and $(P^{-1})^{-1} = P$.

Example. If x', y' is the coordinate system associated to the frame

$$(O(0,0), I'(2,1), J'(-1,3)),$$

we have

$$\begin{aligned} \vec{i}' &= 2\vec{i} + \vec{j} & x &= 2x' - y' \\ \vec{j}' &= -\vec{i} + 3\vec{j} & y &= x' + 3y'. \end{aligned}$$

The transition matrix and its inverse are

$$P = \begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix}, \quad P^{-1} = \begin{bmatrix} 3/7 & 1/7 \\ -1/7 & 2/7 \end{bmatrix}$$

which shows that

$$\begin{aligned} x' &= 3x/7 + y/7 & \vec{i} &= (3/7)\vec{i}' - (1/7)\vec{j}' \\ y' &= -x/7 + 2y/7 & \vec{j} &= (1/7)\vec{i}' + (2/7)\vec{j}'. \end{aligned}$$

Thus, in the $x'y'$ -coordinate system, the coordinates of I and J are respectively

$$(3/7, -1/7), \quad (1/7, 2/7).$$

Note that $\det(P) = 7$ while $\det(P^{-1}) = 1/7$.

If we define the sum of two 1×2 matrices by

$$\begin{bmatrix} a \\ b \end{bmatrix} + \begin{bmatrix} a' \\ b' \end{bmatrix} = \begin{bmatrix} a + a' \\ b + b' \end{bmatrix},$$

the general change of coordinates formula can be written $X = A + PX'$, where $A = [a, b]^t$. Solving for X' , we get $PX' = X - A$. We leave it as an exercise for the reader to show that the 1×2 matrices form a vector space under the above operation of addition and the operation of multiplication by scalars defined by $c[a, b]^t = [ca, cb]^t$. Multiplying both sides by P^{-1} , we get

$$X' = P^{-1}(X - A) = P^{-1}X - P^{-1}A$$

since, as the reader will easily verify,

$$P(A + B) = PA + PB, \quad P(cA) = cPA$$

for any m matrix P , any $2 \times n$ matrices A, B and any scalar c . If we make a second change of coordinates $X' = A' + P'X''$, we have

$$X = A + PX' = A + P(A' + P'X'') = A + (PA' + P(P'X'')) = (A + PA') + (PP')X''.$$

In a change of coordinates $X = A + PX'$, the transition matrix P is the identity matrix iff

$$\begin{aligned} x &= a + x', \\ y &= b + y', \end{aligned}$$

in which case the frame (O', I', J') is obtained from the frame (O, I, J) by translation by the vector $a\vec{i} + b\vec{j}$. Note that, in this case, $\vec{i} = \vec{i}'$, $\vec{j} = \vec{j}'$.

Example. If the equations giving the change of coordinates is are $x = 2 + x'$, $y = -3 + y'$ or, equivalently, $x' = x - 2$, $y' = y + 3$, the new origin is $O'(2, -3)$, the point with $x' = y' = 0$. The x' -axis, having equation $y' = 0$, is the line $y = -3$, and the y' -axis, having equation $x' = 0$, is the line $x = 2$. The unit point on the x' -axis, the point with $x' = 1, y' = 0$, is the point $I'(3, -3)$, and the unit point on the y' -axis, the point with $x' = 0, y' = 1$ is the point $J'(2, -2)$.

In the general case we have $X = A + PX'$, so that $PX' = X - A$. Multiplying both sides by P^{-1} we obtain

$$X' = P^{-1}(X - A) = P^{-1}X - P^{-1}A,$$

which shows that the transition matrix from the $x'y'$ -coordinate system is again the inverse of the transition matrix from the x, y -coordinate system to the $x'y'$ -coordinate system.

Problem 4.2. Find a coordinate system x', y' where the x' -axis is the line $2x - y + 1 = 0$ and the y' -axis is the line $x + 3y - 3 = 0$.

Solution. Let $x' = x + 3y - 3$, $y' = 2x - y + 1$ which is equivalent to the equation

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} -3 \\ 1 \end{bmatrix}.$$

If we multiply both sides of this equation by

$$\begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix}^{-1} = \begin{bmatrix} 1/7 & 3/7 \\ 2/7 & -1/7 \end{bmatrix},$$

we get

$$\begin{bmatrix} 1/7 & 3/7 \\ 2/7 & -1/7 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 1/7 & 3/7 \\ 2/7 & -1/7 \end{bmatrix} \begin{bmatrix} -3 \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \end{bmatrix}.$$

This gives

$$\begin{aligned} x &= x'/7 + 3y'/7 \\ y &= 2x'/7 - y'/7 + 1 \end{aligned}$$

and shows that x', y' is the coordinate system associated to the frame

$$O'(0, 1), I'(1/7, 9/7), J'(3/7, 6/7).$$

The x' -axis is the line $y' = 0$, i.e., $2x - y + 1 = 0$, and the y' -axis is the line $x' = 0$, i.e., $x + 3y - 3 = 0$.

Problem 4.3. Find the equation of a curve in terms of coordinates x, y of a variable point P on it if its equation in terms of the coordinates of P with respect to the frame

$$O'(1, 2), I'(2, 4), J'(2, 1)$$

is $x'y' = 1$. Also find the equations of the asymptotes.

Solution. The equation $x'y' = 1$ is the equation of a hyperbola with asymptotes the x' and y' -axes. We have $x = 1 + x' + y'$, $y = 2 + 2x' - y'$. Solving for x', y' , we get $x' = (x + y - 3)/3$, $y' = (2x - y)/3$ and so the equation $x'y' = 1$ becomes $(x + y - 3)(2x - y)/9 = 1$, which simplifies to

$$2x^2 + xy - y^2 - 6x + 3y = 9.$$

The equations of the asymptotes are $x + y - 3 = 0$, $2x - y = 0$.

In the next problem we use the technique known as **completion of squares** to simplify a quadratic equation in two variables. It is based on the simple identity

$$ax^2 + bx + c = a(x + b/2a)^2 + c - b^2/4a.$$

This technique will be used in the next section to classify the plane sections of a cone.

Problem 4.4. Sketch the curve $2x^2 + xy - y^2 - 6x + 3y - 9 = 0$.

Solution. This is the reverse of the preceding problem. To simplify the given equation, we write it as a polynomial in x , namely,

$$2x^2 + (y - 6)x - y^2 + 3y - 9 = 0$$

and complete the square in x (with $a = 2$, $b = y - 6$, $c = -y^2 + 3y - 9$) to get

$$2(x + y/4 - 3/2)^2 - (y - 6)^2/8 - y^2 + 3y - 9 = 0,$$

which simplifies to

$$2(x + y/4 - 3/2)^2 - 9y^2/8 + 9y/2 - 27/2 = 0$$

and hence to

$$2(x + y/4 - 3/2)^2 - \frac{9}{8}(y - 2)^2 = 9$$

since $-9y^2/8 + 9y/2 - 27/2 = -9(y-2)^2/8 - 9$ on completion of the square in y . Dividing by 9 we get

$$\frac{2}{9}(x + y/4 - 3/2)^2 - \frac{1}{8}(y-2)^2 = 1.$$

Setting $x' = \sqrt{2}(x + y/4 - 3/2)/3$, $y' = (y-2)/2\sqrt{2}$, the equation becomes $x'^2 - y'^2 = 1$ which, on setting $x'' = x' - y'$, $y'' = x' + y'$, becomes $x''y'' = 1$. We have

$$\begin{aligned} x'' &= \frac{\sqrt{2}}{6}(2x - y) \\ y'' &= \frac{\sqrt{2}}{3}(x + y - 3) \end{aligned}$$

If we solve for x, y in terms of x'', y'' , we get

$$\begin{aligned} x &= 1 + \sqrt{2}x'' + \sqrt{2}y''/2 \\ y &= 2 - \sqrt{2}x'' + \sqrt{2}y'', \end{aligned}$$

which shows that x'', y'' is the coordinate system associated to the frame

$$O'(1, 2), I''(1 + \sqrt{2}, 2 - \sqrt{2}), J''(1 + \sqrt{2}/2, 2 + \sqrt{2}).$$

In this coordinate system the equation of the given curve is $x''y'' = 1$ which shows that the given curve is a hyperbola with asymptotes the lines $x'' = 0$ and $y'' = 0$, i.e., the lines $2x - y = 0$ and $x + y - 3 = 0$.

4.1. **Exercises.** 1. Show that the points

$$O'(-2, 1), I'(3, 5), J'(5, 3)$$

form a frame and find the coordinate vector of the point $P(-3, 4)$ with respect to this frame. Find the equations giving the change of coordinates and write them in matrix form. What is the transition matrix from the old system to the new one?

2. Let x, y be a coordinate system in a plane and let $x' = 2x + 3y + 4$, $y' = 3x + 5y - 3$. Show that (x', y') is a coordinate system. Assuming x, y is rectangular, plot the coordinate axes for this coordinate system. What is the associated frame?

3. If x, y is a given coordinate system in a plane and $ax + by + c = 0$, $dx + ey + cz = 0$ are non-parallel lines, show that there is a coordinate system such that the coordinate vector of a point $P(x, y)$ in the new coordinate system is

$$(x', y') = (ax + by + c, dx + ey + y).$$

Find the coordinates of the origin and unit points of this new system. What are the equations of the new coordinate axes?

4. Find the inverses of the following matrices

$$\begin{bmatrix} 3 & 11 \\ -2 & 6 \end{bmatrix}, \quad \begin{bmatrix} 12 & 11 \\ 12 & -16 \end{bmatrix}, \quad \begin{bmatrix} 13 & -21 \\ 21 & -11 \end{bmatrix}.$$

5. Verify that both

$$\begin{aligned} x &= 7x' - 3y' & x' &= 9x'' + 6y'' \\ y &= 6x' - 2y' & y' &= 5x'' + 4y'' \end{aligned}$$

are equations for a change of coordinates and find the equations for the change of coordinates from the xy -coordinate system to the $x''y''$ -coordinate system and from the $x''y''$ -coordinate system to the xy -coordinate system. Do this by direct substitution and an by the use of

matrices. Find the determinant of each transition matrix. Which of the coordinate systems are positively oriented with respect to the xy -coordinate system?

6. Repeat exercise 5 with the equations

$$\begin{aligned} x &= 2 + 3x' + 5y' & x' &= 6 + 9x'' - 6y'' \\ y &= -1 + 7x' + 4y' & y' &= -1 - 2x'' + 2y'' \end{aligned}$$

7. Find 2×2 matrices P, Q with $PQ \neq QP$.

8. If P, Q are 2×2 matrices, prove that $\det(PQ) = \det(P)\det(Q)$.

9. If P, Q are 2×2 matrices with $PQ = I$ or $QP = I$, show that $\det(P) \neq 0$ and $Q = P^{-1}$.

10. If the 2×2 matrices A, B have inverses, show that AB has an inverse and that $(AB)^{-1} = B^{-1}A^{-1}$.

11. Compute

$$\begin{bmatrix} 3 & -2 \\ 4 & 6 \\ -7 & 8 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} 2 & -7 & 5 \\ 3 & 6 & -3 \end{bmatrix}, \quad \begin{bmatrix} 3 & -2 \\ 4 & 6 \\ -7 & 8 \end{bmatrix} \begin{bmatrix} 2 & -7 & 5 & 0 \\ 3 & 6 & -3 & 9 \end{bmatrix}.$$

12. If

$$\begin{aligned} x_1 &= 2r + 3s & r &= 3y_1 + 2y_2 - 7y_3 + 5y_4 \\ x_2 &= -4r + 5s & s &= 2y_1 + 4y_2 + 3y_3 - 2y_4 \\ x_3 &= 3r \end{aligned}$$

show that

$$\begin{aligned} x_1 &= a_1y_1 + a_2y_2 + a_3y_3 + a_4y_4 \\ x_2 &= b_1y_1 + b_2y_2 + b_3y_3 + b_4y_4 \\ x_3 &= c_1y_1 + b_2y_2 + c_3y_3 + b_4y_4 \end{aligned}$$

where

$$\begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ -4 & 5 \\ 3 & 0 \end{bmatrix} \begin{bmatrix} 3 & 2 & -7 & 5 \\ 2 & 4 & 3 & -2 \end{bmatrix}.$$

13. If

$$\begin{aligned} x_1 &= a_1r_1 + b_1r_2 \\ x_2 &= a_2r_1 + b_2r_2 & r_1 &= c_1y_1 + c_2y_2 + \dots c_ny_n \\ \vdots & \quad \quad \quad \quad & r_2 &= d_1y_1 + d_2y_2 + \dots d_ny_n, \\ x_m &= a_mr_1 + b_mr_2 \end{aligned}$$

show that

$$\begin{aligned} x_1 &= a_{11}y_1 + a_{12}y_2 + \dots a_{1n}y_n \\ x_2 &= a_{21}y_1 + a_{22}y_2 + \dots a_{2n}y_n \\ &\quad \dots \\ x_m &= a_{m1}y_1 + a_{m2}y_2 + \dots a_{mn}y_n \end{aligned}$$

where

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ \vdots & \vdots \\ a_m & b_m \end{bmatrix} \begin{bmatrix} c_1 & c_2 & \dots & c_n \\ d_1 & d_2 & \dots & d_n \end{bmatrix}.$$

14. Show that $(AB)^t = B^t A^t$ if A is $m \times 2$ and B is $2 \times n$.

15. Show that

$$\begin{pmatrix} x & y \end{pmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{pmatrix} x & y \end{pmatrix} = \begin{pmatrix} x' & y' \end{pmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{pmatrix} x' & y' \end{pmatrix} = axx' + bxy' + cyx' + dyy'.$$

16. Sketch the curve whose equation in rectangular coordinates is

$$45x^2 + 14xy + 10y^2 + 6x + 8y = 3.$$

Use completion of squares.

17. Sketch the curve whose equation in rectangular coordinates is

$$4x^2 + 12xy + 9y^2 + 3x + 4y = 0.$$

Use completion of squares.

5. CHANGE OF COORDINATES IN SPACE

Let (O, I, J, K) be coordinate frame with associated coordinate system x, y, z and let $\vec{i}, \vec{j}, \vec{k}$ be the position vectors of the unit points I, J, K with respect to the origin O . Let

$$O'(a, b, c), I'(a_1, b_1, c_1), J'(a_2, b_2, c_2), K'(a_3, b_3, c_3).$$

be a second frame and let

$$\begin{aligned} \vec{i}' &= \overrightarrow{O'I'} = (a_1 - a)\vec{i} + (b_1 - b)\vec{j} + (c_1 - c)\vec{k}, \\ \vec{j}' &= \overrightarrow{O'J'} = (a_2 - a)\vec{i} + (b_2 - b)\vec{j} + (c_2 - c)\vec{k}, \\ \vec{k}' &= \overrightarrow{O'K'} = (a_3 - a)\vec{i} + (b_3 - b)\vec{j} + (c_3 - c)\vec{k}. \end{aligned}$$

If P is a point with coordinates (x, y, z) with respect to the first frame and coordinates (x', y', z') with respect to the second frame, we have

$$\begin{aligned} \overrightarrow{OP} &= \overrightarrow{OO'} + \overrightarrow{O'P}, \\ &= \overrightarrow{OO'} + (x'\vec{i}' + y'\vec{j}' + z'\vec{k}'). \end{aligned}$$

Taking coordinate vectors with respect to the first frame, we get

$$\begin{aligned} x &= a + (a_1 - a)x' + (a_2 - a)y' + (a_3 - a)z', \\ y &= b + (b_1 - b)x' + (b_2 - b)y' + (b_3 - b)z', \\ z &= c + (c_1 - c)x' + (c_2 - c)y' + (c_3 - c)z'. \end{aligned}$$

To get the coordinate vector of P with respect to the second frame we have to solve this linear system of equations for x', y', z' . This we can do by Gauss-Jordan elimination or by Cramer's Rule. Notice that, setting (x', y', z') equal to $(0, 0, 0)$, $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$, we get respectively the coordinate vectors of O', I', J', K' with respect to the first frame. A given system of equations

$$\begin{aligned} x &= a + \alpha_1 x' + \alpha_2 y' + \alpha_3 z', \\ y &= b + \beta_1 x' + \beta_2 y' + \beta_3 z', \\ z &= c + \gamma_1 x' + \gamma_2 y' + \gamma_3 z', \end{aligned}$$

are the equations for a change of coordinates iff the vectors

$$\vec{u} = (\alpha_1, \beta_1, \gamma_1), \vec{v} = (\alpha_2, \beta_2, \gamma_2), \vec{w} = (\alpha_3, \beta_3, \gamma_3),$$

represent non-coplanar vectors. A necessary and sufficient condition for this is that none of the vectors $\vec{u}, \vec{v}, \vec{w}$ be a linear combination of the other two or, equivalently, that

$$a\vec{u} + b\vec{v} + c\vec{w} = 0 \Rightarrow a = b = c = 0.$$

Such a sequence of vectors is said to be **linearly independent**. A necessary and sufficient condition for this is that $\vec{u} \times \vec{v} \cdot \vec{w} \neq 0$ or, what is the same thing, that the matrix

$$P = \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{bmatrix}$$

have a non-zero determinant. The matrix P is called the *transition matrix* from the x, y, z coordinate system to the x', y', z' coordinate system. The second frame is said to be **positively oriented** with respect to the first if $\det(P) > 0$. We have

$$\begin{aligned} \vec{i}' &= \alpha_1 \vec{i} + \beta_1 \vec{j} + \gamma_1 \vec{k} \\ \vec{j}' &= \alpha_2 \vec{i} + \beta_2 \vec{j} + \gamma_2 \vec{k} \\ \vec{k}' &= \alpha_3 \vec{i} + \beta_3 \vec{j} + \gamma_3 \vec{k}. \end{aligned}$$

The matrix of coefficients of the vectors $\vec{i}, \vec{j}, \vec{k}$ is

$$\begin{bmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \\ \alpha_3 & \beta_3 & \gamma_3 \end{bmatrix}$$

which is the transpose of P .

Example. Since

$$\begin{vmatrix} 2 & -1 & 2 \\ 1 & 2 & -1 \\ 1 & 1 & 2 \end{vmatrix} = 11 \neq 0,$$

the equations

$$\begin{aligned} x &= 1 + 2x' - y' + 2z' \\ y &= -1 + x' + 2y' - z' \\ z &= 2 + x' + y' + 2z' \end{aligned}$$

give a change of coordinates from the x, y, z -coordinate system to a coordinate system x', y', z' in which the origin is $O'(1, -1, 2)$ and the unit points on the x', y', z' -axes are respectively

$$I'(3, 0, 3), J'(0, 1, 3), K'(3, -2, 4).$$

The transition matrix is the matrix

$$P = \begin{bmatrix} 2 & -1 & 2 \\ 1 & 2 & -1 \\ 1 & 1 & 2 \end{bmatrix}$$

and we have

$$\begin{aligned} \vec{i}' &= 2\vec{i} + \vec{j} + \vec{k} \\ \vec{j}' &= -\vec{i} + 2\vec{j} + \vec{k} \\ \vec{k}' &= 2\vec{i} - \vec{j} + 2\vec{k}. \end{aligned}$$

Suppose now we make two successive changes of coordinates

$$\begin{aligned} x &= \alpha_1 x' + \alpha_2 y' + \alpha_3 z' & x' &= \alpha'_1 x'' + \alpha'_2 y'' + \alpha'_3 z'' \\ y &= \beta_1 x' + \beta_2 y' + \beta_3 z' & y' &= \beta'_1 x'' + \beta'_2 y'' + \beta'_3 z'' \\ z &= \gamma_1 x' + \gamma_2 y' + \gamma_3 z' & z' &= \gamma'_1 x'' + \gamma'_2 y'' + \gamma'_3 z'' \end{aligned}$$

where again, for simplicity, we assume that the origin does not change. To obtain the change of coordinates from the xyz -coordinate system to the $x''y''z''$ -coordinate system, we substitute the expressions for x', y', z' from the second set of equations into the first set to get

$$\begin{aligned} x &= \alpha_1(\alpha'_1 x'' + \alpha'_2 y'' + \alpha'_3 z'') + \alpha_2(\beta'_1 x'' + \beta'_2 y'' + \beta'_3 z'') + \alpha_3(\gamma'_1 x'' + \gamma'_2 y'' + \gamma'_3 z'') \\ y &= \beta_1(\alpha'_1 x'' + \alpha'_2 y'' + \alpha'_3 z'') + \beta_2(\beta'_1 x'' + \beta'_2 y'' + \beta'_3 z'') + \beta_3(\gamma'_1 x'' + \gamma'_2 y'' + \gamma'_3 z'') \\ z &= \gamma_1(\alpha'_1 x'' + \alpha'_2 y'' + \alpha'_3 z'') + \gamma_2(\beta'_1 x'' + \beta'_2 y'' + \beta'_3 z'') + \gamma_3(\gamma'_1 x'' + \gamma'_2 y'' + \gamma'_3 z'') \end{aligned}$$

which, on simplification, becomes

$$\begin{aligned} x &= (\alpha_1 \alpha'_1 + \alpha_2 \beta'_1 + \alpha_3 \gamma'_1) x'' + (\alpha_1 \alpha'_2 + \alpha_2 \beta'_2 + \alpha_3 \gamma'_2) y'' + (\alpha_1 \alpha'_3 + \alpha_2 \beta'_3 + \alpha_3 \gamma'_3) z'' \\ y &= (\beta_1 \alpha'_1 + \beta_2 \beta'_1 + \beta_3 \gamma'_1) x'' + (\beta_1 \alpha'_2 + \beta_2 \beta'_2 + \beta_3 \gamma'_2) y'' + (\beta_1 \alpha'_3 + \beta_2 \beta'_3 + \beta_3 \gamma'_3) z'' \\ z &= (\gamma_1 \alpha'_1 + \gamma_2 \beta'_1 + \gamma_3 \gamma'_1) x'' + (\gamma_1 \alpha'_2 + \gamma_2 \beta'_2 + \gamma_3 \gamma'_2) y'' + (\gamma_1 \alpha'_3 + \gamma_2 \beta'_3 + \gamma_3 \gamma'_3) z''. \end{aligned}$$

If

$$P = \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{bmatrix} \quad P' = \begin{bmatrix} \alpha'_1 & \alpha'_2 & \alpha'_3 \\ \beta'_1 & \beta'_2 & \beta'_3 \\ \gamma'_1 & \gamma'_2 & \gamma'_3 \end{bmatrix},$$

are respectively the transition matrices from the xyz -coordinate system to the $x'y'z'$ system and from the $x'y'z'$ -coordinate system to the $x''y''z''$ -coordinate system, the transition matrix from the xyz -coordinate system to the $x''y''z''$ -coordinate system is the matrix

$$P'' = \begin{bmatrix} \alpha_1 \alpha'_1 + \alpha_2 \beta'_1 + \alpha_3 \gamma'_1 & \alpha_1 \alpha'_2 + \alpha_2 \beta'_2 + \alpha_3 \gamma'_2 & \alpha_1 \alpha'_3 + \alpha_2 \beta'_3 + \alpha_3 \gamma'_3 \\ \beta_1 \alpha'_1 + \beta_2 \beta'_1 + \beta_3 \gamma'_1 & \beta_1 \alpha'_2 + \beta_2 \beta'_2 + \beta_3 \gamma'_2 & \beta_1 \alpha'_3 + \beta_2 \beta'_3 + \beta_3 \gamma'_3 \\ \gamma_1 \alpha'_1 + \gamma_2 \beta'_1 + \gamma_3 \gamma'_1 & \gamma_1 \alpha'_2 + \gamma_2 \beta'_2 + \gamma_3 \gamma'_2 & \gamma_1 \alpha'_3 + \gamma_2 \beta'_3 + \gamma_3 \gamma'_3 \end{bmatrix}.$$

This matrix is called the **product** of P and P' and is denoted by PP' . If we define

$$\begin{bmatrix} a & b & c \end{bmatrix} \begin{bmatrix} a' \\ b' \\ c' \end{bmatrix} = aa' + bb' + cc',$$

then the i, j -th entry of PP' is equal to the i -th row of P times the j -th column of P' . This allows one to define the product of an $m \times 3$ matrix B and a $3 \times n$ matrix C to be the $m \times n$ matrix whose i, j -th entry is the product of the i -th row of B and the j -th column of C . For example, we have

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & k \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} ax + by + cz \\ dx + ey + fz \\ gx + hy + kz \end{bmatrix}.$$

With these definitions, the equations giving the change of coordinates can be written in the form

$$X = PX', \quad X' = P'X''$$

with

$$X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \quad X' = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}, \quad X'' = \begin{bmatrix} x'' \\ y'' \\ z'' \end{bmatrix}.$$

We then have

$$X = PX' = P(P'X'') = (PP')X''$$

which yields the associative law for multiplication of 3×3 matrices. Indeed, it suffices to note that, for 3×3 matrices A, B, C , the i -column of AB is A times the i -column of B and hence that, if C_i is the i -th column of C , the i -column of $A(BC)$ is $A(BC_i)$ since the i -th column of BC is BC_i . But, by the above, $A(BC_i) = (AB)C_i$ which is the i -th column of $(AB)C$.

Example. If

$$\begin{array}{rcl} x & = & 2x' + 3y' + 4z' \\ y & = & 5x' - 7y' + 3z' \\ z & = & 3x' + 7y' - 2z' \end{array} \quad \begin{array}{rcl} x' & = & 5x'' - 3y'' + 2z'' \\ y' & = & 4x'' + 3y'' - 7z'' \\ z' & = & 6x'' + 2y'' + 6z'' \end{array}$$

we have

$$\begin{array}{rcl} x & = & 46x' + 15y' + 7z' \\ y & = & 15x' - 30y' + 77z' \\ z & = & 31x' + 8y' - 55z' \end{array}$$

since

$$\begin{aligned} & \begin{bmatrix} 2 & 3 & 4 \\ 5 & -7 & 3 \\ 3 & 7 & -2 \end{bmatrix} \begin{bmatrix} 5 & -3 & 2 \\ 4 & 3 & -7 \\ 6 & 2 & 6 \end{bmatrix} = \\ & \begin{bmatrix} (2)(5) + (3)(4) + (4)(6) & (2)(-3) + (3)(3) + (4)(2) & (2)(2) + (3)(-7) + (4)(6) \\ (5)(5) + (-7)(4) + (3)(6) & (5)(-3) + (-7)(3) + (3)(2) & (5)(2) + (-7)(-7) + (3)(6) \\ (3)(5) + (7)(4) + (-2)(6) & (3)(-3) + (7)(3) + (-2)(2) & (3)(2) + (7)(-7) + (-2)(6) \end{bmatrix} \\ & = \begin{bmatrix} 46 & 15 & 7 \\ 15 & -30 & 77 \\ 31 & 8 & -55 \end{bmatrix}. \end{aligned}$$

To get the transition matrix from the $x'y'$ -coordinate system to the xy -coordinate system, we solve the equations

$$\begin{array}{rcl} \alpha_1 x' + \alpha_2 y' + \alpha_3 z' & = & x \\ \beta_1 x' + \beta_2 y' + \beta_3 z' & = & y \\ \gamma_1 x' + \gamma_2 y' + \gamma_3 z' & = & z \end{array}$$

for x', y', z' using Cramer's Rule, to get

$$\begin{array}{rcl} \Delta x' & = & (\beta_2 \gamma_3 - \beta_3 \gamma_2)x + (\alpha_3 \gamma_2 - \alpha_2 \gamma_3)y + (\alpha_2 \beta_3 - \alpha_3 \beta_2)z, \\ \Delta y' & = & (\beta_3 \gamma_1 - \beta_1 \gamma_3)x + (\alpha_1 \gamma_3 - \alpha_3 \gamma_1)y + (\beta_1 \alpha_3 - \beta_3 \alpha_1)z, \\ \Delta z' & = & (\beta_1 \gamma_2 - \beta_2 \gamma_1)x + (\alpha_2 \gamma_1 - \alpha_1 \gamma_2)y + (\alpha_1 \beta_2 - \alpha_2 \beta_1)z, \end{array}$$

where $\Delta = \det(P)$. Writing this in matrix form, we get $X' = QX$ with

$$Q = \frac{1}{\Delta} \begin{bmatrix} \left| \begin{array}{cc} \beta_2 & \beta_3 \\ \gamma_2 & \gamma_3 \end{array} \right| & - \left| \begin{array}{cc} \alpha_2 & \alpha_3 \\ \gamma_2 & \gamma_3 \end{array} \right| & \left| \begin{array}{cc} \alpha_2 & \alpha_3 \\ \beta_2 & \beta_3 \end{array} \right| \\ - \left| \begin{array}{cc} \beta_1 & \beta_3 \\ \gamma_1 & \gamma_3 \end{array} \right| & \left| \begin{array}{cc} \alpha_1 & \alpha_3 \\ \gamma_1 & \gamma_3 \end{array} \right| & - \left| \begin{array}{cc} \alpha_1 & \alpha_3 \\ \beta_1 & \beta_3 \end{array} \right| \\ \left| \begin{array}{cc} \beta_1 & \beta_2 \\ \gamma_1 & \gamma_2 \end{array} \right| & - \left| \begin{array}{cc} \alpha_1 & \alpha_2 \\ \gamma_1 & \gamma_2 \end{array} \right| & \left| \begin{array}{cc} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \end{array} \right| \end{bmatrix}.$$

This matrix is called the **inverse** of P and is denoted by P^{-1} . The entry in the i -th row j -th column of P^{-1} is equal to $(-1)^{i+j}/\Delta$ times the determinant of the 2×2 matrix obtained from P by deleting the j -th row and i -th column. Since

$$X = PX' = (PQ)X, \quad X' = QX = (QP)X',$$

we must have $PP^{-1} = P^{-1}P = I$, where

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The 3×3 matrix I is called the **identity matrix**. The reader should verify directly that $PP^{-1} = P^{-1}P = I$.

Example. If

$$\begin{aligned} x &= 2x' + 3y' + 4z' \\ y &= 5x' - 7y' + 3z' \\ z &= 3x' + 7y' - 2z', \end{aligned}$$

we have

$$\begin{aligned} x' &= -7x/267 + 34y/267 + 37z/267 \\ y' &= 19x/267 - 16y/267 - 14z/267 \\ z' &= 56x/267 - 5y/267 - 29z/267 \end{aligned}$$

since

$$\begin{vmatrix} 2 & 3 & 4 \\ 5 & -7 & 3 \\ 3 & 7 & -2 \end{vmatrix} = 267$$

and

$$\begin{aligned} \begin{bmatrix} 2 & 3 & 4 \\ 5 & -7 & 3 \\ 3 & 7 & -2 \end{bmatrix}^{-1} &= \frac{1}{267} \begin{bmatrix} \begin{vmatrix} -7 & 3 \\ 7 & -2 \end{vmatrix} & -\begin{vmatrix} 3 & 4 \\ 7 & -2 \end{vmatrix} & \begin{vmatrix} 3 & 4 \\ -7 & 3 \end{vmatrix} \\ -\begin{vmatrix} 5 & -7 \\ 3 & 7 \end{vmatrix} & \begin{vmatrix} 2 & 4 \\ 3 & -2 \end{vmatrix} & -\begin{vmatrix} 2 & 4 \\ 5 & 3 \end{vmatrix} \\ \begin{vmatrix} 5 & -7 \\ 3 & 7 \end{vmatrix} & -\begin{vmatrix} 2 & 3 \\ 3 & 7 \end{vmatrix} & \begin{vmatrix} 2 & 3 \\ 5 & -7 \end{vmatrix} \end{bmatrix} \\ &= \frac{1}{267} \begin{bmatrix} -7 & 34 & 37 \\ 19 & -16 & -14 \\ 56 & -5 & -29 \end{bmatrix}. \end{aligned}$$

The general change of coordinates formula can be written $X = A + PX'$, where $A = [a, b, c]^t$ and where the operations of addition and multiplication by scalars for 1×3 matrices are defined by

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} + \begin{bmatrix} a' \\ b' \\ c' \end{bmatrix} = \begin{bmatrix} a + a' \\ b + b' \\ c + c' \end{bmatrix}, \quad t \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} ta \\ tb \\ tc \end{bmatrix}.$$

Under these operations, the 1×3 matrices form a vector space. Solving for X' , we get $PX' = X - A$. Multiplying both sides by P^{-1} , we get

$$X' = P^{-1}(X - A) = P^{-1}X - P^{-1}A$$

since, as the reader will easily verify,

$$P(A + B) = PA + PB, \quad P(cA) = cPA$$

for any $m \times 3$ matrix P , any $3 \times n$ matrices A, B and any scalar c .

Example. The equations

$$\begin{aligned} x &= 1 + 2x' + 3y' + 4z' \\ y &= -2 + 5x' - 7y' + 3z' \\ z &= 3 + 3x' + 7y' - 2z', \end{aligned}$$

can be written $X = A + PX'$ with

$$A = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}, \quad P = \begin{bmatrix} 2 & 3 & 4 \\ 5 & -7 & 3 \\ 3 & 7 & -2 \end{bmatrix}.$$

Since $X' = P^{-1}X - P^{-1}A$ and

$$P^{-1} = \frac{1}{267} \begin{bmatrix} -7 & 34 & 37 \\ 19 & -16 & -14 \\ 56 & -5 & -29 \end{bmatrix},$$

we get

$$\begin{aligned} x' &= -7x/267 + 34y/267 + 37z/267 - 172/267 \\ y' &= 19x/267 - 16y/267 - 14z/267 - 9/267 \\ z' &= 56x/267 - 5y/267 - 29z/267 - 21 \end{aligned}$$

If we make a second change of coordinates $X' = A' + P'X''$, we have

$$X = A + PX' = A + P(A' + P'X'') = A + (PA' + P(P'X'')) = (A + PA') + (PP')X''.$$

In a change of coordinates $X = A + PX'$, the transition matrix P is the identity matrix iff

$$\begin{aligned} x &= a + x' \\ y &= b + y' \\ z &= c + z' \end{aligned}$$

in which case the frame (O', I', J') is obtained from the frame (O, I, J) by translation by the vector $a\vec{i} + b\vec{j}$. Note that, in this case, $\vec{i} = \vec{i'}$, $\vec{j} = \vec{j'}$.

Problem 5.1. Find a coordinate system x', y', z' where the $x'y'$ -plane is $2x - y + z = 1$, the $y'z'$ -plane is $x + 3y + z = 3$ and the $x'z'$ -plane is $x - y + z + 1 = 0$.

Solution. Let

$$\begin{aligned} x' &= x + 3y + z - 3 \\ y' &= x - y + z + 1 \\ z' &= 2x - y + z - 1 \end{aligned}$$

which is equivalent to the equation

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 1 & 3 & 1 \\ 1 & -1 & 1 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} -3 \\ 1 \\ -1 \end{bmatrix}.$$

If we multiply both sides of this equation by

$$\begin{bmatrix} 1 & 3 & 1 \\ 1 & -1 & 1 \\ 2 & -1 & 1 \end{bmatrix}^{-1} = \frac{1}{4} \begin{bmatrix} 0 & -4 & 4 \\ 1 & -1 & 1 \\ 1 & 7 & -4 \end{bmatrix},$$

we get

$$\frac{1}{4} \begin{bmatrix} 0 & -4 & 4 \\ 1 & -1 & 1 \\ 1 & 7 & -4 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} -2 \\ 1/2 \\ 2 \end{bmatrix}.$$

This gives

$$\begin{aligned} x &= 2 - y' + z' \\ y &= -1/2 + x'/4 - y'/4 + z'/4 \\ z &= -2 + x'/4 + 7y'/4 - z' \end{aligned}$$

and shows that x', y', z' is the coordinate system associated to the frame

$$O'(2, -1/2, 2), \quad I'(2, -1/4, -7/4), \quad J'(1, -3/4, -1/4), \quad K'(3, -1/4, -3).$$

The $x'y'$ -plane is the plane $z' = 0$, i.e., $2x - y + z - 1 = 0$, the xz'' -plane is the plane $y' = 0$, i.e., $x - y + z + 1 = 0$ and the $y'z'$ -plane is the plane $x' = 0$, i.e., $x + 3y + z - 3 = 0$.

Theorem 5.1. *If A, B are 3×3 matrices then $\det(AB) = \det(A) \det(B)$.*

Proof. Let

$$A = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}, \quad B = \begin{bmatrix} a'_1 & a'_2 & a'_3 \\ b'_1 & b'_2 & b'_3 \\ c'_1 & c'_2 & c'_3 \end{bmatrix}.$$

If $\vec{u}_1 = (a'_1, a'_2, a'_3)$, $\vec{u}_2 = (b'_1, b'_2, b'_3)$, $\vec{u}_3 = (c'_1, c'_2, c'_3)$ are the rows of B , the Corollary to Theorem 3.6 gives

$$\vec{u}_1 \times \vec{u}_2 \cdot \vec{u}_3 = \det(B).$$

If we let

$$\begin{aligned} \vec{v}_1 &= a_1 \vec{u}_1 + a_2 \vec{u}_2 + a_3 \vec{u}_3 \\ \vec{v}_2 &= b_1 \vec{u}_1 + b_2 \vec{u}_2 + b_3 \vec{u}_3 \\ \vec{v}_3 &= c_1 \vec{u}_1 + c_2 \vec{u}_2 + c_3 \vec{u}_3, \end{aligned}$$

the vectors $\vec{v}_1, \vec{v}_2, \vec{v}_3$ are the rows of AB and so

$$\vec{v}_1 \times \vec{v}_2 \cdot \vec{v}_3 = \det(AB).$$

But, by Theorem 3.6,

$$\vec{v}_1 \times \vec{v}_2 \cdot \vec{v}_3 = \det(A) \vec{u}_1 \times \vec{u}_2 \cdot \vec{u}_3$$

which gives the result. **Q.E.D.**

As a consequence, from $PP^{-1} = I$ we get that $\det(P) \det(P^{-1}) = 1$ since $\det(I) = 1$. Hence $\det(P^{-1}) = \det(P)^{-1}$ if $\det(P) \neq 0$. It follows that, if a coordinate system x', y', z' is positively oriented with respect to the coordinate system x, y, z , then the coordinate system x, y, z is positively oriented with respect to the coordinate system x, y, z . If this is so and the coordinate system x', y', z' is positively oriented relative to a coordinate system x'', y'', z'' , then the coordinate system x, y, z is positively oriented relative to the coordinate system x'', y'', z'' . This is so because, if P, P', P'' are the respective transition matrices, $P'' = PP'$ and so

$$\det(P'') = \det(P) \det(P') > 0$$

since $\det(P) > 0$ and $\det(P') > 0$. We say that two coordinate systems have the **same orientation** if they are positively oriented relative to one another. If we choose two coordinate systems which do not have the same orientation, then any coordinate system has the same orientation as one of them. It follows that the coordinate systems divide into two classes, where any two coordinate systems in the same class have the same orientation.

5.1. **Exercises.** 1. Show that the points

$$O'(-2, 1, 3), I'(3, 5, 2), J'(5, 3, 4), K'(3, 5, 7)$$

form a frame and find the coordinate vector of the point $P(-3, 4, 1)$ with respect to this frame. Find the equations giving the change of coordinates and write them in matrix form. What is the transition matrix from the old coordinate system to the new one?

2. Let x, y, z be a given coordinate system and let

$$\begin{aligned}x' &= x + 2y + z + 1, \\y' &= 2x + 5y - z - 2, \\z' &= x + y + 3z + 3.\end{aligned}$$

Find a coordinate frame so that (x', y', z') is the coordinate vector of $P(x, y, z)$ in the new coordinate system.

3. If x, y, z is a given coordinate system and

$$a_1x + b_1y + c_1z + d_1 = 0, a_2x + b_2y + c_2z + d_2 = 0, a_3x + b_3y + c_3z + d_3 = 0$$

are three planes which meet in a single point, show that there is coordinate system such that the coordinate vector of $P(x, y, z)$ in the new coordinate system is

$$(a_1x + b_1y + c_1z + d_1, a_2x + b_2y + c_2z + d_2, a_3x + b_3y + c_3z + d_3).$$

4. Find the inverses of the matrices

$$\begin{bmatrix} 1 & 2 & 4 \\ -1 & 4 & 3 \\ 2 & 3 & 7 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 4 \\ 2 & 4 & 1 \\ 4 & 1 & 2 \end{bmatrix}.$$

5. Verify that both

$$\begin{aligned}x &= 7x' - 3y' + z' & x' &= 9x'' + 6y'' + 2z'' \\y &= 6x' - 2y' - z' & y' &= 5x'' + 4y'' + z'' \\z &= 2x' - y' + 2z' & z' &= x'' + y'' + z''\end{aligned}$$

are equations for a change of coordinates and find the equations for the change of coordinates from the xy -coordinate system to the $x''y''$ -coordinate system and from the $x''y''$ -coordinate system to the xy -coordinate system. Do this by direct substitution and by the use of matrices. Find the determinant of each transition matrix. Which of the coordinate systems are positively oriented with respect to the xy -coordinate system?

6. Repeat exercise 5 with the equations

$$\begin{aligned}x &= 1 + 7x' - 3y' + z' & x' &= -1 + 9x'' + 6y'' + 2z'' \\y &= 2 + 6x' - 2y' - z' & y' &= 2 + 5x'' + 4y'' + z'' \\z &= 3 + 2x' - y' + 2z' & z' &= x'' + y'' + z''\end{aligned}$$

7. If P is a 3×3 matrix, prove that $\det(P^t) = \det(P)$.

8. If P, Q are 3×3 matrices such that $PQ = 1$ or $QP = 1$, show that $\det(P) \neq 0$ and that $Q = P^{-1}$.

9. If the 3×3 matrices A, B have inverses, show that AB has an inverse and that $(AB)^{-1} = B^{-1}A^{-1}$.

10. Compute

$$\begin{bmatrix} 3 & -2 & 5 \\ 4 & 6 & 7 \end{bmatrix} \begin{bmatrix} 2 & -7 & 5 & 0 \\ 3 & 6 & -3 & 9 \\ 1 & 5 & 5 & 0 \end{bmatrix}, \quad \begin{bmatrix} 2 & -3 & 5 \end{bmatrix} \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix}.$$

11. If

$$\begin{aligned} x_1 &= 2r + 3s + 4t & r &= 3y_1 + 2y_2 - 7y_3 + 5y_4 \\ x_2 &= -4r + 5s + t & s &= 2y_1 + 4y_2 + 3y_3 - 2y_4, \\ x_3 &= 3r - 7s + 6t & t &= 3y_1 - 2y_2 - 5y_4 \end{aligned}$$

show that

$$\begin{aligned} x_1 &= a_1y_1 + a_2y_2 + a_3y_3 + a_4y_4 \\ x_2 &= b_1y_1 + b_2y_2 + b_3y_3 + b_4y_4 \\ x_3 &= c_1y_1 + c_2y_2 + c_3y_3 + c_4y_4 \end{aligned}$$

where

$$\begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \end{bmatrix} = \begin{bmatrix} 2 & 3 & 4 \\ -4 & 5 & 1 \\ 3 & -7 & 6 \end{bmatrix} \begin{bmatrix} 3 & 2 & -7 & 5 \\ 2 & 4 & 3 & -2 \\ 3 & -2 & 0 & -5 \end{bmatrix}.$$

12. If

$$\begin{aligned} x_1 &= a_1r_1 + b_1r_2 + c_1r_3 & r_1 &= c_1y_1 + c_2y_2 + \dots c_ny_n \\ x_2 &= a_2r_1 + b_2r_2 + c_2r_3 & r_2 &= d_1y_1 + d_2y_2 + \dots d_ny_n \\ \vdots & & r_3 &= e_1y_1 + e_2y_2 + \dots e_ny_n, \\ x_m &= a_mr_1 + b_mr_2 + c_mr_3 \end{aligned}$$

show that

$$\begin{aligned} x_1 &= a_{11}y_1 + a_{12}y_2 + \dots a_{1n}y_n \\ x_2 &= a_{21}y_1 + a_{22}y_2 + \dots a_{2n}y_n \\ &\dots \\ x_m &= a_{m1}y_1 + a_{m2}y_2 + \dots a_{mn}y_n \end{aligned}$$

where

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ \vdots & \vdots \\ a_m & b_m \end{bmatrix} \begin{bmatrix} c_1 & c_2 & \dots & c_n \\ d_1 & d_2 & \dots & d_n \end{bmatrix}.$$

13. Show that $(AB)^t = B^tA^t$ if A is $m \times 3$ and B is $3 \times n$.

14. Show that

$$\begin{aligned} ([x \ y \ z] \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & k \end{bmatrix}) [x' \ y' \ z'] &= [x \ y \ z] \left(\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & k \end{bmatrix} \right) [x' \ y' \ z'] = \\ &= axx' + bxy' + cxz' + dyx' + eyy' + fyz' + gzx' + hzy' + kzz'. \end{aligned}$$

6. PRODUCTS OF GEOMETRICAL VECTORS

In Euclidean geometry, the rulers on each line can be calibrated, so that it makes sense to say that two line segments have the same length. The length of the geometrical vector $\vec{v} = \overrightarrow{AB}$ is defined to be the length of the line segment AB . If $\vec{v} = \overrightarrow{A'B'}$, we have $|AB| = |A'B'|$ and so the length of \vec{v} is well-defined; it is denoted by $|\vec{v}|$. We have

$$\vec{v} = \vec{0} \iff |\vec{v}| = 0.$$

A vector of length 1 is called a **unit vector**. If c is a scalar,

$$|c\vec{v}| = |c||\vec{v}|.$$

If $c = |\vec{v}| \neq 0$, the vector $c^{-1}\vec{v}$ is a unit vector. There is also a well defined **angle** θ , $0 \leq \theta \leq \pi$, between any two directed line segments with the same initial point. If $\vec{u} = \overrightarrow{AB}$ and $\vec{v} = \overrightarrow{AC}$, the angle between them is, by definition, the angle between the directed line segments with initial point A and terminal points B and C . This is independent of the choice of A . If $\theta = \pi/2$, the vectors are said to be **orthogonal** or **perpendicular**.

If \vec{u} , \vec{v} are vectors with angle θ , the law of cosines can be written in the following form:

$$|\vec{u} - \vec{v}|^2 = |\vec{u}|^2 + |\vec{v}|^2 - 2|\vec{u}||\vec{v}|\cos(\theta).$$

Thus \vec{u} , \vec{v} are orthogonal iff

$$|\vec{u} - \vec{v}|^2 = |\vec{u}|^2 + |\vec{v}|^2,$$

which is just the Pythagorean Theorem. If \vec{u} and \vec{v} are orthogonal, so are \vec{u} and $-\vec{v}$, and so

$$|\vec{u} + \vec{v}|^2 = |\vec{u}|^2 + |\vec{v}|^2.$$

If (O, I, J, K) is a frame for a rectangular coordinate system and P is a point with coordinates (x, y, z) , and

$$\vec{i} = \overrightarrow{OI}, \vec{j} = \overrightarrow{OJ}, \vec{k} = \overrightarrow{OK},$$

we have

$$\begin{aligned} |\overrightarrow{OP}|^2 &= |x\vec{i} + y\vec{j} + z\vec{k}|^2 \\ &= x^2 + |y\vec{j} + z\vec{k}|^2 \\ &= x^2 + y^2 + z^2, \end{aligned}$$

since \vec{i} , \vec{j} , \vec{k} are mutually orthogonal of length 1 and since \vec{i} is orthogonal to the vector $y\vec{j} + z\vec{k}$. It follows the the length of any vector with coordinate vector (a, b, c) is

$$(a^2 + b^2 + c^2)^{1/2}$$

and that of the line segment joining $A(a_1, b_1, c_1)$ and $B(a_2, b_2, c_2)$ is

$$|AB| = ((a_1 - a_2)^2 + (b_1 - b_2)^2 + (c_1 - c_2)^2)^{1/2}.$$

If \vec{u} , \vec{v} are two vectors with coordinate vectors (x_1, y_1, z_1) , (x_2, y_2, z_2) , we have

$$\begin{aligned} |\vec{u}||\vec{v}|\cos(\theta) &= \frac{1}{2}(|\vec{u} + \vec{v}|^2 - |\vec{u}|^2 - |\vec{v}|^2) \\ &= x_1x_2 + y_1y_2 + z_1z_2. \end{aligned}$$

We are thus led to define the *scalar (or dot) product* of the two geometric vectors \vec{u} , \vec{v} to be

$$\vec{u} \cdot \vec{v} = |\vec{u}||\vec{v}|\cos(\theta).$$

Since this definition is independent of the choice of rectangular frame, we have, in any rectangular coordinate system,

$$\vec{u} \cdot \vec{v} = [\vec{u}] \cdot [\vec{v}].$$

In particular, $|\vec{v}|^2 = \vec{v} \cdot \vec{v}$ or $|\vec{v}| = \sqrt{\vec{v} \cdot \vec{v}}$. Hence, for rectangular coordinate systems, the dot product of the two geometric vectors is the same as the dot product of their coordinate vectors. This also shows that the dot product of geometric vectors has the same algebraic properties as the dot product of numerical vectors. More precisely, if $\vec{u}, \vec{v}, \vec{w}$ are geometric vectors and a, b are scalars, we have

- (1) $\vec{u} \cdot \vec{u} \geq 0$ with equality iff $\vec{u} = (0, 0, 0)$;
- (2) $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$;
- (3) $(a\vec{u} + b\vec{v}) \cdot \vec{w} = a\vec{u} \cdot \vec{w} + b\vec{v} \cdot \vec{w}$.

If \vec{i}, \vec{j} are orthogonal unit vectors in a plane Π , then any vector \vec{u} in Π can be uniquely written in the form $x\vec{i} + y\vec{j}$. If $x_1\vec{i} + y_1\vec{j}, x_2\vec{i} + y_2\vec{j}$ are two vectors in Π , then

$$x_1\vec{i} + y_1\vec{j} \cdot x_2\vec{i} + y_2\vec{j} = x_1x_2 + y_1y_2$$

so that the $|x\vec{i} + y\vec{j}| = \sqrt{x^2 + y^2}$. The scalar $x_1x_2 + y_1y_2$ is also called the dot product of the numerical vectors $(x_1, y_1), (x_2, y_2)$. If $ax + by + c = 0$ is the equation of a line in Π in a rectangular coordinate system, the geometric vector with coordinate vector (a, b) is perpendicular to (any vector on) the line since $a(x - x_0) + b(y - y_0) = 0$ for any two points $P(x, y), A(x_0, y_0)$ on the line.

Let us apply the above to the problem of finding the perpendicular distance between a point P and a line L . Suppose that A is a point of L and \vec{v} is a direction vector for L . We are looking for a point Q on L so that

$$\vec{QP} = \vec{AP} - \vec{AQ}$$

is orthogonal to $\vec{AQ} = t\vec{v}$. If $\vec{u} = \vec{AP}$, this is equivalent to

$$(\vec{u} - t\vec{v}) \cdot \vec{v} = 0, \text{ or } \vec{u} \cdot \vec{v} = t\vec{v} \cdot \vec{v},$$

from which we get

$$t = \frac{\vec{u} \cdot \vec{v}}{\vec{v} \cdot \vec{v}} = |\vec{u}| \cos(\theta),$$

where θ is the angle between u and v . Thus $Q = A + t\vec{v}$, where t is given above. The perpendicular distance between P and L is then $|PQ|$. Note that

$$|AQ| = |t\vec{v}| = \frac{|\vec{u} \cdot \vec{v}|}{\vec{v} \cdot \vec{v}} |\vec{v}| = \frac{|\vec{u} \cdot \vec{v}|}{|\vec{v}|} = |\vec{u}| \cdot \frac{|\vec{v}|}{|\vec{v}|}$$

gives a formula for the length of the orthogonal projection AQ of AP on L . The **orthogonal projection** of \vec{u} on L is the vector

$$(\vec{u} \cdot \vec{n}) \vec{n},$$

where \vec{n} is any unit vector parallel to L ; it is independent of the choice of \vec{n} .

For example, let us use this to find the perpendicular distance between $P(1, 2, -1)$ and the line passing through the points $A(-1, 4, 5)$ and $B(5, 2, 1)$, where the coordinates are rectangular. The foot Q of the perpendicular from P to the line L has coordinates

$$(-1, 4, 5) + \frac{(2, -2, -6) \cdot (6, -2, -4)}{(6, -2, -4) \cdot (6, -2, -4)} (6, -2, -4)$$

since $[\overrightarrow{AP}] = (2, -2, -6)$ and $[\overrightarrow{AB}] = (6, -2, -4)$. Thus Q has the coordinate vector $(23/7, 18/7, 15/7)$ and so

$$|PQ| = |(-16/7, -4/7, 22/7)| = \sqrt{108}.$$

Another way to solve the above problem is to note that Q is the intersection of the plane which is perpendicular to L and passing through P . A point R lies on this plane iff

$$\overrightarrow{PR} \cdot \overrightarrow{AB} = 0.$$

If, in a rectangular coordinate system, R , P and \overrightarrow{AB} have coordinate vectors (x, y, z) , (x_1, y_1, z_1) and (a, b, c) we get

$$a(x - x_1) + b(y - y_1) + c(z - z_1) = 0$$

for the equation of the plane. This equation can also be written as

$$ax + by + cz = d$$

with $d = ax_1 + by_1 + cz_1$.

The plane in our example above has equation $6x - 2y - 4z = 6$ and the parametric equations of L are

$$x = -1 + 6t, \quad y = 4 - 2t, \quad z = 5 - 4t.$$

The line and the plane therefore meet when

$$6(-1 + 6t) - 2(4 - 2t) - 4(5 - 4t) = 6,$$

which gives $t = 5/7$ and $(23/7, 18/7, 15/7)$ as the coordinates of Q .

If, in a rectangular coordinate system, we want to find the perpendicular distance of a point $Q(x_1, y_1, z_1)$ to a plane $ax + by + cz + d = 0$ we use the fact that the line through Q and perpendicular to the given plane has the vector equation

$$(x, y, z) = (x_1, y_1, z_1) + t(a, b, c)/\sqrt{a^2 + b^2 + c^2}.$$

Here $|t| = |QP|$ since $(a, b, c)/\sqrt{a^2 + b^2 + c^2}$ is a unit vector. Finding the intersection of the line and plane, we get

$$|t| = \frac{|ax_1 + by_1 + cz_1 + d|}{\sqrt{a^2 + b^2 + c^2}}$$

as the perpendicular distance from Q to the plane. For example, the distance of $O(0, 0, 0)$ to the plane $x + y + z = 1$ is $| -1/\sqrt{3} | = 1/\sqrt{3}$.

Now consider the problem of finding the distance between two distinct lines L_1, L_2 . If the lines are parallel, find the equation of the plane passing through some point P_1 of L_1 and perpendicular to L_1 . If P_2 is the intersection of this plane with L_2 , the required distance is $|P_1P_2|$. If the lines are not parallel, find a vector \vec{n} which is perpendicular to both lines and use this vector to find the equation of the plane containing L_1 and perpendicular to \vec{n} . Since L_2 is parallel to this plane, the required distance is then just the distance from any point of L_2 to the plane.

Example. To find the distance between the non-parallel lines

$$\begin{array}{ll} x &= -1 + 2t, & x &= 2 + t, \\ y &= 2 + t, & y &= 1 - 2t, \\ z &= 3 - t, & z &= 2 + t, \end{array}$$

we use the fact that the vector

$$(1, -2, 1) \times (2, 1, -1) = (1, 3, 5)$$

is the coordinate vector of a geometric vector which is perpendicular to both lines. The given lines are parallel to any plane perpendicular to this vector. The plane

$$x + 3y + 5z = 20,$$

which passes through $(-1, 2, 3)$, contains the first line and the second line is parallel to it. The distance between the two lines is then $|2 + 6 + 10 - 20|/\sqrt{1 + 9 + 25} = 2/\sqrt{35}$, the distance from $P(2, 1, 2)$ to the plane $x + 3y + 5z = 20$.

It is natural to ask if we can define, as in the case of the dot product, a vector product $\vec{u} \times \vec{v}$ of geometrical vectors \vec{u}, \vec{v} without the use of coordinates so that, in any rectangular coordinate system,

$$[\vec{u} \times \vec{v}] = [\vec{u}] \times [\vec{v}].$$

If we choose a rectangular coordinate system with $\vec{i}, \vec{j}, \vec{k}$ as the unit vectors along the x, y and z -axes and let

$$\vec{u} = x_1\vec{i} + y_1\vec{j} + z_1\vec{k}, \quad \vec{v} = (x_2\vec{i} + y_2\vec{j} + z_2\vec{k}),$$

this would force us to define

$$\vec{u} \times \vec{v} = \begin{bmatrix} y_1 & z_1 \\ y_2 & z_2 \end{bmatrix} \vec{i} - \begin{bmatrix} x_1 & z_1 \\ x_2 & z_2 \end{bmatrix} \vec{j} + \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \end{bmatrix} \vec{k}.$$

This vector product has the following algebraic properties:

- (1) $\vec{u} \times \vec{u} = \vec{0}, \quad \vec{u} \times \vec{v} = -\vec{v} \times \vec{u};$
- (2) $(a\vec{u} + b\vec{v}) \times \vec{w} = a\vec{u} \times \vec{w} + b\vec{v} \times \vec{w};$
- (3) $\vec{w} \times (a\vec{u} + b\vec{v}) = a\vec{w} \times \vec{u} + b\vec{w} \times \vec{v};$
- (4) $\vec{u} \times \vec{v} \cdot \vec{w} = \vec{w} \times \vec{u} \cdot \vec{v} = \vec{v} \times \vec{w} \cdot \vec{u};$
- (5) $|\vec{u} \times \vec{v}|^2 + |\vec{u} \cdot \vec{v}|^2 = |\vec{u}|^2 |\vec{v}|^2.$

All of these properties follow immediately from the definition of $\vec{u} \times \vec{v}$ except for property 5 which is left as an exercise. Since

$$|\vec{u} \cdot \vec{v}| = |\vec{u}| |\vec{v}| \cos(\theta),$$

where θ is the angle between \vec{u} and \vec{v} , we obtain from property 5

$$|\vec{u} \times \vec{v}| = |\vec{u}| |\vec{v}| \sin(\theta),$$

which is the area of the parallelogram $ABCD$ with $\vec{u} = \overrightarrow{AB}$ and $\vec{v} = \overrightarrow{AC}$. Since $\vec{u} \times \vec{v} \cdot \vec{u} = 0$ and $\vec{u} \times \vec{v} \cdot \vec{v} = 0$, we also see that $\vec{u} \times \vec{v}$ is orthogonal to \vec{u} and \vec{v} . Thus we get, in the case \vec{u} and \vec{v} are not parallel,

$$\vec{u} \times \vec{v} = |\vec{u}| |\vec{v}| \sin(\theta) \vec{n},$$

where \vec{n} is a unit vector perpendicular \vec{u} and \vec{v} . But there are two possible choices for \vec{n} ; which one do we choose? To see which it is, we use the fact that $\vec{u} \times \vec{v} \cdot \vec{n} > 0$; in other words, $(\vec{u}, \vec{v}, \vec{n})$ is positively oriented with respect to the orientation defined by the given coordinate system. If we change our coordinate system, the coordinate description of the vector product does not change if we choose one with the same orientation; otherwise, the formula for the vector product in terms of coordinates must be preceded by a minus sign. So, our vector product is independent of which rectangular coordinate system we choose, as long as that coordinate system has the same orientation as the the coordinate system we started with. If we change orientation we get a different vector product, differing only in sign from the previous one.

If we designate one orientation of space as being positive and the other negative, then $\vec{u} \times \vec{v}$ is the unique vector such that

$$[\vec{u} \times \vec{v}] = [\vec{u}] \times [\vec{v}],$$

if the coordinate system is positively oriented, and

$$[\vec{u} \times \vec{v}] = -[\vec{u}] \times [\vec{v}],$$

if the coordinate system is negatively oriented.

In practice, the orientation of a coordinate system with frame (O, I, J, K) is determined by the *Right-Hand Rule*: The orientation is positive if when curling and rotating the fingers of your right hand in the direction of rotation of \vec{OI} into \vec{OJ} through the angle θ ($0 \leq \theta \leq \pi$) between them, your thumb points on the same side of the plane through O, A, B as does \vec{OC} ; otherwise, the orientation is negative. We usually choose a rectangular coordinate system in space to be positively oriented.

The magnitude of the triple scalar product $\vec{u} \times \vec{v} \cdot \vec{w}$ is the volume of the box or parallelepiped $\mathcal{B} = (A, \vec{u}, \vec{v}, \vec{w})$ whose vertices are

$$A, A + e_1 \vec{u} + e_2 \vec{v} + e_3 \vec{w},$$

where $e_1, e_2, e_3 = 0$ or 1 . Indeed, the base is a parallelogram $ABCD$ with $\vec{AB} = \vec{u}$, $\vec{AC} = \vec{v}$ and has area

$$|\vec{u}| |\vec{v}| \sin(\theta) = |\vec{u} \times \vec{v}|,$$

where θ is the angle between \vec{u} and \vec{v} . The height of the box is $|\vec{n} \cdot \vec{w}|$, where \vec{n} is a unit normal to the base. Hence the volume is

$$|\vec{u} \times \vec{v}| |\vec{n} \cdot \vec{w}| = |(\vec{u} \times \vec{v}) \cdot \vec{n}| |\vec{w}| = |\vec{u} \times \vec{v} \cdot \vec{w}|.$$

7. PARALLEL PROJECTION ON A PLANE

In this section we give the coordinate description of parallel projection onto a plane. This will prove to be very useful in drawing planar representations of objects in space.

8. QUADRATIC LOCI

We define a *quadratic locus* or *quadric surface* to be a set of points which has the property that there is a real valued function q and a coordinate system x, y, z with

$$q(P(x, y, z)) = ax^2 + bxy + cxz + dy^2 + eyz + fy^2 + gx + hy + kz + m$$

and a, b, c, d, e, f not all zero, such that P is in this set if and only if $q(P) = 0$. If we make a change of coordinates

$$\begin{aligned} x &= a + \alpha_1 x' + \alpha_2 y' + \alpha_3 z', \\ y &= b + \beta_1 x' + \beta_2 y' + \beta_3 z', \\ z &= c + \gamma_1 x' + \gamma_2 y' + \gamma_3 z', \end{aligned}$$

we have

$$q(P) = a'x'^2 + b'x'y' + c'x'z' + d'y'^2 + e'y'z' + f'z'^2 + g'x' + h'y' + k'z' + m'$$

with a', b', c', d', e', f' not all zero. For this reason, we call q a *quadratic function*. The function

$$q_0(P(x, y, z)) = ax^2 + bxy + cxz + dy^2 + eyz + fy^2$$

is called the **quadratic form** associated to q . The intersection of a quadric surface and a plane is called a plane quadratic locus. If we choose our new coordinate system so that the

given plane is the $x'y'$ -plane, the plane quadratic locus will have as equation (in the plane $z' = 0$)

$$a'x'^2 + b'x'y' + d'y'^2 + g'x' + h'y' + m' = 0.$$

The function $q(P(x', y')) = a'x'^2 + b'x'y' + d'y'^2 + g'x' + h'y' + m'$ is a quadratic function on the plane $z' = 0$. The function

$$q_0(P(x', y')) = a'x'^2 + b'x'y' + d'y'^2$$

is the associated quadratic form.

Let us illustrate this with an important example. Consider the quadratic locus $x^2 + y^2 - z^2 = 0$. The intersection of this locus with the plane $z = r$ has equation $x^2 + y^2 = r^2$. If the x, y coordinate system is rectangular, this is the equation of a circle in the plane $z = r$ with center $(0, 0, r)$ and radius $|r|$; otherwise, it is an ellipse. If $r = 0$, you get a point, namely $O = (0, 0, 0)$, which can be viewed as a degenerate ellipse or circle. If we project these curves for varying r 's onto the xy -plane, using parallel projection along the z -axis, we get a family of concentric ellipses or circles $x^2 + y^2 = r^2$ called level curves of the quadratic locus. These curves get larger as a increases in magnitude and gives us a picture of the quadratic locus; namely, that of an elliptical or circular cone with two ends or nappes. This picture is reinforced by noting that the line joining the origin to any other point on the locus lies entirely on the locus. Indeed, if $x^2 + y^2 - z^2 = 0$, then

$$(tx)^2 + (ty)^2 - (tz)^2 = t^2(x^2 + y^2 - z^2) = 0.$$

One can therefore view the locus as the surface traced out by the line joining the origin to a point $(x, y, 1)$ on the quadratic locus as this point moves around the curve $x^2 + y^2 = 1, z = 1$.

If we try to find the intersection of our cone $x^2 + y^2 - z^2$ with the plane $y = a$ we find a curve with equation $z^2 - x^2 = a^2$. If $a = 0$ we get $z^2 - x^2 = (z - x)(z + x) = 0$. Hence $z^2 - x^2 = 0$ iff either $z - x = 0$ or $z + x = 0$, from which we see that the intersection consists of the two lines $z + x = 0, z - x = 0$ in the plane $y = 0$. Let $x' = x + z, y' = y - a, z' = z - x$. This defines a new coordinate system with origin $(0, a, 0)$ and unit points

$$(1/2, a, 1/2), (0, a + 1, 0), (-1/2, a, 1/2).$$

The x' -axis is the line which has equations $x = z, y = a$ in the original coordinate system while the z' -axis has equation $z = -x, y = a$ and the y' -axis is the old y -axis. In the new coordinate system, the equation $z^2 - x^2 = a^2$ becomes $x'z' = a^2$ or $z' = a^2/x'$. Note that, if we change scale on the x' and y' -axes by setting $x'' = x'/a, y'' = y'/a$, the curve of intersection has equation $x''y'' = 1$. Such a curve is called an **hyperbola** with **asymptotes** the lines $x'' = 0$ and $y'' = 0$. From this we can get a sketch for the curve of intersection in the plane $y = a$ as seen from a point $(0, b, 0)$ with $b > 0, b^2 > a^2$. The lines $z = \pm x$ in the plane $y = a$ are the asymptotes of the hyperbola. The hyperbola does not meet the asymptotes but gets arbitrarily close as we get farther from the origin.

Now let's find the curve of intersection of our cone and the plane $z + y = 1$. This plane has parametric equations $x = s, y = 1 - t, z = t$. The parameters s, t are the coordinates of a point $P(x, y, z)$ on this plane in the coordinate system with origin $(0, 1, 0)$ and unit points $(1, 1, 0), (0, 1, 1)$. The curve of intersection has equation $s^2 + (1 - t)^2 - t^2 = 0$ which simplifies to $s^2 = 1 - 2t$. If we let $s_1 = s, t_1 = 1 - 2t$ we get the equation $t_1 = s_1^2$. Such a curve is called a **parabola**. To sketch this curve, we use the fact that s_1, t_1 are the plane coordinates of a point P on the plane $z + y = 1$ with respect to the origin $(0, 1/2, 1/2)$, which corresponds to $s_1 = t_1 = 0$, and unit points $(1, 1/2, 1/2), (0, 1, 0)$, which correspond to $s_1 = 1, t_1 = 0$ and $s_1 = 0, t_1 = 1$ respectively. Note that, if we change the orientation on the line $t_1 = 0$ by replacing s_1 by $-s_1$, the equation of the curve of intersection remains

the same; the unit point on the s_1 -axis becomes $(-1, 1/2, 1/2)$. If we intersect the cone $x^2 + y^2 = z^2$ with the plane $x - z = 0$ we get, as curve of intersection, the line $z = x, y = 0$.

The curves which are obtained in this way as intersections of a plane with the cone $x^2 - y^2 = z^2$ are called **conic sections**. A conic section is said to be **degenerate** if it consists of a single point, or one or two lines. We will show next that a non-degenerate conic section is either an ellipse or circle, a hyperbola or a parabola. More generally, we will show that the intersection of a plane with a quadric surface is either empty, two parallel lines or a conic section. In other words we will show that there is a plane coordinate system x, y so that the section, if it is not empty and does not consist of a single point or of one or two lines, has for equation one of the following

$$\begin{aligned} x^2 + y^2 &= 1 \quad (\text{ellipse or circle}); \\ xy &= 1 \quad (\text{hyperbola}); \\ y &= x^2 \quad (\text{parabola}). \end{aligned}$$

Even though, according to our definition, a quadratic locus which is empty or consists of two parallel lines is not a conic we shall also call such a locus a degenerate conic.

Problem 8.1. *Show that the intersection of the plane $z = x + y + 2$ and the cone $z^2 = 3(x^2 + y^2)$ is an ellipse or circle. If x, y, z is rectangular, show that the curve of intersection is not a circle.*

Solution. The plane $z = x + y + 2$ has parametric equations

$$x = s, \quad y = t, \quad z = s + t + 2.$$

Substituting in $z^2 = 3(x^2 + y^2)$ we get the equation

$$s^2 - st + t^2 - 2s - 2t - 2 = 0$$

which is the equation of the curve of intersection in terms of the coordinate system s, t of the plane $z = x + y + 2$ associated to the frame with origin $(0, 0, 2)$ and unit points $(1, 0, 3)$, $(0, 1, 3)$. Completing the square in s gives

$$(s - t/2 - 1)^2/6 + (t - 2)^2/8 = 1$$

after dividing both sides of the equation by 6. Setting $s' = s - t/2 - 1$, $t' = t - 2$ or, equivalently, $s = s' + t'/2 + 2$, $t = t' + 2$, gives

$$s'^2/6 + t'^2/8 = 1$$

which is the equation of the locus of intersection in the s', t' -coordinate system of the plane $z = x + y + 2$ associated to the frame

$$O'(2, 2), \quad I'(3, 2), \quad J'(5/2, 3)$$

using s, t -coordinates. If we change the scale on the s' and t' -axes by setting $s'' = s'/\sqrt{6}$, $t'' = t'/\sqrt{8}$, the equation of the locus is $s''^2 + t''^2 = 1$. The curve of intersection is either an ellipse or circle with center $(2, 2, 6)$. The points of this curve corresponding to $s'' = 1, t'' = 0$ and $s'' = 0, t'' = 0$ are

$$(\sqrt{6} + 2, 2, \sqrt{6} + 6) \text{ and } (\sqrt{2} + 2, 2\sqrt{2} + 2, 3\sqrt{2} + 6).$$

If x, y, z is a rectangular coordinate system, the distance of the first point from the center is $2\sqrt{3}$ while the distance of the second from the center is $\sqrt{26}$. Since these two distances are not the same, the curve cannot be a circle.

8.1. **Exercises.** 1. Find the curve of intersection of the cone $x^2 + y^2 = z^2$ with the plane $\Pi : 2x + z = 1$ and find a coordinate system for the plane Π so that this conic is in standard form. If x, y, z are rectangular coordinates, find a rectangular coordinate system for Π so that the conic is in standard form.

2. Sketch the quadric surface $x^2 + y^2 = z$ by finding its level curves in each coordinate plane.

8.2. **Affine Classification of Conics.** Consider the plane quadratic locus defined by the equation $q(P) = 0$ where q is the quadratic function

$$q(P(x, y)) = ax^2 + bxy + cy^2 + dx + ey + f$$

with a, b, c not all zero.

Theorem 8.1. *The quadratic locus $q(P) = 0$ is a conic.*

Proof. We first consider the case $a \neq 0$ and write $q(P(x, y))$ as a polynomial in x

$$q(P) = ax^2 + (d + by)x + cy^2 + ey + f,$$

and then complete the square in x to get

$$\begin{aligned} q(P) &= a(x + (d + by)/2a)^2 - (d + by)^2/4a + cy^2 + ey + f \\ &= a(x + (b/2a)y + d/2a)^2 + ((4ac - b^2)/4a)y^2 \\ &\quad + (e - bd/2a)y + f - d^2/4a \\ &= a'x'^2 + c'y'^2 + e'y' + f' \end{aligned}$$

where $x' = x + (b/2a)y + d/2a$, $y' = y$, $a' = a$, $c' = (4ac - b^2)/4a$, $e' = e - bd/2a$, $f' = f - d^2/4a$. The pair (x', y') is the coordinate vector of the point $P(x, y)$ in the plane coordinate system with origin $O'(-d/2a, 0)$ and unit point $I'(1 - d/2a, 0)$ on the x' -axis, and unit point $J'(-(b + d)/2a, 1)$ on the y' -axis. Note that the x' -axis is the line $y' = 0$, the x -axis since $y' = y$, while the y' -axis is the line $x' = 0$, the line $x + (b/2a)y + d/2a$. In this new coordinate system the equation of our locus is simpler: the terms $b'x'y'$ and $d'x'$ have disappeared.

Suppose that $c' \neq 0$. Then, completing the square in y' , we get

$$\begin{aligned} q(P) &= a'x'^2 + c'(y' + e'/2c')^2 + f' - e'^2/4c' \\ &= a''x''^2 + c''y''^2 + f'' \end{aligned}$$

where $x'' = x'$, $y'' = y' + e'/2c'$, $a'' = a'$, $c'' = c'$, $f'' = f' - e'^2/4c'$. The pair (x'', y'') is the coordinate vector of the point $P(x, y)$ in the plane coordinate system with origin $O''((be' - 2c'd)/4a'c', -e'/2c')$ and unit point $I''(1 - (2c'd - be')/4a, -e'/2c')$ on the x'' -axis and unit point $J''((be' - 2c'd + 2b)/4a, 1 - e'/2c')$ on the y'' -axis. Note that the x'' -axis is the line $y = -e'/2c'$ while the y'' -axis is the line $x + (b/2a)y + d/2a = 0$. If $f'' = 0$, locus consists either of a single point or two intersecting lines according as $a''c'' = (4ac - b^2)/4$ is greater or less than zero. If $f'' \neq 0$, then

$$q(x, y) = f''(\epsilon_1 x'''^2 + \epsilon_2 y'''^2 + \epsilon_3),$$

where $x''' = |a''/f''|^{1/2}$, $y''' = |c''/f''|^{1/2}$, $\epsilon_i = \pm 1$, which shows that the locus is either empty, an ellipse or a hyperbola. The first case happens when $4ac - b^2 > 0$, $f'' > 0$, the second when $4ac - b^2 > 0$, $f'' < 0$ and the third when $4ac - b^2 < 0$. If $c' = 0$, $e' \neq 0$, then $q(x, y) = a'x'^2 + e'y' + f' = a'(x''^2 + \epsilon y'')$, where $x'' = x'$, $y'' = e'y' + f'$, $\epsilon = \pm 1$. This shows that the locus is a parabola. If $c' = e' = 0$, the locus is either empty, a single line or two parallel lines. Note that $c' = 0$ iff $4ac - b^2 = 0$.

If $a = 0$ and $c \neq 0$, we can interchange the roles of x and y and proceed as above. If $a = c = 0$, then

$$q(P) = (x + e/b)(by + d) + f - ed/b$$

. If we change coordinates by setting $x' = x + e/b$, $y' = by + d$, we see that our locus is either a hyperbola or a pair of intersecting lines. **Q.E.D.**

Corollary 8.1. *There is a plane coordinate system x, y such that*

$$\begin{aligned} q(P) &= \mu(\epsilon_1 x^2 + \epsilon_2 y^2 + \epsilon_3) \\ &\text{or} \\ &= \mu(\epsilon_1 x^2 - y) \end{aligned}$$

where $\mu > 0$, $\epsilon_i = 0, \pm 1$ and $\epsilon_1 \neq 0$.

Corollary 8.2. *The quadratic function*

$$q(x, y) = ax^2 + bxy + cy^2 + dx + ey + f$$

has a minimum value if $a > 0$, $4ac - b^2 > 0$ and a maximum value if $a < 0$, $4ac - b^2 > 0$.

Corollary 8.3. *Let $\Delta = b^2 - 4ac$. If $\Delta > 0$, the locus $q(P) = 0$ is either a hyperbola or two lines meeting in a point; if $\Delta < 0$, the locus is either an ellipse, a single point or empty; if $\Delta = 0$, the locus is either a parabola, two parallel lines, a single line or empty.*

The quantity $\Delta = b^2 - 4ac$ is called the **discriminant** of q with respect to the given coordinate system. If Δ' is the discriminant of q in some other coordinate system x', y' then $\Delta = k^2 \Delta'$ with $k \neq 0$. To see this we use the easily verified fact that

$$ax^2 + bxy + cy^2 = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

The matrix

$$M = \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix}$$

is called the **matrix** of q in the xy -coordinate system. It is uniquely determined by q . Note that $\det(M) = -\Delta/4$.

Since a translation of coordinate frame does not change the discriminant, we can assume that the change of coordinates is of the form $X = PX'$ with

$$X = \begin{bmatrix} x & y \end{bmatrix}, \quad X' = \begin{bmatrix} x' & y' \end{bmatrix}.$$

Then

$$ax^2 + bxy + cy^2 = X^t M X = X'^t P^t M P X'$$

since $X^t = X'^t P^t$. It follows that the matrix M' of q in the $x'y'$ -coordinate system is $P^t M P$. Taking determinants, we get

$$\det(M') = \det(P^t) \det(M) \det(P) = \det(P)^2 \det(M).$$

Since $\Delta' = -4 \det(M')$ we get $\Delta = k^2 \Delta'$ with $k = \det(P)$.

Problem 8.2. *Identify the plane quadratic locus*

$$4x^2 + 8xy + 3y^2 + 16x + 18y + 14 = 0.$$

Solution. The discriminant is 16 so the locus is either a hyperbola or two intersecting lines. Completing the square in x , we get $4(x + y + 2)^2 - y^2 + 2y - 2 = 0$. Completing the square in y , we get $4(x + y + 2)^2 - (y - 1)^2 = 1$. Let $x' = 2x + 2y + 4$, $y' = y - 1$. Then (x', y') are the coordinates of a point $P(x, y)$ in the plane coordinate system with origin $O'(-3, 1)$ and unit points $A(-5/2, 1)$ on the x' -axis, which is the line $y = 1$, and $B(-4, 2)$ on the

y' -axis, which is the line $x + y + 2 = 0$. In this coordinate system, the locus has equation $x'^2 - y'^2 = 1$, which is the equation of a hyperbola with asymptotes the lines $x' - y' = \pm 1$. These two lines have equations $2x + y + 5 = 0$, $2x + 3y = 3 = 0$ in the original coordinate system.

Problem 8.3. Find the maximum value of the function $q(x, y) = -x^2 - 2xy - 2y^2 + 2x - y + 1$. At what point is this value attained?

Solution. Completing the square in x we get

$$q(x, y) = -(x + y - 1)^2 - y^2 - 3y + 2.$$

Completing the square in y , we get

$$q(x, y) = -(x + y - 1)^2 - (y - 3/2)^2 + 13/4.$$

Thus q has $13/4$ as a maximum value and attains this value for those (x, y) with $x + y - 1 = 0$ and $y - 3/2 = 0$. The only solution to these equations is $x = -1/2$, $y = 3/2$.

8.3. Exercises. 1. Identify and sketch each of the following plane quadratic loci:

- (a): $x^2 + 6xy + y^2 + 2x - y + 1 = 0$;
- (b): $4x^2 + 12xy + 9y^2 + 4x + 6y + 1 = 0$;
- (c): $3x^2 + 14xy + 8y^2 - 2x + 12y = 8$;
- (d): $5x^2 + 6xy + 5y^2 + 10x + 4y + 9 = 0$;
- (e): $9x^2 + 12xy + 4y^2 + 15x + 11y + 8 = 0$.

2. Find the minimum value of the function $q(x, y) = 4x^2 + 3xy + y^2 + 3x - 2y + 2$. At what point is this value attained?

3. Sketch the quadric surface $x^2 + y^2 = z$.

4. Identify and sketch the intersection of the quadric surface $x^2 + y^2 = z$ and the plane $2x + y + z = 10$.

8.4. Affine Properties of Conics. We now use our classification theorem to obtain some important properties of conics which do not need a Euclidean distance for their proofs.

Theorem 8.2. Let C be a non-degenerate conic and let L_0 be a line which intersects C in two distinct points (such a line is called a secant of C). For each secant L of C parallel to L_0 meeting C in two points A_L , B_L , let P_L be the mid-point of the line segment joining A_L and B_L . Then the points P_L all lie on a straight line.

Proof. Case I: C is an ellipse. We choose our coordinates x, y so that C has the equation $x^2 + y^2 = 1$. If L_0 is parallel to the y -axis, then L has equation $x = c$ and intersects C in two points $(c, \sqrt{1 - c^2})$, $(c, -\sqrt{1 - c^2})$ if $-1 < c < 1$. The mid-point of the line segment joining these two points is $(c, 0)$ which lies on the x -axis for any c . If L_0 is not parallel to the y -axis, then L_0 has for equation $y = mx + b_0$ and so L has for equation $y = mx + b$. Substituting $y = mx + b$ in the equation $x^2 + y^2 = 1$, we get the quadratic equation

$$(1 + m^2)x^2 + 2mbx + b^2 - 1 = 0.$$

The discriminant of this quadratic equation is $4(1 + m^2 - b^2)$ which is greater than 0 for $|b| < 1 + m^2$, in which case the quadratic has the distinct roots

$$x_1 = \frac{-mb \pm \sqrt{1 + m^2 - b^2}}{1 + m^2}, \quad x_2 = \frac{-mb \pm \sqrt{1 + m^2 - b^2}}{1 + m^2}.$$

In this case, the line L meets C in the points (x_1, y_1) , (x_2, y_2) where $y_i = mx_i + b$. Since $x_1 + x_2 = -2mb/(1 + m^2)$, the mid-point of the line segment joining these two points is

$$(-mb/(1 + m^2), -m^2b/(1 + m^2) + b) = (-mb/(1 + m^2), b/(1 + m^2)),$$

which lies on the line $x = -my$. So the mid-points all lie on the line $x + my = 0$ which passes through the origin. Note the the line $y = mx$ also passes through the origin and is the locus of mid-points of line segments for the lines parallel to $mx + y = 0$.

Case II: C is a hyperbola. We choose our coordinates x, y so that C has the equation $xy = 1$. The line L_0 must have the equation $y = mx + b_0$ with $m \neq 0$ since any line parallel to the coordinate axes (which are the asymptotes of C) meet C in at most one point. Substituting $y = mx + b$ in $xy = 1$, we get the quadratic equation $mx^2 + bx - 1 = 0$ whose discriminant is $b^2 + 4m$. Thus the quadratic equation has two distinct roots if $b^2 > -4m$, which is always the case if $m > 0$. If x_1, x_2 are these roots we have $x_1 + x_2 = -b/m$. Since L meets C in the points (x_i, y_i) with $y_i = mx_i + b$, the mid-point of the line segment joining these points is $(-bm/2, b/2)$ which lies on the line $x + my = 0$. So the locus of mid-points is $x + my = 0$. Note that, if $m \neq 0$, the line $y = mx$ is the locus of mid-points for the lines parallel to $x + my = 0$.

Case III: C is a parabola. We choose our coordinates so that the equation of C is $y = x^2$. Then L_0 must have equation $y = mx + b_0$ since the lines parallel to the y -axis meet C in exactly one point. Substituting $y = mx + b$ in $y = x^2$, we get $x^2 - mx - b = 0$ which has distinct roots when $m^2 + 4b > 0$, i.e., when $b > -m^2/4$. If x_1, x_2 are these roots, then $x_1 + x_2 = m$. If (x_1, y_1) , (x_2, y_2) are the points of intersection of L with C , the mid-point of the line segment joining these two points lies on the line $x = m/2$. **Q.E.D.**

The locus of the mid-points P_L is called a *diameter* of the conic, more precisely, the diameter conjugate to the secant L_0 . If the conic is an ellipse or hyperbola, the proof of the above theorem shows that the diameters all meet in one point. This point is called the *center* of the conic. Any line which passes through the center, except for the asymptotes in the case of a hyperbola, is a diameter. This implies that, in any coordinate system x, y in which the center is at the origin, a point $P(x, y)$ lies on the conic if $P(-x, -y)$ lies on the conic. In such a coordinate system, the equation of the conic must therefore be of the form

$$ax^2 + bxy + cy^2 + f = 0.$$

The proof of this is left as an exercise for the reader. Because they have centers, ellipses and hyperbolas are called *central conics*.

Since any diameter of a central conic is either a secant or is parallel to one, any diameter has a conjugate diameter. Moreover, the above theorem shows that if L' is the diameter conjugate to the diameter L then the conjugate of L' is L . If x, y are rulers on lines which are conjugate diameters of an ellipse or hyperbola, the equation of the conic has the form

$$ax^2 + by^2 + f = 0.$$

This is because $P(\pm x, \pm y)$ is a point on the conic if $P(x, y)$ is. The proof of this is again left as an exercise for the reader. The coordinate axes $x = 0$ and $y = 0$ are therefore axes of symmetry for the conic. We shall show later that, if a central conic is not a circle, there is exactly one pair of conjugate diameters which are perpendicular. The intersections these axes with the conic are called the **vertices** of the conic.

In the case of an ellipse, there are four vertices which are the endpoints of two line segments on the axes. The axis having the longer line segment is called the **principal** or **major axis** and the other is called the **minor axis**. A hyperbola has two vertices and the axis joining these two vertices is called the **principal axis** of the hyperbola.

In the case of a parabola, the diameters are parallel to each other. Any line parallel to a diameter is again a diameter. We shall show later that there is a unique direction so that the diameter conjugate to a secant with this direction is perpendicular to the secant. This unique diameter is called the **principal axis** of the parabola. The intersection of this axis with the parabola is called the **vertex** of the parabola. If we take a rectangular coordinate system with y -axis the principal axis of the parabola and x -axis passing through the vertex, the equation of the conic has the form $y = cx^2$.

The proof of the above theorem also yields the following result about the intersection of a line and a conic.

Theorem 8.3. *A line meets a non-degenerate conic in at most two points.*

To find the intersection of a line and a non-degenerate conic $q = 0$ with

$$q = ax^2 + bxy + cy^2 + dx + ey + f,$$

we write the line in parametric form $x = x_0 + \alpha t$, $y = y_0 + \beta t$. Then

$$q(x_0 + \alpha t, y_0 + \beta t) = mt^2 + nt + p$$

with $p = q(x_0, y_0)$, $m = a\alpha^2 + b\alpha\beta + c\beta^2$ and

$$n = (bx_0 + 2cy_0 + e)\alpha + (2ax_0 + by_0 + d)\beta.$$

This gives another proof of the above theorem since m, n, p cannot all be 0 as the non-degenerate conic $q = 0$ cannot contain a line. Now suppose that (x_0, y_0) lies on the conic. Then

$$q(x_0 + \alpha t, y_0 + \beta t) = mt^2 + nt$$

with m, n not both zero. If $n \neq 0$ the line crosses the conic at (x_0, y_0) since the sign of $mt^2 + nt$ is different for $t > 0$ and $t < 0$ if $|t|$ is sufficiently small. If $n = 0$ then the line meets the conic but does not cross it since the sign of mt^2 is constant for all $t \neq 0$. In this case the line is said to be **tangent** to the conic at the point (x_0, y_0) . Since $2ax_0 + by_0 + d$ and $bx_0 + 2cy_0 + e$ are not both zero (otherwise, any line through (x_0, y_0) would meet the conic in a single point), the conic has a unique tangent line with equation

$$(2ax_0 + by_0 + d)(x - x_0) + (bx_0 + 2cy_0 + e)(y - y_0) = 0.$$

Problem 8.4. *Find the equation of the tangent line to the parabola $y^2 = 4x$ at the point $(1, 2)$.*

Solution. Writing the equation of the conic in the form $y^2 - 4x = 0$, the equation of the tangent line is $-4(x - 1) + 4(y - 2) = 0$ or $y = x + 1$.

To see how two conics intersect we will need to use a parametrization for one of these curves. Let C be a non-degenerate conic and let x, y be a coordinate system so that C has equation $y = x^2$ or $y = 1/x$ or $x^2 + y^2 = 1$. Note that, in the latter case, it is possible to choose our coordinate system so that a given point A of C has coordinates $(-1, 0)$. Therefore, if C is a parabola, the point $P(x, y)$ is on C iff $x = t, y = t^2$ for some real number t . If C is a hyperbola, we get the rational parametrization $x = t, y = 1/t$ with $t \neq 0$. The case of an ellipse is a little more complicated—consider the line $y = m(x + 1)$ which passes through $A(-1, 0)$. This line intersects C in another point $P(x, y)$ where

$$(1 + m^2)x^2 + 2m^2x + m^2 - 1 = 0.$$

Since one root of this quadric is -1 and the sum of the roots is $-2m/(1+m^2)$, we get $x = 1 - 2m^2/(1+m^2) = (1-m^2)/(1+m^2)$, which gives us the rational parametrization

$$x = \frac{1-t^2}{1+t^2}, \quad y = \frac{2t}{1+t^2}$$

for the points of C other than A . Note that the numerator and denominators in this parametrization are polynomials in m of degree at most 2.

If C and C' are distinct non-degenerate conics, then (after possibly interchanging C and C') there is a point A of C which does not lie on C' . We choose a coordinate system x, y so that the equation for C is as above with the exception point A having coordinates $(-1, 0)$ in the case C is an ellipse. The equation of C' is

$$ax^2 + bxy + cy^2 + dx + ey + f = 0.$$

Using the above parametric representation of a point $P(x, y)$ on C , substituting it in the equation for C' , and clearing denominators by multiplying by $1+t^2$, we get a polynomial equation in t of degree at most 4 whose roots give precisely the points $P(x, y)$ in the intersection of C and C' . Since this polynomial is not identically zero (otherwise C is a subset of C' , contradicting the fact that A is not a point of C'), and hence has at most four roots, we see that there are at most four points in the intersection. We thus obtain the following result.

Theorem 8.4. *Two distinct non-degenerate conics intersect in at most four points.*

The same result holds if one conic is degenerate or if the two conics are degenerate and do not have a line in common.

Suppose $P_i(x_i, y_i)$ are five distinct points in a plane with coordinate system x, y and we want to find a conic which passes through each of these points. Since the general equation of this conic is

$$ax^2 + bxy + cy^2 + dx + ey + f = 0,$$

we can substitute the coordinates of the points P_i in it to give a system of five linear homogeneous equations in the six unknowns a, b, c, d, e, f . But a homogeneous system of linear equations always has a non-trivial solution, one in which not all the unknowns are zero. Indeed, using Gauss-Jordan elimination, there will be at least one non-pivot variable at the end of the elimination process. So our system has a non-trivial solution. If $a = b = c = 0$ in this solution then $(dx + ey + f)^2 = 0$ is a conic passing through the five points.

If four or more of the points lie on a line, there are many solutions by degenerate conics. If three of the points lie on a line and the other two do not lie on this line, there is a unique conic which is the union of the above line and the line joining the other two points. If no three of the P_i are collinear there is a unique conic passing through the five points as any solution of the system of equations gives a non-degenerate conic and two distinct non-degenerate conics passing through five distinct points is impossible. We thus obtain the following result:

Theorem 8.5. *Given five distinct points, there is a conic which passes through them. If no four of the five points lie on a line, this conic is unique. If no three of the five points lie on a line this conic is non-degenerate.*

We now give a different proof of the last part of this theorem which will us give a more precise description of the equations of the conics which pass through five given points, no three of which are collinear. Let \mathcal{P} be the set of conics which pass through four of the given points, say A, B, C, D . After possibly permuting these points, we can assume that

the line joining A, B meets the line joining C, D in a unique point $O \neq A, B, C, D$. Choose O, A, C as our coordinate frame. Then, with respect to this frame, the points A, B, C, D have coordinate vectors $(1, 0), (r, 0), (0, 1), (0, s)$. If $q(P) = 0$ is the equation of a conic, we

$$q(P(x, y)) = ax^2 + bxy + cy^2 + dx + ey + f.$$

If this conic is in our family, we have $q(A) = q(B) = q(C) = q(D) = 0$ and so

$$\begin{aligned} a + d + f &= 0, \\ r^2a + rd + f &= 0, \\ c + e + f &= 0, \\ s^2c + se + f &= 0. \end{aligned}$$

Solving these equations, we find $a = f/r, d = -f(r+1)/r, c = f/s, e = -f(s+1)/s$ and $q(P) = fq_1(P) + bq_2(P)$, where

$$\begin{aligned} q_1(P(x, y)) &= x^2/r + y^2/s - rx/(r+1) - sy/(s+1) + 1, \\ q_2(P(x, y)) &= xy \end{aligned}$$

Moreover, the conics $q_1 = 0$ and $q_2 = 0$ lie in \mathcal{P} and intersect precisely in the points A, B, C, D . If we want to have $q = 0$ to pass through a fifth point $E \neq A, B, C, D$, it suffices to choose f, b so that $f q_1(E) + b q_2(E) = 0$. This is possible since $q_1(E)$ and $q_2(E)$ are not both zero. Since the pair (f, b) is determined up to multiplication by a scalar, there is a unique conic passing through the five points A, B, C, D, E . Moreover, the equation of this conic is uniquely determined up to multiplication by a constant.

Corollary 8.4. *A conic having at least two distinct points has, up to multiplication by a constant, a unique equation.*

The following result, which comes out of the above proof, gives a simple way to find the equation of a conic passing through five points.

Theorem 8.6. *Let C be a conic passing through four distinct points A, B, C, D no three of which are collinear. If $q_1 = 0$ and $q_2 = 0$ are distinct conics passing through A, B, C, D then there are scalars a, b such that C has the equation $aq_1 + bq_2 = 0$.*

Proof. Let E be any point of C which is different from A, B, C, D . Then, since $q_1 = 0$ and $q_2 = 0$ are distinct and have no line in common, $q_1(E)$ and $q_2(E)$ cannot be both zero. We can therefore choose scalars a, b not both zero with $aq_1(E) + bq_2(E) = 0$. Hence $aq_1 + bq_2 = 0$ is a conic passing through A, B, C, D, E and so must be C since such a conic is unique. **Q.E.D.**

Problem 8.5. *Find the equation of the conic which passes through the points*

$$(-1, 2), (1, 3), (4, 9), (2, 1), (3, -1).$$

Solution. The equation of the line joining $(-1, 2)$ and $(2, 1)$ is $x + 3y = 5$. This line does not pass through any of the other points. The equation of the line through $(1, 3)$ and $(4, 9)$ is $2x - y = -1$ and also does not pass through any of the other points. It follows that no three of the first four points are collinear. If

$$q_1 = (x + 3y - 5)(2x - y + 1) = 2x^2 + 5xy - y^2 - 9x + 8y - 5$$

then $q_1 = 0$ is the equation of a conic passing through the first four points. The line joining $(1, 3)$ and $(-1, 2)$ has equation $x - 2y = -5$ and the line joining $(4, 9)$ and $(2, 1)$ has equation $4x - y = 7$. If

$$q_2 = (x - 2y + 5)(4x - y - 7) = 4x^2 - 9xy + 2y^2 + 13x + 9y - 35$$

then $q_2 = 0$ is another conic passing through the first four points which is distinct from the conic $q_1 = 0$. Let $q = aq_1 + bq_2$ where a, b are scalars, not both zero, with $aq(3, -1) + bq_2(3, -1) = 0$. Since $q_1(3, -1) = -40$ and $q_2(3, -1) = 60$ we have $2a = 3b$. Choosing $b = 2$ gives $a = 3$ and so

$$3q_1 + 2q_2 = 14x^2 + 3xy + y^2 - x + 42 - 85 = 0$$

is the equation of the conic.

8.5. Exercises. 1. Find the diameter of the ellipse $x^2 + 2y^2 = 5$ which is conjugate to the diameter $y = 2x$. Find the equation of the tangent to the ellipse at the point $(1, 2)$. Show that it is parallel to the diameter found above. Can you explain why this must be so?

2. Find the diameter of the conic $3x^2 - y^2 = -1$ which is conjugate to the diameter $y = 2x$. Show that the tangent to this hyperbola at the point $(1, 2)$ is parallel to the diameter found above.

3. Find the diameter of the parabola $y = 2x^2 + 3x + 1$ which is conjugate to the secant $y = 2x + 1$. Show that the tangent to this parabola at the point of intersection with the above diameter is parallel to the secant $y = 2x + 1$.

4. Find the diameter of the conic $ax^2 + by^2 = c$ which is conjugate to the diameter $y = mx$. We assume that $abc \neq 0$.

5. Find the diameter of the parabola $y = ax^2 + bx + c$ which is conjugate to the secant $y = mx + d$.

6. Find the equation of the conic passing through the points $(3, 2)$, $(3, 5)$, $(5, 3)$, $(2, 3)$, $(2, 2)$.

7. Find the points of intersection of the two conics

$$\begin{aligned} 2x^2 + y^2 &= 3 \\ xy &= 1. \end{aligned}$$

8. Find the equation of the conic which passes through $A(1, 2)$ and the points of intersection of the two conics in the previous problem. Sketch all three curves together.

9. Find the equation of a conic which passes through $A(1, 1)$ and the points of intersection of the conics

$$\begin{aligned} x^2 - 3xy + 2y^2 + 1 &= 0 \\ 3x^2 - 10xy + 9y^2 &= 3. \end{aligned}$$

Is this conic unique?

8.6. Euclidean Classification of Conics. In this section we are given a rectangular coordinate system x, y in a plane Π and a quadratic locus

$$q = Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0.$$

If $B = 0$ then we, after completing the square in x, y separately and translating the axes, we obtain a new rectangular coordinate system x', y' in which the equation of the locus has

one of the following standard forms:

$$\begin{aligned}
x'^2/a^2 + y'^2/b^2 &= 1 \text{ (ellipse),} \\
x'^2/a^2 + y'^2/b^2 &= 0 \text{ (single point),} \\
x'^2/a^2 + y'^2/b^2 &= -1 \text{ (empty locus),} \\
x'^2/a^2 - y'^2/b^2 &= 1 \text{ (hyperbola),} \\
x'^2/a^2 - y'^2/b^2 &= -1 \text{ (hyperbola),} \\
x'^2/a^2 - y'^2/b^2 &= 0 \text{ (two intersecting lines),} \\
x'^2 - ay' &= 0 \text{ (parabola),} \\
y'^2 - ax' &= 0 \text{ (parabola),} \\
x'^2/a^2 &= 1 \text{ (two parallel lines),} \\
x'^2 &= 0 \text{ (a single line),} \\
x'^2/a^2 &= -1 \text{ (empty locus),} \\
y'^2/a^2 &= 1 \text{ (two parallel lines),} \\
y'^2 &= 0 \text{ (a single line),} \\
y'^2/a^2 &= -1 \text{ (empty locus).}
\end{aligned}$$

If $B \neq 0$, we want to find a rectangular coordinate system x', y' such that, when q is expressed in terms of the new coordinates, the coefficient of $x'y'$ is zero. The equations

$$\begin{aligned}
x &= a + \alpha_1 x' + \alpha_2 y', \\
y &= b + \beta_1 x' + \beta_2 y'
\end{aligned}$$

are the equations for a change of rectangular coordinates iff the vectors

$$\begin{aligned}
\vec{i}' &= \alpha_1 \vec{i} + \beta_1 \vec{j}, \\
\vec{j}' &= \alpha_2 \vec{i} + \beta_2 \vec{j}
\end{aligned}$$

are orthogonal and of unit length. This is equivalent to

$$\alpha_1^2 + \beta_1^2 = \alpha_2^2 + \beta_2^2 = 1 \text{ and } \alpha_1\beta_1 + \alpha_2\beta_2 = 0$$

or to $PP^t = 1$, where P is the transition matrix. Thus P is the transition matrix for a change of rectangular coordinates iff $P^t = P^{-1}$. Such a matrix is called an **orthogonal matrix**. In this case, if θ is the oriented angle between \vec{i} and \vec{i}' , we have

$$\begin{aligned}
\vec{i}' &= \cos(\theta)\vec{i} + \sin(\theta)\vec{j}, \\
\vec{j}' &= -\sin(\theta)\vec{i} + \cos(\theta)\vec{j}
\end{aligned}$$

if the orientation is not changed and

$$\begin{aligned}
\vec{i}' &= \cos(\theta)\vec{i} + \sin(\theta)\vec{j}, \\
\vec{j}' &= -\sin(\theta)\vec{i} + \cos(\theta)\vec{j}
\end{aligned}$$

if the orientation is reversed. In particular, $\det(P) = \pm 1$ which one could also see from

$$1 = \det(PP^t) = \det(P)\det(P^t) = \det(P)^2.$$

Thus, for a rectangular change of coordinates in which the orientation is preserved, we have

$$\begin{aligned}
x &= a + \cos(\theta)x' - \sin(\theta)y', \\
y &= b + \sin(\theta)x' + \cos(\theta)y'
\end{aligned}$$

and we see that the new axes are obtained from the old ones by a translation and then a rotation through an angle θ .

If P and P' are the transition matrices for a rotation of axes through θ and θ' respectively, then PP' is the transition matrix for a rotation of axes through $\theta + \theta'$. This is equivalent to

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} \cos(\theta') & -\sin(\theta') \\ \sin(\theta') & \cos(\theta') \end{bmatrix} = \begin{bmatrix} \cos(\theta + \theta') & -\sin(\theta + \theta') \\ \sin(\theta + \theta') & \cos(\theta + \theta') \end{bmatrix}.$$

which, after multiplying the matrices and equating corresponding entries, is seen to be equivalent to

$$\begin{aligned} \cos(\theta + \theta') &= \cos(\theta) \cos(\theta') - \sin(\theta) \sin(\theta'), \\ \sin(\theta + \theta') &= \sin(\theta) \cos(\theta') + \cos(\theta) \sin(\theta'). \end{aligned}$$

These two identities are the **addition laws** for the sine and cosine functions. If we set $\theta = \theta'$, we get

$$\begin{aligned} \cos(2\theta) &= \cos^2(\theta) - \sin^2(\theta), \\ \sin(2\theta) &= 2 \sin(\theta) \cos(\theta). \end{aligned}$$

This yields

$$\tan(2\theta) = \frac{\sin(2\theta)}{\cos(2\theta)} = \frac{2 \tan(\theta)}{1 - \tan^2(\theta)}.$$

If we now express q in terms of the $x'y'$ -coordinates, the coefficients A' , B' , C' of x'^2 , $x'y'$, y'^2 respectively are

$$\begin{aligned} A' &= A \cos^2(\theta) + B \cos(\theta) \sin(\theta) + C \sin^2(\theta), \\ B' &= -2A \sin(\theta) \cos(\theta) + B(\cos^2(\theta) - \sin^2(\theta)) + 2C \sin(\theta) \cos(\theta), \\ C' &= A \sin^2(\theta) - B \cos(\theta) \sin(\theta) + \cos^2(\theta). \end{aligned}$$

Note that

$$A' + C' = A + C.$$

Using the identities for $\sin(2\theta)$ and $\cos(2\theta)$, we get

$$B' = (C - A) \sin(2\theta) + B \cos(2\theta).$$

If $A = C$ then $B' = 0$ if $\theta = \pi/4$ or 45 degrees. If $A \neq C$, we have $B' = 0$ if and only if

$$\tan(2\theta) = \sin(2\theta) / \cos(\theta) = \frac{B}{A - C}.$$

If we set $m = \tan(\theta)$, this is the same as

$$\frac{B}{A - C} = \frac{2m}{1 - m^2}$$

or, equivalently,

$$Bm^2 + 2(A - C)m - B = 0.$$

If $B \neq 0$, this equation has two real roots m_1, m_2 with $m_1 m_2 = -1$. These are the slopes of the new coordinate axes. Since m determines θ up to a multiple of π we choose θ such that $|\theta| < \pi/2$. We then have

$$\sin(\theta) = \frac{m}{\sqrt{1 + m^2}}, \quad \cos(\theta) = \frac{1}{\sqrt{1 + m^2}}$$

where m can be either of m_1, m_2 .

Example. Let x, y be a rectangular coordinate system and consider the conic

$$2x^2 + 4xy + 5y^2 + 2x + y = 9.$$

The discriminant $\Delta = 14 - 40 = -24$ and so the locus is of elliptic type (ellipse, circle, point or empty). To simplify the equation of the locus we introduce a new rectangular coordinate system x', y' where the slopes of the coordinate axes are the roots of the equation

$$4m^2 + 6m - 4 = 0.$$

The roots of this equation are $m = 2$ and $m = -1/2$. If we choose $m = -1/2$, we have

$$\sin(\theta) = -1/\sqrt{5}, \quad \cos(\theta) = 2/\sqrt{5}$$

and the transformation equations are

$$\begin{aligned} x &= 2x'/\sqrt{5} + y'/\sqrt{5}, \\ y &= -x'/\sqrt{5} + 2y'/\sqrt{5}. \end{aligned}$$

Substituting this in the equation of the conic, we get

$$\frac{2}{\sqrt{5}}(2x' + y')^2 + \frac{4}{\sqrt{5}}(2x' + y')(-x' + 2y') + (-x' + 2y')^2 + \frac{2}{\sqrt{5}}(2x' + y') + \frac{1}{\sqrt{5}}(-x' + 2y') = 9$$

which, on simplifying, becomes

$$x'^2 + 6y'^2 + \frac{3}{\sqrt{5}}x' + \frac{4}{\sqrt{5}}y' = 9.$$

Completing the square in x' , we get

$$(x' + 3/2\sqrt{5})^2 + 6(y' + 1/3\sqrt{5})^2 = 9 + 9/20 + 2/15 = 115/12.$$

Setting $x'' = x' + 3/2\sqrt{5}$, $y'' = y' + 1/3\sqrt{5}$ and dividing both sides by $115/12$, we get

$$\frac{x''^2}{115/12} + \frac{y''^2}{115/72} = 1$$

which is the standard equation of an ellipse with major axis the line $y'' = 0$ and minor axis the line $x'' = 0$. Since

$$\begin{aligned} x &= -2/3 + 2x''/\sqrt{5} + y''/\sqrt{5}, \\ y &= 1/6 - x''/\sqrt{5} + 2y''/\sqrt{5}, \end{aligned}$$

and

$$\begin{aligned} x'' &= 3/2\sqrt{5} + 2x/\sqrt{5} - y/\sqrt{5}, \\ y'' &= 1/3\sqrt{5} + x/\sqrt{5} + 2y/\sqrt{5}, \end{aligned}$$

the major axis has the equation $3x + 6y + 1 = 0$, the minor axis has the equation $4x - 2y + 3 = 0$ and the center of the ellipse is $(-2/3, 1/6)$.

Another way of doing this problem is to first translate the axes to the center of the conic. The coordinates of the center can be found by the method of completion of squares but there is another way of finding it which avoids completing the square. Namely, we have the following result:

Theorem 8.7. Let $q = Ax^2 + Bxy + Cy^2 + Dx + Ey + F$ and suppose that $\Delta = B^2 - 4AC \neq 0$. If (x_0, y_0) is the unique solution of the equations

$$\begin{aligned} 2Ax + By + D &= 0, \\ Bx + 2Cy + E &= 0, \end{aligned}$$

and $x' = x - x_0$, $y' = y - y_0$ then

$$q = Ax'^2 + Bx'y' + Cy'^2 + q(x_0, y_0).$$

Proof. Since the determinate of the coefficient matrix of the system

$$\begin{aligned} 2Ax + By &= -D, \\ Bx + 2Cy &= -E, \end{aligned}$$

is $-\Delta$, it has a unique solution (x_0, y_0) . If we make a change of coordinates $x = x_0 + x'$, $y = y_0 + y'$, we have

$$q = Ax'^2 + Bx'y' + Cy'^2 + (2Ax_0 + By_0 + D)x' + (Bx_0 + 2Cy_0 + E)y' + q(x_0, y_0).$$

and hence

$$q = Ax'^2 + Bx'y' + Cy'^2 + q(x_0, y_0).$$

Q.E.D.

Corollary 8.5. If $\Delta > 0$ and $q(x_0, y_0) \neq 0$, the conic $q = 0$ is a hyperbola with center (x_0, y_0) . If $\Delta > 0$ and $q(x_0, y_0) = 0$, the locus is a pair of lines meeting at (x_0, y_0) . If $\Delta < 0$ and $Aq(x_0, y_0) < 0$, the locus is an ellipse (or circle) with center (x_0, y_0) . If $\Delta < 0$ and $Aq(x_0, y_0) > 0$, the locus is empty. If $\Delta < 0$ and $q(x_0, y_0) = 0$, the locus consists of the single point (x_0, y_0) .

Example. In the previous example, the center is the solution of the system

$$\begin{aligned} 4x + 4y + 2 &= 0, \\ 4x + 10y + 1 &= 0, \end{aligned}$$

which is $x = -2/3$, $y = 1/6$. Evaluating

$$q = 2x^2 + 4xy + 5y^2 + 2x + y - 9$$

at the center, we get $q(-2/3, 1/6) = -115/12$ which shows that the locus is an ellipse with equation

$$2x'^2 + 4x'y' + 5y'^2 = 115/12$$

where $x' = x + 2/3$, $y' = y - 1/6$ is the coordinate system obtained by translating the coordinate frame by $(-2/3, 1/6)$. The $x'y'$ term can now be eliminated by a rotation of axes, exactly as in the previous problem.

The above simplification works only when $\Delta \neq 0$. If $B^2 - 4AC = 0$, we have $A, C \neq 0$ if $B \neq 0$ and

$$q = \frac{1}{4A}(2Ax + By)^2 + Dx + Ey + F.$$

Setting $x' = (2Ax + By)/\sqrt{4A^2 + B^2}$, $y' = (-Bx + 2Ay)/\sqrt{4A^2 + B^2}$, we obtain a rectangular coordinate system x', y' with

$$\begin{aligned} x &= (2Ax' - By')/\sqrt{4A^2 + B^2}, \\ y &= (Bx' + 2Ay')/\sqrt{4A^2 + B^2}. \end{aligned}$$

In this coordinate system we have

$$q = \frac{1}{4A}x'^2 + D'x' + E'y' + F'.$$

The locus $q = 0$ can then be identified after completing the square in x' .

Example. Consider the locus $q = 0$ with

$$q = 3x^2 - 12xy + 12y^2 + 3x - y + 10.$$

We have

$$q = 3(x - 2y)^2 + 3x - y + 10.$$

Setting $x' = (x - 2y)/\sqrt{5}$, $y' = (2x + y)/\sqrt{5}$, we get a rectangular coordinate system with x' -axis the line $2x + y = 0$ and y' -axis the line $x - 2y = 0$. Since

$$\begin{aligned} x &= (x' + 2y')/\sqrt{5}, \\ y &= (-2x' + y')/\sqrt{5}, \end{aligned}$$

the unit point on the x' -axis is $I'(1/\sqrt{5}, -2/\sqrt{5})$ and the unit point on the y' -axis is $J'(2/\sqrt{5}, 1/\sqrt{5})$. The equation of the locus in the $x'y'$ -coordinate system is

$$3x'^2 + \sqrt{5}x' + \sqrt{5}y' + 10 = 0.$$

Completing the square in x' , we get

$$3(x' + \sqrt{5}/6)^2 + \sqrt{5}y' + 115/12 = 0.$$

Dividing by $\sqrt{5}$, we get

$$\frac{3}{\sqrt{5}}(x' + \sqrt{5}/6)^2 + y' + 23\sqrt{5}/12 = 0.$$

Setting $x'' = x' + \sqrt{5}/6$, $y'' = y' + 23\sqrt{5}/12$, we get

$$y'' = -3/\sqrt{5}x''^2$$

which is the equation of a parabola with principal axis the line $y'' = 0$ and vertex $x'' = 0$, $y'' = 0$. In the original coordinate system, the equation of the principal axis is $2x + y + 115/12 = 0$ and, since

$$\begin{aligned} x &= -4 + (x'' + 2y'')/\sqrt{5}, \\ y &= -19/12 + (-2x'' + y'')/\sqrt{5}, \end{aligned}$$

the vertex has coordinates $(-4, -19/12)$. The y'' axis has equation $x - 2y + 5/6 = 0$. The unit points on the x'' -axis and y'' -axis are respectively

$$I''(-4 + 1/\sqrt{5}, 19/12 - 2/\sqrt{5}), \quad J''(-4 + 2/\sqrt{5}, -19/12 + 1/\sqrt{5}).$$

We now give an alternate way of computing P , A' and C' . If we have a quadratic form $q = Ax^2 + Bxy + Cy^2$ and we write it in matrix form, we have

$$q = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} A & B/2 \\ B/2 & C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

If we make a change of coordinates

$$\begin{bmatrix} x \\ y \end{bmatrix} = P \begin{bmatrix} x' \\ y' \end{bmatrix}.$$

with

$$P = \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix},$$

we have

$$q = \begin{bmatrix} x' & y' \end{bmatrix} \begin{bmatrix} A' & B'/2 \\ B'/2 & C' \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix}$$

with

$$\begin{bmatrix} A' & B'/2 \\ B'/2 & C' \end{bmatrix} = P^t \begin{bmatrix} A & B/2 \\ B/2 & C \end{bmatrix} P.$$

If P is orthogonal and $B' = 0$ we have

$$\begin{bmatrix} A & B/2 \\ B/2 & C \end{bmatrix} P = P \begin{bmatrix} A' & 0 \\ 0 & C' \end{bmatrix}.$$

which is equivalent to the two equations

$$\begin{bmatrix} A & B/2 \\ B/2 & C \end{bmatrix} \begin{bmatrix} a_1 \\ b_1 \end{bmatrix} = A' \begin{bmatrix} a_1 \\ b_1 \end{bmatrix} \text{ and } \begin{bmatrix} A & B/2 \\ B/2 & C \end{bmatrix} \begin{bmatrix} a_2 \\ b_2 \end{bmatrix} = C' \begin{bmatrix} a_2 \\ b_2 \end{bmatrix}$$

Given P , this gives a way of computing A' , B' without having to substitute for x, y . It also gives a new way of computing P . Indeed, setting $\lambda_1 = A'$, $\lambda_2 = C'$, these two equations can be written

$$\begin{bmatrix} \lambda_i - A & -B/2 \\ -B/2 & \lambda_i - C \end{bmatrix} \begin{bmatrix} a_i \\ b_i \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

with $i = 1, 2$. Then (a_i, b_i) is a non-zero solution of the homogeneous system

$$\begin{aligned} (\lambda_i - A)x - B/2y &= 0, \\ -B/2x + (\lambda_i - C)y &= 0, \end{aligned}$$

where $\lambda = \lambda_1, \lambda_2$ satisfy the equation

$$\begin{vmatrix} \lambda - A & -B \\ -B & \lambda - C \end{vmatrix} = \lambda^2 - (A + C)\lambda + AC - B^2/4 = 0.$$

This equation is called the **characteristic equation** of the matrix

$$M = \begin{bmatrix} A & B/2 \\ B/2 & C \end{bmatrix}$$

and its roots are called **eigenvalues** of the matrix M . Note that the roots of this equation are always real. Note also that

$$\lambda_1 \lambda_2 = AC - B^2/4$$

is the determinant of M and that

$$\lambda_1 + \lambda_2 = A + C$$

is the sum of the diagonal elements of M . The sum of the diagonal elements of an $n \times n$ matrix is called the **trace** of the matrix. Thus each λ_i can be found by finding the roots of the characteristic equation of M and the corresponding unit vector (a_i, b_i) is then a solution of the system

$$\begin{aligned} (\lambda_i - A)x - B/2y &= 0, \\ -B/2x + (\lambda_i - C)y &= 0. \end{aligned}$$

which determines it uniquely up to multiplication by ± 1 in the case that $\lambda_1 \neq \lambda_2$. The signs are chosen to make $\det(P) = 1$. Solutions of this system are called **eigenvectors** of M corresponding to the eigenvalue λ_i . To show that P is orthogonal, we have to show that

$$d = (a_1, b_1) \cdot (a_2, b_2) = 0.$$

This follows from

$$\lambda_1 d = ([a_1, b_1]M) \begin{bmatrix} a_2 \\ b_2 \end{bmatrix} = [a_1, b_1](M \begin{bmatrix} a_2 \\ b_2 \end{bmatrix}) = \lambda_2 d$$

which yields $(\lambda_1 - \lambda_2)d = 0$ and hence $d = 0$ since $\lambda_1 \neq \lambda_2$.

The case $\lambda_1 = \lambda_2$ arises iff $A = C$ and $B = 0$ (proof left as an exercise).

Example. The matrix of the quadratic function $q = 2x^2 + 4xy + 5y^2$ is

$$\begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix}$$

and its characteristic equation is

$$\lambda^2 - 7\lambda + 6 = 0.$$

The roots of this equation are $\lambda = 1, 6$. The unit vectors corresponding to $\lambda = 1$ is a solution of

$$\begin{aligned} -x - 2y &= 0, \\ -2x - 4y &= 0. \end{aligned}$$

which has for solutions $c(2, -1)$. The corresponding unit vectors are therefore

$$\pm(2/\sqrt{5}, -1/\sqrt{5}).$$

The unit vectors corresponding to $\lambda = 6$ are solutions of

$$\begin{aligned} 4x - 2y &= 0, \\ -2x + 1y &= 0. \end{aligned}$$

which has for solutions $c(1, 2)$. The corresponding unit vectors are thus

$$\pm(1/\sqrt{5}, 2/\sqrt{5}).$$

The transition matrix P can be taken to be

$$\begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}.$$

Therefore, setting

$$\begin{aligned} x &= (2x' + y')/\sqrt{5}, \\ y &= (-x' + 2y')/\sqrt{5}, \end{aligned}$$

we have

$$q = x'^2 + 6y'^2.$$

Note that if we had chosen $\lambda_1 = 6, \lambda_2 = 1$ then

$$\begin{aligned} x &= (x' - 2y')/\sqrt{5}, \\ y &= (2x' + y')/\sqrt{5}, \end{aligned}$$

would give

$$q = 6x'^2 + 6y'^2.$$

The corresponding transition matrix is obtained from the one above by interchanging the columns and then multiplying the second column by -1 to make $\det(P) = 1$.

Problem 8.6. *Identify the locus with equation*

$$2x^2 + 12xy - 3y^2 - 14x + 6y = 5$$

in a rectangular coordinate system x, y and find a rectangular coordinate system in which the locus is in standard form.

Solution. Since the discriminant of $q = 2x^2 + 12xy - 3y^2 - 14x + 6y - 17$ is 168, the locus is of hyperbolic type (hyperbola or two intersecting lines). The center is the solution of the system

$$\begin{aligned} 2x + 12y - 14 &= 0, \\ 12x - 6y - 6 &= 0, \end{aligned}$$

namely, the point $(1, 1)$. We therefore introduce the coordinate system x', y' where $x = 1 + x', y = 1 + y'$. In this coordinate system,

$$q = 2x'^2 + 12x'y' - 3y'^2 - 14$$

so the locus is a hyperbola. Since the matrix of q in the $x'y'$ -coordinate system is

$$\begin{bmatrix} 2 & 6 \\ 6 & -3 \end{bmatrix},$$

the characteristic equation of q is

$$\lambda^2 + \lambda - 42 = 0$$

whose roots are 6 and -7 . Solving the system

$$\begin{aligned} (\lambda - 2)x - 6y &= 0, \\ -6x + (\lambda + 3)y &= 0, \end{aligned}$$

for $\lambda = 6$ and $\lambda = -7$, we respectively get the general solutions

$$c(3, 2), c(-2, 3).$$

Therefore, the solutions of unit length are respectively

$$\pm(3/\sqrt{5}, 2/\sqrt{5}), \pm(-2/\sqrt{5}, 3/\sqrt{5}).$$

Hence, if we make the following change of coordinates

$$\begin{aligned} x' &= (3x'' - 2y'')/\sqrt{5}, \\ y' &= (-2x'' + 3y'')/\sqrt{5}, \end{aligned}$$

with transition matrix the orthogonal matrix

$$P = \begin{bmatrix} 3/\sqrt{5} & -2/\sqrt{5} \\ 2/\sqrt{5} & 3/\sqrt{5} \end{bmatrix},$$

the equation of our locus becomes

$$6x''^2 - 7y''^2 = 14.$$

Dividing by 14, we get the standard form of a hyperbola

$$\frac{x''^2}{7/3} - \frac{y''^2}{2} = 1$$

with principal axis the x'' -axis and asymptotes the lines

$$\sqrt{6}x'' \pm \sqrt{7}y'' = 0.$$

As an exercise, the reader should find the equations of these lines in the original coordinate system.

8.7. Exercises.

8.8. Focus-Directrix Description of Conics. A parabola has the equation $y^2 = cx$ in a rectangular coordinate system x, y if the principal axis is the x -axis and the vertex is the origin $O(0, 0)$. Since

$$cx = (x + c/4)^2 - (x - c/4)^2,$$

the equation of the parabola can be written

$$y^2 = (x + c/4)^2 - (x - c/4)^2.$$

Adding $(x - c/4)^2$ to both sides, we get

$$(x - c/4)^2 + y^2 = (x + c/4)^2.$$

Taking the positive square root of both sides, we get the equivalent equation

$$\sqrt{(x - c/4)^2 + y^2} = |x + c/4|$$

which shows that our parabola is the locus of points $P(x, y)$ that are equidistant from the point $F(c/4, 0)$ and the line L with equation $x = -c/4$. The constant c has for absolute value twice the distance of F to L . The line through F perpendicular to L is the principal axis of the parabola whose vertex is halfway between F and L . It follows that the point F and the line L are uniquely determined by the parabola. They are respectively called the **focus** and **directrix** of the parabola.

Example. In a rectangular coordinate system x, y the curve $y^2 = 4x$ is a parabola with focus $F(1, 0)$ and directrix the line $x = -1$. The curve $y^2 = -x$ has focus $F(-1/4, 0)$ and directrix $x = 1/4$.

Problem 8.7. In the rectangular coordinate system x, y find the equation of the parabola with focus $(1, 2)$ and directrix the line $x + y + 1 = 0$. Write the equation of this parabola in standard form. What are the coordinates of the vertex and what is the equation of the principal axis?

Solution. The parabola has for equation

$$(x - 1)^2 + (y - 2)^2 = \left| \frac{x + y + 1}{\sqrt{2}} \right|^2$$

which, on simplifying becomes

$$x^2 - 2xy + y^2 - 6x - 10y + 9 = 0.$$

Since the distance of the focus to the directrix is $2\sqrt{2}$, a standard form for the equation of the parabola is $y'^2 = 4\sqrt{2}x'$. The principal axis is the line through $(1, 2)$ perpendicular to the line $x + y + 1 = 0$. Its equation is $x - y + 1 = 0$. The principal axis and directrix intersect at the point $(-1, 0)$. The vertex is the point halfway between this point and the focus; it has coordinates $(0, 1)$.

We now show that any parabola has an important **reflection property**. Namely, if P is a point on the parabola, the acute angle made by the line through P parallel to the

principal axis and the tangent line at P is equal to the acute angle made by the tangent at P and the line through P and the focus F .

After possibly changing the orientation on the x and y -axes, we can assume that the equation of our parabola is $y^2 = 2px$ with $p > 0$ being the distance of the focus $F(p/2, 0)$ to the directrix $x = -p/2$. If $P(a, b)$ is a point on this parabola, the tangent to the parabola at this point has (b, p) as direction vector. If θ is the angle between this vector and the vector $(1, 0)$, we have

$$\sqrt{b^2 + p^2} \cos(\theta) = (b, p) \cdot (1, 0) = b,$$

and so θ is the acute angle that the line $y = b$ makes with the tangent to the parabola at $P(a, b)$. If θ' is the angle between the vector $(-b, -p)$ and the vector $(p/2 - a, -b)$, which is the coordinate vector of \overrightarrow{PF} , we have

$$\sqrt{b^2 + p^2} \cos(\theta') = \frac{(p/2 - a, -b)}{\sqrt{(p/2 - a)^2 + b^2}} \cdot (-b, -p) = \frac{b(a + c/2)}{\sqrt{(a - p/2)^2 + b^2}} = b$$

since $\sqrt{(a - p/2)^2 + b^2} = a + c/2$ by the focus-directrix property of the parabola. It follows that θ' is the acute angle between the tangent line at $P(a, b)$ and the line joining P and F and that $\theta = \theta'$.

Let us now, more generally, investigate the locus of points P such that the ratio of the distances from a fixed point F and a fixed line L not passing through F is a constant. If we choose a coordinate system x, y with origin F and x -axis perpendicular to L so that L has the equation $x = d$ with $d > 0$, our locus has the equation

$$\sqrt{x^2 + y^2} = e|x - d| \quad \text{or} \quad x^2 + y^2 = e^2(x - d)^2$$

with $e \geq 0$, a constant called the **eccentricity** of the locus. Simplifying, we get the equivalent equation

$$(1 - e^2)x^2 + y^2 + 2e^2dx - e^2d^2 = 0$$

which is the equation of a conic with discriminant $4(e^2 - 1)$. If $e = 1$ we get a parabola. Suppose that $e \neq 1$. Then, completing the square in x , we get

$$(1 - e^2)\left(x + \frac{e^2d}{1 - e^2}\right)^2 + y^2 = \frac{e^2d^2}{1 - e^2},$$

which shows that the conic is a non-degenerate central conic with center $(-e^2d/(1 - e^2), 0)$. It is a circle, ellipse or hyperbola according as $e = 0$, $e < 1$, $e > 1$. If we translate the axes so that the new origin is the center of the conic and divide both sides of the equation by $\frac{e^2d^2}{1 - e^2}$, we get a new coordinate system x', y' in which the conic has the equation

$$\frac{x'^2}{\frac{e^2d^2}{(1 - e^2)^2}} + \frac{y'^2}{\frac{e^2d^2}{1 - e^2}} = 1.$$

If $0 < e < 1$, the equation can be written

$$\frac{x'^2}{a^2} + \frac{y'^2}{b^2} = 1$$

with $a = ed/(1 - e^2)$, $b = ed/\sqrt{1 - e^2} > a$ so that the x' -axis (the line through F perpendicular to L) is the major axis. The original equation becomes

$$\sqrt{(x' - c)^2 + y'^2} = e|x' - \ell|$$

with $c = e^2 d / (1 - e^2)$, $\ell = d + c$. We also have the relations $c^2 = a^2 - b^2$, $e = c/a$, $d = b^2/c$, $\ell = a/e$. By symmetry, we have

$$\sqrt{(x' + c)^2 + y'^2} = e|x' + \ell|.$$

The points F, F' with x' -coordinates $(c, 0), (-c, 0)$ respectively are the **focii** of the ellipse with corresponding **directrices** $x' = \ell, x' = -\ell$.

If $e > 1$, the equation can be written

$$\frac{x'^2}{a^2} - \frac{y'^2}{b^2} = 1$$

with $a = ed/(e^2 - 1)$, $b = ed/\sqrt{e^2 - 1}$ so that the principal axis of the hyperbola is the x' -axis (the line through F perpendicular to L). The original equation becomes

$$\sqrt{(x' + c)^2 + y'^2} = e^2|x' + \ell|$$

with $c = e^2 d / (e^2 - 1) > d$, $\ell = c - d$. We also have the relations $c^2 = a^2 + b^2$, $e = c/a$, $d = b^2/c$, $\ell = a/e$. By symmetry, we have

$$\sqrt{(x' - c)^2 + y'^2} = e|x' - \ell|.$$

The points F, F' with x' -coordinates $(c, 0), (-c, 0)$ respectively are the **focii** of the hyperbola with corresponding **directrices** $x' = \ell, x' = -\ell$.

The ellipse and hyperbola have another description involving only their focii. Consider first the ellipse whose equation in a rectangular coordinate system x, y is

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

with $a > b$. By the above, we have

$$\sqrt{(x - c)^2 + y^2} = e|x - a/e| = e(a/e - x) = a - ex$$

since $a/e > a \geq x$. Similarly,

$$\sqrt{(x + c)^2 + y^2} = e|x + a/e| = e(x + a/e) = ex + a$$

since $x \geq -a > -a/e$. Hence,

$$\sqrt{(x - c)^2 + y^2} + \sqrt{(x + c)^2 + y^2} = 2a$$

and so an ellipse is the locus of points the sum of whose distances from 2 fixed points (the focii) is a constant.

In the case of the hyperbola

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

we have

$$\sqrt{(x - c)^2 + y^2} = e|x - a/e| = \begin{cases} ex - a & \text{if } x \geq a \\ a - ex & \text{if } x \leq -a \end{cases}.$$

Similarly,

$$\sqrt{(x + c)^2 + y^2} = e|x + a/e| = \begin{cases} ex + a & \text{if } x \geq a \\ -ex - a & \text{if } x \leq -a \end{cases}.$$

$$\sqrt{(x - c)^2 + y^2} - \sqrt{(x + c)^2 + y^2} = \pm 2a$$

and so the hyperbola is the locus of points the difference of whose distances from two fixed points (the focii) is, up to sign, a constant.

We leave as an exercise for the reader the proof of the fact that, if F, F' are the foci of an ellipse or hyperbola and P is any point on the curve, then the vectors $\overrightarrow{PF}, \overrightarrow{PF'}$ make equal acute angles with the tangent line to the curve at P .

8.9. Exercises. In these exercises, the given coordinate system is rectangular.

1. Find the focus and directrix of the parabola whose equation is $y = 8x^2$. Sketch the curve showing clearly the focus and directrix.
2. Find the equation of the parabola with focus $(-1, 4)$ and directrix the line $2x - y = 1$. Find an equation for the parabola which is in standard form. What are the coordinates of the vertex and what is the equation of the principal axis? Sketch the curve.
3. Find the foci and directrices of the ellipse $3x^2 + 2y^2 = 12$. Sketch the curve showing the foci, directrices and vertices.
4. Find the equation of the ellipse which, in a rectangular coordinate system, has focus $(1, 2)$, corresponding directrix $2x + y + 1 = 0$ and eccentricity $1/2$. Sketch this curve showing clearly the foci, directrices and vertices.
5. Find the foci and directrices of the hyperbola $3x^2 - 2y^2 = 1$. Sketch the curve showing clearly the foci, directrices, vertices and asymptotes.
6. Find the equation of the hyperbola with focus $(1, 2)$, corresponding directrix $2x + y + 1 = 0$ and eccentricity $\sqrt{5}$. Sketch this curve, showing clearly the foci, directrices, vertices and asymptotes.

8.10. Affine Classification of Quadric Surfaces. In this section we give the affine classification of quadric surfaces. This is obtained from the following theorem.

Theorem 8.8. *If q is a quadratic function on Euclidean space, there is a coordinate system x', y', z' such that*

$$\begin{aligned} q(P(x', y', z')) &= \mu(\epsilon_1 x'^2 + \epsilon_2 y'^2 + \epsilon_3 z'^2 + \epsilon_4) \\ &\text{or} \\ &= \mu(\epsilon_1 x'^2 + \epsilon_2 y'^2 - z'), \end{aligned}$$

where $\mu > 0$, $\epsilon_i = 0, \pm 1$ and $\epsilon_1 \neq 0$.

Proof. If x, y, z is any coordinate system, we have

$$q = ax^2 + bxy + cxz + dy^2 + eyz + fz^2 + gx + hy + kz + m.$$

Writing q as a polynomial in x , we get

$$q = ax^2 + (by + cz + g)x + dy^2 + eyz + fz^2 + hy + kz + m.$$

If $a \neq 0$, we can complete the square in x to get

$$q(P(x, y, z)) = a(x + (b/2a)y + (c/2a)z + g/2a)^2 + q_1(Q(0, y, z)),$$

with q_1 a function on the y, z -plane which is quadratic, linear or constant. Let $y'z'$ be a coordinate system in the y, z -plane so that

$$\begin{aligned} q_1(Q(0, y, z)) &= \mu(\epsilon_2 y'^2 + \epsilon_3 z'^2 + \epsilon_4) \\ &\text{or} \\ &= \mu(\epsilon_2 y'^2 - z'), \end{aligned}$$

where $\mu > 0$, $\epsilon_i = 0, \pm 1$. Setting $x' = \lambda(x + (b/2a)y + (c/2a)z + g/2a)$ with $\lambda = |a/\mu|^{1/2}$, we get a coordinate system x', y', z' with $q(P)$ in the required form.

If $a = 0$ and $d \neq 0$, let $x' = y, y' = x, z' = z$. Then the coefficient of x'^2 in $q(P)$ is d . Similarly, if $a = 0$ and $f \neq 0$ let $x' = z, y' = y, z' = x$ to get the same result. If $a = d = f = 0$ and $b \neq 0$, the change of coordinates $x' = x, y' = -x + y, z' = z$ yields b as the coefficient of x'^2 ; if $a = b = d = f = 0$ and $c \neq 0$, the change of coordinates $x' = x, y' = y, z' = -x + z$ yields c as the coefficient of x'^2 . If $a = b = c = d = f = 0$, then the change of coordinates $x = y', y = x', z' = -y + z$ yields e as the coefficient of x'^2 . We are thus reduced to the previous case $a \neq 0$. **Q.E.D.**

Corollary 8.6. *If $\epsilon_1 = \epsilon_2 = \epsilon_3 = 1$, the quadratic function q has a minimum value μ_{ϵ_4} at $x' = y' = z' = 0$. If $\epsilon_1 = \epsilon_2 = \epsilon_3 = -1$, it has a maximum value of μ_{ϵ_4} there.*

As in the two variable case, there are criteria for the existence of maxima and minima of quadratic functions in terms of the coefficients of these functions. We will derive them in a later chapter.

Problem 8.8. *Show that the quadratic function*

$$q(x, y, z) = x^2 + 2xy + 2xz + 2y^2 + 3z^2 + 6x + 2z + 2$$

has a minimum value. Where is this minimum attained?

Solution. Completing the square in x , we get

$$q(x, y, z) = (x + y + z + 3)^2 + y^2 - 2yz + 2z^2 - 6y - 4z - 7.$$

Now, completing the square in y , we get

$$q(x, y, z) = (x + y + z + 3)^2 + (y - z - 3)^2 + z^2 - 10z - 16.$$

Finally, completing the square in z , we get

$$q(x, y, z) = (x + y + z + 3)^2 + (y - z - 3)^2 + (z - 5)^2 - 41$$

which shows that -41 is the minimum value of q and that this value is attained when $x + y + z + 3 = y - z - 3 = z - 5 = 0$. This happens exactly when $x = -16, y = 8, z = 5$.

A quadric surface is said to be degenerate if its locus is empty or reduces to a point, line, plane or pair of planes. Using the above theorem we obtain the following classification of the non-degenerate quadric surfaces:

Theorem 8.9. *A non-degenerate quadric surface has one of the following equations in a suitable coordinate system:*

- (1) $x^2 + y^2 + z^2 = 1$ (Ellipsoid);
- (2) $x^2 + y^2 - z^2 = 1$ (Hyperboloid of one sheet);
- (3) $x^2 + y^2 - z^2 = -1$ (Hyperboloid of two sheets);
- (4) $x^2 + y^2 = z^2$ (Cone);
- (5) $x^2 + y^2 = z$ (Paraboloid);
- (6) $x^2 - y^2 = z$ (Hyperbolic Paraboloid);
- (7) $x^2 + y^2 = 1$ (Elliptic Cylinder);
- (8) $x^2 - y^2 = 1$ (Hyperbolic Cylinder);
- (9) $x^2 = y$ (Parabolic Cylinder).

Problem 8.9. *Identify the quadric surface*

$$x^2 + 2xy + 2xz + 2y^2 + 3z^2 + 6x + 2z + 2 = 0.$$

Solution. Using the previous problem, we see that the change of coordinates

$$\begin{aligned}x' &= (x + y + 3)/\sqrt{41} \\y' &= (y - z - 3)/\sqrt{41} \\z' &= (z - 5)/\sqrt{41}\end{aligned}$$

brings the equation of the quadric to the standard form

$$x'^2 + y'^2 + z'^2 = 1.$$

The quadric surface is an ellipsoid.

8.11. Euclidean Classification of Quadric Surfaces. Let x, y, z be a rectangular coordinate system and consider the quadric surface $q(P) = k$ with

$$q(P(x, y, z)) = Ax^2 + 2Bxy + 2Cxz + Dy^2 + 2Eyz + Fz^2.$$

Writing this in matrix form, we have $q(P(x, y, z)) = X^t M X$ with

$$X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \quad M = \begin{bmatrix} A & B & C \\ B & D & E \\ C & E & F \end{bmatrix}.$$

The matrix M satisfies $M = M^t$; such a matrix is called a **symmetric matrix**. We want to find a change of coordinates $X = PX'$ so that the new coordinate system x', y', z' is rectangular and

$$q(P(x', y', z')) = X'^t P^t M P X' = \lambda_1 x'^2 + \lambda_2 y'^2 + \lambda_3 z'^2$$

with $M' = P^t M P$ a diagonal matrix, i.e.,

$$M' = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}.$$

If

$$P = \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{bmatrix},$$

we have

$$\begin{aligned}x &= \alpha_1 x' + \alpha_2 y' + \alpha_3 z' \\y &= \beta_1 x' + \beta_2 y' + \beta_3 z' \\z &= \gamma_1 x' + \gamma_2 y' + \gamma_3 z' .\end{aligned}$$

The coordinate system x', y', z' is rectangular if and only if the vectors

$$\begin{aligned}\vec{i}' &= \alpha_1 \vec{i} + \beta_1 \vec{j} + \gamma_1 \vec{k} \\ \vec{j}' &= \alpha_2 \vec{i} + \beta_2 \vec{j} + \gamma_2 \vec{k} \\ \vec{k}' &= \alpha_3 \vec{i} + \beta_3 \vec{j} + \gamma_3 \vec{k}\end{aligned}$$

are mutually orthogonal unit vectors, i.e.,

$$\begin{aligned}\alpha_i^2 + \beta_i^2 + \gamma_i^2 &= 1 & \text{for } i = 1, 2, 3 \\ \alpha_i \alpha_j + \beta_i \beta_j + \gamma_i \gamma_j &= 0 & \text{for } i < j.\end{aligned}$$

Such vectors are also called **orthonormal**. This is equivalent to $PP^t = I$ (the identity matrix) or, equivalently $P^t = P^{-1}$. Such a matrix P is called an **orthogonal matrix**. The condition $P^tMP = M'$ can then be written $MP = PM'$ which is equivalent to

$$MP_i = \lambda_i P_i,$$

where P_i is the i -th column of P . This is a system of homogeneous equations in the coordinates of P_i which can be written

$$(\lambda_i I - M)P_i = 0.$$

A necessary and sufficient condition that this system has a non-zero solution is that

$$\det(\lambda_i I - M) = 0$$

. Evaluating

$$\det(\lambda I - M) = \begin{vmatrix} \lambda - A & -B & -C \\ -B & \lambda - D & -E \\ -C & -E & \lambda - F \end{vmatrix},$$

we get the polynomial

$$\lambda^3 + c_1\lambda^2 + c_2\lambda + c_3,$$

where

$$\begin{aligned} c_1 &= -(A + D + F) = -\text{trace of } M \\ c_2 &= AD + DF + AF - B^2 - C^2 - E^2 = \begin{vmatrix} A & B \\ B & D \end{vmatrix} + \begin{vmatrix} A & C \\ C & F \end{vmatrix} + \begin{vmatrix} D & E \\ E & F \end{vmatrix} \\ c_3 &= -\det(M). \end{aligned}$$

This polynomial is called the **characteristic polynomial** of the matrix M . Its roots $\lambda_1, \lambda_2, \lambda_3$ are called **eigenvalues** of M . The solutions of the homogeneous system

$$(\lambda_i I - M)X = 0$$

are called **eigenvectors** of M corresponding to the eigenvalue λ_i . Thus the problem of finding a rectangular coordinate system x', y', z' such that

$$q(P(x', y', z')) = \lambda_1 x'^2 + \lambda_2 y'^2 + \lambda_3 z'^2$$

is equivalent to finding an orthonormal basis of eigenvectors of the symmetric matrix M .

Since the characteristic polynomial of M is a cubic, it has at least one real root λ_1 . Let $(\alpha_1, \beta_1, \gamma_1)$ be an eigenvector of M , of unit length, corresponding to the eigenvalue λ_1 and let $(\alpha_2, \beta_2, \gamma_2)$, $(\alpha_3, \beta_3, \gamma_3)$ be orthogonal unit vectors each of which is orthogonal to $(\alpha_1, \beta_1, \gamma_1)$. If we make the orthogonal change of coordinates $X = PX'$ with

$$P = \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{bmatrix},$$

we have

$$M' = P^t M P = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & D' & E' \\ 0 & E' & F' \end{bmatrix}$$

since M' is a symmetric matrix. Indeed, $M'^t = P^t M^t (P^t)^t = P^t M P = M'$. We therefore have

$$q(P(x', y', z')) = \lambda_1 x'^2 + D' y'^2 + 2E' y' z' + F' z'^2$$

and we are reduced to diagonalizing the quadratic form

$$q' = D'y'^2 + 2E'y'z' + F'z'^2.$$

If we let $\vec{j''}, \vec{k''}$ be the unit vectors of a rectangular coordinate system in the $x'y'$ -plane (with origin O) that diagonalizes q' and replace the second and third columns of P by the column matrices $[\vec{j''}]^t, [\vec{k''}]^t$ with the coordinates taken with respect to the xyz -coordinate system, we obtain an orthogonal matrix that diagonalizes q . We thus obtain the following result:

Theorem 8.10. *If M is a symmetric 3×3 matrix, there is an orthogonal 3×3 matrix P such that $P^{-1}MP$ is a diagonal matrix.*

Corollary 8.7. *The roots of the characteristic polynomial of a real symmetric 3×3 matrix are all real.*

Another important fact about eigenvectors of real symmetric matrices is the following:

Theorem 8.11. *Let M be a real symmetric 3×3 (or 2×2) matrix. Then eigenvectors of M corresponding to distinct eigenvalues are orthogonal.*

Proof. Let X, Y be eigenvectors of M with eigenvalues λ, μ respectively. Then $MX = \lambda X$ and $MY = \mu Y$ and hence

$$\lambda X^t Y = (\lambda X)^t Y = (MX)^t Y = X^t MY = X^t (\mu Y) = \mu X^t Y.$$

This gives $(\lambda - \mu)X^t Y = 0$ from which we get $X^t Y = 0$ if $\lambda \neq \mu$. Since $X^t Y$ is equal to the dot product of X and Y we obtain the result. **Q.E.D.**

Example 1. If we apply the above to the quadric surface with equation

$$x^2 - 2xy + 2xz + y^2 + 2yz - z^2 = 1$$

the matrix M is equal to

$$\begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & 1 \\ 1 & 1 & -1 \end{bmatrix}.$$

The characteristic polynomial of this matrix is

$$\lambda^3 - \lambda^2 - 4\lambda + 4 = (\lambda^2 - 4)(\lambda - 1).$$

The eigenvalues of M are therefore $\lambda = 1, \pm 2$. The eigenvectors of M corresponding to $\lambda = 1$ are solutions X of $(I - M)X = 0$. This equation is equivalent to

$$\begin{aligned} y - z &= 0 \\ x - z &= 0 \\ -x - y + 2z &= 0 \end{aligned}$$

which has the solutions $x = y = z = t$ with t arbitrary. The eigenvectors with eigenvalue 2 are solutions X of $(2I - M)X = 0$ giving the equations

$$\begin{aligned} x + y - z &= 0 \\ x + y - z &= 0 \\ -x - y + 3z &= 0 \end{aligned}$$

which has the solutions $x = -y = t$, $z = 0$ with t arbitrary. Finally, the eigenvectors corresponding to $\lambda = -1$ are solutions of $(-2I - M)X = 0$ giving the equations

$$\begin{aligned} -3x + y - z &= 0 \\ x - 3y - z &= 0 \\ -x - y - z &= 0 \end{aligned}$$

which has the solutions $x = y = t$, $z = -2t$ with t arbitrary. If we choose for the columns of the matrix P , eigenvectors of length 1 corresponding to the distinct eigenvalues we get as one possible such P the matrix

$$P = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{bmatrix}.$$

If we make the change of coordinates $X = PX'$, the given quadric has equation

$$x'^2 + 2y'^2 - 2z'^2 = 1$$

in the new coordinate system. This new system is positively oriented since $\det(P) = 1$. The given quadric is a hyperboloid of one sheet with center of symmetry at the origin.

Example 2. If our quadric has the equation

$$x^2 - 2xy + 2yz + y^2 + 2yz - z^2 + x - y = 0,$$

we first diagonalize the associated quadratic form

$$q = x^2 - 2xy + 2yz + y^2 + 2yz - z^2.$$

The above example shows that this can be done with the change of coordinates

$$\begin{aligned} x &= x'/\sqrt{3} + y'/\sqrt{2} + z'/\sqrt{6} \\ y &= x'/\sqrt{3} - y'/\sqrt{2} + z'/\sqrt{6} \\ y &= x'/\sqrt{3} - 2z'/\sqrt{6}. \end{aligned}$$

If we make this change of coordinates the given quadric has equation

$$x'^2 + 2y'^2 - 2z'^2 + \sqrt{2}y' = 0$$

in the new coordinate system. Completing the square in y' , we get the equation

$$x'^2 + 2(y' + \sqrt{2}/4)^2 - 2z'^2 = 1/4.$$

If we multiply both sides by 4 and make the change of coordinates

$$x' = x'', \quad y' = y'' - \sqrt{2}/4, \quad z' = z'',$$

we get the equation

$$4x''^2 + 8y''^2 - 8z''^2 = 1$$

in the $x''y''z''$ -coordinate system. This is quadric surface is a hyperboloid of one sheet with center of symmetry O'' . The center of symmetry has coordinates $(-1/4, +1/4, 0)$ in the xy -coordinate system. The unit points I'' , J'' , K'' have xy -coordinates

$$(1/\sqrt{3}-1/4, -1/\sqrt{3}+1/4, 0), (1/\sqrt{2}-1/4, 1/\sqrt{2}+1/4, 0), (1/\sqrt{6}-1/4, 1/\sqrt{6}+1/4, -2/\sqrt{6}).$$

We could have eliminated completion of squares from the above example by first translating the origin to the center of the quadric. As the case for central conics, there is a system of linear equations which have the center of a central quadric as the unique solution. Let

$$q = Ax^2 + 2Bxy + 2Cxz + Dy^2 + 2Eyz + Fz^2 + Gx + Hy + Kz + L = 0$$

be a quadric surface with associated quadratic form

$$q_0 = Ax^2 + 2Bxy + 2Cxz + Dy^2 + 2Eyz + Fz^2$$

and let

$$\Delta = \begin{vmatrix} A & B & C \\ B & D & E \\ C & E & F \end{vmatrix}$$

be the determinant of the matrix M of q_0 . Since $\det(M)$ is the product of the eigenvalues of M , the quadric surface $q = 0$ is central if and only if $\Delta \neq 0$. If $\Delta > 0$, it is of elliptic type (ellipsoid, point or empty) and if $\Delta < 0$, it is a hyperboloid.

Theorem 8.12. *If $q = 0$ is a central quadric, (x_0, y_0, z_0) is the unique solution of the system of equations*

$$\begin{aligned} Ax + By + Cz &= -G/2 \\ Bx + Dy + Ez &= -H/2 \\ Cx + Ey + Fz &= -K/2 \end{aligned}$$

and $x' = x - x_0$, $y' = y - y_0$, $z' = z - z_0$ then

$$q = q_0(x', y', z') + q(x_0, y_0, z_0).$$

In particular, (x_0, y_0, z_0) are the coordinates of the centre of the quadric surface $q = 0$.

Proof. If we make the change of coordinates $x = x_0 + x'$, $y = y_0 + y'$, $z = z_0 + z'$ we have

$$\begin{aligned} q &= q_0(x', y', z') + 2(Ax_0 + By_0 + Cz_0 + G/2)x' + 2(Bx_0 + Dy_0 + Ez_0 + H/2)y' + \\ &\quad 2(Cx_0 + Ey_0 + Fz_0 + K/2)z' + q(x_0, y_0, z_0). \end{aligned}$$

Q.E.D.

Example. In example 2 above we have

$$\Delta = \begin{vmatrix} 1 & -1 & 1 \\ -1 & 1 & 1 \\ 1 & 1 & -1 \end{vmatrix} = -4.$$

So the surface is a hyperboloid, or a cone, with center the unique solution of the system

$$\begin{aligned} x - y + z &= -1/2 \\ -x + y + z &= 1/2 \\ x + y - 1 &= 0. \end{aligned}$$

This solution is $(-1/4, 1/4, 0)$. Translating our coordinate system to this point, we get a coordinate system x', y', z' in which our quadric has equation

$$x'^2 - 2x'y' + y'^2 + 2y'z' - z'^2 = 1/4$$

which shows that the surface is a hyperboloid. The change of coordinates

$$\begin{aligned}x' &= x''/\sqrt{3} + y''/\sqrt{2} + z''/\sqrt{6} \\y' &= x''/\sqrt{3} - y''/\sqrt{2} + z''/\sqrt{6} \\y' &= x''/\sqrt{3} - 2z''/\sqrt{6}\end{aligned}$$

transforms the above equation into

$$x''^2 + 2y''^2 - 2z''^2 = 1/4$$

showing that the surface is a hyperboloid of one-sheet.