Notes and Questions for Geometry (640:435:01)

1 Conics

You encountered conics (ellipses, hyperbolas, and parabolas) in pre-Calculus. You probably learned their analytic representations (i.e. their canonical equations) when they are placed in special positions in a rectangular coordinate system. We will do a little bit of review of such basic facts of conics: their canonical equations when placed in special positions in a rectangular coordinate system, and their main geometric features and properties (parameters such as foci, eccentricity, and directrices that determine their shape, size, and positions). We will also learn more about them: geometrically they all arise as sections from a plane slicing through a double right circular cone, and algebraically they can all be described by a quadratic equation in two variables and all satisfy a unifying focus-directrix relation.

Basic tools for this chapter include completion of squares, tangents to curves (derivatives, parametric and implicit differentiation), basic trigonometry, 2×2 matrix algebra, eigenvalues and eigenvectors of such matrices, and diagonalization of 2×2 real symmetrical matrices by orthogonal matrices.

1.1 Conic Sections and Conics

1.1.1 Conic sections

This section describes how different conic sections arise from a plane slicing through a double right circular cone. You may have treated ellipses, hyperbolas, and parabolas as different geometric shapes. However, this slicing description indicates that they are somehow related. Note also the possibility of exceptional cases in this slicing process: single point, single straight line, and pair of straight lines. These are referred to as *degenerate conic sections*.

The algebraic relations between this group of geometric objects can be most easily seen through the *algebraic equations* that describe them. We recall that any straight line in $\mathbb{R}^2$ can be described in terms of a *linear equation* in x and y in the form: $ax + by + c = 0$, for some coefficients a, b, and c which are not all zero (what does the equation mean when all a, b, and c are zero? Straight lines in other settings may have different forms of equations, so a natural question to review is: what properties of straight lines in $\mathbb{R}^2$ are used to establish that they are represented by equations of the form $ax + by + c = 0$?). Conversely, any linear equation $ax + by + c = 0$, in which a, b, and c are not all zero, represent a straight line. So we can say that in some sense straight lines (in the Euclidean plane) and linear equations in two variables represent the same thing. What are the algebraic relations among ellipses, hyperbolas, and parabolas? We saw their geometric relations in the slicing description above. Algebraically, all of them can be represented by a *quadratic equation* in two variables x and y. Recall that

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \qquad \text{(ellipse)}$$

for some $a, b > 0$, represents an ellipse;

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1, \quad \text{or} \quad \frac{x^2}{a^2} - \frac{y^2}{b^2} = -1, \qquad \text{(hyperbola)}$$

for some $a, b > 0$, represents a hyperbola; and

$$y^2 = \pm 4ax, \quad \text{or} \quad x^2 = \pm 4ay, \qquad \text{(parabola)}$$

for some $a > 0$, represents a parabola. All these equations are quadratic in x and y.

Question. *Given an arbitrary quadratic equation, which, in its most general form, can be written as*

$$Ax^2 + Bxy + Cy^2 + Fx + Gy + H = 0. \quad (A, B, \text{ and } C \text{ are not all zero.}) \qquad (1)$$

Does the equation always represent one of the ellipses, hyperbolas, and parabolas? If so, how to determine their shapes and positions from the equation? If not, what are all the possibilities?

It turns out that any such equation corresponds either to an ellipse, a hyperbola, a parabola, or to one of the degenerate conic sections. We work this out in **1.3**, while in **1.1-2**, we study some concrete properties of these conic sections.

1.1.2 Circles

We start with the simplest conic section: circles. The most geometric way of writing down an equation for a circle is

$$(x - a)^2 + (y - b)^2 = r^2. \qquad (2)$$

From the equation, one can read off the centre of the circle to be (a, b) and radius of the circle to be r. Expanding the above equation, one sees that it can be written in the form of

$$x^2 + y^2 + fx + gy + h = 0, \qquad (3)$$

for some f, g, and h.

Question. *Given an equation like* (3)*. Does it always represent a circle? If so, how to find its centre and radius?*

The answer is: almost yes (come up with your own example of an equation of type (3) which does not represent a circle). And the way to carry out the answer is by completing the squares.

The rest of this subsection discusses how to use algebra to describe special geometric relations between circles: orthogonal circles and the family of circle through two given points. The text's treatment for the latter is worth learning: see the argument on the lower part of p. 10 to learn how it avoids deriving the equation in a traditional way.

1.1.3 Focus-Directrix Definition of the Non-Degenerate Conics

Here we describe a conic in terms of (i) a point F representing a *focus*; (ii) a straight line d called a *directrix*; (iii) a number $e \geq 0$ representing the *eccentricity*; and (iv) a relation in terms of the eccentricity e and distances from the points on the conic to the directrix and to the focus.

For any point P on the conic. Let M be the foot of the perpendicular straight line from P to the directrix d. Then the distance relation is

$$PF = e \cdot PM.$$

Then what we do in this section is to demonstrate that, *in appropriately chosen coordinates*, the set of points satisfying this relation is described by one of the conics in standard forms.

A general pattern for the choice of coordinates is that we choose coordinates to make $(ae, 0)$ a focus, and $x = \frac{a}{e}$ to be a directrix (therefore the straight line perpendicular to the directrix and through the focus will be our x-axis). Of course when $e = 0$ or 1 this has to be modified: when $e = 0$, the directrix in some sense moves off to infinity and the focus is at the origin; while for $e = 1$, we still take $(ae, 0) = (a, 0)$ to be a focus, but take $x = -\frac{a}{e} = -a$ to be the directrix.

We then discuss the geometric parameters for the graphs of the conics: major and minor axes, vertices, a second focus-directrix pair for ellipse and hyperbola, and the asymptotes for hyperbola — see the side margin diagram summaries on pp. 15 and 17. We also discuss the use of parametric equations and polar equations for conics. Lastly we discuss the *focal distance property*, which, for an ellipse, says that *the sum of distances from any point on an ellipse to its two foci is a constant equal to the length of its major axis*, and for a hyperbola, says that *the absolute value of the difference of distances from any point on a hyperbola to its two foci is a constant equal to the length of its major axis.*

1.2 Properties of Conics

The properties of conics discussed in this subsection are all related to the tangents to conics. Even though some proofs of this section require clever set ups and fair amount of computation, they all boil down to making use of the the most basic information about the tangents: their slopes are computed as $\frac{dy}{dx}$ *at the point of tangency*, which can be computed either (i) directly, if one can solve for y in terms of x easily, or (ii) parametrically as $\frac{dy}{dt}/\frac{dx}{dt}$ at the point of tangency, if both x and y are given in terms of a parameter t and $\frac{dx}{dt} \neq 0$, or (iii) by implicit differentiation.

Example. *Given the circle $x^2 + y^2 = R^2$ and a point $P = (x_0, 0)$ outside of it. Find equations of the tangent(s) from P to the given circle.*

Proof. The key is to find the coordinates for the point of tangency. Suppose $(\bar{x}, \bar{y})$ is a point of tangency from P to the circle. Then we can write down an equation for the tangent line to given the circle through $(\bar{x}, \bar{y})$ as $\bar{x}x + \bar{y}y = R^2$—this is worked out according to the principle above and is summarized as **Theorem 2** on p. 25. Since $(x_0, 0)$ is on this line, we must have $\bar{x}x_0 = R^2$. Thus $\bar{x} = R^2/x_0$, and we can solve for $\bar{y}$ from $\bar{x}^2 + \bar{y}^2 = R^2$. This is a quadratic equation in $\bar{y}$, and we expect to find two solutions: $\bar{y} = \pm\sqrt{R^2 - (R^2/x_0)^2}$, provided $R^2 - (R^2/x_0)^2 > 0$. This is the case as we assumed that $(x_0, 0)$ outside of the given circle, so $x_0^2 + 0^2 > R^2$. Now that we have $(\bar{x}, \bar{y}) = (R^2/x_0, \pm R\sqrt{1 - R^2/x_0^2})$, we know the two tangent lines from P to the circle are

$$\frac{R^2}{x_0}x \pm \left(\frac{R}{x_0}\sqrt{x_0^2 - R^2}\right) y = R^2.$$

Simplifying the equation a bit, we find two tangents to the given circle through $(x_0, 0)$ as

$$Rx \pm \sqrt{x_0^2 - R^2}\, y = Rx_0.$$

□

Remark (on the Proof of Reflection Property of the Ellipse)**.** *Notice that the proof in the text does not make direct use of the reflected ray at P. Instead, the proof joins P and F' by a straight line and verifies that the line $F'P$ makes equal angle with the tangent at P as FP makes with it, thereby proving that $F'P$ has to coincide with the reflected ray. The same strategy is also used in the proof of Reflection Property of the Parabola. Also note that the proof in the text does not try to compute directly the angles formed by the incoming and reflected rays with the tangent, but makes an indirect argument cleverly using the Sine Formula in two pairs of triangles. A more direct argument directly aiming to prove that the incoming angle equals the reflected angle can be done but involves more trigonometry: let θ and λ denote the angles of inclination of the tangent PT and the segment FP, respectively. Then*

$\tan\theta =$ *the slope of* PT*, which can be computed as* $b\cos t/(-a\sin t)$*, assuming* $P = (a\cos t, b\sin t)$*;* $\tan\lambda =$ *slope of* $FP = b\sin t/(a\cos t - ae)$*, and the angle between* FP *and* PT *is* $\theta - \lambda$ *by the exterior angle theorem for triangles. Using a trigonometric identity, we compute*

$$
\begin{aligned}
&\tan(\theta-\lambda)\\
=&\frac{\tan\theta-\tan\lambda}{1+\tan\theta\tan\lambda}\\
=&\frac{b\cos t/(-a\sin t)-b\sin t/(a\cos t-ae)}{1+\frac{b\cos t}{-a\sin t}\frac{b\sin t}{a\cos t-ae}}\\
=&\frac{b\cos t(a\cos t-ae)+ab\sin^2 t}{-a\sin t(a\cos t-ae)+b^2\sin t\cos t}\\
=&\frac{ab(1-e\cos t)}{(b^2-a^2)\sin t\cos t+a^2e\sin t}\\
=&\frac{ab(1-e\cos t)}{-a^2e^2\sin t\cos t+a^2e\sin t} \quad \textit{using } b^2-a^2=-a^2e^2\\
=&\frac{b}{ae\sin t};
\end{aligned}
$$

while if μ *denotes the angle of inclination of* $F'P$*, then* $\tan\mu = b\sin t/(a\cos t + ae)$*, and the angle between* $F'P$ *and* PT *is* $\mu + \pi - \theta$*. Finally, we need to verify* $\tan(\mu + \pi - \theta) = \tan(\theta - \lambda)$ *to prove that* $\mu + \pi - \theta = \theta - \lambda$*. But*

$$
\begin{aligned}
&\tan(\mu+\pi-\theta)\\
=&\tan(\mu-\theta)\\
=&-\tan(\theta-\mu)\\
=&-\frac{b\cos t/(-a\sin t)-b\sin t/(a\cos t+ae)}{1+\frac{b\cos t}{-a\sin t}\frac{b\sin t}{a\cos t+ae}}\\
=&-\frac{b\cos t(a\cos t+ae)+ab\sin^2 t}{-a\sin t(a\cos t+ae)+b^2\sin t\cos t}\\
=&-\frac{ab(1+e\cos t)}{(b^2-a^2)\sin t\cos t-a^2e\sin t}\\
=&-\frac{ab(1+e\cos t)}{-a^2e^2\sin t\cos t-a^2e\sin t} \quad \textit{using } b^2-a^2=-a^2e^2\\
=&\frac{b}{ae\sin t}\\
=&\tan(\theta-\lambda).
\end{aligned}
$$

Thus we conclude that PF *and* PF' *form equal angles with the tangent to the ellipse at* P*.*

1.3 Recognizing Conics

We will discuss four aspects:

(i). If we move a conic in its standard form to a new position, by a rotation or translation, then it is still the same geometric figure (with the same size and shape: same lengths of major/minor axes, same eccentricity, etc), but its equation may look different, though still quadratic.

(ii). If we move $P(x_P, y_P)$ to $Q(x_Q, y_Q)$ by a rotation or translation, we need to figure out the relation between the coordinates (x_P, y_P) of P and the coordinates (x_Q, y_Q) of Q.

(iii). Given a quadratic equation in x and y, we expect to be able to set up a new coordinate system $x'O'y'$ in which the conic takes its standard form. A first step is to choose a new coordinate system (x', y') in which the $x'y'$ term disappears (namely, its coefficient is made to be zero). One can then recognize the geometry of this equation by completing squares.

Example. *Identify the geometry of the equation* $x^2 - 2y^2 + 2x + 4y - 3 = 0$.

Solution. *We group the terms involving* x *together, and those involving* y *together, and complete squares.*

$$x^2 - 2y^2 + 2x + 4y - 3 = x^2 + 2x + 1 - 2(y^2 - 2y + 1) - 2 = (x+1)^2 - 2(y-1)^2 - 2.$$

Thus we have $(x+1)^2 - 2(y-1)^2 = 2$, *which can be written as*

$$\frac{(x+1)^2}{2} - \frac{(y-1)^2}{1} = 1.$$

Thus if we set

$$\begin{aligned} x' &= x + 1 \\ y' &= y - 1 \end{aligned} \tag{4}$$

then we obtain

$$\frac{(x')^2}{2} - \frac{(y')^2}{1} = 1,$$

a hyperbola in standard form. The transformation (4) *represents a translation by vector* $(-1, 1)$, *since* $x = -1$ *and* $y = 1$ *corresponds to the new origin* $x' = 0$ *and* $y' = 0$. *So the equation* $x^2 - 2y^2 + 2x + 4y - 3 = 0$ *represents a hyperbola with center at* $(-1, 1)$, *major semi-axis* $a = \sqrt{2}$, *and minor semi-axis* $b = 1$. *Using the associated right triangle relation to solve* $c = \sqrt{a^2 + b^2} = \sqrt{3}$, *we know that* *the eccentricity of the hyperbola is* $e = c/a = \sqrt{3}/\sqrt{2}$, *the foci are* $\sqrt{3}$ *units to two sides of the center* $(-1, 1)$ *along the major axis, so are* $(-1+\sqrt{3}, 1)$ *and* $(-1-\sqrt{3}, 1)$. *The vertices are* $a = \sqrt{2}$ *units to the two sides of the center* $(-1, 1)$ *along the major axis, so are* $(-1+\sqrt{2}, 1)$ *and* $(-1-\sqrt{2}, 1)$. *The asymptotes are the straightlines through the center* $(-1, 1)$ *with* $\pm b/a = \pm 1/\sqrt{2}$ *as slopes. In the* x'-y' *coordiates, they are given by* $\frac{x'}{\sqrt{2}} + \frac{y'}{1} = 0$ *or* $\frac{x'}{\sqrt{2}} - \frac{y'}{1} = 0$. *In terms the* x-y *coordinates, we have* $\frac{x+1}{\sqrt{2}} + (y-1) = 0$, *or* $\frac{x+1}{\sqrt{2}} - (y-1) = 0$. *It turns out that one could also obtain the asymptotes by setting*

$$\frac{(x+1)^2}{2} - \frac{(y-1)^2}{1} = 0,$$

which reduces through factorization

$$\frac{(x+1)^2}{2} - \frac{(y-1)^2}{1} = (\frac{x+1}{\sqrt{2}} + (y-1))(\frac{x+1}{\sqrt{2}} - (y-1))$$

to $\frac{x+1}{\sqrt{2}} + (y-1) = 0$, *or* $\frac{x+1}{\sqrt{2}} - (y-1) = 0$, *which are the two asymptotes.*

(iv). In order to tackle the problem raised in (iii), namely, choosing a new coordinate system (x', y') in which the $x'y'$ term disappears, we need to solve the following problems

(a). The same point P has two sets of coordinates: its coordinates (x, y) in the original coordinate system, and its coordinates (x', y') in the new rotated/translated coordinates $x'O'y'$. Find their relation —note the difference and connection between this problem and that in (ii).

(b). For the given quadratic equation in x and y,

$$Ax^2+Bxy+Cy^2+Fx+Gy+H = 0. \quad (A, B, \text{ and } C \text{ are not all zero.}) \quad (1')$$

substitute the coordinate transformation in (a) into the equation to determine what rotation/translation would make the equation as close to its standard form as possible in the new coordinates (x', y'). In practice, we first express the original given equation in terms of matrices: set

$$\mathbf{A} = \begin{pmatrix} A & B/2 \\ B/2 & C \end{pmatrix} \quad \text{and } \mathbf{J} = \begin{pmatrix} F \\ G \end{pmatrix},$$

then the original equation $(1')$ can be written in terms of the vector $\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}$ as

$$\mathbf{x}^T\mathbf{A}\mathbf{x} + \mathbf{J}^T\mathbf{x} + H = 0. \quad (5)$$

We will find that the new axes that simplify (5) correspond to the eigenvectors of the matrix $\mathbf{A}$!

Let's first record the solutions to (ii) and (a) here. For (ii), if we rotate every point by angle θ, and suppose that $P(x_P, y_P)$ is rotated to its new position $Q(x_Q, y_Q)$, then

$$\begin{pmatrix} x_Q \\ y_Q \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_P \\ y_P \end{pmatrix}. \tag{6}$$

One can check this as follows: we expect $\begin{pmatrix} x_P \\ y_P \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ to be rotated to $\begin{pmatrix} x_Q \\ y_Q \end{pmatrix} = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$, and $\begin{pmatrix} x_P \\ y_P \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ to be rotated to $\begin{pmatrix} x_Q \\ y_Q \end{pmatrix} = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}$. These are consistent with the formula above. But spot-checking is not the same as a proof. A proof for (6) can be given by using polar coordinates for P and Q, noting that both P and Q have the same radii r, and if ϕ is the polar angle of P, then that for Q is $\phi + \theta$. Thus

$$\begin{cases} x_P = r\cos\theta \\ y_P = r\sin\theta \end{cases} \quad \text{and} \quad \begin{cases} x_Q = r\cos(\theta+\phi) \\ y_Q = r\sin(\theta+\phi) \end{cases}$$

It follows now

$$\begin{cases} x_Q = r\left(\cos\theta\cos\phi - \sin\theta\sin\phi\right) = x_P\cos\theta - y_P\sin\theta \\ y_Q = r\left(\sin\theta\cos\phi + \cos\theta\sin\phi\right) = x_P\sin\theta + y_P\cos\theta, \end{cases}$$

which becomes (6) when written in matrix form.

For (a), if we rotate the original coordinate axes by angle θ to obtain a new coordinate system $x'O'y'$, and a point P has coordinates $\begin{pmatrix} x \\ y \end{pmatrix}$ in the xOy coordinate system and coordinates $\begin{pmatrix} x' \\ y' \end{pmatrix}$ in the $x'Oy'$ coordinate system, then

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}. \tag{7}$$

This can also be checked as follows: the point $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$ should have coordinates $\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ in the new coordinates, and the point $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}$ should have coordinates $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ in the new coordinates. Notice that $\begin{pmatrix} x \\ y \end{pmatrix}$ now appears on the left side of the equation, in contrast to (6). But (7) is closely linked to (6), and can in fact be derived from (6): think of P as obtained from a point P' through rotation of angle θ; then the coordinates of P' in the xOy coordinate system would be $\begin{pmatrix} x' \\ y' \end{pmatrix}$, so we can apply (6) to P' to obtain (7).

Remark. *The matrix* $\mathbf{P} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ *satisfies*

(o1). each of its columns is a unit vector.

(o2). the different columns are orthogonal to each other.

(det). its determinant equals 1.

Any square matrix $\mathbf{P}$ *satisfying (o1) and (o2) is called an* orthogonal *matrix. Such a matrix* $\mathbf{P}$ *satisfies* $\mathbf{P}^T\mathbf{P} = \mathbf{P}\mathbf{P}^T = I$ *and* $\mathbf{P}^{-1} = \mathbf{P}^T$. *A* 2×2 *orthogonal matrix satisfying also (det) is called a rotation matrix. Any* 2×2 *rotation matrix can be written in the form* $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ *for some* θ *— see Theorem 2 on p. 38.*

Back to (b): can we find a rotation matrix $\mathbf{P}$ such that after the substitution $\mathbf{x} = \mathbf{P}\mathbf{x}'$ into (5), the equation in x' and y' would be simplified to be close to a standard form for conics? After the substitution, (5) looks like

$$(\mathbf{x}')^T\mathbf{P}^T\mathbf{A}\mathbf{P}\mathbf{x}' + \mathbf{J}^T\mathbf{P}\mathbf{x}' + H = 0. \tag{8}$$

So we want $\mathbf{P}^T\mathbf{A}\mathbf{P}$ to be a diagonal matrix, say $\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ for some λ_1, λ_2:

$$\mathbf{P}^T\mathbf{A}\mathbf{P} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}. \tag{9}$$

If this can be done, then

$$(\mathbf{x}')^T\mathbf{P}^T\mathbf{A}\mathbf{P}\mathbf{x}' = (\mathbf{x}')^T \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \mathbf{x}' = \lambda_1(x')^2 + \lambda_2(y')^2,$$

so (8) would not contain $x'y'$ term and a further simplification can be carried out by completing squares in x' and y'. But the rotation matrix $\mathbf{P}$ satisfies $\mathbf{P}^T = \mathbf{P}^{-1}$, so (9) is equivalent to

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix},$$

which is equivalent to

$$\mathbf{A}\mathbf{P} = \mathbf{P}\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \tag{10}$$

i.e., the columns of $\mathbf{P}$ are *eigenvectors* of $\mathbf{P}$! But can we use eigenvectors of $\mathbf{A}$ to form a matrix $\mathbf{P}$ which in the mean time is also a rotation matrix? The answer is yes. This is because $\mathbf{A}$ is a symmetric matrix, and we recall that we can choose a set of orthonormal basis consisting of eigenvectors of a symmetric matrix. The only additional thing we need to make sure is to form a matrix with these vectors such that its determinant is 1. This can be achieved by switching the positions of two column vectors, if necessary.

The strategy for recognizing a conic from a given quadratic equation is described in the box on p.41. Let me just expand on step 2: after rewriting the equation into matrix form (5) and have the symmetric matrix $\mathbf{A}$ to work with,

(2a). Use the characteristic equation for matrix $\mathbf{A}$ to find its eigenvalues.

(2b). Solve for eigenvectors of $\mathbf{A}$ for each of its eigenvalues.

(2c). Because $\mathbf{A}$ is symmetric, its eigenvectors corresponding to distinct eigenvalues are automatically orthogonal to each other, and because $\mathbf{A}$ is 2×2, we only need to normalize its eigenvectors, *i.e.*, to make the eigenvectors of length 1.

(2d). Use the eigenvectors as columns to form a matrix $\mathbf{P}$ and arrange its order to make sure that $\det \mathbf{P} = 1$—note that the choice of $\mathbf{P}$ is not unique, as the role of x'-axis and y'-axis can be switched. This $\mathbf{P}$ will be what we wanted, and after the substitution $\mathbf{x} = \mathbf{P}\mathbf{x}'$, the quadratic terms involving $\mathbf{x}' = \begin{pmatrix} x' \\ y' \end{pmatrix}$ will take the form $\lambda_1 x'^2 + \lambda_2 y'^2$, where λ_1, λ_2 are the two eigenvalues corresponding to the first and second column of $\mathbf{P}$, respectively. Note from the relation $\mathbf{x} = \mathbf{P}\mathbf{x}'$ that the x'-axis, which points to the same direction as $\mathbf{x}' = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, is then in the direction of $\mathbf{x} = \mathbf{P}\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, which is the first column of $\mathbf{P}$; and that y'-axis will be pointing in the direction of $= \mathbf{P}\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, which is the second column of $\mathbf{P}$.

Remark (on Theorem 3 on p.42)**.** *The reason for Theorem 3 is due to the relation*

$$\mathbf{P}^T\mathbf{A}\mathbf{P} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix},$$

which implies that

$$\lambda_1\lambda_2 = \det \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \det(\mathbf{P}^{-1}\mathbf{A}\mathbf{P}) = \det \mathbf{P}^{-1} \det \mathbf{A} \det \mathbf{P} = \det \mathbf{A}.$$

So λ_1 and λ_2 have the same (opposite) signs depending on whether $\det \mathbf{A} > (< 0)$, *etc.*

2 Affine Geometry

2.1 Geometry and Transformations

An **isometry** of $\mathbb{R}^2$ is a function which maps $\mathbb{R}^2$ onto $\mathbb{R}^2$ and *preserves distances.* Examples of isometries of $\mathbb{R}^2$ include

translation by a vector in $\mathbb{R}^2$;

rotation about a point in $\mathbb{R}^2$;

reflection in a line in $\mathbb{R}^2$;

compositions of the above.

We know that a translation by a vector $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ is represented by

$$t\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + x_0 \\ y + y_0 \end{pmatrix},$$

and a rotation about the origin with angle θ is represented by

$$rot\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

A question arises naturally:

Question. *how to represent rotation about an arbitrary point in $\mathbb{R}^2$? how to represent a reflection in $\mathbb{R}^2$? and how to represent the compositions of these isometries?*

Geometrically, a reflection ref in a line l through the origin in $\mathbb{R}^2$ has the property

$$ref(\vec{v}) = \vec{v}, \quad ref(\vec{w}) = -\vec{w},$$

where $\vec{v}$ is a direction vector for the line l, and $\vec{w}$ is a vector orthogonal to $\vec{v}$. We can choose $\vec{v}$ and $\vec{w}$ such that they are both unit vectors and $\vec{w}$ is obtained from $\vec{v}$ by a counterclockwise rotation of $\pi/2$, namely,

$$\vec{v} = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}, \quad \text{and} \quad \vec{w} = \begin{pmatrix} \cos(\theta + \frac{\pi}{2}) \\ \sin(\theta + \frac{\pi}{2}) \end{pmatrix}\begin{pmatrix} \cos(\theta + \frac{\pi}{2}) \\ \sin(\theta + \frac{\pi}{2}) \end{pmatrix} = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix},$$

where θ is the angle of inclination of the line l. Then for any vector $\begin{pmatrix} x \\ y \end{pmatrix}$, we can find scalars a and b such that

$$\begin{pmatrix} x \\ y \end{pmatrix} = a\vec{v} + b\vec{w} = a\begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} + b\begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix},$$

from which we find

$$ref\begin{pmatrix}x\\y\end{pmatrix} = ref(a\vec{v}+b\vec{w}) = a\,ref(\vec{v})+b\,ref(\vec{w}) = a\,\vec{v}-b\,\vec{w} = \begin{pmatrix}\cos\theta & \sin\theta\\ \sin\theta & -\cos\theta\end{pmatrix}\begin{pmatrix}a\\b\end{pmatrix}.$$

Solving

$$\begin{pmatrix}a\\b\end{pmatrix} = \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}^{-1}\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}$$

and substituting back into the formula for ref, we find

$$\begin{aligned} ref\begin{pmatrix}x\\y\end{pmatrix} &= \begin{pmatrix}\cos\theta & \sin\theta\\ \sin\theta & -\cos\theta\end{pmatrix}\begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}\\ &= \begin{pmatrix}\cos^2\theta-\sin^2\theta & 2\sin\theta\cos\theta\\ 2\sin\theta\cos\theta & -\cos^2\theta+\sin^2\theta\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}\\ &= \begin{pmatrix}\cos(2\theta) & \sin(2\theta)\\ \sin(2\theta) & -\cos(2\theta)\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}, \end{aligned}$$

where we used $\cos^2\theta - \sin^2\theta = \cos(2\theta)$ and $2\sin\theta\cos\theta = \sin(2\theta)$. So we conclude that

$$ref\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}\cos(2\theta) & \sin(2\theta)\\ \sin(2\theta) & -\cos(2\theta)\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}$$

represents a reflection in the line l that passes through origin and has an inclination angle θ. Note that the matrix

$$\begin{pmatrix}\cos(2\theta) & \sin(2\theta)\\ \sin(2\theta) & -\cos(2\theta)\end{pmatrix}$$

is an orthogonal matrix, but with its determinant equal to -1.

Another way to derive the formula which represents the reflection across a straight-line through the origin with inclication angle θ is *to assume* that it is given by a multiplication by matrix

$$\begin{pmatrix}a & c\\ b & d\end{pmatrix},$$

then we expect

$$\begin{pmatrix}a\\b\end{pmatrix} = \begin{pmatrix}a & c\\ b & d\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} = \text{ the unit vector with inclination angle } 2\theta = \begin{pmatrix}\cos(2\theta)\\ \sin(2\theta)\end{pmatrix},$$

which is the reflection of $(1,0)$ with respect to the line through origin with inclination angle θ. Furthermore, we expect

$$\begin{pmatrix}a & c\\ b & d\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}$$

to be the unit vector whose angle of inclination angle is $2\theta - \frac{\pi}{2}$, as we expect it to be orthogonal to the reflection of $(1,0)$ but whose angle of inclination angle is $\frac{\pi}{2}$ less. Thus

$$\begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \cos(2\theta - \frac{\pi}{2}) \\ \sin(2\theta - \frac{\pi}{2}) \end{pmatrix} = \begin{pmatrix} \sin 2\theta \\ -\cos 2\theta \end{pmatrix},$$

and

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} \cos(2\theta) & \sin 2\theta \\ \sin(2\theta) & -\cos 2\theta \end{pmatrix}.$$

Recall that any 2×2 orthogonal matrix can be written in the form of

$$U = \begin{pmatrix} \cos\theta & \cos\phi \\ \sin\theta & \sin\phi \end{pmatrix}$$

for some angles θ and ϕ. When $\det U = -1$, we find that $\cos\theta \sin\phi - \sin\theta\cos\phi = -1$, which implies that $\phi - \theta = -\pi/2$ up to integer multiples of 2π. Therefore we can identify

$$U = \begin{pmatrix} \cos\theta & \cos(\theta - \pi/2) \\ \sin\theta & \sin(\theta - \pi/2) \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$$

to be a reflection in the line through origin with inclination angle $\theta/2$.

The composition of a translation t and a rotation rot can be worked out according to matrix rules:

$$rot \circ t \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x + x_0 \\ y + y_0 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix}$$

where

$$\begin{pmatrix} e \\ f \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix},$$

and

$$t \circ rot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}.$$

Note that in general $rot \circ t \neq t \circ rot$. The compositions between translations and reflections, and between rotations and reflections can also be worked in a similar way. The general pattern is that any such compositions can be written in the form of

$$U \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix}$$

for some orthogonal matrix U and vector $\begin{pmatrix} e \\ f \end{pmatrix}$. A transformation of $\mathbb{R}^2$ of this form is called a *Euclidean transformation*. It turns out that all isometries of $\mathbb{R}^2$ are

Euclidean transformations, so isometries and Euclidean transformations, as well as rigid motions, are synonyms.

Here are the key points of this section:

(1). All translations, rotations, reflections, and their compositions can be represented in the form of

$$t\begin{pmatrix}x\\y\end{pmatrix} = \mathbf{U}\begin{pmatrix}x\\y\end{pmatrix} + \begin{pmatrix}e\\f\end{pmatrix}$$

for some 2×2 orthogonal matrix $\mathbf{U} = \begin{pmatrix}a & b\\c & d\end{pmatrix}$. Transformations of this form are called Euclidean transformations.

(2). A 2×2 orthogonal matrix is either a rotation matrix or a reflection matrix.

(3). The compositions of Euclidean transformations are still Euclidean transformations.

(4). All Euclidean transformations are invertible, and their inverses are also Euclidean transformations.

(5). Every Euclidean transformation represents either a translation, or a rotation, or a reflection, or a glide reflection.

Remark. *The simplest glide reflection is given by*

$$gr\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}x+g\\-y\end{pmatrix} = \begin{pmatrix}1 & 0\\0 & -1\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix} + \begin{pmatrix}g\\0\end{pmatrix}$$

which represents reflection in the x-axis followed by a glide *along the x-axis by g (here the effect of gliding along the x-axis by g then followed by reflection in the x-axis produces the same effect).*

In general, both reflection and glide reflection are given by $\mathbf{Ux} + \mathbf{b}$, where $\mathbf{U}$ is a reflection matrix (an orthogonal matrix with determinant -1. check for yourself that a reflection matrix has 1 and -1 as eigenvalues*); but a reflection would leave the points along the axis of reflection invariant, namely there would exist a family of solutions $\mathbf{x}$ to $\mathbf{x} = \mathbf{Ux} + \mathbf{b}$, equivalently,*

$$(\mathbf{I} - \mathbf{U})\,\mathbf{x} = \mathbf{b}$$

*has a family of solutions. This happens when 1 is an eigenvalue of $\mathbf{U}$ and $\mathbf{b}$ is in the column space of $\mathbf{I} - \mathbf{U}$ (*review the linear algebra here!*). When $\mathbf{U}$ is a reflection matrix and $\mathbf{b}$ is* not *in the column space of $\mathbf{I} - \mathbf{U}$, the transformation above defines a glide reflection.*

Theorem 2 of **2.1** uses the concept of a *group*. Its definition is as follows: *any set G endowed with a "multiplication rule" $\circ$ is called a group, if (i) there is an element $e \in G$ such that $e \circ g = g \circ e = g$ for any $g \in G$, (ii) $(g_1 \circ g_1) \circ g_3 = g_1 \circ (g_2 \circ g_3)$ for all $g_1, g_2, g_3 \in G$, and (iii) for any $g \in G$, there is an element $g^{-1} \in G$ such that $g \circ g^{-1} = g^{-1} \circ g = e$.* Properties (3) and (4), together with the fact that the identity transformation

$$I \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$$

satisfies the property $I \circ t = t \circ I = t$ for any Euclidean transformation t shows that the set of all Euclidean transformations forms a group.

The following example illustrates how to compute the compositions and inverses of Euclidean transformations.

Example. *Let*

$$t_1 \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 1 \\ -2 \end{pmatrix},$$

$$t_2 \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} -3 \\ 0 \end{pmatrix}.$$

Then we can make the following conclusions:

- $\mathbf{U}_1 = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}$ *is a rotation matrix with rotation angle $\frac{\pi}{3}$, and* $\mathbf{U}_2 = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$ *is a reflection matrix, representing reflection in the line $y = (\tan(\frac{\pi}{6}))x$.*

- *We can compute*

$$\begin{aligned} t_2 \circ t_1 \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} t_1 \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} -3 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \left(\begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 1 \\ -2 \end{pmatrix} \right) + \begin{pmatrix} -3 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 \\ -2 \end{pmatrix} + \begin{pmatrix} -3 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \frac{1-2\sqrt{3}}{2} - 3 \\ \frac{\sqrt{3}+2}{2} + 0 \end{pmatrix}, \end{aligned}$$

and

$$\begin{aligned}
t_1 \circ t_2 \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} t_2 \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 1 \\ -2 \end{pmatrix} \\
&= \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \left(\begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} -3 \\ 0 \end{pmatrix} \right) + \begin{pmatrix} 1 \\ -2 \end{pmatrix} \\
&= \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} -3 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ -2 \end{pmatrix} \\
&= \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} -\frac{3}{2} + 1 \\ -\frac{3\sqrt{3}}{2} - 2 \end{pmatrix}.
\end{aligned}$$

So $t_1 \circ t_2 \neq t_2 \circ t_1$.

- *To compute* t_1^{-1}, *we solve* $\begin{pmatrix} x \\ y \end{pmatrix}$ *in terms of* $\begin{pmatrix} u \\ v \end{pmatrix}$ *through the relation*

$$\begin{pmatrix} u \\ v \end{pmatrix} = t_1 \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 1 \\ -2 \end{pmatrix},$$

to find

$$\begin{aligned}
\begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}^{-1} \left(\begin{pmatrix} u \\ v \end{pmatrix} - \begin{pmatrix} 1 \\ -2 \end{pmatrix} \right) \\
&= \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} - \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 \\ -2 \end{pmatrix} \\
&= \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} - \begin{pmatrix} \frac{1-2\sqrt{3}}{2} \\ -\frac{\sqrt{3}-2}{2} \end{pmatrix}.
\end{aligned}$$

It is customary to exchange of the roles of (x, y) *and* (u, v) *at this point, and conlude that*

$$t_1^{-1} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} \frac{1-2\sqrt{3}}{2} \\ -\frac{\sqrt{3}-2}{2} \end{pmatrix}.$$

- *To see that* t_1 *represents a rotation about some point* $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$, *we first note that* $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ *must be invariant under* t_1*:*

$$t_1 \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix},$$

which amounts to solving

$$\begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + \begin{pmatrix} 1 \\ -2 \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix},$$

or equivalently

$$\left[\begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} - I \right] \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = - \begin{pmatrix} 1 \\ -2 \end{pmatrix}.$$

The solvability of this system can be verified directly. In fact, there is a general argument for it: the rotation matrix with angle θ has eigenvalues $\cos\theta \pm i \sin\theta$, so, unless $\theta = 0$, 1 is not an eigenvalue of U and $U - I$ is invertible.

Now we can verify that t_1 is a rotation about $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$, if we can show that the action of t_1 satisfies

$$t_1 \begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = U \left(\begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \right)$$

for some rotation matrix U. This is indeed the case, as we have

$$\begin{aligned} t_1 \begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} &= \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 1 \\ -2 \end{pmatrix} - \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} - \begin{pmatrix} 1 \\ -2 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \left(\begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \right), \end{aligned}$$

i.e., t_1 represents a rotation with $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ as center of rotation.

Remark. *A word on the statement in the middle of p. 63,* every isometry [of $\mathbb{R}^2$] has one of the following forms:

- *a translation along a line in* $\mathbb{R}^2$
- *a reflection in a line in* $\mathbb{R}^2$*.*
- *a rotation about a point in* $\mathbb{R}^2$*.*
- *a glide reflection in* $\mathbb{R}^2$*.*

Let t be a Euclidean transformation of the form

$$x \mapsto U \begin{pmatrix} x \\ y \end{pmatrix} + \mathbf{b}$$

If U is the identity matrix, then t is a translation. If U is a rotation matrix but not the identity matrix, then, as discussed in the example above, 1 is not an eigenvalue of U, so there exists a vector $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ satisfying

$$(U - I) \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + \mathbf{b} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \text{ namely, } U \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + \mathbf{b} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}.$$

Then t is a rotation about this $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$. If U is a reflection matrix, and the system $U \begin{pmatrix} x \\ y \end{pmatrix} + \mathbf{b} = \begin{pmatrix} x \\ y \end{pmatrix}$ has a solution (then it will have infinitely many solutions. Give an argument by yourself), then t will be a reflection. Finally if U is a reflection matrix, but the system $U \begin{pmatrix} x \\ y \end{pmatrix} + \mathbf{b} = \begin{pmatrix} x \\ y \end{pmatrix}$ has no solution, then t is a glide reflection.

2.2 Affine Transformations and Parallel Projections

Orientation. *From a contemporary view point, we often need to take images of objects, process and store the images, and sometimes use the images for object recognition. What is crucial here is to understand the relation between an object and its image, in particular to understand what geometric features are preserved between the object and its image. We can think of the relation as a transformation t from the plane (containing the original) to the plane (containing the image), mapping the object to its image. Obviously t will not be a Euclidean transformation most of the time (how often do we take images of objects that keep the exact size as the objects?). Of course the description of t depends on how we take the images. We will study two methods of projections in this course: parallel projections and central projections. A parallel projection is constructed using a beam of parallel lights passing through two planes π_1 and π_2, sending each point P on π_1 to its image on π_2 under the light ray through P. A central projection is constructed using a beam of lights emitting from a central point which projects (almost) every point P on π_1 to a point on π_2. Our objectives in this and the next chapter are to understand (a) how to represent these projections algebraically? (b) the algebraic properties of such transformations, such as their compositions, inverses; (c) their geometric properties, in particularly, what kind of geometric features of the original are preserved by the image. For parallel projections, we will find that they (i) map straight lines to straight lines; (ii) map parallel* *straight lines to parallel straight lines; and (iii) preserve ratios of lengths along a given straight line. We will also find that they send conics to conics, but may send an ellipse to one with a different eccentricity, and may send a hyperbola to one with a different eccentricity, etc. For central projections, we will find that they still have property (i), but not (ii) and (iii); that they still send conics to conics, but may send an ellipse to a hyperbola, or a parabola, etc.*

Here we will study transformations which can be represented in the form of

$$t\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{A}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix},$$

for some invertible matrix $\mathbf{A}$. Such transformations are called affine transformations. Parallel projections and their compositions are examples of such transformations. In fact a parallel projection can be represented in the above form where the matrix $\mathbf{A}$ has a more strict requirement: it must be invertible and preserves the length of some non-zero vector. The reason is that if the two planes π_1 and π_2 are parallel themselves, then a parallel projection will be in fact a Euclidean transformation (think of π_2 as an identical copy of π_1 moved to the position of π_2), as it preserves the distances of any pair of points. When π_1 and π_2 are not parallel, they meet along a straight line l, then if we choose a point on l as the origin, the vectors in the direction of l are invariant under a parallel projection, so must be eigenvectors of the matrix $\mathbf{A}$ with eigenvalues 1. However, as soon as we do compositions of two parallel projections, we easily lose the property of preserving the length of some non-zero vector, but still maintain the property that $\mathbf{A}$ is invertible. The focus of this section is to study geometric properties which are invariant under affine transformations. These are listed on p. 73.

Note the relation and difference between Euclidean transformations and Affine transformations: all Euclidean transformations are Affine transformations, but *not* vice versa. Thus all Euclidean transformations also satisfy the mapping properties of affine transformations on p. 73. However, Euclidean transformations are much more rigid, as they preserve all aspects of size and shape, including angles, while a non-Euclidean affine transformation may not preserve the size and shape of geometric figures, though at least it will map a triangle to a triangle, a parallelogram to a parallelogram thanks to properties 1 and 2 on p. 73 (it could map an equilateral triangle into one which is not, or map a rectangle into a parallelogram which is not a rectangle). In addition, ratios of lengths *along any given straight line* is preserved. For example, suppose ΔABC is an equilateral triangle with vertices A, B, and C, and D is the midpoint between A and B. Let t be an affine transformation, then t maps ΔABC to $\Delta t(A)t(B)t(C)$. Since $AD : DB = 1$, it follows that $t(A)t(D) : t(D)t(B) = 1$, implying that $t(D)$ is the midpoint of $t(A)t(B)$. Note that even though $AB : AC = 1$, we can not conclude that $t(A)t(B) : t(A)t(C) = 1$, because AB and AC are not along the same straight line, so the ratio of their lengths may not be preserved.

2.2.1 Affine Transformations

The algebraic operational properties of affine transformations (computations of compositions and inverses) are very similar to those of the Euclidean transformations, with the exception of how to compute the inverse of the matrix in the representation of an affine transformation.

2.2.2 Parallel Projections

In this subsection, parallel projections and their mapping properties are described geometrically. Their mapping properties on p.75 are the same as those for affine transformations on p.73, and as explained in 2.2.3, they are affine transformations.

2.2.3 Affine Geometry

This subsection introduces a new and important method in geometry: if a property is *invariant* under affine transformations, and figure F_1 is mapped into figure F_2 by an affine transformation, then figure F_1 has this property iff F_2 has this property .

This method is illustrated by a result on Ellipses: Pick any chord l of an ellipse E, construct all chords of E parallel to l and construct the midpoints of these chords. Conclusions: these midpoints form a straight line going through the center of E.

Under an affine transformation t, the chords E parallel to l are mapped to chord of $t(E)$, and the midpoint of each of these chord becomes the midpoint of the corresponding image chord. The midpoints of the original family of parallel chords form a straight line iff the midpoints of the image chords form a straight line. Suppose there is an affine transformation t that maps E into a configuration where the desired property can be checked easily, then we can conclude that the same property holds for the original E. Here we want to make $t(E)$ into a circle, for which we can argue easily that the midpoints of parallel chords form a straight line passing through the center of the circle. So the last ingredient we need to check is: can any given ellipse be mapped to a circle by an affine transformation? The answer is yes. As a consequence, we can also state that any ellipse can be mapped into any other ellipse by an affine transformation.

This method will be used more in 2.4.

On pp.80–82, the text gives an argument to show that indeed any parallel projection is an affine transformation. Note, however, the **Observation** on p.82: An affine transformation is not necessarily a parallel projection. On the other hand, **Theorem 6** on p.84 says that *any affine transformation can be expressed as the composite of* *(at most) two parallel projections.* As a consequence, to verify properties for affine transformations, it suffices to verify them for parallel projections.

2.3 Properties of Affine Transformations

2.3.1 Images of Sets Under Affine Transformations

We went over the geometric descriptions already when we introduced affine transformations. To compute the images of sets under a concrete affine transformation

$$t\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{A}\begin{pmatrix} x \\ y \end{pmatrix} + \mathbf{b},$$

the key is to solve for $\begin{pmatrix} x \\ y \end{pmatrix}$ from

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \mathbf{A}\begin{pmatrix} x \\ y \end{pmatrix} + \mathbf{b}$$

to get

$$\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{A}^{-1}\begin{pmatrix} x' \\ y' \end{pmatrix} - \mathbf{A}^{-1}\mathbf{b},$$

so we could express both x and y in terms of x' and y'. Then if a set (a straight line, a conic, etc) S is expressed in terms of an equation in x and y, through substitution, we would find the equation that points (x', y') on the image of S must satisfy.

Example. *Suppose that*

$$t(\mathbf{x}) = \begin{pmatrix} 4 & 1 \\ 2 & 1 \end{pmatrix}\mathbf{x} + \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$

Note that

$$t\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \quad t\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \end{pmatrix} + \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \quad t\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$

Thus

$$\begin{pmatrix} 4 \\ 2 \end{pmatrix} = t\begin{pmatrix} 1 \\ 0 \end{pmatrix} - t\begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} = t\begin{pmatrix} 0 \\ 1 \end{pmatrix} - t\begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

namely, the first column vector of the coefficient matrix of t is the vector pointing from $t\begin{pmatrix}0\\0\end{pmatrix}$ to $t\begin{pmatrix}1\\0\end{pmatrix}$, and the second column vector is the vector pointing from $t\begin{pmatrix}0\\0\end{pmatrix}$ to $t\begin{pmatrix}0\\1\end{pmatrix}$. The inverse of t can be found as

$$t^{-1}(\mathbf{x}) = \begin{pmatrix}4 & 1\\2 & 1\end{pmatrix}^{-1}\mathbf{x} - \begin{pmatrix}4 & 1\\2 & 1\end{pmatrix}^{-1}\begin{pmatrix}2\\-1\end{pmatrix} = \begin{pmatrix}1/2 & -1/2\\-1 & 2\end{pmatrix}\mathbf{x} - \begin{pmatrix}3/2\\-4\end{pmatrix}.$$

To find the image of the line $y = 0$ under t, we write t^{-1} in terms of components

$$\begin{pmatrix}x'\\y'\end{pmatrix} = t^{-1}\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}\frac{1}{2}x - \frac{1}{2}y - \frac{3}{2}\\-x + 2y + 4\end{pmatrix}.$$

A point (x, y) is on the image of $y = 0$ under t iff (x', y') is on the line $y = 0$, namely, $y' = 0$, where $y' = -x + 2y + 4$ from above. Thus every point (x, y) on the image of $y = 0$ under t satisfies $-x + 2y + 4 = 0$.

A second method for finding the image of the line $y = 0$ under t is to note that $(0, 0)$ and $(1, 0)$ are two points on the line $y = 0$, and we know the image of the line $y = 0$ under t must be a straight line, so it must be the straight line passing through $t\begin{pmatrix}0\\0\end{pmatrix} = \begin{pmatrix}2\\-1\end{pmatrix}$ and $t\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}6\\1\end{pmatrix}$. We can therefore write down the equation of the image line through the two points formula as

$$(y + 1)/(x - 2) = (1 - (-1))/(6 - 2),$$

which turns out to be the same as $-x + 2y + 4 = 0$.

2.3.2 The Fundamental Theorem of Affine Geometry

The essence of **Theorem 1** on p. 88 is that an affine transformation is determined by its action on a set of three non-collinear points. As a consequence, if a photo is processed by an affine transformation, and we know its action on three non-collinear points, then the image of all other points are already determined and can be calculated. **Example 2** on p. 86 explains how this is done directly. **Example 3** on p. 89 explains how to redo this through intermediate affine transformations which map vertices of a given triangle to those of the standard triangle. An advantage of this latter method is that one can write out the solution in a fairly routine way.

2.3.3 Proofs of the Basic Properties of Affine Transformations

If we accept **Theorem 6** on p.84 that any affine transformation can be expressed as the composite of two parallel projections, then the properties for affine transformations follow from those for parallel projections, for which we have provided geometric proofs.

To give direct proofs for the properties for affine transformations, we need to work with the algebraic expression for the inverse t^{-1} of an affine transformation t. Suppose that

$$t^{-1}\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x' \\ y' \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix}.$$

Then, if l is the straight line given by $\alpha x + \beta y + \gamma = 0$, and (x', y') is the image of a point (x, y) on the line l, x' and y' would satisfy the equation

$$\alpha(ax' + by' + e) + \beta(cx' + dy' + f) + \gamma = 0,$$

which represents an equation for a straight line. Thus the image of a straight line under an affine transformation is still a straight line. The other properties can be proved along similar lines.

Another approach is to use more of vector operations. One useful fact is the following: *if C is a point on the straight line through points A and B, then we can represent $C = (1 - \lambda)A + \lambda B$, where we think of A, B and C as vectors; moreover,* $\vec{AC} : \vec{CB} = \lambda : (1 - \lambda)$. For example, the point C that lies on the straight line through points A and B and divides AB into $2 : 1$ portion must corresponds to the point $C = (1 - \lambda)A + \lambda B$, where $\lambda : (1 - \lambda) = 2 : 1$. Thus $\lambda = 2/3$, and $C = \frac{1}{3}A + \frac{2}{3}B$. The same formula can also be interpreted when C lies on the straight line through points A and B but not between A and B. For example, if C is on the extension of AB such that $|AC| = 2|BC|$, because the vectors $\vec{AC}$ and $\vec{CB}$ now point in opposite directions, we write the relation as $AC : CB = 2 : (-1)$. In order for $\lambda : (1 - \lambda) = 2 : (-1)$, we must choose $\lambda = 2$. Thus $C = -A + 2B$. A more geometric way to interpret $C = (1 - \lambda)A + \lambda B$ is to rewrite it as

$$\vec{AC} = C - A = (1 - \lambda)A + \lambda B - A = \lambda(B - A) = \lambda\vec{AB},$$

which means that $\vec{AC}$ is λ multiples of $\vec{AB}$.

With this algebraic fact, we can easily prove that if t is an affine transformation, A, B and C are three points on a straight line, and if $A' = t(A)$, $B' = t(B)$, and $C' = t(C)$, then $AC : CB = A'C' : C'B'$. This can be done after we assume $t(\mathbf{x}) = \mathbf{U}x + \mathbf{b}$ for some invertible matrix U and vector $\mathbf{b}$, and $AC : CB = \lambda : (1 - \lambda)$ for some λ, then $C = (1 - \lambda)A + \lambda B$, so

$$C' = t(C) = (1-\lambda)\mathbf{U}A + \lambda\mathbf{U}B + \mathbf{b} = (1-\lambda)\left[\mathbf{U}A + \mathbf{b}\right] + \lambda\left[\mathbf{U}B + \mathbf{b}\right] = (1-\lambda)A' + \lambda B',$$

which implies that $C' = t(C)$ lies on the straight line through $A' = t(A)$ and $B' = t(B)$, and $A'C' : C'B' = \lambda : (1 - \lambda) = AC : CB$.

2.4 Using the Fundamental Theorem of Affine Geometry

We have seen some basic usage of affine transformations to deal with geometric problems in **2.2.3**. This subsection gives more examples of such applications. Although proofs for **Theorems** 2 and 4 are long, their basic ideas are simple: (i) The properties being studied are affine invariant, so we hope to reduce a general situation to a special situation where we can carry out the study by either concrete computations or symmetry arguments; (ii) For problems involving triangles, the Fundamental Theorem of Affine Geometry guarantees that any triangle can be mapped to the *standard* triangle with $(0,0)$, $(1,0)$, and $(0,1)$ as vertices, thus it suffices to examine only the situation with this triangle.

One needs to note that the ratios in this subsection are all *signed ratios*, i.e., if P is a point along the straight line through A and B, but not between A and B, then AP and PB are pointing in opposite directions, so we take $AP : PB$ as a negative number.

Here is an example of an application of the converse to Ceva's Theorem.

Example. *For any triangle ΔABC, let the angle bisectors through A, B, and C intersect their respective opposite sides at P, Q, and R. Then AP, BQ, and CR are concurrent.*

The property to be proved here are apparently relevant to affine transformations. Note that even though the angle bisecting property may not be preserved by affine transformations, the concurrency of AP, BQ, and CR would be guaranteed by the converse to Ceva's Theorem if we can verify that

$$\frac{BP}{PC} \cdot \frac{CQ}{QA} \cdot \frac{AR}{RB} = 1.$$

Note that ΔABP and ΔAPC share the same heights, so $BP : PC = Area(\Delta ABP) : Area(\Delta APC)$. On the other hand,

$$Area(\Delta ABP) = \frac{1}{2}|AB||AP| \sin\angle BAP, \quad \text{and} \quad Area(\Delta APC) = \frac{1}{2}|AP||AC| \sin\angle PAC.$$

Since AP is an angle bisector of $\angle BAC$, we know $\angle BAP = \angle PAC$, thus $Area(\Delta ABP) : Area(\Delta APC) = |AB| : |AC|$, and $BP : PC = |AB| : |AC|$. By the same reasoning, $CQ : QA = |BC| : |BA|$, and $AR : RB = |AC| : |BC|$, so we find

$$\frac{BP}{PC} \cdot \frac{CQ}{QA} \cdot \frac{AR}{RB} = \frac{|AB|}{|AC|} \cdot \frac{|BC|}{|BA|} \cdot \frac{|AC|}{|BC|} = 1,$$

and conclude that AP, BQ, and CR are concurrent.