

## Chapter 2. Quadric figures

In this chapter want to outline quickly how to decide what figure associated in  $2D$  and  $3D$  to quadratic equations look like. In any number of dimensions, a **quadratic** equation is one of the form

$$\sum a_{i,j}x_i x_j + \sum b_i x_i + c = 0$$

or in other words one where all terms have degree at most two. In addition we shall ignore the degenerate case where there are no terms of degree two. In the first Chapter, I have defined the conic sections in terms of simple quadratic equations, and in dealing with the characterization in terms of focus and directrix we have looked at translations of these equations. The key to our discussion in this Chapter will be *making coordinate changes so the equation has a particularly simple form*.

As an application, we will also discuss **conic sections**, the curves we get by intersecting circular cones with planes. One application will be to see why it is that when you look at a circle in space in any position at all, what you really see is an ellipse. This is not at all an evident fact, as simple sketches should demonstrate.

For the sake of simplicity, I shall often assume we are working in  $2D$ . Only the algebra becomes more complicated in higher dimensions—the basic concepts remain the same.

### 1. Coordinate changes

One of the main basic points is that quadratic equations become simple if we make a suitable change of coordinates. Since we don't want to change the shape of things we are looking at, we shall usually restrict ourselves to **orthogonal** changes of coordinates. This means that if  $(x, y)$  are the original coordinates and  $(x_*, y_*)$  the new coordinates for the same point, the distances between two points are given by

$$\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} = \sqrt{(x_{*,2} - x_{*,1})^2 + (y_{*,2} - y_{*,1})^2}.$$

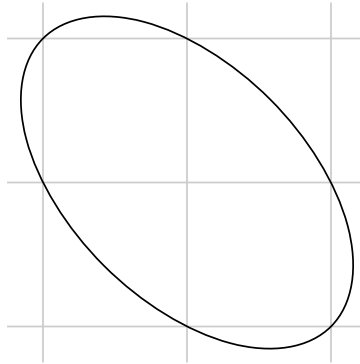
Even with this restriction we can always simplify quadratic equations to a number of simple types, so that in the end we shall see that quadratic equations, with a few simple exceptions, define ellipses, parabolas, or hyperbolas. I have essentially *defined* these to be the curves defined in an orthogonal coordinate system by equations

$$\begin{aligned} \left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 &= 1 \\ \left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 &= 0 \\ \left(\frac{x}{a}\right)^2 - \left(\frac{y}{b}\right)^2 &= 1 \end{aligned}$$

Let me start with an example. Consider the curve

$$x^2 + xy + y^2 = 1$$

You can guess by plotting a few points that this curve looks like this:



In other words, it looks like an ellipse whose major axis is along the line  $y = -x$  and whose minor axis is along the line  $y = x$ . Some calculation will tell you that the size of the semi-major axis is  $\sqrt{2}$  and that of the semi-minor axis is  $\sqrt{2/3}$ . We can therefore say that this is the ellipse

$$\frac{3x_*^2}{2} + \frac{y_*^2}{2} = 1$$

if we choose our coordinate system to match the axes  $y = \pm x$ . This is the sort of thing we want to generalize.

The first question to be answered is *what happens to quadratic equations when we make a coordinate change?* We will allow at first the most general kind of coordinate change. Such a change is effected by an **affine transform**, and can be expressed as a combination of a **linear** change of coordinates

$$\begin{bmatrix} x \\ y \end{bmatrix} = A \begin{bmatrix} x_* \\ y_* \end{bmatrix}$$

where  $A$  is an invertible  $2 \times 2$  matrix, with a **translation**

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_* \\ y_* \end{bmatrix} + \begin{bmatrix} \tau_x \\ \tau_y \end{bmatrix}.$$

In the example above, we make a linear change where  $A$  is the rotation matrix

$$\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}, \quad \theta = 45^\circ.$$

Why is this? Because the transformation of **coordinate frames** from one coordinate system to the other is by this rotation.

So we analyze how quadratic equations change in two phases. In both phases *the main trick is to write a homogeneous quadratic expression in terms of matrix multiplication:*

$$ax^2 + bxy + cy^2 = [x \ y] \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = [x \ y] Q \begin{bmatrix} x \\ y \end{bmatrix}$$

as you can calculate yourself. Here the important thing is that  $Q$  is a **symmetric** matrix.

Similarly we can also write the linear expression in a quadratic function in terms of matrix multiplication:

$$ax + by = [a \ b] \begin{bmatrix} x \\ y \end{bmatrix} = \ell \begin{bmatrix} x \\ y \end{bmatrix}.$$

where  $\ell$  is a row vector. If we change coordinates we change  $Q$  to

$$Q_* = {}^tAQA$$

and we change  $\ell$  to

$$\ell_* = \ell A .$$

Suppose we translate coordinates

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_* \\ y_* \end{bmatrix} + \begin{bmatrix} \tau_x \\ \tau_y \end{bmatrix} .$$

Then the homogeneous function corresponding to  $Q$  becomes the expression with the same quadratic term  $Q$ , but linear term

$$[\tau_x \quad \tau_y]$$

and constant term

$$[\tau_x \quad \tau_y] Q \begin{bmatrix} \tau_x \\ \tau_y \end{bmatrix}$$

**Exercise 1.1.** Under translation, what happens to a linear function corresponding to the row vector  $\ell$ ?

## 2. The basic classification

Since  $Q$  is a symmetric matrix we can **diagonalize** it. Recall from your course on linear algebra that it has **real** eigenvalues (as opposed to ones which are possibly complex) and that we can choose its eigenvectors to be perpendicular to each other. Furthermore we can normalize the eigenvectors to have length 1. In other words we can find an **orthogonal** matrix  $A$  such that

$$A^{-1}QA = \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}$$

where  $\lambda$  and  $\mu$  are the eigenvalues of  $Q$  and the columns of  $A$  are its eigenvectors. Recall that an orthogonal matrix  $A$  is one such that  ${}^tA = A^{-1}$ . Equivalently, its columns are all perpendicular to each other and have length 1. In other words we can make a coordinate change so the quadratic expression becomes

$$\lambda x_*^2 + \mu y_*^2 .$$

The matrix associated to

$$x^2 + xy + y^2$$

is

$$\begin{bmatrix} 1 & 1/2 \\ 1/2 & 1 \end{bmatrix}$$

with eigenvalues  $1/2, 3/2$  and eigenvectors

$$\begin{bmatrix} 1/\sqrt{2} \\ \pm 1/\sqrt{2} \end{bmatrix} .$$

As a result of the first phase, we have changed our quadratic expression into one of the form

$$ax^2 + by^2 + cx + dy + e .$$

What we do next depends on what the expression we get. We are assuming that we are not dealing with a linear expression, in which case it cannot be that both  $a$  and  $b$  vanish. We can swap  $x$  and  $y$  if necessary to make  $b \neq 0$ . If we set  $y = y_* + t$ , we get an expression

$$\begin{aligned} ax^2 + b(y_* + t)^2 + cx + d(y_* + t) + e &= ax^2 + by_*^2 + 2bty_* + bt^2 + cx + dy_* + e \\ &= ax^2 + by_*^2 + cx + (2bt + d)y_* + bt^2 + dt + e \end{aligned}$$

and if we choose

$$2bt + d = 0, \quad t = -d/2b$$

(possible since  $b \neq 0$ ) we can make the linear term in  $y$  vanish. If  $a = 0$  we now have an expression of the form

$$cx + by^2 + d.$$

And if  $a \neq 0$  we can similarly translate  $x$  as well to obtain an expression of the form

$$ax^2 + by^2 + d.$$

We can summarize:

- If

$$f(x, y) = ax^2 + bxy + cy^2 + dx + ey + f$$

is any quadratic expression, then we combine an orthogonal change of coordinates with a translation of coordinates to make the expression into one of the two forms

$$a_*x_*^2 + b_*y_* + c_*$$

$$a_*x_*^2 + b_*y_*^2 + c_*.$$

As a corollary, we see that any curve of the form

$$f(x, y) = 0$$

is equivalent to one of the form

$$ax^2 + by + c = 0$$

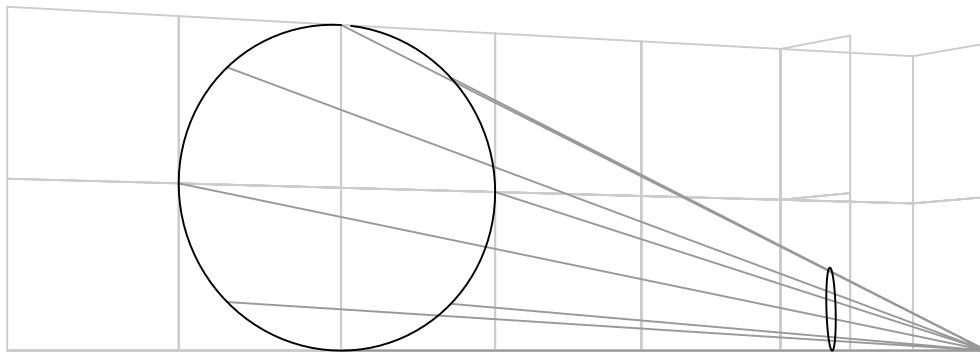
or

$$ax^2 + by^2 + c = 0.$$

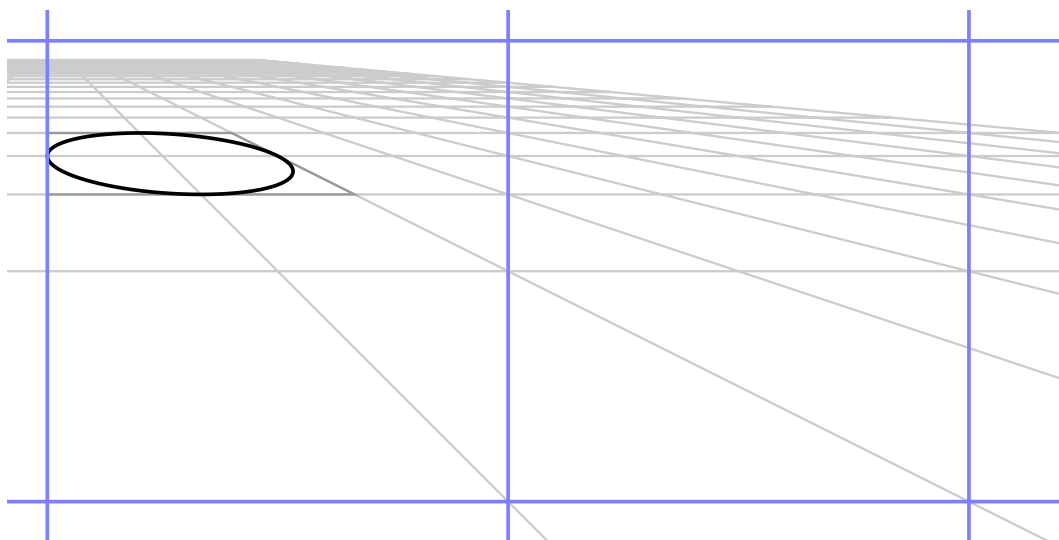
**Exercise 2.1.** Among the curves we get will be parabolas, ellipses, and hyperbolas. What else?

### 3. Circles in perspective

Suppose we place a circle somewhere in space, placing the extra condition on it that the plane containing it does not pass through the origin. Collect together all lines passing through pairs  $O, P$  where  $O$  is the origin and  $P$  is a point on the circle. These lines will make up a kind of oblique cone whose vertex is at the origin. For example, if the circle we start with is the circle  $x^2 + y^2 = R^2$  of radius  $R$  on the plane  $z = 1$  we get the *right circular cone*  $z^2 = (x^2 + y^2)/R^2$ . This cone has the property that if we slice it with any 'horizontal' plane  $z = c$  we get a circle (whose radius grows proportionally to  $c$ ). The problem we want to discuss here is: *if in the general case we slice the cone by the plane  $z = 1$ , what geometrical figure do we get?*



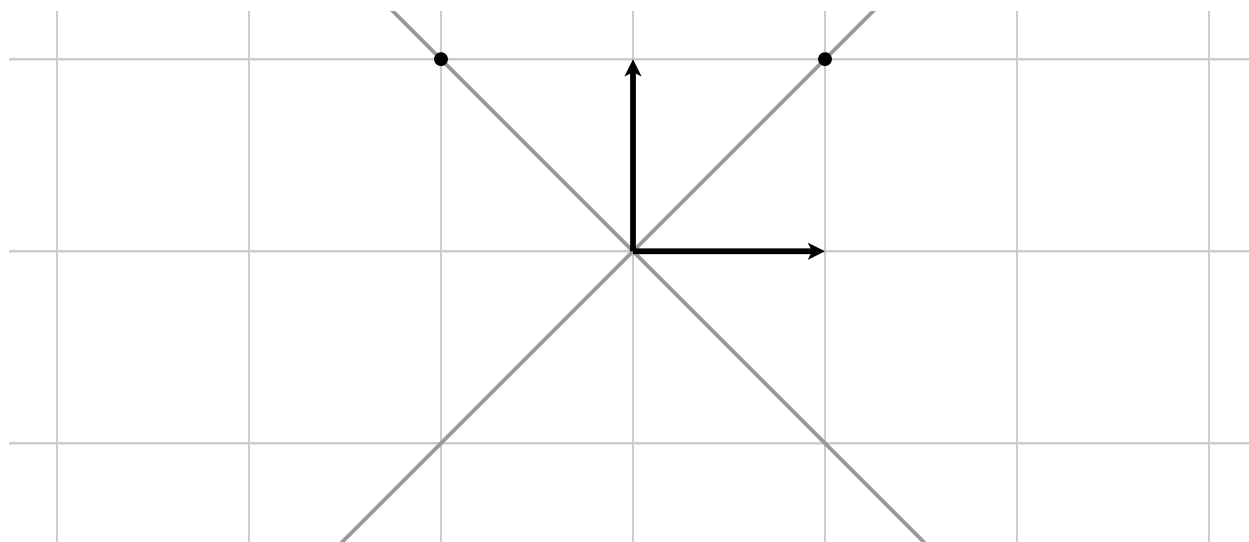
This corresponds very closely to the psychological problem of determining what a circle located somewhere in space looks like, since the image of the circle is the figure traced out by all the light rays from the circle passing through the lens of the eye (the origin) hitting the back of the eye (roughly the plane  $z = -h$  for some small  $h$ ). This optical image is essentially the figure on the plane  $z = 1$ , scaled down by  $h$ , and inverted.



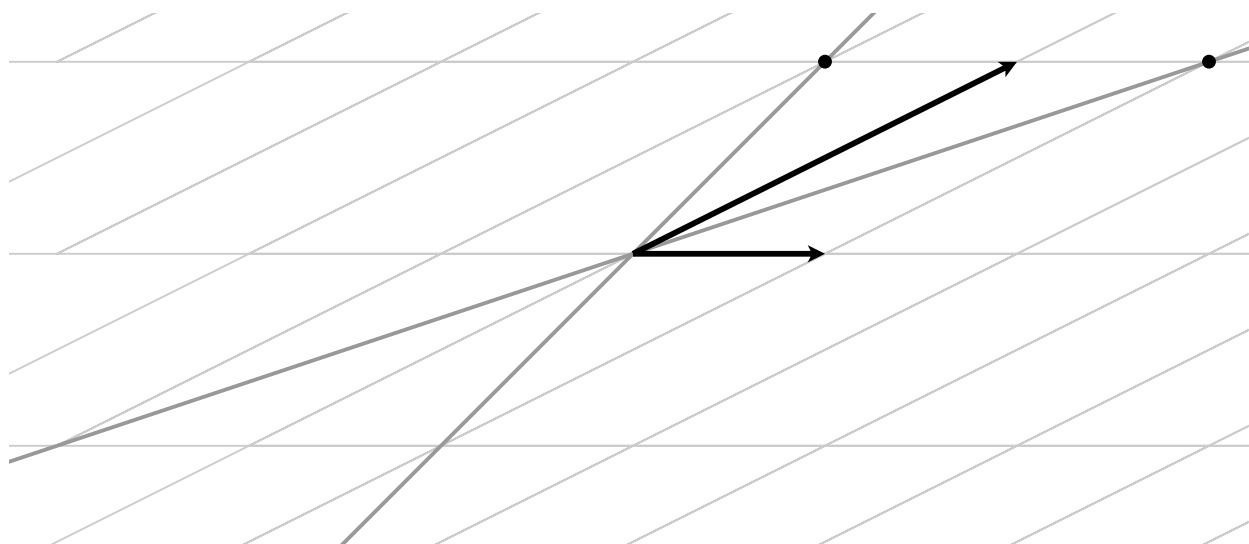
There are several steps in figuring out the solution. The point of the first few steps is to obtain the equation of the cone. We then set  $z = 1$  in this equation to get the equation of the slice, and apply the techniques described above (eigenvalues, etc.) to see what this curve looks like.

- We find the equation of the cone by applying an extremely fundamental principle which can be applied in far more general circumstances. The principle is to *choose a coordinate system in which the equation of the cone is simple*, and then to figure out how these coordinates relate to the ones we start with. This principle has been implicitly encountered already, where we chose a coordinate system related to eigenvectors to see the real shape of a quadric curve. But the new feature here, one that is in fact quite natural in spite of being new, is that the coordinate system we choose is *oblique*, that is to say not at all orthogonal.

An analogous example in  $2D$  should clarify things. Consider the  $2D$  'cone' which looks like this:



We think of it as being constructed in the following way: we choose the two points  $(-1, 1)$  and  $(1, 1)$  and make our 'cone' by drawing lines to the origin through them. Of course the lines we draw are  $y = \pm x$  and the equation of the cone made up of these two lines is the quadratic equation  $x^2 - y^2 = 0$ . Now suppose we choose two different points, say  $(1, 1)$  and  $(3, 1)$ . What is the equation of this cone? This is such a simple problem that there are several answers, but the one that turns out to be most useful is to choose a coordinate system in which these points have the same values as in the previous problem. That coordinate system is indicated in this picture:



Let  $(x_*, y_*)$  be the coordinates in this picture. The frame for this coordinate system is made up of the vectors  $f_1 = (1, 0)$  and  $f_2 = (2, 1)$ . The relationship between the two coordinate systems is given, according to what we have seen before, by the formula

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_* \\ y_* \end{bmatrix} = F \begin{bmatrix} x_* \\ y_* \end{bmatrix}$$

where the columns of the matrix  $F$  are the vectors  $f_1$  and  $f_2$ . Now the quadratic expression  $x_*^2 - y_*^2$  can be written in terms of matrix multiplication as

$$x_*^2 - y_*^2 = [x_* \quad y_*] \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_* \\ y_* \end{bmatrix}$$

which is the same as

$$[x \quad y] {}^tF^{-1} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} F^{-1} \begin{bmatrix} x \\ y \end{bmatrix}$$

since

$$\begin{bmatrix} x \\ y \end{bmatrix} = F^{-1} \begin{bmatrix} x_* \\ y_* \end{bmatrix}, \quad [x \quad y] = [x_* \quad y_*] {}^tF^{-1}.$$

Here

$$F = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, \quad F^{-1} = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$$

so the expression in  $(x, y)$  is

$$\begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix}$$

which corresponds to the quadratic expression

$$x^2 - 4xy + 3y^2.$$

So the equation of our 'cone' in the  $(x, y)$  coordinate system is

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

You can check this by seeing if the proper points are on it.

This is a lot of work to write down the equation for a pair of lines through the origin, but the principles remain the same in any number of dimensions. Here is what we do for a cone through a circle of radius  $R$  in space: (1) Choose a coordinate system in which the equation of the cone is  $z^2 = x^2 + y^2$ . The frame for this coordinate system will be made up of three vectors  $f_1, f_2, f_3$  with these properties: (a) The vector  $f_3$  has its head at the centre of the circle. (b) The two vectors  $f_1$  and  $f_2$  have length  $R$ , are perpendicular to each other, and lie in the plane through the origin parallel to the plane in which the circle lies. (2) If

$$F = [f_1 \quad f_2 \quad f_3]$$

Then the symmetric matrix corresponding to the quadratic expression we are looking for is

$${}^tF^{-1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} F^{-1}$$