A rotated ellipse from three points

Jerry R. Van Aken

September 28, 2018

The Demo1 program in this repository calls the DrawEllipse function to draw a rotated ellipse. The function takes as its input arguments the following three points:

> $P_0 = (x_0, y_0)$ is the center of the ellipse.
>
> $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$ are the end points of two conjugate diameters of the ellipse.

These three points are sufficient to describe an ellipse of any shape and orientation.

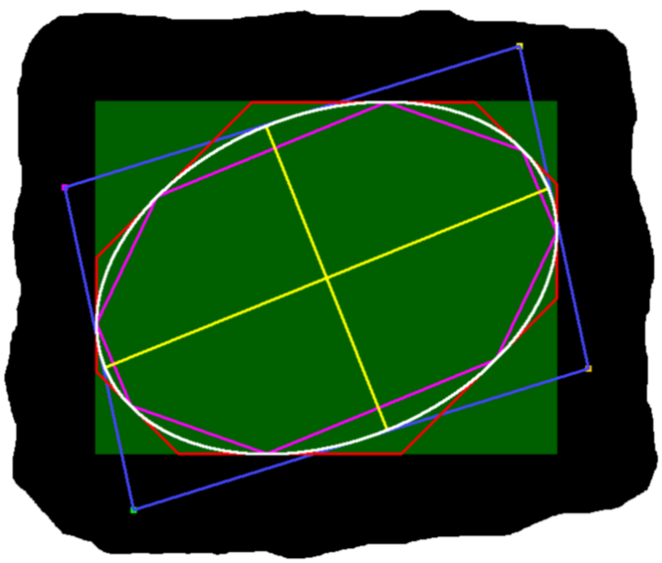

Figure 1: Demo1 screenshot

Figure 1 is a screenshot of an ellipse drawn by the Demo1 program. The ellipse is surrounded by several additional figures, each of which is constructed from the same three points, P_0, P_1, and P_2, that are used to construct the ellipse. These figures are as follows:

- A parallelogram in which the ellipse is inscribed. In Figure 1, the parallelogram is outlined in blue, and its vertexes are highlighted. The ellipse touches the parallelogram at the midpoint of each side of the parallelogram and is tangent to the side at that point.

- The major and minor axes of the ellipse. In Figure 1, these axes are the perpendicular yellow lines that pass through the center of the ellipse. A graphics library might need to determine the orientation of the major axis, for example, to properly orient text drawn inside the ellipse. The bounding rectangle for the ellipse is easily determined from the major and minor axes.

- The minimum bounding box for the ellipse. In Figure 1, the bounding box is shown as a filled green rectangle. The sides of the box are axis-aligned.

- An 8-sided minimum bounding polygon for the ellipse with two horizontal sides, two vertical sides, and four ± 45-degree diagonal sides. In Figure 1, this polygon is outlined in red. The ellipse touches the polygon at one point on each of its sides and is tangent to the side at that point.

- An 8-sided polygon inscribed inside the ellipse. In Figure 1, this polygon is outlined in magenta. Each vertex of the (magenta) inscribed polygon is a point on the ellipse at which the ellipse touches and is tangent to a side of the (red) bounding polygon.

This document explains how to construct each of these figures from the three points, P_0, P_1, and P_2, that specify the ellipse. It also explains how to derive the coefficients of the implicit equation for the ellipse from these three points. In the following sections, these coefficients are used to calculate the parameters for the figures listed above.

1 Drawing octants

In Figure 1, the triangular spaces between the (red) bounding polygon and (magenta) inscribed polygon neatly frame the octants traversed by Pitteway's algorithm, which the Demo1 program in this repository uses to draw ellipses. As described by Pitteway [1] and Foley, et al [2], the inner loop of this algorithm traverses a single octant without intervention, but must detect octant changes so that the algorithm can exit the inner loop temporarily to set up the inner loop parameters that will be used to draw the next octant.

Octant changes occur at points on the ellipse where the tangent is vertical, horizontal, or at a ± 45-degree angle. Figure 2 shows the numbers assigned to the eight octants in the Demo1 source code. Octant changes occur at the points shown in the figure as small hollow circles. In Figure 1, these points are the vertexes of the (magenta) inscribed polygon; and they are the points where the ellipse touches the (red) bounding polygon.

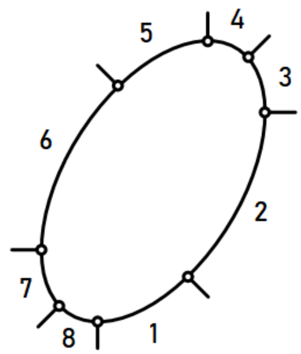

Figure 2: Octants traversed by ellipse-drawing algorithm

To detect octant changes, the inner loop of Pitteway's algorithm uses simple integer operations that are fast but that can fail to accurately track the boundary of an ellipse that is extremely thin [1][2]. This behavior potentially limits its usefulness in general-purpose graphics environments that require, for example, precise clipping of ellipses to drawing regions.

However, the information that is used to construct the bounding and inscribed polygons in Figure 1 might also be used between inner loops of the graphics algorithm to determine in advance the extent of the next octant to be traversed by the algorithm. In this case, a simple loop counter could be used to inform the inner loop of a pending octant change. This scheme might succeed in reliably keeping the algorithm confined to the triangular regions between the bounding and inscribed polygons.

This is an area for future research.

2 Bounding parallelogram from three points

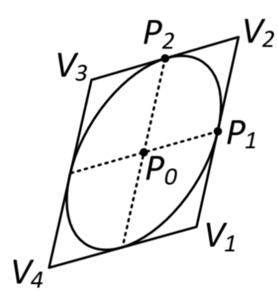

Figure 3: Bounding parallelogram

Figure 3 shows a rotated ellipse inscribed in a parallelogram. The ellipse is specified by three points: its center point P_0 and two conjugate diameter end points, P_1 and P_2. The ellipse touches and is tangent to the bounding parallelogram at the midpoint of each of the four sides of the parallelogram. P_1 and P_2 are the midpoints of two adjacent sides of the parallelogram.

Specifying an ellipse with the three points P_0, P_1, and P_2 is equivalent to specifying the ellipse in terms of the corresponding bounding parallelogram. In Figure 3, the four parallelogram vertexes, labeled V_1 through V_4, can be calculated from P_0, P_1, and P_2 using simple vector arithmetic:

$$
\begin{aligned}
V_1 &= P_1 + (P_0 - P_2) \\
V_2 &= P_1 + (P_2 - P_0) \\
V_3 &= P_2 + (P_0 - P_1) \\
V_4 &= V_1 + (V_3 - V_2) \\
&= 3P_0 - P_1 - P_2
\end{aligned}
$$

3 Implicit equation for ellipse

In this section, we will solve for the implicit equation of an ellipse given the end points P and Q of two conjugate diameters of the ellipse. To simplify our calculations, we translate the center of the ellipse to the origin. In terms of the three points P_0, P_1, and P_2 previously discussed, the translated end points of the conjugate diameters are $P = P_1 - P_0$ and $Q = P_1 - P_0$.

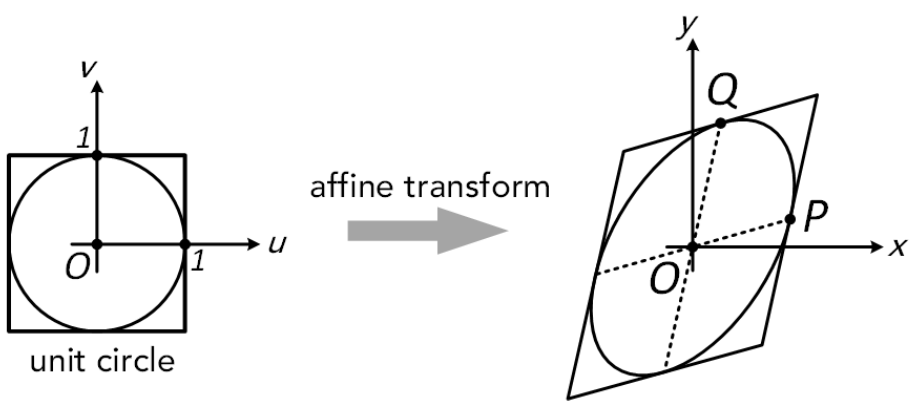

Figure 4: Affine transformation of unit circle to ellipse

Figure 4 shows that an ellipse and its bounding parallelogram are affine transformations of a unit circle and its bounding square. We use u-v coordinates for the unit circle, and x-y coordinates for the ellipse. Points $(1, 0)$ and $(0, 1)$ on the unit circle are transformed to points $P = (x_P, y_P)$ and $Q = (x_Q, y_Q)$ on the ellipse.

The parameteric equations for the unit circle are

$$
\begin{aligned}
u(t) &= \cos t \\
v(t) &= \sin t
\end{aligned}
$$

for $0 \leq t < 2\pi$.

The first step is to solve for the coefficients m_{ij} of the matrix $\mathbf{M}$ that performs the affine transformation from a point (u, v) on the unit circle to a point (x, y) on the ellipse:

$$
\begin{aligned}
\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} &= \mathbf{M} \begin{bmatrix} u(t) \\ v(t) \end{bmatrix} \\
&= \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} \begin{bmatrix} \cos t \\ \sin t \end{bmatrix}
\end{aligned}
$$

To solve for matrix $\mathbf{M}$, we use the equations $P = \mathbf{M} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $Q = \mathbf{M} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, which we expand as follows:

$$
\begin{aligned}
\begin{bmatrix} x_P \\ y_P \end{bmatrix} &= \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} m_{11} \\ m_{21} \end{bmatrix}
\end{aligned}
$$

and

$$
\begin{aligned}
\begin{bmatrix} x_Q \\ y_Q \end{bmatrix} &= \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\
&= \begin{bmatrix} m_{12} \\ m_{22} \end{bmatrix}
\end{aligned}
$$

From inspection, we see that

$$
\mathbf{M} = \begin{bmatrix} x_P & x_Q \\ y_P & y_Q \end{bmatrix}
$$

We can now express the parametric equations for the ellipse as

$$
\begin{aligned}
x(t) &= x_P \cos t + x_Q \sin t \\
y(t) &= y_P \cos t + y_Q \sin t
\end{aligned}
$$

We use these two equations to solve for $\cos t$ and $\sin t$, and obtain

$$
\begin{aligned}
\cos t &= \frac{x(t)/x_Q - y(t)/y_Q}{x_P/x_Q - y_P/y_Q} \\
\sin t &= \frac{x(t)/x_P - y(t)/y_P}{x_Q/x_P - y_Q/y_P}
\end{aligned}
$$

To obtain the implicit equation for the ellipse, we substitute these two expressions into the trigonometric identity $\sin^2 t + \cos^2 t = 1$. After simplifying and rearranging terms, we have the following form of the implicit equation for the ellipse:

$$f(x, y) = Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0 \tag{1}$$

$$\text{where} \qquad \begin{aligned} A &= y_P^2 + y_Q^2 \\ B &= -2(x_P y_P + x_Q y_Q) \\ C &= x_P^2 + x_Q^2 \\ D &= E = 0 \\ F &= -(x_P y_Q - x_Q y_P)^2 \end{aligned}$$

If we are given the values of coefficients A and C for a particular ellipse, and also given an end point $P = (x_P, y_P)$ of a diameter of that ellipse, we can use the preceding expressions for A and C to solve for the end points $Q = (x_Q, y_Q)$ of the corresponding conjugate diameter.

4 Minimum bounding box

In this section, we determine the x and y limits of the minimum bounding box of an ellipse. The bounding box is an axis-aligned rectangle. As in the previous section, we are given the end points P and Q of a pair of conjugate diameters of the ellipse. To simplify the calculations, the center of the ellipse is located at the origin.

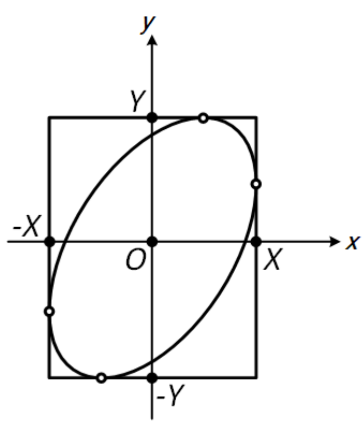

Figure 5: Minimum bounding box for an ellipse

Figure 5 shows the minimum bounding box for a rotated ellipse. The ellipse touches the box at one point on each of the four sides of the box. At each of these points (shown as small hollow circles), the ellipse is tangent to the corresponding horizontal or vertical edge of the box.

To find the x coordinates at the left and right sides of the bounding box, we solve for the y coordinates at the intersection of the ellipse with a vertical line $x = X$, where X is the x coordinate at the left or right side of the bounding box that we wish to find. We plug X into equation (1) to obtain:

$$f(X, y) = AX^2 + BXy + Cy^2 + F = 0$$

We rearrange this equation as follows:

$$(C)y^2 + (BX)y + (AX^2 + F) = 0$$

To solve for y, we plug these coefficients (in parentheses in the preceding equation) into the quadratic formula to obtain:

$$y = \frac{-BX \pm \sqrt{B^2X^2 - 4C(AX^2 + F)}}{2C} \tag{2}$$

For some vertical line $x = X$ that traverses the interior of the ellipse, equation (2) yields two intersections, at y_+ and y_-, with the line. For the bounding box, however, we're interested only in the single point at which the line touches the ellipse boundary (so that $y_+ = y_-$). Equation (2) yields a single y value only if the expression inside the square-root operation is zero; that is, if

$$B^2X^2 - 4C(AX^2 + F) = 0$$

Solving this equation for X, we obtain an equation for the x coordinates at the left and right sides of the bounding box:

$$X = \pm\sqrt{\frac{4CF}{B^2 - 4AC}}$$

For convenience, we can express X directly in terms of the conjugate diameter end points P and Q. By substituting the expressions for coefficients A, B, and C from equation (1) into the preceding equation, and simplifying, we obtain

$$X = \pm\sqrt{x_P^2 + x_Q^2} \tag{3}$$

This equation provides two solutions, X_+ and X_-, which are the x coordinates at the right and left edges of the bounding box.

As previously discussed, equation (2) yields a single value for y at the point where the ellipse touches the right or left side of the bounding box, in which case the quantity in the square-root operation is zero. The y coordinate at this point is $y = -\frac{BX}{2C}$. Or, expressed in terms of conjugate diameter end points P and Q, we have

$$y = \frac{x_P y_P + x_Q y_Q}{X} \tag{4}$$

where we have made use of the fact that $C = X^2$.

Using similar methods to obtain the y coordinates at the top and bottom of the bounding box, we get the result

$$Y = \pm\sqrt{y_P^2 + y_Q^2} \tag{5}$$

This equation provides two solutions, Y_+ and Y_-, which are the y coordinates at the bounding box's top and bottom edges, respectively.

Similarly, the x coordinate at the point where the ellipse touches the top or bottom of the bounding box is $x = -\frac{BY}{2A}$. Or, expressed in terms of conjugate diameter end points P and Q, we have

$$x = \frac{x_P y_P + x_Q y_Q}{Y} \tag{6}$$

5 Bounding polygon

In this section, we determine the line equations of the eight sides of the minimum bounding polygon for an ellipse, given the end points P and Q of two conjugate diameters of the ellipse. To simplify the calculations, the center of the ellipse is located at the origin. The sides of the polygon are constrained to be vertical, horizontal, or tilted at ±45-degrees. The ellipse touches the bounding polygon at one point on each side of the polygon, and is tangent to the side at this point.

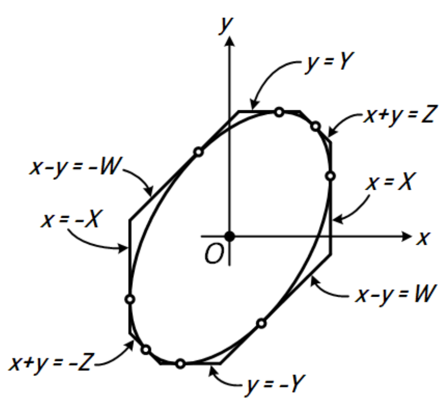

Figure 6: Line equations for bounding polygon

Figure 6 shows the bounding polygon for an ellipse. The point at which the ellipse touches and is tangent to each side of the polygon is shown as a small hollow circle. The general form of the line equation for each side of the polygon is also shown. The values X, Y, Z, and W in the line equations for a particular ellipse can be determined from the end points, P and Q, of two conjugate diameters of the ellipse.

The vertical and horizontal sides of the bounding polygon shown in Figure 6 must, of course, coincide with the corresponding sides of the bounding box described in the preceding section. Thus, we already have solutions for X and Y; see equations (3) and (5). In addition, we already have solutions for the y coordinates where the ellipse touches the two vertical sides, and for the x coordinates where the ellipse touches the two horizontal sides; see equations (4) and (6).

We still need to solve for the values of Z and W.

A diagonal line at a -45-degree angle is described by an equation of the form $x+y=Z$. Our goal is to find the Z values for the two diagonal lines that touch the ellipse at a single point on either side of the ellipse. For any given x coordinate on this line, the corresponding y coordinate is $y=Z-x$. To find the intersection of the line with the ellipse, we substitute this expression for y into equation (1) to get

$$f(x, Z-x) = Ax^2 + Bx(Z-x) + C(Z-x)^2 + F = 0$$

To solve for x, we first rearrange this equation as follows:

$$(A-B+C)x^2 + Z(B-2C)x + (CZ^2+F) = 0$$

Next, we plug these coefficients (the quantities in parentheses in the preceding equation) into the quadratic formula, and obtain

$$x = \frac{-Z(B-2C) \pm \sqrt{Z^2(B-2C)^2 - 4(A-B+C)(CZ^2+F)}}{2(A-B+C)} \tag{7}$$

For some diagonal line $x+y=Z$ that traverses the interior of the ellipse, equation (7) yields two intersections, at x_+ and x_-, with the line. However, we're interested in the single point at which the line touches the ellipse boundary (so that $x_+ = x_-$). Equation (7) yields a single value for x only if the expression inside the square-root operation is zero; that is, if

$$Z^2(B-2C)^2 - 4(A-B+C)(CZ^2+F) = 0$$

Solving this equation for Z, we obtain

$$Z = \pm\sqrt{\frac{4F(A-B+C)}{B^2-4AC}}$$

We can express Z directly in terms of the conjugate diameter end points P and Q. By substituting the expressions for coefficients A, B, C, and F from equation (1) into the preceding equation, and simplifying, we obtain

$$Z = \pm\sqrt{(x_P+y_P)^2 + (x_Q+y_Q)^2} \tag{8}$$

We now have the line equations $x+y=Z_+$ and $x+y=Z_-$ for the two -45-degree tangents to the ellipse.

Equation (7) yields the x coordinates at these two tangents, given that the quantity inside the square-root operation is zero, as previously explained. For this case, we have

$$x = \frac{-Z(B - 2C)}{2(A - B + C)}$$

The x and y coordinates at the tangents can be conveniently expressed in terms of conjugate diameter end points P and Q. By substituting the expressions for coefficients A, B, C, and F from equation (1) into the preceding equation, and simplifying, we obtain

$$x = \frac{x_P(x_P + y_P) + x_Q(x_Q + y_Q)}{Z} \tag{9}$$

$$y = Z - x \tag{10}$$

where we have made use of the fact that $Z^2 = A - B + C$.

Using similar methods to solve for W in the line equation $x - y = W$, we obtain

$$W = \pm\sqrt{(x_P - y_P)^2 + (x_Q - y_Q)^2} \tag{11}$$

This expression gives us the line equations $x - y = W_+$ and $x - y = W_-$ for the two +45-degree tangents to the ellipse.

The x and y coordinates at these two tangents are

$$x = \frac{x_P(x_P - y_P) + x_Q(x_Q - y_Q)}{W} \tag{12}$$

$$y = x - W \tag{13}$$

The vertexes of the minimum bounding polygon in Figure 6 occur at the points at which the diagonal lines $x + y = \pm Z$ and $x - y = \pm W$ intersect the axis-aligned lines $x = \pm X$ and $y = \pm Y$, where X and Y are given by equations (3) and (5).

6 Inscribed polygon

In this section, we determine the vertexes of an 8-sided polygon inscribed in an ellipse. The vertexes of this polygon are constrained to be points on the ellipse at which the tangent is vertical, horizontal, or at a ±45-degree angle. We are given the end points P and Q of two conjugate diameters of the ellipse. For simplicity, the ellipse center is located at the origin.

Figure 7 shows an example of such an inscribed polygon. The eight tangent points are shown as small hollow circles in the figure.

In fact, expressions for the x-y coordinates at these vertexes were derived in the preceding two sections. Here, we simply summarize these results, which are expressed in terms of the conjugate diameter end points, $P = (x_P, y_P)$ and $Q = (x_Q, y_Q)$.

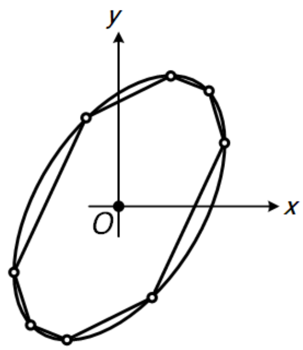

Figure 7: Polygon inscribed in ellipse

First, the points at which the ellipse touches and is tangent to the two vertical lines $x = X_+$ and $x = X_-$ are given by equations (3) and (4), which are repeated here:

$$\begin{aligned} x &= X \\ y &= \frac{x_P y_P + x_Q y_Q}{X} \\ \text{where } X &= \pm\sqrt{x_P^2 + x_Q^2} \end{aligned}$$

Next, the points at which the ellipse touches and is tangent to the two horizontal lines $y = Y_+$ and $y = Y_-$ are given by equations (5) and (6), which we repeat here:

$$\begin{aligned} x &= \frac{x_P y_P + x_Q y_Q}{Y} \\ y &= Y \\ \text{where } Y &= \pm\sqrt{y_P^2 + y_Q^2} \end{aligned}$$

The points at which the ellipse touches and is tangent to the two -45-degree diagonal lines $x + y = Z_+$ and $x + y = Z_-$ are given by equations (9) and (10), which are repeated here:

$$\begin{aligned} x &= \frac{x_P(x_P + y_P) + x_Q(x_Q + y_Q)}{Z} \\ y &= Z - x \\ \text{where } Z &= \pm\sqrt{(x_P + y_P)^2 + (x_Q + y_Q)^2} \end{aligned}$$

Finally, the points at which the ellipse touches and is tangent to the two $+45$-degree diagonal lines $x - y = W_+$ and $x - y = W_-$ are given by equations

(12) and (13), which we repeat here:

$$x = \frac{x_P(x_P - y_P) + x_Q(x_Q - y_Q)}{W}$$
$$y = x - W$$
$$\text{where } W = \pm\sqrt{(x_P - y_P)^2 + (x_Q - y_Q)^2}$$

7 Major and minor axes

In this section, we solve for the major and minor axes of an ellipse. In Figure 1, the two perpendicular yellow lines that cross in the center of the ellipse are major and minor axes of the ellipse. Our goal is to calculate the x-y coordinates at the end points of these two axes. To simplify the calculations, the ellipse center is located at the origin.

We are given the end points, $P = (x_P, y_P)$ and $Q = (x_Q, y_Q)$, of two conjugate diameters of the ellipse. We will make use of equation (1), which expresses the coefficients A, B, C, and F of the ellipse equation in terms of P and Q.

The major and minor axes pass through the origin (the center of the ellipse) and can be described by line equations of the form $y = mx$, where m is the slope of the line. We begin by solving for the values of m that correspond to the major and minor axes.

The distance from the ellipse center to a point (x, y) on the ellipse is $\sqrt{x^2 + y^2}$. This distance is maximum for an end point of the major axis, and is minimum for an end point of the minor axis.

The calculations can be simplified a bit by using distance squared instead of distance. To find the value of slope m at the points on the ellipse where the maximum and minimum distances occur, we set the derivative (with respect to slope m) of the distance squared to zero:

$$\frac{d}{dm}(x^2 + y^2) = 0$$

To evaluate this derivative, we need expressions for x^2 and y^2. But we know that $y^2 = m^2x^2$, so we only need to determine x^2.

Because $y = mx$ for an ellipse diameter of slope m, we can substitute this expression for y into equation (1), in which case the implicit equation for the ellipse becomes

$$f(x, mx) = Ax^2 + Bx(mx) + C(mx)^2 + F = 0$$

We solve this equation for x^2 and obtain

$$x^2 = \frac{-F}{A + Bm + Cm^2}$$

which is valid for $F \neq 0$, and $A + Bm + Cm^2 > 0$. (If $F = 0$, the ellipse has zero area. If $A + Bm + Cm^2 \leq 0$, the ellipse is ultra thin; that is, the area is nonzero but negligible. In either case, the minor axis is of zero or negligible length.)

And because $y^2 = m^2x^2$, we have

$$x^2 + y^2 = \frac{-F(1+m^2)}{A + Bm + Cm^2}$$

We can now evaluate the derivative as follows:

$$\begin{aligned}\frac{d}{dm}(x^2 + y^2) &= -F\left(\frac{2m}{A + Bm + Cm^2} - (1+m^2)\frac{B + 2Cm}{(A + Bm + Cm^2)^2}\right) \\ &= 0\end{aligned}$$

After simplifying the expression on the right and rearranging, we obtain

$$(B)m^2 + (2A - 2C)m + (-B) = 0$$

To solve for m, we plug these coefficients (in parentheses in the preceding equation) into the quadratic formula to obtain:

$$\begin{aligned}m &= \frac{-(2A - 2C) \pm \sqrt{(2A - 2C)^2 - 4(B)(-B)}}{2B} \\ &= \beta \pm \sqrt{\beta^2 + 1} \qquad \text{where } \beta = \frac{C - A}{B}.\end{aligned}$$

The expressions for m and β are valid for $B \neq 0$. (If $B = 0$, the ellipse is in standard position; that is, the major and minor axes are aligned with the x-y coordinate axes. Or, the ellipse is a circle. These can be handled as special cases.)

To summarize, we are given the end points P and Q of two conjugate diameters of an ellipse. We use equation (1) to obtain the coefficients A, B, C, and F of the ellipse equation in terms of P and Q. The x-y coordinates at the end points of the major and minor axes are calculated as follows:

$$\begin{aligned}x &= \pm\sqrt{\frac{-F}{A + Bm + Cm^2}} \qquad \text{for } F \neq 0, \text{ and } A + Bm + Cm^2 > 0 \\ y &= mx, \qquad \text{where } m = \beta \pm \sqrt{\beta^2 + 1}, \text{ and } \beta = \frac{C - A}{B} \text{ for } B \neq 0\end{aligned}$$

The equation for m has two solutions, m_+ and m_-, which are the slopes of the major and minor axes. To decide which of these two slopes belongs to the major axis, simply determine which set of corresponding x-y coordinates is further from the center of the ellipse.

By inspecting equation (1), we see that coefficient F is always negative or zero. It can never have a positive, nonzero value.

Because the major and minor axes are perpendicular, their slopes must satisfy the relation $m_+ = -1/m_-$, as is easily verified.

8 References

1. Pitteway, M.L.V., "Algorithm for Drawing Ellipses or Hyperbolae with a Digital Plotter," Computer Journal, 10(3), November 1967, 282-289.

2. Foley, J., A. van Dam, S. Feiner, and J. Hughes, Computer Graphics: Principles and Practice, 2nd ed., Addison-Wesley, 1990, 945-961.