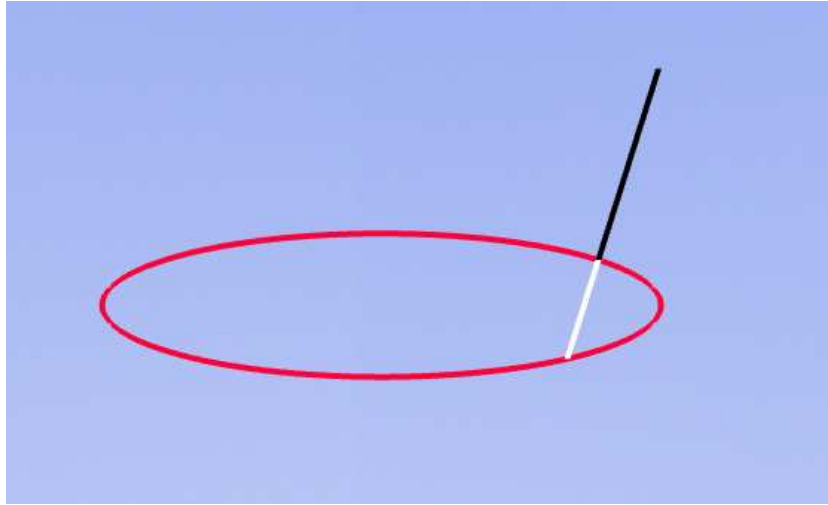


An Analytic Solution for Ellipse and Line Intersection

Andy Giese

March 21, 2015



Introduction

If you have a line and an ellipse, how can you tell where they intersect? This is a relatively simple problem that has worked-out examples all over the web if you Google for "line ellipse intersection". However, what I've come to find is that nobody will actually *give* you the solution for an *arbitrary* line and an *arbitrary* ellipse. I'm here to do just that.

Derivation

An ellipse is defined by a long axis and a short axis, called the semi-major and semi-minor axes, respectively. Usually people use the variable a to represent the length of the semi-major axis, and b to represent the length of the semi-minor axis. In this article I'll use a to represent only the **horizontal** axis and b to represent only the **vertical** axis.

That said, the formal equation for an ellipse is this:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad (1)$$

And the equation for a line is this:

$$y = mx + b$$

To avoid confusion about what b means, I'll use the term c to represent the y-intercept instead.

$$y = mx + c \quad (2)$$

To find your potentially two intersecting points, you need to solve for x and then use the values you found for x (there will be two) to find corresponding values for y . That is, you need to simultaneously solve equations 1 and 2. But first, let's discuss our line.

Equation of a Line for Two Points P_0 and P_1

You're given two points P_0 , P_1 and you need to find values for slope and y-intercept like in Eqn. 2.

Well slope, m , is simply the change in y over the change in x .

$$\frac{P_1.y - P_0.y}{P_1.x - P_0.x} \tag{3}$$

The actual order of P_0 and P_1 in Eqn. 3 doesn't matter — you can have $P_1 - P_0$ or vice versa and you'll get the same slope.

Now to find the y intercept, which we're referring to as c , we just take one of our points (arbitrarily choose P_1) and plug it into our equation to solve for c :

$$y = mx + c \Rightarrow P_1.y = m(P_1.x) + c$$

subtracting $m(P_1.x)$ from both sides leaves us with our y-intercept c :

$$c = P_1.y - m(P_1.x) \tag{4}$$

Now that we know our values for a , b , m , and c , we are ready to solve for the intersection points between the line and the ellipse.

First, substitute the line equation (Eqn. 2) into the ellipse equation (Eqn. 1) so that we can solve for x :

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \Rightarrow \frac{x^2}{a^2} + \frac{(mx + c)^2}{b^2} = 1$$

Expanding the square:

$$\frac{x^2}{a^2} + \frac{m^2x^2 + 2cmx + c^2}{b^2} = 1$$

We want to have a common denominator for both fractions on the left-hand side, so we'll multiply the first term by $\frac{b^2}{b^2}$ and the second term by $\frac{a^2}{a^2}$:

$$\frac{x^2b^2}{a^2b^2} + \frac{a^2m^2x^2 + 2a^2cmx + a^2c^2}{a^2b^2} = 1$$

Now we can multiply both sides of the equation by a^2b^2 so we don't have a fraction on the left-hand side:

$$x^2b^2 + a^2m^2x^2 + 2a^2cmx + a^2c^2 = a^2b^2$$

Notice how the first two terms on the left-hand side have a common term: x^2 , let's factor that out:

$$x^2(b^2 + a^2m^2) + 2a^2cmx + a^2c^2 = a^2b^2$$

Now let's notice that the terms $(b^2 + a^2m^2)$ and $2a^2cm$ both consist of only our known constants. To make the rest of our solution simpler, let's label these constants A and B . That is:

$$A = (b^2 + a^2m^2) \tag{5}$$

$$B = 2a^2cm \tag{6}$$

With our constant-naming out of the way, let's re-examine our equation:

$$Ax^2 + Bx + a^2c^2 = a^2b^2$$

That's much cleaner isn't it? Okay, next let's move the term a^2c^2 to the other side:

$$Ax^2 + Bx = a^2b^2 - a^2c^2$$

The left-hand side is very clean now: just a quadratic equation. I'm going to use a trick called completing the square to help us solve for x .

If we first divide everything by A we get:

$$x^2 + \frac{B}{A}x = \frac{a^2b^2 - a^2c^2}{A}$$

Which is of the form $x^2 + bx + c$. b and c in this case refer to constants of a quadratic equation, not the same variables we're using. Because we have it of this form, we know that if we add $(\frac{1}{2} \times \frac{B}{A})^2$ to both sides, then we can easily factor the left side:

$$x^2 + \frac{B}{A}x + \left(\frac{B}{2A}\right)^2 = \frac{a^2b^2 - a^2c^2}{A} + \left(\frac{B}{2A}\right)^2$$

Becomes:

$$\left(x + \frac{B}{2A}\right)^2 = \frac{a^2b^2 - a^2c^2}{A} + \left(\frac{B}{2A}\right)^2$$

Now, since we're interested in finding the value of x we need to take a square root of both sides:

$$\pm\sqrt{\left(x + \frac{B}{2A}\right)^2} = \pm\sqrt{\frac{a^2b^2 - a^2c^2}{A} + \left(\frac{B}{2A}\right)^2}$$

Evaluating the left hand side:

$$\pm\left(x + \frac{B}{2A}\right) = \pm\sqrt{\frac{a^2b^2 - a^2c^2}{A} + \left(\frac{B}{2A}\right)^2}$$

Now we want to find the value for x , not $-x$ so let's only keep the positive root:

$$x + \frac{B}{2A} = \pm\sqrt{\frac{a^2b^2 - a^2c^2}{A} + \left(\frac{B}{2A}\right)^2}$$

Let's get x by itself on the left-hand side by subtracting $\frac{B}{2A}$ from both sides:

$$x = -\frac{B}{2A} \pm \sqrt{\frac{a^2b^2 - a^2c^2}{A} + \left(\frac{B}{2A}\right)^2}$$

Because things are getting kind of messy with that big square root, I'm going to notice that it's simply a square root of constants that we already know, and label the whole thing D . That is,

$$D = \sqrt{\frac{a^2b^2 - a^2c^2}{A} + \left(\frac{B}{2A}\right)^2} \tag{7}$$

This makes our equation much cleaner:

$$x = -\frac{B}{2A} \pm D$$

Later I'll resubstitute for those constants, but bear with me as I use them to solve for x and y .

We now know that x has two solutions:

$$x = -\frac{B}{2A} + D$$

$$x = -\frac{B}{2A} - D$$

If we take these values for x along with our equation for a line (Eqn. 2), then we can solve for y :

$$y = mx + c \Rightarrow y = m\left(-\frac{B}{2A} \pm D\right) + c$$

which yields solutions for y :

$$y = m\left(-\frac{B}{2A} + D\right) + c$$

$$y = m\left(-\frac{B}{2A} - D\right) + c$$

This gives us our final intersection points of $\left(-\frac{B}{2A} + D, m\left(-\frac{B}{2A} + D\right) + c\right)$ and $\left(-\frac{B}{2A} - D, m\left(-\frac{B}{2A} - D\right) + c\right)$ \square

If we resubstitute back in for D we can simplify it ever so slightly. From Eqn. 7:

$$D = \sqrt{\frac{a^2b^2 - a^2c^2}{A} + \left(\frac{B}{2A}\right)^2}$$

Let's substitute back in for A and B on the first term (Refer to Eqns. 5 and 6):

$$D = \sqrt{\frac{a^2b^2 - a^2c^2}{b^2 + a^2m^2} + \frac{B^2}{4A^2}}$$

Notice how we have a common a^2 term in the numerator, let's factor it out:

$$D = \sqrt{\frac{a^2(b^2 - c^2)}{b^2 + a^2m^2} + \frac{B^2}{4A^2}}$$

Now let's substitute on the second term:

$$D = \sqrt{\frac{a^2(b^2 - c^2)}{b^2 + a^2m^2} + \frac{4a^4c^2m^2}{4(b^2 + a^2m^2)^2}}$$

Notice again how there is a common a^2 term in the numerator. Now we can factor an a^2 out of both, and get an a term outside the radical:

$$D = a\sqrt{\frac{b^2 - c^2}{b^2 + a^2m^2} + \frac{a^2c^2m^2}{(b^2 + a^2m^2)^2}}$$

Finally, to resubstitute everything back into our point equations, our two potential intersection points are:

$$-\frac{2a^2cm}{2(b^2+a^2m^2)} + a\sqrt{\frac{b^2-c^2}{b^2+a^2m^2} + \frac{a^2c^2m^2}{(b^2+a^2m^2)^2}},$$

$$m\left(-\frac{2a^2cm}{2(b^2+a^2m^2)} + a\sqrt{\frac{b^2-c^2}{b^2+a^2m^2} + \frac{a^2c^2m^2}{(b^2+a^2m^2)^2}}\right) + c \quad (8)$$

And

$$-\frac{2a^2cm}{2(b^2+a^2m^2)} - a\sqrt{\frac{b^2-c^2}{b^2+a^2m^2} + \frac{a^2c^2m^2}{(b^2+a^2m^2)^2}},$$

$$m\left(-\frac{2a^2cm}{2(b^2+a^2m^2)} - a\sqrt{\frac{b^2-c^2}{b^2+a^2m^2} + \frac{a^2c^2m^2}{(b^2+a^2m^2)^2}}\right) + c \quad (9)$$

Conclusions

The final equation for the points isn't really the cleanest is it? I myself prefer to keep the constants A B and D that I defined.

Note that the points you've discovered won't necessarily lie on the ellipse if the line doesn't intersect the ellipse at all; you should be able to substitute your discovered x and y values into equation 1 and see if it still equals 1.

So there you have it, an analytic solution for the intersection points of a line with an ellipse in a convenient equation for you to translate into code. Thanks for reading!