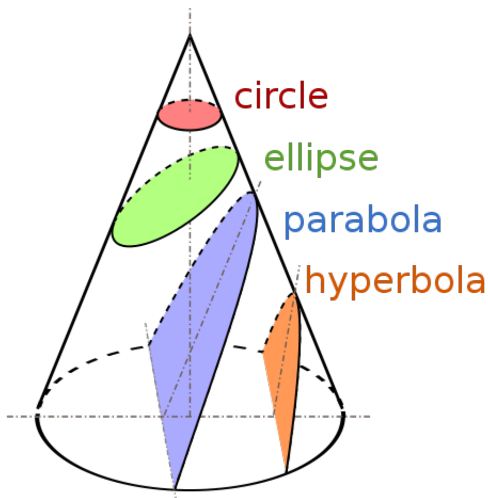
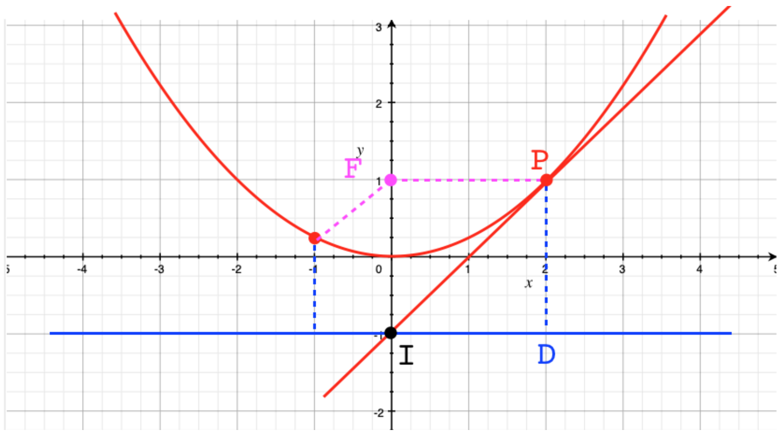
Parabola

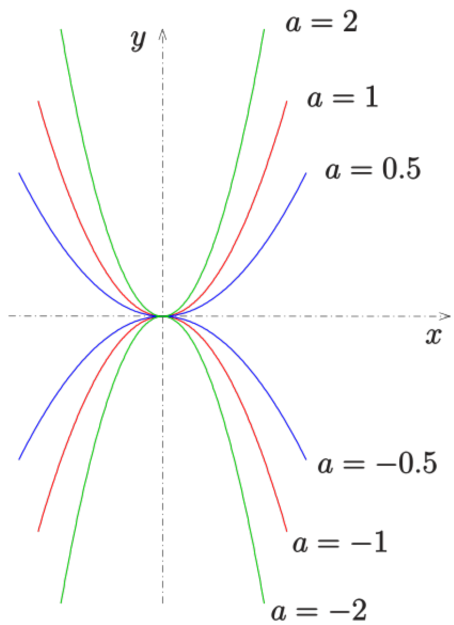
The parabola is one of a larger class of geometric figures called the conic sections.



It is pretty complicated to look at parabolas in the way that the Greeks did, so we will fudge a little and use analytic geometry to get a general formula in Cartesian coordinates, so we can draw the curve.



The equation of a general parabola is simply $y = ax^2$, where a is called the shape factor of the parabola.

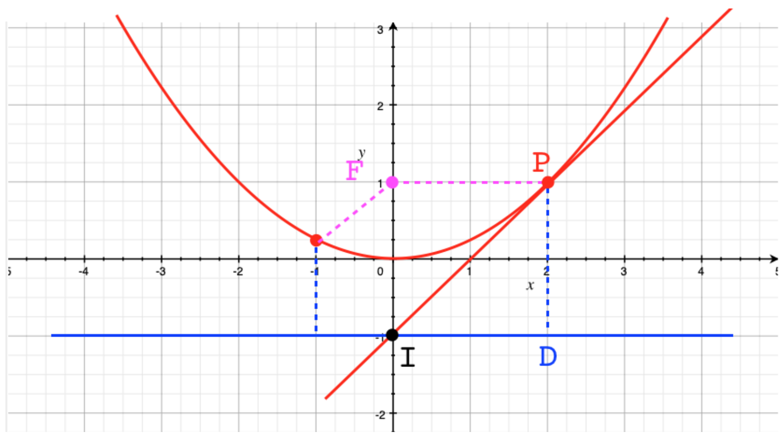


The one in the figure is a little flatter than we usually draw, we'll see the reason for that choice in a minute. At $x = 1$, $y = 1/4$, so $a = 1/4$. Consistent with that, at $x = 2$ we have $y = ax^2 = 1$.

focus and directrix

We need one more idea to start our geometric look at the parabola.

The geometric definition is this. Pick a point on the y-axis a distance p up from the origin, colored magenta in the figure. This point is called the focus (F).



Then draw a line parallel to the x-axis which intersects the y-axis the same distance p below the origin. This line is called the directrix. It is colored blue and its equation is $y = -p$.

The parabola consists of all those points whose *distance to the focus is equal to the vertical distance to the directrix.*

It is another fact that we will establish later that the distance p is related to a by the equation:

$$4ap = 1$$

which explains our choice of a. We need the parabola flat enough to see F and the line $y = -p$ clearly.

If we consider the point $P = (2, 1)$ we can compute the distance to the focus as simply $\Delta x = 2$ and to the directrix as $\Delta y = 1 + 1 = 2$.

slope of the tangent

One last fact we will import from the future (to be justified in a bit). At any point (x, y) on the parabola, the slope is $2ax$. Therefore, the slope of the tangent to the curve at $x = 2$ is

$$m = 2 \cdot 1/4 \cdot 2 = 1$$

By inspection of the graph we see that the tangent line goes through $I = (0, -1)$ which gives a point slope formula of $y = x - 1$. It is easy to verify that $(2, 1)$ and $(0, -1)$ are both on the line, and of course the slope is 1, as advertised.

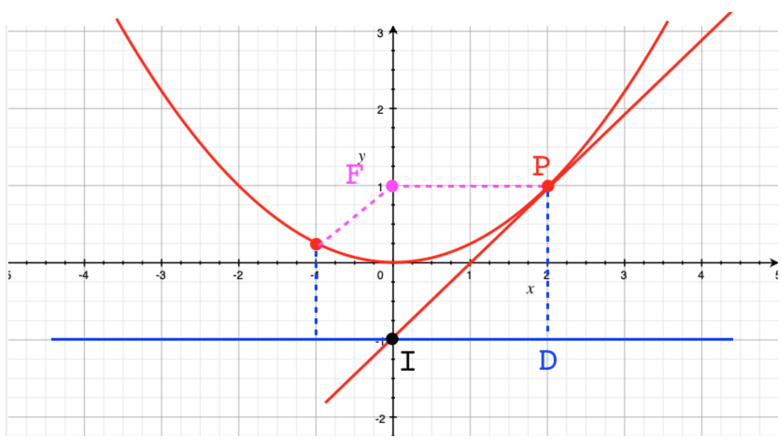
Notice that the x-intercept, x_0 is

$$0 = x - 1, \qquad x = 1$$

This is exactly halfway on the x-axis between P and I.

another point

From the equation of the curve we can get that $(x = \pm 1, y = 1/4)$ is on the curve, for either value of x. The distance from each point to the directrix is just 1 1/4.



The distance to the focus is

$$\sqrt{1^2 + (\frac{3}{4})^2} = \sqrt{\frac{25}{16}} = \frac{5}{4}$$

which checks.

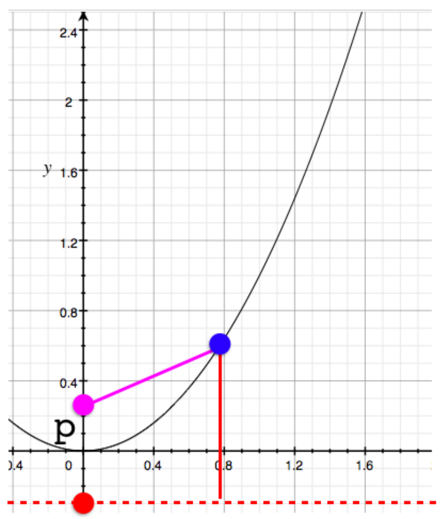
And this should not be a surprise. A look at the figure will show that in units of 1/4, we have a 3-4-5 right triangle.

computing p

Pick an arbitrary point on a parabola (in blue), with coordinates (x, ax^2). The squared distance to the focus (magenta point) is

$$\Delta x^2 + \Delta y^2 = x^2 + (ax^2 - p)^2$$

while the squared distance to the directrix (red line) is $(ax^2 + p)^2$ because Δx is zero.



For the correct choice of p these distances are to be equal:

$$(ax^2 - p)^2 + x^2 = (ax^2 + p)^2$$

$$a^2x^4 - 2apx^2 + p^2 + x^2 = a^2x^4 + 2apx^2 + p^2$$

Canceling two terms on each side

$$-2apx^2 + x^2 = +2apx^2$$

Divide by x^2

$$-2ap + 1 = 2ap$$

$$4ap = 1$$

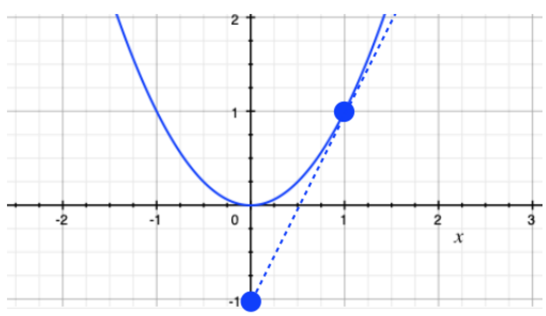
$$ap = \frac{1}{4}$$

The shape factor a determines the distance of the focus from the origin, which is p, and from the directrix, which is $2p$.

slope of the tangent

It will turn out that the slope of the tangent to $y = ax^2$ at any fixed point x is equal to $2ax$.

This is literally the first result from differential calculus, but we will also see a way to find it using analytical geometry, as well as a vector approach later on.



Thus, the equation of a line passing through the point (x, ax^2) with the given slope is

$$y' - ax^2 = 2ax(x' - x)$$

where (x', y') is any other point on the line.

What *that* means is that the x-intercept of the tangent line ($y' = 0$, $x' = x_0$) is:

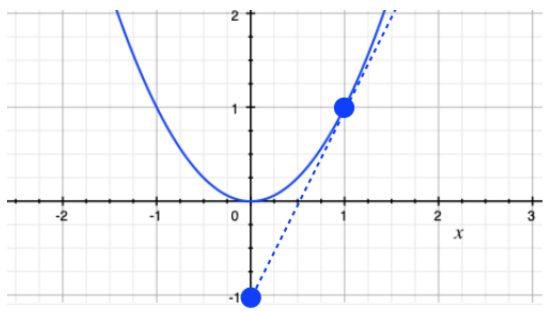
$$-ax^2 = 2axx_0 - 2ax^2$$

$$ax^2 = 2axx_0$$

$$x = 2x_0$$

$$x_0 = \frac{1}{2}x$$

The tangent line passes through the x axis halfway back toward the origin.

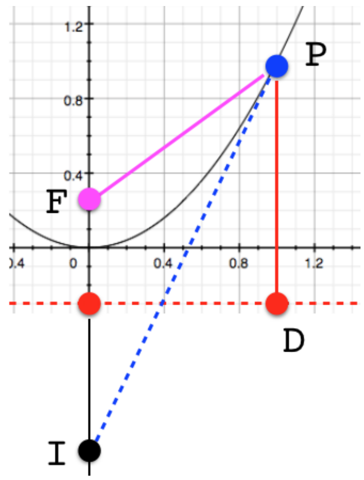


And what *that* means is that the y-intercept is symmetrical with the original point (as far below the x-axis as the point is above it). Here's the algebra:

$$y_0 - ax^2 = 2ax(0 - x)$$

$$y_0 = -ax^2$$

And then finally, if the point on the parabola is P, the focus F, the intersection with the directrix D, and the y-intercept I



the quadrilateral $FPDI$ is a regular parallelogram with all four equal sides, and its long diagonal (the tangent line) makes equal angles with FP and PD.

If PD is extended vertically, the angle it makes with the tangent line is equal to the angle between FP and the tangent line, so that for example, all vertical light rays entering a parabola will reflect and then come together at the focus.

analytic geometry

A general formula for a parabola with its vertex at the point (h, k) is

$$y - k = a(x - h)^2$$

where a is called the *shape factor*. It governs how steeply the curve rises (and by its sign, in which direction it opens). Multiplying out:

$$y - k = a(x^2 - 2xh + h^2)$$
$$y = ax^2 - 2ahx + ah^2 + k$$

In this form the cofactors are usually simplified as

$$y = ax^2 + bx + c$$

where

$$b = -2ah; \qquad c = ah^2 + k$$

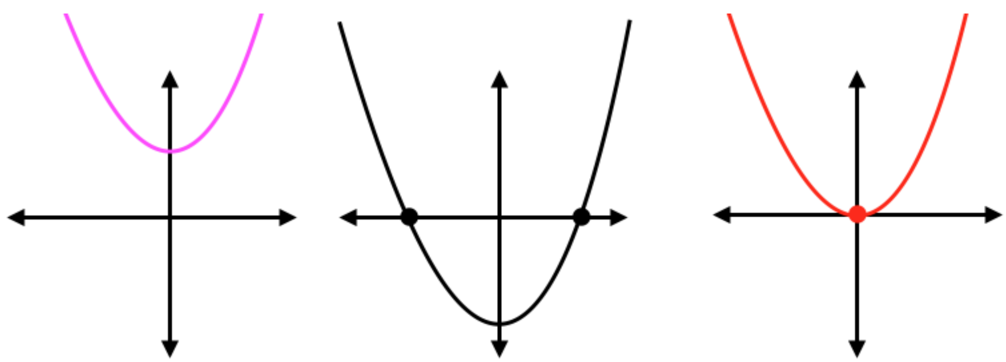
This means that any parabola's shape is solely governed by the value of a.

If the equation is given in the second form then we can find:

$$\begin{aligned} h &= -\frac{b}{2a} \\ k &= c - ah^2 \\ &= c - \frac{b^2}{4a} \end{aligned}$$

Probably the most common thing we're asked to do with a quadratic equation like this is to find the roots, the values of x for which $y = 0$ is a solution. These are the points where the graph of the curve crosses the x-axis.

It is possible to have 0, 1 or 2 roots. The black curve has two roots, the red curve has one. The latter's equation is $y = x^2$, the former is $y = x^2 - 1$ and then we can see that $x^2 = 1$ has two real solutions $x = \pm 1$.



On the left, the magenta curve does not cross the y-axis. Its equation is $y = x^2 + 1$, and there are no (real) solutions, no values of x that solve the equation when $y = 0$.

$$0 = x^2 + 1$$

$$x^2 = -1$$

To find the roots of

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

We can guess solutions by trying to factor into a form like:

$$(x - s)(x - t) = 0$$

The case of a single root occurs when $s = t$ so we have $a(x - s)^2 = 0$. A common example of that is a parabola with its vertex at the origin, so $s = 0$ and $y = ax^2$ (right panel, above).

Roots do not have to be integers (or even rational). An arguably more productive and certainly more general approach to finding them is the process of *completing the square*.

First, multiply through by $1/a$ and rearrange:

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

The key insight is to recognize that if we add $(b/2a)^2$ to both sides, the left-hand side will become a perfect square:

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

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

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

Multiplying top and bottom of the first term under the square root gives a common factor:

$$x + \frac{b}{2a} = \pm\sqrt{-\frac{4ac}{4a^2} + (\frac{b}{2a})^2}$$

which can come out of the square root and then matches what's in the second term on the left-hand side:

$$x + \frac{b}{2a} = \pm\frac{\sqrt{-4ac + b^2}}{2a}$$

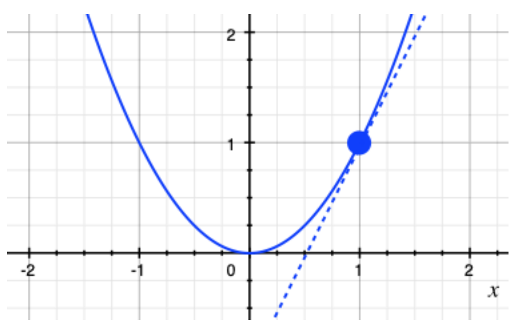
which we rearrange slightly to give the standard *quadratic formula*:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

part 1

Consider the simplest parabola: $y = x^2$.

The point $(1, 1)$ is on the curve, because $(x = 1, y = 1)$ satisfies the equation $y = x^2$.



Suppose we know that the slope of the tangent to the curve at the point $(1, 1)$ is equal to 2.

(Using calculus to find this result is trivial, we'll also show a non-calculus method in part three, below).

The equation of the tangent line is

$$y' - y = m(x' - x)$$

Plugging in for $(x', y') = (1, 1)$:

$$y - 1 = 2(x - 1)$$

$$y = 2x - 1$$

Now suppose that we knew only the parabola and this slope, but we did not know the point where the tangent meets the curve, and so do not know the y-intercept.

We have the equation of a line:

$$y = 2x + y_0$$

We seek points which are simultaneously on the line and the curve. They must satisfy both equations.

Since this is a tangent line, we seek the value for which this expression has only a single solution. The tangent "touches" the curve at a single point.

So, substitute for y from the equation for the curve:

$$x^2 = 2x + y_0$$

$$x^2 - 2x - y_0 = 0$$

Now look at the quadratic formula we would use to solve this equation for x:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

There is a single solution when the part under the square root (called the discriminant) is equal to zero.

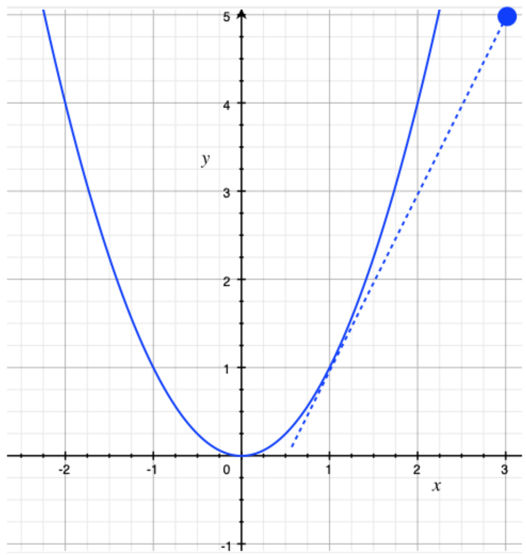
$$
\begin{aligned}
b^2 - 4ac &= 0 \\
b^2 &= 4ac \\
(-2)^2 &= 4(-y_0) \\
y_0 &= -1
\end{aligned}
$$

Therefore, the equation of the tangent line is $y = 2x - 1$, which matches what we had before.

In general, $y = 2x + y_0$ is a *family* of lines. For $y_0 = -1$, there is a single solution for x to be both on the line and the parabola. For $y_0 < -1$, there are no solutions, while for $y_0 > -1$ there are two solutions, because the line actually traces out a secant of the parabola, passing through the curve at two points.

part 2

Now suppose we have the same parabola and a point not on the parabola, but in the plane and outside of the "cup" of the parabola, such as $(3, 5)$. We seek the equations of tangent lines to the parabola that go through this point.



There will be two of them. We show just one in the figure.

The equations of lines passing through this point, with different slopes m are given by:

$$(y' - y) = m(x' - x)$$

Here, let (x', y') be $(3, 5)$ and then multiply by -1:

$$5 - y = m(3 - x)$$
$$y - 5 = m(x - 3)$$

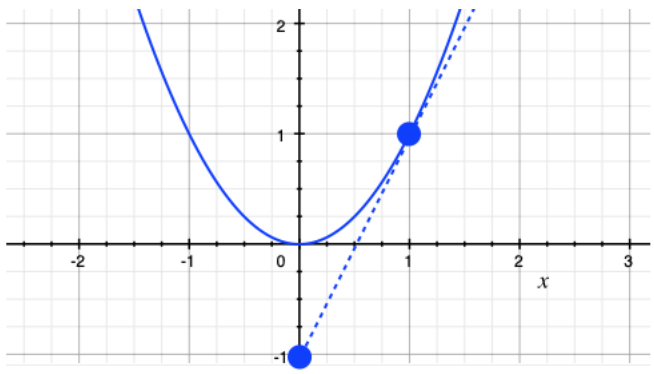
Since values of (x, y) are both on the line and the parabola $y = x^2$, we can plug in for y:

$$x^2 - 5 = mx - 3m$$
$$x^2 - mx + (3m - 5) = 0$$

As before, solutions are given by the quadratic equation. The value of the slope m giving a single solution (zero discriminant) is:

$$(-m)^2 - 4(3m-5) = 0$$
$$m^2 - 12m + 20 = 0$$
$$(m-2)(m-10) = 0$$
$$m = 2, \quad m = 10$$

We knew the first one already, because the point $(3,5)$ is on the line $y = 2x - 1$. This is the tangent to the curve at $(1,1)$, which has slope $m = 2$.



Actually, there is always another solution which we haven't found explicitly and isn't shown on the graph either. Any vertical line (with infinite slope) passes through only a single point on the parabola.

Basically what this amounts to is that in the equation

$$x = \frac{m \pm \sqrt{(-m)^2 - 4(3m-5)}}{2}$$

as m gets very large, only the term $(-m)^2$ matters under the square root, so we have

$$x = \frac{m \pm \sqrt{(-m)^2}}{2}$$

if we choose the negative root, then as $m \to \infty$, $m - \sqrt{m^2} \to 0$.

part 3

Now suppose we are given the same parabola again and also a point on it such as (x_1, y_1).

Any line through that point has the equation:

$$y - y' = m(x - x')$$

To find the equation of a tangent line through that point we need the slope m.

If there is a point (x, y) that is on the line and *also* on the parabola, it must satisfy $y = ax^2$ as well, so:

$$ax^2 - ax'^2 = m(x - x')$$
$$ax^2 - mx - ax'^2 + mx' = 0$$

Certainly $x = x'$ is a solution.

The value of m must be such that there are *no other solutions*.

Write the quadratic equation to solve for x:

$$x = \frac{m \pm \sqrt{m^2 - 4a(mx' - ax'^2)}}{2a}$$

There is a single solution when the discriminant is zero, that is, when

$$x = \frac{m}{2a}$$
$$m = 2ax$$

Since $x = x'$ for the tangent line

$$m = 2ax'$$

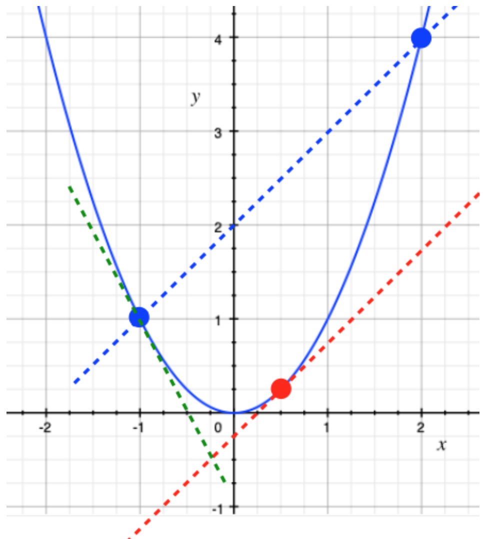
as expected.

The slope of the tangent line is $2ax'$ and in particular, at the point $(1, 1)$, the slope is equal to 2.

further comment

The slope of the parabola has some simple interesting properties. For example, pick any two points (x, y) and (x', y') on our standard parabola.

The slope of the line that connects those two points is equal to the slope of the parabola at the point whose x-value is halfway in between.



For the first part:

$$m = \frac{y' - y}{x' - x}$$

$$= \frac{ax'^2 - ax^2}{x' - x}$$

$$= a \left[\frac{x'^2 - x^2}{x' - x} \right]$$

$$= a(x' + x)$$

For the midpoint

$$x_m = \frac{1}{2}(x' + x)$$

and the slope is

$$\begin{aligned} & 2a \cdot \frac{1}{2}(x' + x) \\ & = a(x' + x) \end{aligned}$$

A similar result is that if we pick any two points (x, y) and (x', y'), and draw their slopes, the point where the two slope lines meet has its x-value exactly halfway in between x and x'.