1. Quadratic Equations

1.1 Factorizing quadratics

We start algebra by expanding expressions like

$$(a+b)(c+d) = ac + bc + ad + bd.$$

If we multiply linear expressions in a variable x we collect together terms in powers of x as in

$$(ax+b)(cx+d) = acx^2 + (bc+ad)x + bd.$$

Factorizing quadratics is the reverse process: we are given $acx^2 + (bc+ad)x + bd$ and try to recover $(ax+b)$ and $(cx+d)$. Why? Well because if $(ax+b)(cx+d) = 0$ then either $(ax+b) = 0$ or $(cx+d) = 0$, that is $x = -b/a$ or $x = -d/c$.

So we have "solved the quadratic equation" or we have "found its roots".

Typically we spend some time getting used to simple examples:

$$x^2 + 4x + 3 = (x+3)(x+1) \qquad x^2 + 2x - 3 = (x+3)(x-1)$$
$$x^2 - 4x + 3 = (x-3)(x-1) \qquad x^2 - 2x - 3 = (x-3)(x+1)$$

We get used to patterns involving small numbers and special cases like:

$$x^2 + 4x + 4 = (x+2)^2 \qquad x^2 - 4 = (x+2)(x-2)$$
$$x^2 - 4x + 4 = (x-2)^2$$

Exercise: Find all integers k such that $x^2 + kx + 12$ factorizes in whole numbers. Do the same for $x^2 + kx - 12$. (Here k can be negative)

1.2 The quadratic formula

We factorize in order to solve quadratic equations and the quadratic formula gives a general form for solutions. The general quadratic equation

$$ax^2 + bx + c = 0 \quad \text{with } a \neq 0$$

has the solutions

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

ASIDE: If $(b^2 - 4ac) < 0$ then we do not have real square roots: the equation has no real solutions. There are solutions in the complex numbers but we do not go into all that here.

The quadratic formula is rather horrid but we do not really need to remember or use it. However we should know how to derive it as that tells us how to solve quadratic equations.

Very often when faced with $ax^2 + bx + c = 0$ people start by dividing by $a \neq 0$. I think it more instructive instead to multiply by $4a$.

Well we get $4a^2x^2 + 4abx + 4ac = 0$, and if we are used to squared expressions we recognize $4a^2x^2 + 4abx + b^2 = (2ax + b)^2$, so we get $(2ax + b)^2 - b^2 + 4ac = 0$, so $(2ax + b)^2 = b^2 - 4ac$.

Taking $\pm$ square roots and rearranging we get

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

the quadratic formula.

Of course this is all a matter of taste. If you know and like the formula that is just fine!

Completing the square

What we just did is a process called "completing the square". It has wide application but in particular it gives an effective way to solve quadratic equations.

Illustrative examples:

1. $x^2 - 4x - 1 = 0 \implies (x-2)^2 - 5 = 0 \implies x = 2 \pm \sqrt{5}$
   $x^2 - 4x + 1 = 0 \implies (x-2)^2 - 3 = 0 \implies x = 2 \pm \sqrt{3}$
   (What about $x^2 - 4x + 5 = 0$?)

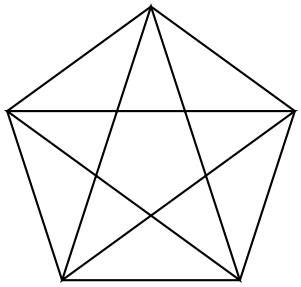
2. $x^2 - 3x + 1 = 0$
   Multiply by 4: $4x^2 - 12x + 4 = 0 \implies (2x-3)^2 - 5 = 0 \implies x = \frac{3 \pm \sqrt{5}}{2}$

3. $3x^2 - 2x - 2 = 0$
   Multiply by 3: $9x^2 - 6x - 6 = 0 \implies (3x-1)^2 - 7 = 0 \implies x = \frac{1 \pm \sqrt{7}}{3}$
   (N.B. The coefficient of x is even so we don't need that 4.)

   Of course, this is all a matter of taste. If you know and like the formula, that is just fine!

1.3 Exercise: The regular pentagon

Take a regular pentagon with chords as shown in the diagram.



Our aim is to find the ratio of the chord to the side. So we suppose the side has length 1 and ask for the length λ of the chord.

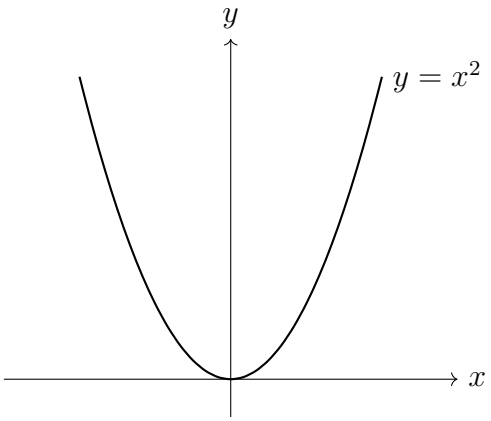
Hints

- There is some not quite obvious symmetry.
- What are the angles in the figure? (They are all multiples of what angle?)
- What is special about all (!) the Δs in the figure? That should give the length of some of the lines.
- Now use similar Δs to write down a quadratic equation for λ. (There is a lot of choice!)
- Solve the equation. The λ which you should find is a famous number.

---

2. Transformations and Change of Variables

We are going to look at the graphs in the x-y plane of quadratic functions $y = ax^2 + bx + c$ ($a \neq 0$). The simplest such is $y = x^2$. Its graph looks like

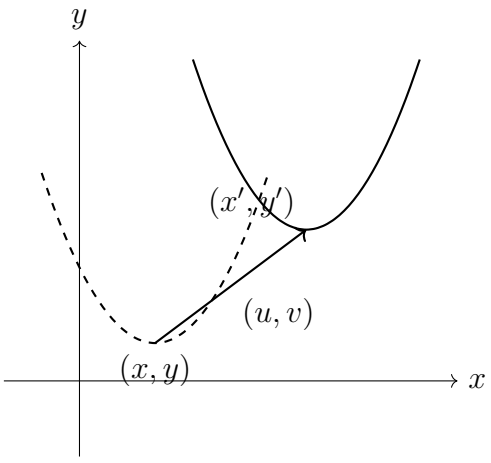


It is a parabola. It has an evident line of reflective symmetry ($x \leftrightarrow -x$) and that line intersects the parabola at the vertex which in this case is the origin.

We shall see that any quadratic function has a graph similar to that one: they are all parabolas. In this section we consider the transformations (translations and scalings) which we need to show that.

2.1 Translations

Suppose we have the graph of the equation $y = f(x)$ in the x-y plane. Imagine translating it by the vector (u, v) so that each point (x, y) is moved to a new point (x', y') with $x' = x + u$, $y' = y + v$.



What is the equation of the new graph?

Well we have $x = x' - u$, $y = y' - v$ and substituting in $y = f(x)$ gives $y' - v = f(x' - u)$ or $y' = f(x' - u) + v$.

So (dropping the primes) the new equation is

$$y = f(x - u) + v$$

N.B. Our use of primed variables is not just for clarity. It suggests a different point of view. Rather than translate the graph through (u, v) we could instead move the origin to $(-u, -v)$. Then (x', y') are the new coordinates of (x, y) with respect to the new origin. It is good to get used to the two ways of thinking. Sadly it is easy to confuse them!

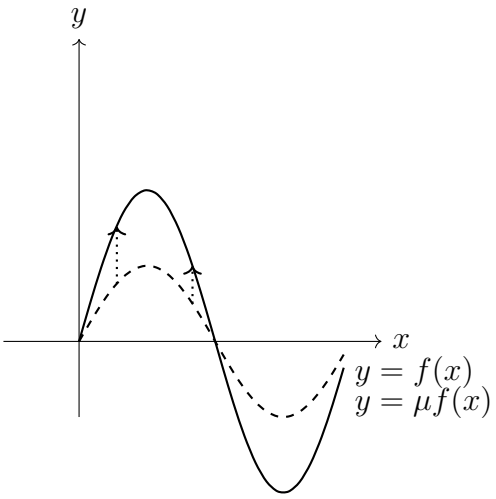
Here is something which may help. The new equation $y = f(x - u) + v$ means: move x from new argument to old argument $x - u$; apply f to get old value $f(x - u)$; move that to new value $f(x - u) + v$. (This pattern is ubiquitous.)

Exercises

1. What is the equation which results from translating $y = x^2$ by $(1, -1)$? by $(2, -1)$? by $(3, -1)$? Do you see the pattern? Did you expect the quadratics would all have roots? Why?
2. The quadratic $y = x^2$ has vertex at the origin. Translate so that the vertex is at each of the four points $(\pm 1, \pm 1)$. What in each case is the new equation? In which cases do you expect to be able to solve the new equation (i.e. it has real roots)? Check that you are right.
3. Show that if we translate the quadratic $y = ax^2 + bx + c$ $(a \neq 0)$ by (u, v) then we get an equation of the form $y = ax^2 + b'x + c'$, that is, a quadratic with the same leading term.
4. What translation transforms the quadratic $y = ax^2 + bx + c$ to the form $y = ax^2$?
5. Given $y = ax^2 + bx + c$ and $y = ax^2 + b'x + c'$, can we translate one into the other? (Please do not bother with the algebra!)

2.2 Scaling

Take the graph of $y = f(x)$ in the x-y plane. Imagine stretching it by a factor $\lambda \neq 0$ in the x-direction and by $\mu \neq 0$ in the y-direction. Every point (x, y) is moved to a new point (x', y') with $x' = \lambda x$, $y' = \mu y$.



To get the equation of the new graph we substitute $x = x'/\lambda$, $y = y'/\mu$ in $y = f(x)$ which gives

$$y'/\mu = f(x'/\lambda) \quad \text{or} \quad y' = \mu f(x'/\lambda).$$

Again dropping the primes the new graph is

$$y = \mu f(x/\lambda)$$

Again there is a different point of view. We could rescale the axes by $1/\lambda$ in the x-direction and $1/\mu$ in the y-direction. Then (x', y') are the new coordinates of (x, y) with respect to the rescaled axes.

Again to get the new equation $y = \mu f(x/\lambda)$ we can take a new argument x; map it to an old argument x/λ; apply f, $f(x/\lambda)$ gives the old value and then $\mu f(x/\lambda)$ is the new value.

Exercises

1. What scaling takes the graph of $y = f(x)$ to the graph of $y = 2f(2x)$?
2. What scalings take the graph of $y = x^2$ to itself?

We shall be mostly interested in uniform scalings, that is, when $\lambda = \mu$ in the above. Scaling by (λ, μ) with $\lambda \neq \mu$ changes shapes. For example the square with vertices $(0, 0)$, $(1, 0)$, $(1, 1)$ and $(0, 1)$ is taken to the rectangle with vertices $(0, 0)$, $(\lambda, 0)$, (λ, μ) and $(0, \mu)$.

Exercises

3. What happens to the line $y = ax$ if we scale uniformly by λ? And what happens to the straight line $y = ax + b$?
4. What uniform scaling takes the graph for $y = ax^2$ to that for $y = bx^2$?

2.3 Scaling and translating

(This gives a taste of group theory)

Take the graph of $y = f(x)$.

(1) First scale by (λ, μ) and then translate by (u, v). What is the equation of the new graph?

(2) First translate by (u, v) and then scale by (λ, μ). What is the equation of the new graph?

You should get different answers in (1) and (2).

(3) Can you write a translation followed by a scaling as a scaling followed by a translation?

Perhaps it is good to have some notation for these transformations. Write $\sigma_{(\lambda,\mu)}$ for scaling by (λ, μ) and $\tau_{(u,v)}$ for translating by (u, v).
Then using the usual notation for functions or transformations we write (1) above as $\tau_{(u,v)}\sigma_{(\lambda,\mu)}$, that is, first do $\sigma_{(\lambda,\mu)}$ then $\tau_{(u,v)}$.

(4) Show that

$$\sigma_{(\lambda,\mu)}\tau_{(u,v)} = \tau_{(\lambda u,\mu v)}\sigma_{(\lambda,\mu)}$$

Explain intuitively what this means.

(5) Show that the result of any combination of translations and scalings can be written in the form $\tau_{(u,v)}\sigma_{(\lambda,\mu)}$, i.e. as first scale and then translate.

---

3 Quadratic functions

The general quadratic function is

$$y = ax^2 + bx + c \quad (a \neq 0)$$

In Exercises 4 and 5 of 2.1 we showed that the graph is congruent by a translation to the graph of $y = ax^2$.
Then in Exercise 4 of 2.2 we showed that any two such graphs are similar by a (uniform) scaling.
So any two parabolas = graphs of quadratic functions are similar.

In this section we revisit this from various points of view.

3.1 Deriving the quadratic formula again

Suppose we manipulate $y = ax^2 + bx + c \quad (a \neq 0)$ as we derived the quadratic formula. We get

$$\begin{aligned} 4ay &= 4a^2x^2 + 4abx + 4ac \\ &= (2ax + b)^2 - (b^2 - 4ac) \end{aligned}$$

and so

$$4ay + (b^2 - 4ac) = (2ax + b)^2$$

If we choose new variables $X = 2ax + b$, $Y = 4ay + (b^2 - 4ac)$ then this is just the equation $Y = X^2$.

Immediate consequences

(1) $Y = X^2$ is symmetric about $X = 0$. So $y = ax^2+bx+c$ is symmetric about $x = -b/2a$.

(2) Assuming $a > 0$, $y = ax^2 + bx + c$ will have a minimum when $x = -b/2a$ and that minimum value is

$$-\frac{(b^2 - 4ac)}{4a}$$

What have we done to the graph to get into the form $Y = X^2$?

- **Scaling:** we stretched by $2a$ in the x-direction and by $4a$ in the y-direction.
- **Translation:** we translated the (scaled) curve by $(b, b^2 - 4ac)$.

That is not a uniform scaling but Exercise 2 of 2.2 makes that a trifling difficulty. We only have to scale further by $(1/2, 1/4)$.

3.2 The direct manipulation

What we just saw suggests that it was a bit too clever to multiply by $4a$. If instead we simply write

$$\begin{aligned} ay &= a^2x^2 + abx + ac \\ &= \left(ax + \frac{b}{2}\right)^2 - \frac{1}{4}(b^2 - 4ac) \end{aligned}$$

and just take $X = ax + \frac{b}{2}$, $Y = ay + \frac{b^2-4ac}{4}$;
and then we get $Y = X^2$ with a uniform scaling by a and a translation.

3.3 Revisiting the argument from the §2

Take the general quadratic $y = ax^2 + bx + c \quad (a \neq 0)$
Translate so that its vertex is at the origin. Can you see at once (without any calculation) that the new equation must be $y = ax^2$?
Then all such parabolas are similar by the simple calculation of Exercise 4 of 2.2.

---

4. The parabola in standard form

4.1 The standard form for a parabola is

$$4ay = x^2$$

which is parametrized by $(x = 2at, y = at^2)$. How might one stumble upon this idea?

Start with $y = ax^2$. Differentiating twice gives $\frac{d^2y}{dx^2} = 2a$ constant. It suggests kinematics. Replace the variable x by t for time and read $y = at^2$ as a solution to $\frac{d^2y}{dt^2} = 2a$, the equation which says y is subject to constant acceleration $2a$. When deriving the quadratic formula, we had the bright idea to multiply by $4a$. Do it again and we get $4ay = 4a^2t^2 = (2at)^2$. So scaling we set $x = 2at$ and we get

$$4ay = x^2$$

with $x = 2at$ and $y = at^2$ points on the parabola. In the kinematics, t is time a particle traverses the parabola with constant velocity $2a$ in the x direction and constant acceleration $2a$ in the y direction.

Exercise All parabolas are similar. So $y = ax^2$ and $4ay = x^2$ must be similar. They have the same vertex. What is needed to change $y = ax^2$ into $4ay = x^2$?

4.2 Focus and directrix

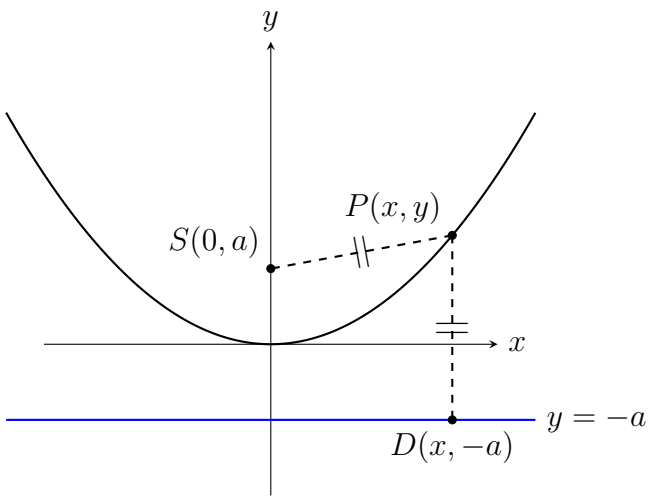
You may have heard of dimensional analysis. Let's do a bit of that.

First look at $y = ax^2$. y and x are lengths L; so the dimension of the constant a should be L^{-1}, the inverse of length. Now look at $4ay = x^2$. Again, y and x are lengths L; but now the dimension of a is also L. So a must be a length. What length is it?

In the standard equation $4ay = x^2$ we can recognize $4ay = (y + a)^2 - (y - a)^2$ so we can rewrite to give $(y + a)^2 = x^2 + (y - a)^2$, that is,

$$|y + a| = +\sqrt{x^2 + (y - a)^2}.$$

The LHS is the perpendicular distance from (x, y) to the line $y = -a$. The RHS is the distance from (x, y) to $(0, a)$.



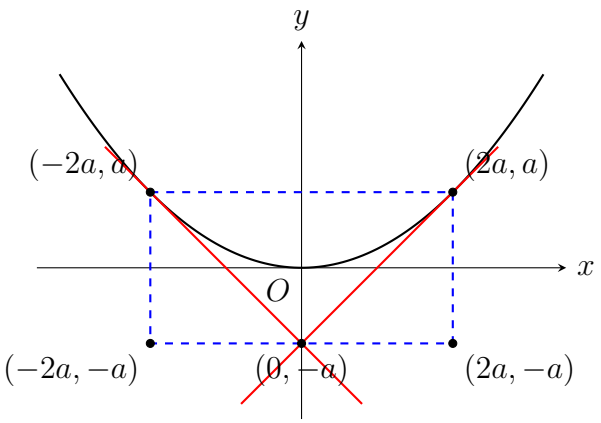
So in the diagram we have $|PS| = |PD|$. The parabola $4ay = x^2$ is the locus of points P such that $PS = PD$.

In the diagram the point $S = (0, a)$ is the focus of the parabola; the line $y = -a$ is the directrix of the parabola. For a parabola the distance from the focus equals the distance from the directrix. Of course this property is invariant under translation and under uniform scaling from the vertex.

Exercise What are the focus and directrix of $y = ax^2$?

4.3 Latus rectum

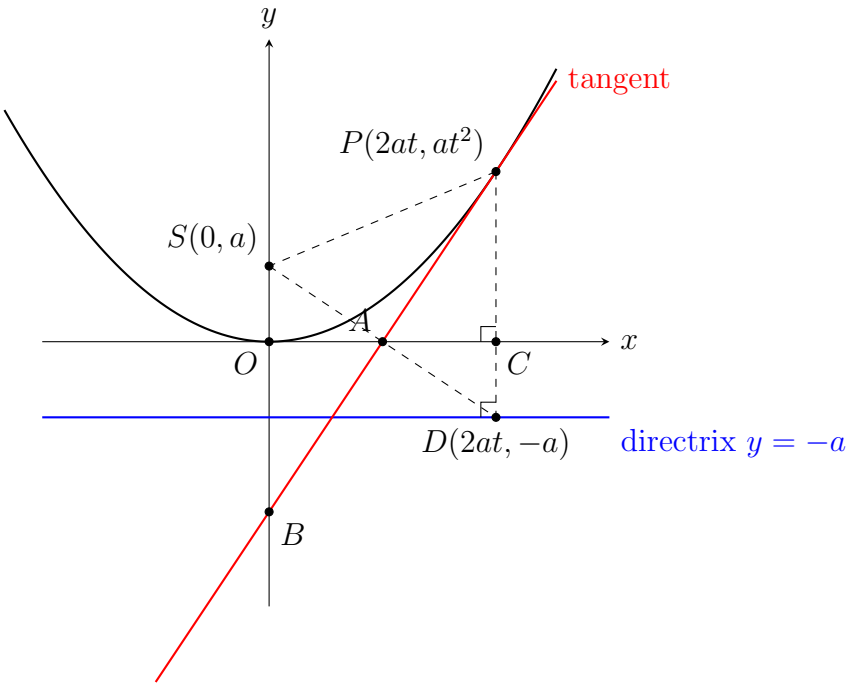
(This is not essential but of historical interest.) In $4ay = x^2$ with $x = 2at, y = at^2$, the points $(\pm 2a, a)$ corresponding to $t = \pm 1$ lie on the line $y = a$ through the focus $(0, a)$. The chord between them is called the latus rectum (latus = wide, rectus = straight). The length is $4a$, twice the distance between focus and directrix. And the slope of the parabola at the points is ± 1. So we have a diagram.



4.4 The tangent to the parabola

Take the standard parabola $4ay = x^2$ with parametrization $(2at, at^2)$. Consider the tangent to the parabola at $P = (2at, at^2)$. The slope of the tangent is t.

(EITHER $\frac{dy}{dx} = \frac{1}{4a} \cdot 2x = \frac{4at}{4a} = t$. OR $\frac{dy}{dx} = \frac{dy/dt}{dx/dt} = \frac{2at}{2a} = t$.)



In the diagram O is the origin, S is the focus $(0, a)$, $D = (2at, -a)$ is the perpendicular from $P = (2at, at^2)$ to the directrix.

Exercises

1. The tangent crosses the x-axis at A. What is A? It crosses the y-axis at B. What is B?

2. In 4.2 we saw that $|PS| = |PD|$. And from the previous exercise you can deduce $|SB| = |PD|$. Deduce that SAD is a straight line with $|SA| = |AD|$ and that the lines BP and SD are perpendicular.

3. Deduce also that $\angle SPA = \angle PAD$. It follows from this that a light ray from infinity along $x = 2at$ (i.e. parallel to the y-axis) incident at P will be reflected along PS to the focus. This is the focal property of the parabola.

4. Suppose the line PS from P to the focus meets the parabola again at P'. (Here we assume $P \neq O$, that is $t \neq 0$.) Show that the tangents at P and P' are perpendicular. (You can do this algebraically, but the focal property makes this simple!)

5. Observe that the Δs PSA, PAD, SBA and ABD are all congruent. Note also that Δs PAC and OBA are similar and that Δs SOA and ADC are congruent. Show that all these Δs are similar.
   Write down the lengths of all the sides (in terms of t).

---

5. Further directions

5.1 The affine symmetries of the parabola

An affine transformation of the plane is given in matrix notation by

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} r & s \\ t & u \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

where $\begin{pmatrix} r & s \\ t & u \end{pmatrix}$ is an invertible matrix (equivalently where $\det \begin{pmatrix} r & s \\ t & u \end{pmatrix} = ru - st \neq 0$).

Show that such transformations take straight lines to straight lines and parallel straight lines to parallel straight lines.

Which affine transformations map the parabola $y = x^2$ to itself?

Explain geometrically what these transformations do in particular to straight lines $x =$ constant.

(These phenomena were understood by Archimedes, who used them in his calculation of areas.)

5.2 A problem in Physics

The following is a favourite of our friends in Physics. It is problem 37 in P. Gnädig, G. Honyek and K.F. Riley. 200 Puzzling Physics Problems CUP 2001.

I quote their formulation:

"What is the maximal angle to the horizontal at which a stone can be thrown and always be moving away from the thrower?"

First, explain why you think there should be such an angle.

Now you may assume that the stone describes a parabola with a vertical axis of symmetry. On that basis alone, solve the problem.

Finally, a PDF of the book can be found online. Compare your solution with that given in the book. Do you regard this as a problem in Physics?