Solving the Cubic and Drawing the Cusp Surface

JWR

February 3, 2013

§1. The goal is to solve the cubic equation[1]

$$x^3 - 3ax + b = 0. \tag{1}$$

where a and b are real numbers. If $x = u + v$ then

$$x^3 - 3ax + b = u^3 + v^3 + 3(uv - a)(u + v) + b$$

so any solution (u, v) of the equations

$$uv = a, \qquad u^3 + v^3 = -b \tag{2}$$

gives a solution $x = u + v$ of (1). By Bezout a conic and a cubic intersect in six points but if (u, v) satisfies (2) so does (v, u) and these give the same value for x.

§2. Now (2) implies that the product of u^3 and v^3 is a^3 and the sum is $-b$ which in turn implies that u^3 and v^3 are the roots of the quadratic equation

$$W^2 + bW + a^3 = 0 \tag{3}$$

Conversely, if w is any solution of

$$w^6 + bw^3 + a^3 = 0 \tag{4}$$

then, because

$$\left(\frac{a}{w}\right)^6 + b\left(\frac{a}{w}\right)^3 + a^3 = \left(\frac{a^3}{w^6}\right)\left(w^6 + bw^3 + a^3\right)$$

a/w is another solution and because

$$w^6 + bw^3 + a^3 = w^3 \left(w^3 + \left(\frac{a}{w}\right)^3 + b \right),$$

the pair $(u, v) = (w, a/w)$ satisfies (2). This means that if $w_1, \ldots, w_6$ are the solutions of (4), then $(u_i, v_i) = (w_i, a/w_i)$ are the solutions of (2) and $x_i = w_i + a/w_i$ are the solutions of (1) each appearing twice in the list.

§3. It is amusing that the construction yields non-real values for u^3 and v^3 precisely when the roots $x = u + v$ of the original equation are real. We see this by graphing $y = f(x)$ where $f(x) = x^3 - 3ax + b$ so $f'(x) = 3(x^2 - a)$. In case $a < 0$ the derivative $f'(x)$ is always positive so there is only one real root and $b^2 - 4a^3 > 0$ so the values of u^3 and v^3 are real. In case $a > 0$ the function $f(x)$ has a local minimum at $x = \sqrt{a}$ and a local maximum at $x = -\sqrt{a}$ so there are three real roots if and only if

$$f(\sqrt{a}) < 0 < f(-\sqrt{a}). \tag{5}$$

But $f(\sqrt{a}) = -2a^{3/2} - b$ and $f(-\sqrt{a}) = 2a^{3/2} - b$ so condition (5) is equivalent to the condition $-2a^{3/2} + b < 0 < 2a^{3/2} + b$ i.e. to $b^2 - 4a^3 < 0$ which is the condition that the solutions u^3 and v^3 to (3) be non-real.

§4. Here's another way to look at it. The solutions of (4) are the cube roots of the solutions

$$U = \frac{-b + \sqrt{b^2 - 4a^3}}{2}, \qquad V = \frac{-b - \sqrt{b^2 - 4a^3}}{2}.$$

of (3). The two numbers U and V satisfy $U + V = -b$ and $UV = a^3$. The solutions of (4) are therefore

$$U^{1/3}, \;\; U^{1/3}\omega, \;\; U^{1/3}\omega^2, \;\; V^{1/3}, \;\; V^{1/3}\omega, \;\; V^{1/3}\omega^2.$$

Here $U^{1/3}$ and $V^{1/3}$ are cube roots of U and V respectively and ω is a primitive cube root of unity (so $\bar{\omega} = \omega^2 = \omega^{-1}$ is the other other one). Define

$$u_j = U^{1/3}\omega^j, \qquad v_j = V^{1/3}\omega^j$$

for $j = 0, 1, 2$. Then $u_j^3 + v_k^3 = -b$. If we choose the cube roots so that $U^{1/3}V^{1/3} = a$, then $u_j v_k = a\omega^{j+k}$ so $u_0 v_0 = u_1 v_2 = u_2 v_1 = a$ and the three solutions of the cubic are

$$x_0 = u_0 + v_0, \qquad x_1 = u_1 + v_2, \qquad x_2 = u_2 + v_1.$$

§5. We distinguish two cases.

(I) If $b^2 > 4a^3$, the roots U and V are real and distinct. The real cube root $W^{1/3}$ of a real number W is real so W and $W^{1/3}$ have the same sign. This assures that $U^{1/3}V^{1/3} = a$. In this case

$$x_0 = U^{1/3} + V^{1/3}$$

is the only real root by §3.

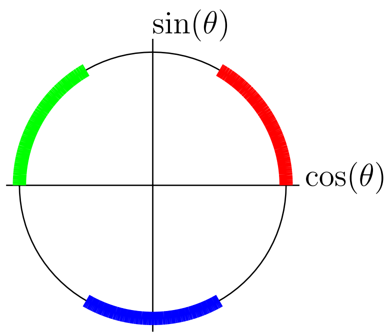
(II) If $b^2 < 4a^3$, the roots U and V are complex conjugate. In this case a is positive and $U = a^{3/2}e^{i\theta}$ and $V = a^{3/2}e^{-i\theta}$ where θ is the unique solution of $\cos\theta = b/(2a^{3/2})$ with $0 < \theta < \pi$. The solutions of (4) are as in case (I) but with $U^{1/3} = \sqrt{a}e^{i\theta/3}$ and $V^{1/3} = \sqrt{a}e^{-i\theta/3}$. Again $U^{1/3}V^{1/3} = a$. The three roots x_0, x_1, x_2 are

$$\begin{aligned} x_0 &= \sqrt{a}(e^{i\theta/3} + e^{-i\theta/3}) &&= 2\sqrt{a}\cos\tfrac{\theta}{3}, \\ x_1 &= \sqrt{a}(e^{i\theta/3}\omega + e^{-i\theta/3}\bar{\omega}) &&= 2\sqrt{a}\cos\tfrac{\theta+2\pi}{3}, \\ x_2 &= \sqrt{a}(e^{i\theta/3}\bar{\omega} + e^{-i\theta/3}\omega) &&= 2\sqrt{a}\cos\tfrac{\theta-2\pi}{3}. \end{aligned}$$

As in §3 the x_i are real. As $0 < \theta < \pi$ we have

$$0 < \frac{\theta}{3} < \frac{\pi}{3}, \qquad \frac{2\pi}{3} < \frac{\theta+2\pi}{3} < \pi, \qquad -\frac{2\pi}{3} < \frac{\theta-2\pi}{3} < -\frac{\pi}{3}.$$

Thus x_0 is positive and x_1 is negative and $x_1 < x_2 < x_0$.



§6. The following `sage`[2] program implements the above.[3]

```
def solve(a,b):
    if a==0 and b==0: return [0]
    if a==0: return [-b/abs(b)^(2/3)]
    d=b^2-4*a^3;
    U=(-b+sqrt(abs(d)))/2; V=(-b-sqrt(abs(d)))/2;
    if d>0:
        u_0=U/(abs(U)^(2/3)); v_0=V/(abs(V)^(2/3))
        x_0=u_0+v_0
        return [x_0]
    else: # d<=0:
        theta=acos(-(b/2)/a^(3.0/2))
        x_0=2*sqrt(a)*cos(theta/3)
        x_1=2*sqrt(a)*cos((theta+2*pi)/3)
        x_2=2*sqrt(a)*cos((theta-2*pi)/3)
        return [x_1,x_2,x_0]
```

§7. Let p, q, r be the roots of (1). The constant term b is the product pqr of the roots and the sum of $p+q+r$ of the roots is zero because (1) contains no x^2 term. Thus

$$x^3 - 3ax + b = (x-p)(x-q)(x-r) = x^3 - (pq+pr+qr)x + pqr.$$

The discriminant $b^2 - 4a^3$ vanishes when there is a double root, say when $p = q = -\frac{1}{2}r$ and in this case

$$x^3 - 3ax + b = (x + \tfrac{1}{2}r)^2(x - r).$$

The surface in (a, b, x)-space defined by (1) is invariant under the symmetry $(a, b, x) \mapsto (a, -b, -x)$ and this symmetry has the effect of reversing the signs of the three roots. The curve $b^2 - 4a^3 = 0$ has parametric equations $a = t^2$, $b = 2t^3$. Over this curve we have the factorization

$$(x-t)^2(x+2t) = (x^2 - 2tx + t^2)(x+2t) = x^3 - 3t^2x + 2t^3$$

and the roots lie on the twisted cubic $a = t^2$, $b = 2t^3$, $x = t$. The root $x = -2t$ also lies on the twisted cubic as

$$a = (-2t)^2, \quad b = 2(-2t)^3 \implies 4a^3 = b^2.$$

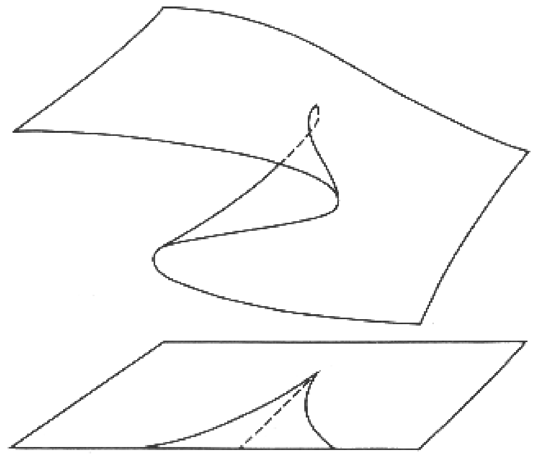


Figure 1: A hand drawn graph of the cusp surface $x^3 - 3ax + b = 0$

§ 8. Figure 1 shows a hand drawn graph of the cusp surface (1). When drawing the cusp surface, the vertical coordinate is the x-axis, the horizontal coordinate is the b-axis, and the a-axis drawn as a line of positive slope; the left face has equation $b = -1$, the right face has equation $b = 1$, the back face has equation $a = -1$,and the front face has equation $a = 1$.[4] To draw the surface with the computer, we can use the vertical coordinate x to parameterize curves in the surface as follows.[5]

§ 9. The cusp surface (1) intersects the plane a = constant in the graph of the polynomial

$$b = -x^3 + 3ax. \tag{6}$$

The following table determines the qualitative behavior of the graph when $a = 1$.

x	$b = -x^3 + 3x$	$\dfrac{db}{dx} = -3(x^2 - 1)$	$\dfrac{d^2b}{dx^2} = -6x$
2	−2	−9	−12
1	2	0	−6
0	0	3	0
−1	−2	0	6
−2	2	9	12

For $a < 0$ the polynomial (6) is strictly decreasing. The endpoints of this curve in the interval $-2 \leq x \leq 2$ are the points $(a, b, x) = (1, -2, 2)$ and $(a, b, x) = (1, 2, -2)$. (See Figure 2.)

§10. The following table determines the qualitative behavior of the intersection of the surface (1) with the plane $b = -2$.

x	$a = \dfrac{x^2}{3} - \dfrac{2}{3x}$	$\dfrac{da}{dx} = \dfrac{2x}{3} + \dfrac{2}{3x^2}$
$-\infty$	$+\infty$	
-1	1	0
$0-$	$+\infty$	
$0+$	$-\infty$	
$+\infty$	$+\infty$	

The symmetry $(a, b, x) \mapsto (-a, -b, -x)$ sends the intersection with the plane $b = -2$ to the intersection with the plane $b = +2$.

x	$a = \dfrac{x^2}{3} + \dfrac{2}{3x}$	$\dfrac{da}{dx} = \dfrac{2x}{3} - \dfrac{2}{3x^2}$
$-\infty$	$+\infty$	
$0-$	$-\infty$	
$0+$	$+\infty$	
1	1	0
$+\infty$	$+\infty$	

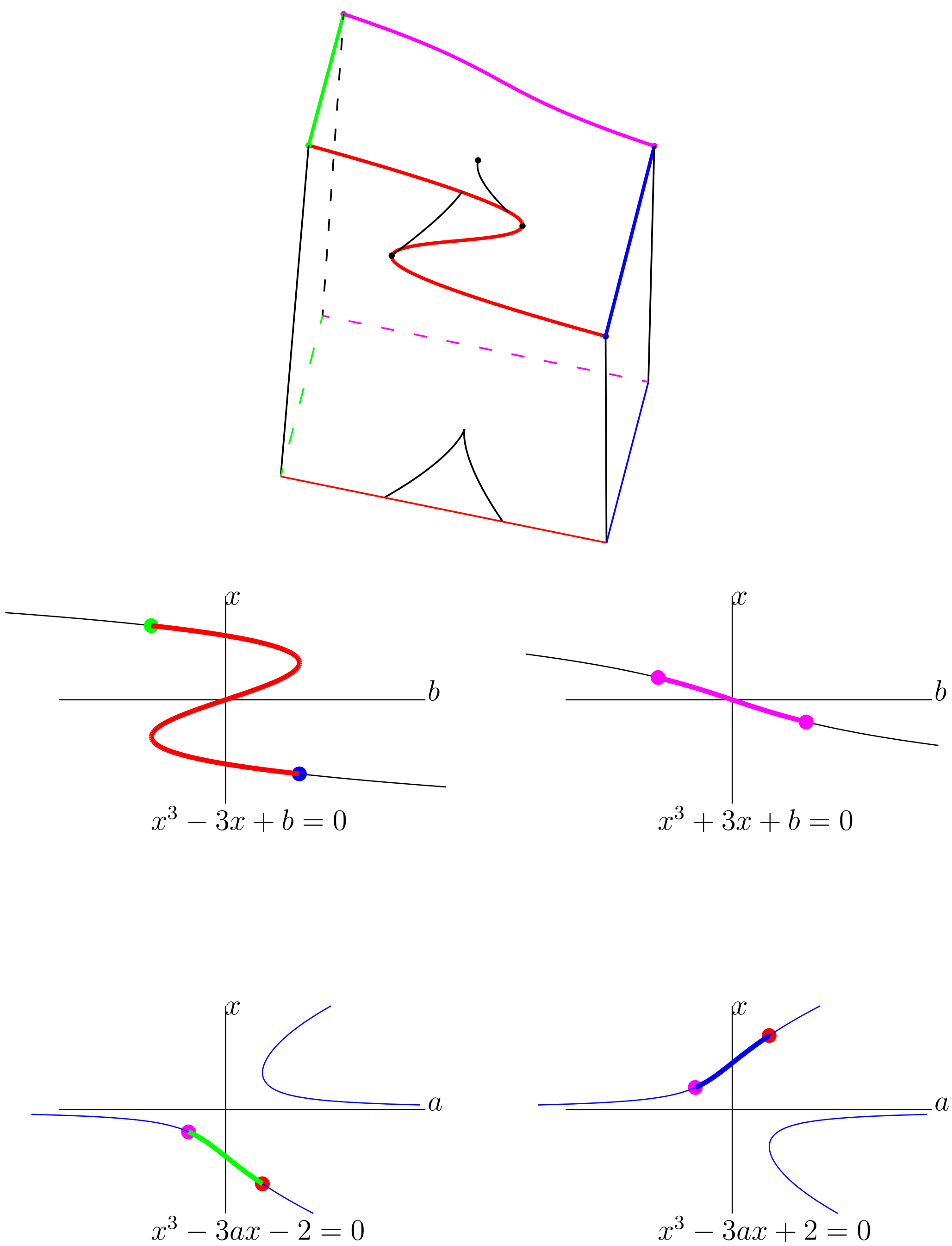
Each of these two curves has two branches, but only one branch is indicated in the picture of the cusp surface shown in Figure 2.

§11. The picture of the surface $x^3 - 3ax + b^2 = 0$ drawn in Figure 2 it is a little wonky. The quadrilateral in the base is almost, but not exactly, a parallelogram. Here's why that happened. The picture is drawn by projecting the boundary curves in (a, b, x)-space to curves in (X, Y)-space via a linear transformation

$$\begin{aligned} X &= c_{11}\, a + c_{12}\, b + c_{13}\, x \\ Y &= c_{21}\, a + c_{22}\, b + c_{23}\, x \end{aligned} \tag{7}$$

and drawing the image. If $c_{13} = 0$ this linear transformation maps a vertical line $a = a_0$, $b = b_0$ to a vertical line $X = X_0$. This feature is desirable

Figure 2: The cusp surface $x^3 - 3ax + b = 0$



but doesn't produce a good picture because the image of the back curve $x^3 + 3x + b = 0$ overlaps the image of the front curve $x^3 - 3x + b = 0$. Hence we must choose a nonzero value for c_{13}. But now the image of a vertical line in (a, b, x)-space is not vertical in (X, Y)-space. To compensate for this the projection of a point (a, b, x) onto the (a, b) plane was drawn by subtracting a fixed length from the Y coordinate of the image (X, Y) of that point. Since the four corners in the surface are almost the vertices of a parallelogram (but not exactly: they are not even coplanar) the resulting picture is not too bad.

§12. To draw a better picture we could use two transformations

$$(X, Y) = P(a, b, x), \qquad (X, Y) = Q(a, b, x);$$

the first transformation P to draw a parallelogram in the base and the second transformation Q to draw the intersection of the preimage $Q^{-1}(\ell)$ of a side ℓ of the parallelogram with the surface $x^3 - 3ax + b^2 = 0$. We impose the following two properties

(I) The restriction of P to the plane $x = 0$ is linear.

(II) If $(X_0, Y_0) = P(a, b, 0)$ then the image $(X, Y) = Q(a, b, x)$ of any point on the vertical line through $(a, b, 0)$ lies on the vertical line $X = X_0$ through $P(a, b, 0)$.

Condition (I) assures that the image under P of a parallelogram in the plane $x = 0$ is a parallelogram in the view plane.

Notes

1 Of course, the most general cubic equation is

$$c_3 y^3 + c_2 y^2 + c_1 y + c_0 = 0$$

where $c_3 \neq 0$. However it is easy to transform this equation equation to equation (1). After dividing by c_3 we may assume w.l.o.g. that $c_3 = 1$. Then, after performing the substitution $y = x - c_2/3$, the equation takes the desired form (1). (The substitution $y = x - c_2/3$ is an example of a *Tschirnhaus transformation*.)

2 `sage` is a free open-source mathematics software system licensed under the GPL. It combines the power of many existing open-source packages into a common Python-based interface. The URL is `http://www.sagemath.org`.

3 To run this program in `python` indicate raising to a power by `x**p` and be sure that at least one of the arguments in a division is not an integer (say by multiplying by1.0). In either language the command `from math import cos` must also be included in the program.

4 In calculus courses the letters a, b, x correspond to the letters x, y, z so the equation for the cusp surface would be $z^3 - 3xz + y = 0$.

5 In precalculus and first semester calculus the x-axis is the horizontal axis but in the drawings in Figure 2 it is the vertical axis.