Rational points on plane curves

Degree d plane curve is given by a *degree d polynomial in two variables*

$$\begin{aligned} f(x,y) = \quad & a_{d,0}x^d + a_{d-1,1}x^{d-1}y + \cdots + a_{0,d}y^d \\ & + a_{d-1,0}x^{d-1} + a_{d-2,1}x^{d-2}y + \cdots + a_{0,d-1}y^{d-1} \\ & \qquad \vdots \\ & + a_{1,0}x + a_{0,1}y + a_{0,0}. \end{aligned}$$

We will mostly be concerned with curves defined over the integers $\mathbb{Z}$, which means that *all the coefficients a_{ij} are integers.* A *rational point* on the curve is a pair (x,y) of rational numbers such that $f(x,y)=0$. The collection of *rational points* is

$$C_f(\mathbb{Q}) = \{(x,y) \in \mathbb{Q}^2 \mid f(x,y)=0\}.$$

The collection of *real points* is

$$C_f(\mathbb{R}) = \{(x,y) \in \mathbb{R}^2 \mid f(x,y)=0\},$$

and the *complex points* are

$$C_f(\mathbb{C}) = \{(x,y) \in \mathbb{C}^2 \mid f(x,y)=0\}.$$

Something not discussed as much in Chapter 5 of the text is *points modulo p*

$$C_f(\mathbb{Z}/p\mathbb{Z}) = \{(x,y) \in (\mathbb{Z}/p\mathbb{Z})^2 \mid f(x,y) \equiv 0 \pmod{p}\},$$

for a prime number p.

The *gradient of f at (x_0,y_0)* is

$$\nabla f(x_0,y_0) = \left(\frac{\partial f}{\partial x}(x_0,y_0), \frac{\partial f}{\partial y}(x_0,y_0)\right).$$

Curve is *smooth* if the gradient is nonzero at every complex point:

$$f(x_0,y_0)=0 \implies \nabla f(x_0,y_0) \neq 0 \qquad (x_0,y_0) \in \mathbb{C}^2.$$

The *tangent line to C_f* at a smooth point (x_0,y_0) is

$$(x-x_0)\frac{\partial f}{\partial x}(x_0,y_0) + (y-y_0)\frac{\partial f}{\partial y}(x_0,y_0) = 0.$$

(Notice that this really is the equation of a line through the point (x_0,y_0).)

In general the line of slope m through (x_0,y_0) has the equation

$$y - y_0 = m(x-x_0),$$

or the parametric equation

$$t \mapsto (x_0+t, y_0+mt).$$

In case of "infinite slope" (a vertical line) these equations need to be rewritten as

$$0 = (x - x_0), \qquad s \mapsto (x_0, y_0 + s).$$

The *homogeneous equation for the curve* is obtained by introducing a third variable z, and multiplying each monomial by a power of z so that it has total degree d:

$$\begin{aligned} F(x,y,z) = \quad & a_{d,0}x^d + a_{d-1,1}x^{d-1}y + \cdots + a_{0,d}y^d \\ & + a_{d-1,0}x^{d-1}z + a_{d-2,1}x^{d-2}yz + \cdots + a_{0,d-1}y^{d-1}z \\ & \qquad \vdots \\ & + a_{1,0}xz^{d-1} + a_{0,1}yz^{d-1} + a_{0,0}z^d. \end{aligned}$$

Solutions (x, y) to f correspond to solutions $(x, y, 1)$ of F. Because F is homogeneous, any multiple of a solution is again a solution:

$$F(x,y,z) = 0 \iff F(tx,ty,tz) = 0 \qquad (0 \neq t).$$

We therefore identify all these solutions: whenever $(x, y, z) \neq 0$, we write

$$[x : y : z] = \{(tx, ty, tz) \mid t \neq 0\}.$$

Therefore

$$[x : y : z] = [x' : y' : z'] \iff (x', y', z') = t(x, y, z) \qquad (\text{some } t \neq 0).$$

So the "finite points" of the curve are the points $[x : y : 1]$, and the "points at infinity" are

$$\{[x : y : 0] \mid (x, y) \neq (0, 0), F(x, y, 0) = 0\} = \{[x' : 1 : 0] \mid F(x', 1, 0)\}$$

together with perhaps

$$\{[x : 0 : 0] \mid x \neq 0, F(x, 0, 0) = 0\} = \{[1 : 0 : 0] \mid F(1, 0, 0) = 0\}.$$

Another way to say this is that the points at infinity of C_f include

$$\{[x : 1 : 0] \mid \text{degree } d \text{ terms of } f \text{ vanish at } (x, 1)\}$$

and perhaps

$$\{[1 : 0 : 0] \mid \text{if coefficient } a_{d,0} \text{ of } x^d \text{ is zero.}$$

This notion of "points at infinity" makes sense over any field where the curve makes sense: for us, $\mathbb{C}$, $\mathbb{R}$, $\mathbb{Q}$, and $\mathbb{Z}/p\mathbb{Z}$. Do you see why *a degree d curve has at most d points at infinity*? (This is a bit tricky: the points $[x : 1 : 0]$ are roots of a degree d polynomial, in x, so it seems that there could be d of them; but then there is one more possible point $[1 : 0 : 0]$, which seems to make $d + 1$.)

The curve is *smooth at infinity* (really *smooth at both finite and infinite points* if

$$F(x_0, y_0, z_0) = 0 \implies \nabla F(x_0, y_0, z_0) \neq 0 \qquad 0 \neq (x_0, y_0, z_0) \in \mathbb{C}^3.$$

Elliptic curves

An *elliptic curve* is a degree three smooth (also at infinity) curve f together with a chosen point $P_0 = [x_0 : y_0 : z_0]$ (over whatever field you are working). A moderately difficult theorem says that an elliptic curve can always be written in *Weierstrass form*

$$\text{(WEIERSTRASS)} \qquad y^2 = x^3 + bx + c, \qquad f(x, y) = y^2 - x^3 - bx - c,$$

with a, b, and c rational numbers. (This is not true for elliptic curves over the finite fields $\mathbb{Z}/2\mathbb{Z}$ and $\mathbb{Z}/3\mathbb{Z}$; to include those cases you need to allow equations

$$y^2 - a_1xy - a_3y = x^3 + a_2x^2 + a_4x + a_6,$$

but we won't worry about that.) The homogeneous equation is

$$F(x, y, z) = zy^2 - x^3 - bxz^2 - cz^3,$$

so the unique point at infinity is

$$\{[x : y : 0] \mid -x^3 = 0\} = \{[0 : 1 : 0]\}.$$

This point has rational coefficients. The *discriminant* of the curve is the discriminant of the cubic polynomial in x:

$$D = -4b^3 - 27c^2.$$

The curve is smooth if and only if $D \neq 0$. If p is a prime, the equation of the *curve makes sense modulo* p as long as the denominators of b and c don't involve p; that is, *for all but finitely many* p. It remains *smooth modulo* p if and only if $D \neq 0$ modulo p. So for example

$$y^2 = x^3 - 2$$

has $D = -108$. It is defined modulo p for all p, and smooth for $p \neq 2, 3$.

Big fact is that if f defines an elliptic curve, then *the points* $C_f(F)$ *over any field form an abelian group, with identity element the chosen base point* $P_0 = [x_0 : y_0 : z_0]$. If P_1 and P_2 are any two distinct points on C_f, then the line $L(P_1, P_2)$ through P_1 and P_2 meets C_f in a third point Q. The group law is characterized by

$$P_1 + P_2 + Q = P_0 \qquad \text{((when } P_1\text{, } P_2\text{, and } Q \text{ lie on a single line).}$$

Now assume the equation is in Weierstrass form; I'll write some formulas for the group law. The *inverse* of a point $P_1 = (x_1, y_1)$ is

$$\text{(INVERSE)} \qquad -P_1 = (x_1, -y_1).$$

Notice that the line through P_1 and $-P_1$ is the vertical line $x = x_1$; its third point of intersection with the curve is the point at infinity $[0 : 1 : 0]$. This means that

$$P_1 + (-P_1) + P_0 = P_0,$$

which is the equation defining inverses.

If $P_1 = (x_1, y_1) \neq (x_2, y_2) = P_2$, then $P_1 + P_2 = (x_3, y_3)$ has coordinates

$$\text{(PLUS)} \qquad x_3 = \left(\frac{y_2 - y_1}{x_2 - x_1}\right)^2 - x_1 - x_2, \quad y_3 = -y_1 - \left(\frac{y_2 - y_1}{x_2 - x_1}\right)(x_3 - x_1).$$

The only caveat in this formula is the case $x_2 = x_1$, when the line through P_1 and P_2 is vertical, and the third intersection point is the point $[0 : 1 : 0]$ at infinity.

Finally, there is formula for $P_4 = 2P_1 = (x_4, x_3)$, computed using not the secant but the tangent line at P_1:

$$\text{(DOUBLE)} \qquad x_4 = \left(\frac{3x_1^2 + b}{2y_1}\right)^2 - 2x_1, \quad y_4 = -y_1 - \left(\frac{3x_1^2 + b}{2y_1}\right)(x_4 - x_1).$$

The caveat is the case $y_1 = 0$, when the tangent is vertical and the third intersection point is $[0 : 1 : 0]$.

A *point of order 2* is $P = (x_0, y_0)$ such that $2P = 0$; that is, $y_0 = 0$. So

$$\{\text{points of order } 2\} = \{(x_0, 0) \mid x_0^3 + bx_0 + c = 0\}.$$

A *point of order 3* is $Q = (x_1, y_1)$ such that $3Q = 0$, or equivalently $2Q = -Q$. This means that the third curve point on the tangent line at Q is Q; that is, that the tangent line at Q is actually a *quadratic* approximation. From calculus you know that this is equivalent to $\frac{d^2y}{dx^2} = 0$. Easy calculus tells you that $\frac{dy}{dx} = (3x^2 + b)/y$, and slightly messier calculus gives

$$\frac{d^2y}{dx^2} = (3x^4 + 6bx^2 + 12cx - b^2)/y^3,$$

so

$$\{\text{points of order } 3\} = \{(x_1, y_1) \mid 3x_1^4 + 6bx_1^2 + 12cx_1 - b^2 = 0, y_1 = \pm\sqrt{x_1^3 + bx + c}\}.$$

You may prefer to find this formula by solving $x_4 = x_1$ in (DOUBLE) above.