18.782 Introduction to Arithmetic Geometry **Fall 2013**
Lecture #2 **09/10/2013**

2.1 Plane conics

A *conic* is a plane projective curve of degree 2. Such a curve has the form

$$C/k\colon ax^2 + by^2 + cz^2 + dxy + exz + fyz$$

with $a, b, c, d, e, f \in k$. Assuming the characteristic of k is not 2, we can make $d = e = f = 0$ via an invertible linear transformation. First, if $a = b = c = 0$ we can make one of them nonzero by replacing a variable by its sum with another; in this case one of d, e, f must be nonzero, say d, and then replacing y with $x + y$ yields an equation with $a \neq 0$. So assume without loss of generality that $a \neq 0$. Replacing x with $x - \frac{d}{2a}y$ kills the xy term, and we can similarly kill the xz term by replacing x with $x - \frac{e}{2a}z$ (we are just completing the square). Finally, if $f \neq 0$ we can make b nonzero and then replace y with $y - \frac{f}{2b}z$ to eliminate the yz term. Each of these substitutions corresponds to an invertible linear transformation of the projective plane, as does their composition.

So we now assume $\operatorname{char}(k) \neq 2$, and that C has the diagonal form

$$ax^2 + by^2 + cz^2 = 0. \tag{1}$$

If any of the coefficients a, b, c are zero, then this curve is not irreducible.[1] [2] For example, if the coefficient c is zero, we can factor the LHS of (1) over $\overline{k}$:

$$ax^2 + by^2 = (\sqrt{a}x + \sqrt{-b}y)(\sqrt{a}x - \sqrt{-b}y) = 0.$$

In this case $C(\overline{k})$ is the union of two projective lines that intersect at $(0 : 0 : 1)$ (but $C(k)$ might contain only one point, as when $k = \mathbb{Q}$ and $a, b > 0$, for example).

We now summarize this discussion with the following theorem.

Theorem 2.1. *Over a field whose characteristic is not 2, every geometrically irreducible conic is isomorphic to a diagonal curve $ax^2 + by^2 + cz^2 = 0$ with $abc \neq 0$.*

Remark 2.2. This does not hold in characteristic 2.

2.2 Parameterization of rational points on a conic

Suppose $(x_0 : y_0 : z_0)$ is a rational point on the diagonal conic $C\colon ax^2 + by^2 + cz^2 = 0$. Without loss of generality, we assume $z_0 \neq 0$ and consider the substitution

$$x = x_0W + U, \qquad y = y_0W + V, \qquad z = z_0W \tag{2}$$

[1] In Lecture 1 we defined a plane projective curve $f(x, y, z) = 0$ to be reducible if $f = gh$ for some $g, h \in \overline{k}[x, y, z]$, where $\overline{k}$ is the algebraic closure of k. Some authors distinguish between irreducibility over k versus $\overline{k}$, referring to the latter as *geometric* (or *absolute*) irreducibility. In this course we will always work with the notion of geometrically irreducibility.

[2] Out definition of a plane projective curve $f(x, y, z) = 0$ requires f to have no repeated factors in $\overline{k}[x, y, z]$, which precludes the case where two of a, b, c are zero. In more general settings, curves defined by a polynomial with repeated factors are said to be *non-reduced*. We require our curves to be reduced. Curves that are reduced and geometrically irreducible are also said to be *geometrically integral*.

where U, V, W denote three new variables. We then have

$$\begin{aligned} a(x_0W + U)^2 + b(y_0W + V)^2 + c(z_0W)^2 &= 0 \\ (ax_0^2 + by_0^2 + cz_0^2)W + 2(ax_0U + by_0V)W + aU^2 + bV^2 &= 0 \\ 2(ax_0U + by_0V)W &= -aU^2 - bV^2, \end{aligned}$$

where we have used $ax_0^2 + by_0^2 + cz_0^2 = 0$ to eliminate the quadratic term in W. After rescaling by $2(ax_0u + by_0v)$ and substituting for W in (2) we obtain the parameterization

$$\begin{aligned} x &= x_0(-aU^2 - bV^2) + 2(ax_0U + by_0V)U = ax_0U^2 + 2by_0UV - bx_0V^2 = Q_1(U, V) \\ y &= y_0(-aU^2 - bV^2) + 2(ax_0U + by_0V)V = -ay_0U^2 + 2ax_0UV + by_0V^2 = Q_2(U, V) \\ z &= z_0(-aU^2 - bV^2) = -az_0U^2 - bz_0V^2 = Q_3(U, V) \end{aligned}$$

Thus $(Q_1(U, V) : Q_2(U, V) : Q_3(U, V))$ is a polynomial map defined over k that sends each projective point $(U : V)$ on $\mathbb{P}^1$ to a point on the curve C. Moreover, we can recover the point $(U : V)$ via the inverse map from C to $\mathbb{P}^1$ defined by

$$U = x - \frac{x_0}{z_0}z, \qquad V = y - \frac{y_0}{z_0}z.$$

Thus we have an invertible map from C to $\mathbb{P}^1$ that is given by rational (in fact polynomial) functions that are defined at every point (such a map is said to be *regular*). In this situation we regard C and $\mathbb{P}^1$ as isomorphic curves. This yields the following theorem.

Theorem 2.3. *Let C/k be a geometrically irreducible conic with a k-rational point and assume that* $\operatorname{char}(k) \neq 2$. *Then C is isomorphic over k to the projective line $\mathbb{P}^1$.*

Remark 2.4. This theorem also holds when $\operatorname{char}(k) = 2$, but we will not prove this.

2.3 Conics over $\mathbb{Q}$

We now consider the case $k = \mathbb{Q}$. Given a diagonal conic

$$ax^2 + by^2 + cz^2 = 0$$

with $abc \neq 0$, we wish to either find a rational point (which we can then use to parameterize all the rational points), or prove that there are none. After clearing denominators we can assume a, b, c are nonzero integers, and we note that if they all have the same sign then there are clearly no rational points. So let us assume that this is not the case, and without loss of generality suppose that $a > 0$ and $b, c < 0$. Multiplying both sides by a and setting $d = -ab$ and $n = -ac$, we can put our curve in the form

$$x^2 - dy^2 = nz^2, \tag{3}$$

where d and n are positive integers that we may assume are square-free. Solving this equation is equivalent to expressing $n = (\frac{x}{z} + \frac{y}{z}\sqrt{d})(\frac{x}{z} - \frac{y}{z}\sqrt{d})$ as the norm of an element of the real quadratic field $\mathbb{Q}(\sqrt{d})$.

We now present a recursive procedure for doing this, based on Legendre's method of descent; the algorithm we give here is adapted from [1, Alg. I]. The basic idea is to either determine that there are no integer solutions to (3) (and hence no rational solutions), or to reduce the problem to finding a solution to a similar equation with smaller values of d or n (this is why it is called a *descent*). In order to facilitate the recursion, we allow d and n to also take negative values (but still insist that they be square-free).

Given square-free integers d and n, the procedure $\textsc{Solve}(d, n)$ either returns an integer solution to (3), or determines that no solution exists; we use the notation **`fail`** to indicate that the latter has occured.

```
Solve$(d,n)$
  1. If $d,n<0$ then fail.
  2. If $|d|>|n|$ then let $(x_0,y_0,z_0)=$ Solve$(n,d)$ and return $(x_0,z_0,y_0)$.
  3. If $d=1$ return $(1,1,0)$; if $n=1$ return $(1,0,1)$; if $d=-n$ return $(0,1,1)$.
  4. If $d=n$ then let $(x_0,y_0,z_0)=$ Solve$(-1,d)$ and return $(dz_0,x_0,y_0)$.
  5. If $d$ is not a quadratic residue modulo $n$ then fail.
  6. Let $x_0^2\equiv d \bmod n$, with $|x_0|\le |n|/2$, and let $t=t_1t_2^2=(x_0^2-d)/n$ with $t_1$ square-free.
  7. Let $(x_1,y_1,z_1)=$ Solve$(d,t_1)$ and return $(x_0x_1+dy_1, x_0y_1+x_1, t_1t_2z_1)$.
```

It is clear that if the algorithm **fail**s in steps 1 or 5 then (3) has no solutions, and that the solutions returned in step 3 are all correct. Assuming the algorithm works correctly when $|d|\le|n|$, then the solution returned in step 3 is clearly correct, and in step 4 with $d=n$, if Solve$(-1,d)$ succeeds then we have

$$\begin{aligned} x_0^2+y_0^2&=dz_0^2\\ dx_0^2+dy_0^2&=(dz_0)^2\\ (dz_0)^2-dx_0^2&=dy_0^2=ny_0^2, \end{aligned}$$

and therefore the solution (dz_0,x_0,y_0) is correct (note that -1 and d are both square-free, assuming the input d is, so our square-free constraint is preserved in the recursive call).

It remains to show that the solution returned in step 7 is correct, and that the algorithm is guaranteed to terminate. If we reach step 6 then we have $|d|<|n|$, and since $x_0^2-d=nt$, we have

$$|t|\le\frac{|x_0^2-d|}{|n|}\le\frac{|x_0|^2+|d|}{|n|}\le\frac{|d|^2}{4|n|}+\frac{|d|}{|n|}<\frac{|n|}{4}+\frac{|d|}{|n|}\le\frac{|n|}{2},$$

where the last inequality is justified by checking each of the cases $|n|=2$, $|n|=3$, and $|n|\ge 4$, remembering that the integer $|d|$ is at least 1 and strictly smaller than $|n|$. It follows that $|t_1|\le|t|<|n|$, which ensures that the algorithm will terminate, since either $|d|$ or $|n|$ is reduced in every recursive call; indeed, the number of recursive calls is clearly bounded by a logarithmic function of $\max(|d|,|n|)$.

To see that the solution returned in step 7 is correct, we first note that t_1 is square-free as required, and if Solve(d,t_1) succeeds then we may inductively assume that $x_1^2-dy_1^2=t_1z_1^2$. Multiplying the LHS by x_0^2-d and the RHS by $x_0^2-d=nt$ yields

$$\begin{aligned} (x_0^2-d)(x_1^2-dy_1^2)&=ntt_1z_1^2\\ x_0^2x_1^2-dx_0^2y_1^2-dx_1^2+d^2y_1^2&=nt_1t_2^2t_1z_1^2\\ (x_0x_1)^2+(dy_1)^2-d\left((x_0y_1)^2+x_1^2\right)&=n(t_1t_2z_1)^2\\ (x_0x_1+dy_1)^2-d(x_0y_1+x_1)^2&=n(t_1t_2z_1)^2, \end{aligned}$$

which shows that $(x_0x_1+dy_1, x_0y_1+x_1, t_1t_2z_1)$ is indeed a solution to (3), as desired.

Computationally, the most expensive step of the algorithm (by far) is the computation of x_0 in step 6. As we will see in the next lecture, it is easy to compute square-roots modulo primes, but in general n may be composite, and the only known algorithm for computing square-roots modulo a square-free composite integer n is to compute square-roots modulo each of its prime factors and use the Chinese remainder theorem to get a square-root modulo n. This requires factoring the integer n, a problem for which no polynomial-time algorithm is known.

As described in [1], the algorithm $\textsc{Solve}(d, n)$ can be modified to avoid factorization in any of its recursive steps so that only one initial factorization is required. This does not yield a polynomial-time algorithm, but it greatly speeds up the process, and in practice it is now feasible to find rational solutions to $ax^2 + by^2 + cz^2 = 0$ even when the coefficients a, b, and c are as large as 10^{100}.

Another deficiency of the algorithm $\textsc{Solve}(d, n)$ is that the solutions it finds are typically much larger than necessary. There is a theorem due to Holzer that gives us an upper bound on the size of the smallest solution to (1), and hence of the smallest solution to (3).

Theorem 2.5 (Holzer)**.** *Let a, b, c be square-free integers that are pairwise coprime and suppose that the equation $ax^2 + by^2 + cz^2 = 0$ has a nonzero rational solution. Then there exists a nonzero integer solution (x_0, y_0, z_0) with*

$$|x_0| \le \sqrt{|bc|}, \qquad |y_0| \le \sqrt{|ac|}, \qquad |z_0| \le \sqrt{|ab|}.$$

Proof. See [2] for a short and elementary proof. □

On Problem Set 1 you will implement a simple improvement to algorithm $\textsc{Solve}(d, n)$ that significantly reduces the size of the solutions it finds (and reduces the number of recursive calls), and generally comes close to achieving the Holzer bounds.

Finally, we note that there is a simple criterion for determining whether or not a diagonal conic has a rational solution that does not require actually looking for one.

Theorem 2.6 (Legendre)**.** *Let a, b, c be square-free integers that are pairwise coprime and whose signs are not all the same. The equation $ax^2 + by^2 + cz^2 = 0$ has a rational solution if and only if the congruences*

$$X^2 \equiv -bc \bmod a, \qquad Y^2 \equiv -ca \bmod b, \qquad Z^2 \equiv -ab \bmod c$$

can be simultaneously satisfied.

The necessity of the condition given in Theorem 2.6 is easy to check; if we look at the equation modulo a, for example, we have $by^2 \equiv -cz^2 \bmod a$, and it follows that $-b/c$ and therefore $-bc$ must be a quadratic residue modulo a. The sufficiency can be proved by showing that if the condition holds than $\textsc{Solve}(d, n)$ will succeed in finding a solution to the corresponding norm equation $x^2 - dy^2 = nz^2$. This is basically how Legendre proved the theorem, but we will prove a more general statement after we have developed the theory of p-adic numbers.

It is worth noting that while the congruences in Legendre's theorem apparently give a very simple criterion for determining whether a conic has a rational point, in order to apply them we need to know the factorization of the integers a, b, c. This means that, in general, the problem of determining the existence of a rational solution is not significantly easier than actually finding one, and we still do not have a polynomial-time algorithm for determining the existence of a rational solution to a conic over $\mathbb{Q}$.

References

[1] J.E. Cremona and D. Rusin, *Efficient solution of rational conics*, Mathematics of Computation **72** (2003), 1417–1441.

[2] T. Cochrane and P. Mitchell, *Small solutions of the Legendre equation*, Journal of Number Theory **70** (1998), 62–66.

18.782 Introduction to Arithmetic Geometry
Fall 2013