The Local-Global Principle for Conics

Jonathan Bloom

May 31, 2004

A conic is defined by a nonsingular quadratic form in three variables.

Definition. A *conic $\mathcal{C}$ over $\mathbb{Q}$ is given by an equation*

$$F(\mathbf{X}) = \sum f_{ij} X_i X_j = 0$$

where $\mathbf{X} = (X_1, X_2, X_3)$, $f_{ij} = f_{ji} \in \mathbb{Q}$, and $\det f_{ij} \neq 0$.

When does a conic contain a rational point (a solution to $F(\mathbf{X}) = 0$ with rational coordinates)? The Local-Global Principle says that a global solution exists (in $\mathbb{Q}$) if and only if local solutions exist (in $\mathbb{R}$ and all $\mathbb{Q}_p$, which are the completions of $\mathbb{Q}$ with respect to the distinct valuations). It was first proven in the 1920s by Helmut Hasse.

Theorem (The Local-Global Principal for Conics). *Let $\mathcal{C}$ be a conic defined over $\mathbb{Q}$. Then $\mathcal{C}$ has a point in $\mathbb{Q}$ if and only if $\mathcal{C}$ has points in $\mathbb{R}$ and $\mathbb{Q}_p$ for every prime p.*

The *only if* direction is trivial, as solutions in $\mathbb{Q}$ also lie in each of the completions of $\mathbb{Q}$. In this sense, a solution in $\mathbb{Q}$ is global.

To prove the *if* direction, we will use Minkowski's Theorem from the geometry of numbers (restricting ourselves to subsets S that are Lebesgue measurable).

Theorem (Minkowski). *Let Λ be a subgroup of $\mathbb{Z}^n$ of index m. Let $S \subset \mathbb{R}^n$ be a symmetric convex set of volume*

$$V(S) > 2^n m.$$

Then S and Λ have a common point other than $\mathbf{0} = (0, ..., 0)$.

In fact, if S is closed, then we can weaken the condition to $V(S) \geq 2^n m$ using a limit argument. We will need the following lemma, which is essentially the pigeonhole principle.

Lemma. *Let S be a bounded subset of $\mathbb{R}^n$ with $V(S) > m$. Then there exists some $w \in \mathbb{R}^n/\mathbb{Z}^n$ with $|\pi^{-1}(w)| > m$, where $\pi : S \to \mathbb{R}^n/\mathbb{Z}^n$ is projection.*

Proof. We first give a conceptual proof. Imagine cutting S up along the lattice hyperplanes and integrally translating each piece to the fundamental cube $I = \{0 \leq x_i < 1 \ \forall \ 1 \leq i \leq n\}$, which we identify with $\mathbb{R}^n/\mathbb{Z}^n$. The number of preimages via π of a point $w \in I$ is equal to the number of translated pieces of S lying above w. Since $V(S) > mV(I)$, intuitively some point w must lie below more than m pieces of S, i.e. $|\pi^{-1}(w)| > m$.

We could formalize this argument further by assigning multiplicities to each "region" of I, but if there are infinitely many such regions we may run into trouble when summing. Instead follow Cassels lead and make the argument rigorous by hiding the limits within integrals. Let $\psi(x)$ by the characteristic function of S. Every $x \in \mathbb{R}^n$ can be uniquely expressed as $w + z$ with $w \in I$ and $z \in \mathbb{Z}^n$. Therefore

$$m < V(S) = \int_{\mathbb{R}^n} \psi(x)dx = \int_I \psi(x) \left(\sum_{z\in\mathbb{Z}^n} \psi(w+z)\right) dw.$$

If $\sum_{z\in\mathbb{Z}^n} \psi(w+z) \leq m$ for every $w \in I$, then we would have the contradiction

$$V(s) = \int_I \psi(x) \left(\sum_{z\in\mathbb{Z}^n} \psi(w+z)\right) dw \leq m \int_I \psi(x)dw = m.$$

Therefore, there exists some $w \in I$ with $|\pi^{-1}(w)| = \sum_{z\in\mathbb{Z}^n} \psi(w+z) > m$. □

Proof of Minkowski. $V(S) > 2^n m$ implies $V(S/2) > m$. By the Lemma, there exists $w \in \mathbb{R}^n/\mathbb{Z}^n$ with $|\pi^{-1}(w)| > m$, where $\pi : S/2 \to \mathbb{R}^n/\mathbb{Z}^n$. So we have distinct points $p_0, p_1, ..., p_m \in S$ with $\frac{p_i}{2} - \frac{p_j}{2} \in \mathbb{Z}^n$ for every $0 \leq i, j \leq m$. Then $T = \{\frac{p_0-p_i}{2} \mid 0 \leq i \leq m\} \subset \mathbb{Z}^n$ is a set of size $m + 1$. Furthermore, $T \subset S$ since $p_i \in S$ by symmetry and then $\frac{p_0+(-p_i)}{2} \in S$ by convexity. So we have $m + 1$ distinct points in $S \bigcap \mathbb{Z}^n$ (in fact, a little more care gives 2m + 1 distinct points).

Now Λ has index m in $\mathbb{Z}^n$, so by the pigeonhole principle, there exist distinct points $x, y \in S \bigcap \mathbb{Z}^n$ which are in the same coset of Λ. Thus, $x - y$ is a non-zero point of $S \bigcap \Lambda$. □

Proof of the Local-Global Principle for Conics. Our proof fleshes out that of Cassels in *Lectures on Elliptic Curves*. The first step is to normalize $F(\mathbf{X})$ to a diagonal form, which we will prove for an arbitrary number of variables.

Suppose we change coordinates by sending X_i to $X'_i = \sum_i t_{ij} X_i$, with $\det t_{ij} \neq 0$. Then $T = (t_{ij})$ gives a bijection between the rational points of $F(\mathbf{X})$ and the rational points of $G(\mathbf{X}')$. Similarly, T gives a bijection between the points in $\mathbb{R}$ on $F(\mathbf{X})$ and $G(\mathbf{X}')$, as well as between such points in $\mathbb{Q}_p$. Clearly we can also multiply through by any non-zero rational without affecting the solution set.

So suppose by induction that there exist coordinates $X_1, ..., X_n$ such that

$$F(X_1, ..., X_n) = f_1 X_1^2 + ... + f_{r-1} X_{r-1}^2 + \sum_{i,j=r}^{n} f_{ij} X_i X_j$$

$f_{ij} = f_{ji} \in \mathbb{Q}$ and the matrix $\det f_{ij} \neq 0$. The base case $r = 0$ is given.

For the induction step, we first show that we can make $f_{rr}(0) \neq 0$ by a non-singular linear transformation on the last $n - r + 1$ coordinates. The proof works the same for any r, so for simplicity let r = 1. If we have $f_{ii}(0) \neq 0$ for some $1 \leq i \leq n$ then we are done by transposing X_1 and X_i. Otherwise, since $(f_{ij}(0))$ is non-singular, there exists some $f_{ij}(0) \neq 0$ with $i \neq j$. Through a pair of transpositions, we can assume $f_{11}(0) = 0$ and $f_{12}(0) = f_{21}(0) \neq 0$. We define a new set of coordinates $X'_1, ..., X'_n$ by

$X'_1 = \frac{1}{2}(X_1 + X_2)$ $\qquad$ $X'_2 = \frac{1}{2}(X_1 - X_2)$ $\qquad$ $X'_i = X_i$ for $i > 2$

This linear transformation is invertible with inverse given by

$X_1 = (X'_1 - X'_2)$ $\qquad$ $X_2 = (X'_1 + X'_2)$ $\qquad$ $X_i = X'_i$ for $i > 2$

Substituting in these new coordinates and regrouping terms, we have $F(\mathbf{X}) = \Sigma_{i,j=1}^{n} f'_{ij} X'_i X'_j$ with $f'_{11}(0) = f_{12}(0) + f_{21}(0) = 2f_{12} \neq 0$.

So without loss of generality we assume $f_{rr} \neq 0$. After dividing through $F(\mathbf{X})$ by f_{rr}, we may assume $f_{rr} = 1$. We now define a new set of coordinates $X'_1, ..., X'_n$ by

$X'_i = X_i$ for $i \neq r$.

$X'_r = X_r + \sum_{i>r} f_{ir} X_i$

Then

$$\begin{aligned} F(\mathbf{X}) &= f_1X_1^2 + ... + f_{r-1}X_{r-1}^2 + \sum_{i,j=r}^{n} f_{ij}X_iX_j \\ &= f_1X_1^2 + ... + f_{r-1}X_{r-1}^2 + \left[X_r^2 + 2X_r\sum_{i>r} f_{ir}X_i + (\sum_{i>r} f_{ir}X_i)^2\right] \\ &\quad -(\sum_{i>r} f_{ir}X_i)^2 + \sum_{i,j>r, i\neq j} f_{ij}X_iX_j. \end{aligned}$$

The term in brackets is $X_r'^2$ so it is clear that we can choose f_{ij}' for $i, j > r$ so that

$$F(\mathbf{X}) = \textstyle\sum_{i=1}^{r} f_i'X_i'^2 + \sum_{i,j>r}^{n} X_i'X_j'f_{ij}'.$$

with $f_{ij}' = f_{ji}'$. Furthermore, $f_{ij}' = 1/f_{rr}P^t f_{ij}P$ (for some non-singular P) is non-singular. This completes the induction step.

So now $F(\mathbf{X}) = f_1X_1^2 + f_2X_2^2 + f_3X_3^2$, with each $f_i \neq 0$. Multiply through by the product of the denominators to get each $f_i \in \mathbb{Z}$. Now we will further reduce to the case where $f_1f_2f_3 \in \mathbb{Z}$ is square-free. If a prime p divides two coefficients, send the other variable to p times itself, thus introducing two factors of p into $f_1f_2f_3$. Follow up by dividing through by p, removing three factors of p from $f_1f_2f_3$. Since $|f_1f_2f_3|$ is reduced in $\mathbb{Z}$ on each cycle, this process will eventually terminate, leaving the f_i pair-wise coprime. Finally, if $p^2 \mid f_i$, send $X_i \to \frac{1}{p}X_i$. Repeat until each f_i is square-free, and thus $f_1f_2f_3$ is square-free.

Our goal is to define a subgroup Λ of index $m = 4|f_1f_2f_3|$ in $\mathbb{Z}^3$ such that $F(\mathbf{x}) \equiv 0\ (4|f_1f_2f_3|)$ for $\mathbf{x} \in \Lambda$. Given this, we can apply Minkowski's Theorem to Λ and the convex symmetric set

$$S\ :\ |f_1|x_1^2 + |f_2|x_2^2 + |f_3|x_3^2 < 4|f_1f_2f_3|.$$

Then $V(S) = (\pi/3)2^3|4f_1f_2f_3| > 2^3|4f_1f_2f_3| = 2^3m$. So by Minkowski, there is an $\mathbf{x} \in S \bigcap \Lambda$ for which

$$F(\mathbf{x}) \equiv 0\ \ (4|f_1f_2f_3|)$$

and

$$|F(\mathbf{x})| \leq |f_1|x_1^2 + |f_2|x_2^2 + |f_3|x_3^2 < 4|f_1 f_2 f_3|.$$

Therefore, $F(\mathbf{x}) = 0$. A rational point exists!

As satisfying as that was, we must now actually construct said Λ. If $\mathbf{a} = (a_1, a_2, a_3) \neq (0, 0, 0)$ is a point in $\mathbb{Q}_p$ with $F(\mathbf{a}) = 0$, then all multiples of $\mathbf{a}$ are solutions as well. So without loss of generality, we may assume that

$$\max |a_j|_p = 1.$$

We have pushed $\mathbf{a}$ into $\mathbb{Z}_p$ with minimal force.

We now divide out attack into three cases.

Case 1: $p \neq 2$, $p \mid f_1 f_2 f_3$. Without loss of generality, $p \mid f_1$, so $p \nmid f_2, f_3$. Then $|f_1 a_1^2|_p < 1$. If it were the case that $|a_2|_p < 1$, then

$$|f_3 a_3^2|_p = |f_1 a_1^2 + f_2 a_2^2|_p < 1$$

by the ultra-metric inequality, and thus $|a_3|_p < 1$. Thus,

$$|f_1 a_1^2|_p = |f_2 a_2^2 + f_3 a_3^2|_p < p^{-2}$$

and so $|a_1|_p < 1$ as well, contradicting the normalization. Therefore $|a_2|_p = |a_3|_p = 1$. This gives

$$|f_2 a_2^2 + f_3 a_3^2|_p < 1$$

and we can divide by the unit a_2 to deduce that there is some $r_p \in \mathbb{Z}$ such that

$$f_2 + r_p^2 f_p \equiv 0 \pmod{p}.$$

We impose the condition

$$x_3 \equiv r_p x_2 \pmod{p}.$$

Then

$$\begin{aligned} F(\mathbf{x}) &\equiv f_1 x_1^2 + f_2 x_2^2 + f_3 x_3^2 \\ &\equiv (f_2 + r_p^2 f_3) x_2^2 \\ &\equiv 0 \pmod{p}. \end{aligned}$$

Case 2: $p = 2$, $2 \nmid f_1 f_2 f_3$. Working modulo 2, we see that exactly two of the a_i must be units, say a_1 and a_2. $a^2 \equiv 0$ or 1 (4) for $a \in \mathbb{Z}$, giving

$$f_2 + f_3 \equiv 0 \pmod 4.$$

We impose the conditions

$$x_1 \equiv 0 \pmod 2$$
$$x_2 \equiv x_3 \pmod 2$$

Then $F(\mathbf{x}) \equiv 0 \pmod 4$.

Case 3: $p = 2$, $2 \mid f_1 f_2 f_3$. Without loss of generality, $2 \mid f_1$ so $2 \nmid f_2, f_3$. As in Case 1, we have $|a_2|_2 = |a_3|_2 = 1$. For $a \in \mathbb{Z}$ odd, we have $a^2 \equiv 1$ (8). So

$$f_2 + f_3 \equiv 0 \text{ (8) if } 2 \mid a_1 \quad (\text{set } s = 0)$$

or

$$f_1 + f_2 + f_3 \equiv 0 \text{ (8) if } 2 \nmid a_1 \quad (\text{set } s = 1)$$

We impose the conditions

$$x_2 \equiv x_3 \pmod 4$$
$$x_1 \equiv s x_3 \pmod 2$$

Then $F(\mathbf{x}) \equiv 0 \pmod 8$.

We define Λ to be the subgroup of $\mathbb{Z}^3$ satisfying these equivalence relations. Λ has index $m = 4|f_1 f_2 f_3|$ and by construction $F(\mathbf{x}) \equiv 0$ $(4|f_1 f_2 f_3|)$ for $\mathbf{x} \in \Lambda$. □

The quadratic result can be extended from $\mathbb{Q}$ to number fields (and more generally, global fields). It would be convenient if the Local-Global Principle extended to curves of higher degree as well. However, it is known to fail even in the cubic case. For example, the curve defined by $3X^3 + 4Y^3 + 5Z^3 = 0$ has non-zero points in $\mathbb{R}$ and $\mathbb{Q}_p$ for every prime p, but no non-zero rational points (see §18 of Cassels). Yet if the number of variables is bumped up to 9, the cubic result holds as well. According to †: "The 'large number of variables' results depend on the Hardy-Littlewood circle

method, which was extended to all number fields by C. L. Siegel (quadratic case) and C. P. Ramanujam (in general)."

†http://encyclopedia.thefreedictionary.com/ local-global%20principle