

FINITENESS OF SOLUTIONS TO THE  $S$ -UNIT EQUATION, PART II  
 February 25, 2003

Recall that our goal is to prove the following two results:

**Proposition 1.** (*Gap principle*) *There exists a constant  $M_1 > 1$  such that if  $u, v \in Z_{>L}$  are distinct and  $\text{angle}(u, v) \leq \delta$ , then  $\|v\| \geq \|u\|$  implies  $\|v\| \geq M_1 \|u\|$ .*

**Proposition 2.** (*Cap principle*) *There exists a constant  $M_2 > 1$  such that if  $u, v \in Z_{>L}$  and  $\text{angle}(u, v) \leq \delta$ , then  $\|v\| \leq M_2 \|u\|$ .*

Here  $Z$  is the image in  $\Lambda := \Gamma/\Gamma^{\text{tor}}$  of the set of solutions  $(x, y) \in \Gamma$  to the unit equation  $x + y = 1$ . Recall also that  $L$  is a suitably large real number,  $\delta$  is a suitably small positive real number, and  $Z_{>L}$  denotes the set of elements of  $Z$  with norm greater than  $L$ . The norm of an element  $u$  of  $\Lambda$  is defined to be

$$\|u\| = h(x) + h(y),$$

where  $(x, y)$  is any element of  $\Gamma$  representing the class  $u \in \Gamma/\Gamma^{\text{tor}}$ .

1. THE GAP PRINCIPLE

We begin with some preliminary simple results concerning heights of algebraic numbers.

First, some notation: Let  $K$  be a number field. For  $v \in M_K$ , we let

$$\epsilon_v := \begin{cases} 0 & v \in M_K^0 \\ 1 & v \in M_K^\infty, v \text{ real} \\ 2 & v \in M_K^\infty, v \text{ complex.} \end{cases}$$

Then we have the following lemma, whose proof is immediate.

**Lemma 1.** *If  $x, y \in K$  and  $v \in M_K$ , then*

$$\|x + y\|_v \leq 2^{\epsilon_v} \max\{\|x\|_v, \|y\|_v\}.$$

From this, we obtain the following estimate:

**Lemma 2.** *If  $a, b, a', b' \in K$  and  $v \in M_K$ , then*

$$\|ab' - a'b\|_v \leq 2^{\epsilon_v} \max\{\|a\|_v, \|b\|_v\} \max\{\|a'\|_v, \|b'\|_v\}.$$

*Proof.*

$$\begin{aligned} \|ab' - a'b\|_v &\leq 2^{\epsilon_v} \max\{\|ab'\|_v, \|a'b\|_v\} \\ &= 2^{\epsilon_v} \max\{\|a\|_v \|b'\|_v, \|a'\|_v \|b\|_v\} \\ &\leq 2^{\epsilon_v} \max\{\|a\|_v, \|b\|_v\} \max\{\|a'\|_v, \|b'\|_v\}. \end{aligned}$$

□

The proofs of both principles will involve the following simple lemma which bounds the heights of solutions to linear equations in two variables.

**Lemma 3.** *Let  $a, b, a', b', c, c' \in \overline{\mathbf{Q}}$ , and suppose that  $\Delta := ab' - a'b \neq 0$ . Then if  $x, y \in \overline{\mathbf{Q}}$  are such that*

$$\begin{aligned} ax + by &= c \\ a'x + b'y &= c', \end{aligned}$$

*we have*

$$h(x : y : 1) \leq h(a : b : c) + h(a' : b' : c') + \log 2.$$

*Proof.* By Cramer's rule,  $x = \frac{bc' - cb'}{ab' - a'b}$  and  $y = \frac{ac' - ca'}{ab' - a'b}$ .

Therefore, if  $K$  is a number field containing all of the algebraic numbers involved in the lemma, we have

$$\begin{aligned} h(x : y : 1) &= h(bc' - cb' : ac' - ca' : ab' - a'b) \\ &= \frac{1}{[K:\mathbf{Q}]} \sum_{v \in M_K} \log \max\{\|bc' - cb'\|_v, \|ac' - ca'\|_v, \|ab' - a'b\|_v\}. \end{aligned}$$

The previous lemma now gives

$$\begin{aligned} h(x : y : 1) &\leq \frac{1}{[K:\mathbf{Q}]} \sum_{v \in M_K} \log \max\{\|a\|_v, \|b\|_v, \|c\|_v\} \\ &\quad + \log \max\{\|a'\|_v, \|b'\|_v, \|c'\|_v\} + \epsilon_v \log 2 \\ &= h(a : b : c) + h(a' : b' : c') + \log 2. \end{aligned}$$

□

If  $x, y \in \overline{\mathbf{Q}}$ , we let  $h(x, y) = h(x : y : 1)$  be the height of  $(x, y) \in \mathbf{A}^2(\overline{\mathbf{Q}})$  as a point of  $\mathbf{P}^2(\overline{\mathbf{Q}})$ . We thus have two notions of height on  $\mathbf{A}^2(\overline{\mathbf{Q}})$ ,  $h$  and  $\hat{h}$ , and we leave it as an exercise to see that for all  $P \in \mathbf{A}^2(\overline{\mathbf{Q}})$ , we have

$$\frac{1}{2} \hat{h}(P) \leq h(P) \leq \hat{h}(P).$$

Lemma 3 has the following corollary, which is closely related to the gap principle:

**Corollary 1.** *Suppose  $P_1 = (x_1, y_1)$  and  $P_2 = (x_2, y_2)$  are in  $\overline{\mathbf{Q}}^* \times \overline{\mathbf{Q}}^*$ , with  $P_1 \neq P_2$ . Suppose furthermore that  $x_1 + y_1 = 1$  and  $x_2 + y_2 = 1$ . Then*

$$h(P_1) \leq h(P_2 P_1^{-1}) + \log 2.$$

*Proof.* We have the system of linear equations

$$\begin{aligned} 1 \cdot x_1 + 1 \cdot y_1 &= 1 \\ \frac{x_2}{x_1} \cdot x_1 + \frac{y_2}{y_1} y_1 &= 1. \end{aligned}$$

Note that  $\Delta := \frac{y_2}{y_1} - \frac{x_2}{x_1} \neq 0$ , for otherwise  $x_1 y_2 = x_2 y_1$ , which implies that  $x_1(1 - y_1) = (1 - x_1)y_1$ , i.e.,  $x_1 = y_1$ , and therefore  $x_2 = y_2$  as well.

Now apply Lemma 3 with  $(x, y) = (x_1, y_1)$ ,  $(a, b) = (1, 1)$ , and  $(a', b') = (x_2/x_1, y_2/y_1)$ .  $\square$

This corollary, together with the estimates

$$\frac{1}{2} \hat{h}(P) \leq h(P) \leq \hat{h}(P),$$

implies the following important geometric constraint on  $Z$ :

**Corollary 2.** *For all  $u, v \in Z$ ,  $u \neq v$ , we have*

$$\|u\| \leq 2\|v - u\| + 2 \log 2.$$

This estimate implies the gap principle. Intuitively, this seems reasonable: if  $\|u\|$  and  $\|v\|$  both lie in a cone of small angle and if  $\|v\|$  is only slightly bigger than  $\|u\|$ , then  $\|v - u\|$  will be small, so Lemma 2 gives a bound on the size of  $\|u\|$ . Therefore if  $\|u\|$  is known to be large, then  $\|v\|$  will have to be “significantly” larger than  $\|u\|$ .

To make this heuristic reasoning precise, we will use the following simple geometric inequality:

**Lemma 4.** *Suppose  $u, v \in V$  are nonzero vectors with  $\|u\| \leq \|v\|$ . Then*

$$\|v - u\| \leq \|v\| - \|u\| + \|u\| \text{angle}(u, v).$$

*Proof.* By the triangle inequality, if we set

$$\eta = \frac{\|v\|}{\|u\|} - 1,$$

then we have

$$\begin{aligned} \|v - u\| &= \|u\| \left( \|\nu(v) \frac{\|v\|}{\|u\|} - \nu(u)\| \right) \\ &= \|u\| (\|\nu(v) - \nu(u) + \eta \nu(v)\|) \\ &\leq \|u\| (\|\nu(v) - \nu(u)\| + \eta) \\ &= \|v\| - \|u\| + \|u\| \text{angle}(u, v). \end{aligned}$$

$\square$

We now prove the following more precise version of the gap principle:

**Proposition 3.** *If  $u, v \in Z$  are distinct elements with  $\|v\| \geq \|u\| \geq 14$  and  $\text{angle}(u, v) \leq \frac{1}{10}$ , then  $\|v\| > \frac{5}{4}\|u\|$ .*

*Proof.* Using Corollary 2 and Lemma 4, we find (since  $2 \log 2 \approx 1.386 < 1.4$ ) that

$$\begin{aligned} \|u\| &\leq 2\|v - u\| + 2 \log 2 \\ &\leq 2(\|v\| - \|u\| + \|u\| \text{angle}(u, v)) + 2 \log 2 \\ &= 2\|v\| - \frac{9}{5}\|u\| + 1.4 \\ &\leq 2\|v\| - \frac{9}{5}\|u\| + \frac{1}{10}\|u\| \end{aligned}$$

so that

$$\|u\| \leq \frac{20}{27}\|v\| < \frac{20}{25}\|v\|$$

as desired.  $\square$

## 2. THE CAP PRINCIPLE

One way to get a cap principle would be to show that if  $u, v \in Z$  then for each positive integer  $m$ , the size of  $\|u\|$  is small compared to the size of  $\|v - mu\|$ . For then if  $\text{angle}(u, v)$  is small and  $\|v\| \approx m\|u\|$ , the quantity  $\|v - mu\|$  will be small and therefore  $\|u\|$  must be small.

We will obtain a bound of this type for *even* integers  $m$  only, which will suffice. The key idea is to use the polynomials  $(A_n, B_n, C_n)$  introduced in our lectures on the ABC conjecture.

We need a few lemmas before we can get started.

**Lemma 5.** *Let  $P = (\alpha_1, \dots, \alpha_n) \in \mathbf{A}^n(\overline{\mathbf{Q}})$ , and let  $h(P) := h(\alpha_1 : \dots : \alpha_n : 1)$ . Let  $f_1, \dots, f_m \in \mathbf{Z}[x_1, \dots, x_n]$ . Then*

$$h(f_1(\alpha_1, \dots, \alpha_n) : \dots : f_m(\alpha_1, \dots, \alpha_n)) \leq \deg(f)h(P) + \log \max\{|f_1|_{L^1}, \dots, |f_m|_{L^1}\}.$$

*Proof.* If  $f \in \mathbf{Z}[x]$  and  $\delta := \deg(f)$ , then for any number field  $K$  containing  $\alpha_1, \dots, \alpha_n$  and any  $v \in M_K$ , we have

$$\|f(\alpha_1, \dots, \alpha_n)\|_v \leq \max\{\|\alpha_1\|_v, \dots, \|\alpha_n\|_v, 1\}^d |f|_{L^1}^{e_v}.$$

The result follows easily.  $\square$

We leave the proof of the next lemma as an exercise.

**Lemma 6.** *If  $f(t) \in \mathbf{Z}[t]$ , then*

- $|fg|_{L^1} \leq |f|_{L^1}|g|_{L^1}$ .
- $|f'(t)|_{L^1} \leq \deg(f)|f|_{L^1}$ .
- $|f(1-t)|_{L^1} \leq 2^{\deg(f)}|f|_{L^1}$ .

The next lemma will allow us to get around a potential stumbling block in the proof of Proposition 4 below.

**Lemma 7.** *Let  $n \geq 1$  be a positive integer. Then any  $\mathbf{C}$ -linear combination of  $A_n(t)$ ,  $B_n(t)$ , and  $C_n(t)$  is either identically zero, or else has simple roots outside  $\{0, 1, \infty\}$ .*

*Proof.* Since  $B_n(t)$  is a linear combination of  $A_n(t)$  and  $C_n(t)$ , we may assume that  $h(t) = \alpha A_n(t) + \gamma C_n(t)$  for some complex numbers  $\alpha, \gamma$ . We may assume that  $\alpha$  and  $\gamma$  are not both zero, and then it follows that  $h(t) \neq 0$ , since  $A_n(t)$  and  $C_n(t)$  are relatively prime. Note also that  $-\gamma/\alpha$  is well-defined as an element of  $\mathbf{P}^1$ .

If we define  $\phi(t) = \frac{A(t)}{C(t)}$ , then we have

$$h(t_0) = 0 \Leftrightarrow \phi(t_0) = -\frac{\gamma}{\alpha},$$

and  $h(t)$  has a multiple zero at  $t_0$  if and only if  $\phi(t_0) = -\frac{\gamma}{\alpha}$  and  $e_{t_0}(\phi) \geq 2$ .

But  $e_{t_0} \geq 2$  implies that  $t_0 \in \{0, 1, \infty\}$ , since  $\phi$  is a Belyi function. Therefore if  $t_0 \notin \{0, 1, \infty\}$ , then  $h(t)$  has at most a simple zero at  $t_0$ .  $\square$

**Proposition 4.** *Suppose  $P_1 = (x_1, y_1)$  and  $P_2 = (x_2, y_2)$  are in  $\overline{\mathbf{Q}}^* \times \overline{\mathbf{Q}}^*$ , with  $P_1 \neq P_2$ . Suppose furthermore that  $x_1 + y_1 = 1$  and  $x_2 + y_2 = 1$ . Then for any integer  $n \geq 2$ , we have*

$$h(P_1) \leq \frac{1}{n-1} h(P_2 P_1^{-2n}) + M.$$

where  $M > 0$  is a real constant independent of  $P_1, P_2$ .

*Proof.* Let  $A = A_n(t), B = B_n(t), C = C_n(t)$ , and similarly define  $\tilde{A}, \tilde{B}, \tilde{C}$ , so that  $A(t) = t^{2n+1} \tilde{A}(t)$  and  $B(t) = (1-t)^{2n+1} \tilde{B}(t)$ .

Since  $A(t) + B(t) = C(t)$ , we have

$$x_1^{2n+1} \tilde{A}(x_1) + y_1^{2n+1} \tilde{B}(x_1) = C(x_1),$$

so that we have the system of linear equations

$$\begin{aligned} \frac{x_2}{x_1^{2n}} \cdot x_1^{2n} + \frac{y_2}{y_1^{2n}} \cdot y_1^{2n} &= 1 \\ x_1 \tilde{A}(x_1) \cdot x_1^{2n} + y_1 \tilde{B}(x_1) \cdot y_1^{2n} &= C(x_1). \end{aligned}$$

Let  $a = \frac{x_2}{x_1^{2n}}, b = \frac{y_2}{y_1^{2n}}, a' = x_1 \tilde{A}(x_1), b' = y_1 \tilde{B}(x_1)$ .

Define  $\Delta := ab' - a'b$ , and suppose that  $\Delta \neq 0$ .

Then by Lemma 3, we have

$$\begin{aligned} h(P_1^{2n}) &= 2nh(P_1) \\ &\leq h\left(\frac{x_2}{x_1^{2n}}, \frac{y_2}{y_1^{2n}}, 1\right) + h(x_1 \tilde{A}(x_1), y_1 \tilde{B}(x_1), C(x_1)) + \log 2 \\ &= h(P_2 P_1^{-2n}) + h(x_1 \tilde{A}(x_1), y_1 \tilde{B}(x_1), C(x_1)) + \log 2. \end{aligned}$$

Since  $\tilde{A}(t) = (-1)^n \tilde{B}(1-t)$  and  $\tilde{C}(t) = (-1)^n t^n \tilde{B}(1 - \frac{1}{t})$ , we have

$$\tilde{A}(x_1) = (-1)^n \tilde{B}(y_1), \quad \tilde{C}(x_1) = (-1)^n x_1^n \tilde{B}\left(-\frac{y_1}{x_1}\right),$$

so that by Lemma 5, we have

$$\begin{aligned} 2nh(P_1) &\leq h(P_2 P_1^{-2n}) + h(\pm x_1 \tilde{B}(y_1), y_1 \tilde{B}(x_1), \pm x_1^n \tilde{B}\left(-\frac{y_1}{x_1}\right)) + \log 2. \\ &\leq h(P_2 P_1^{-2n}) + (n+1)h(P_1) + \log |\tilde{B}|_{L^1} + \log 2 \\ &\leq h(P_2 P_1^{-2n}) + (n+1)h(P_1) + (n-1) \log M_1 \end{aligned}$$

for some constant  $M_1 > 0$ . It follows that

$$h(P_1) \leq \frac{1}{n-1} h(P_2 P_1^{-2n}) + \log M_1$$

as desired.

If  $\Delta = 0$ , then it looks as if we might be stuck. However, notice that  $\Delta = 0$  means that  $x_2 y_1^{2n+1} \tilde{B}(x_1) = y_2 x_1^{2n+1} \tilde{A}(x_1)$ , i.e., that the polynomial  $h(t) := y_2 A(t) - x_2 B(t)$  vanishes at  $t = x_1$ . Since  $x_1 \notin \{0, 1, \infty\}$ , it follows from Lemma 7 that  $x_1$  is a *simple* zero of  $h(t)$ . Therefore  $h'(x_0) \neq 0$ . Let's differentiate the identity  $A(t) + B(t) = C(t)$  and see if we can use this information.

Recalling that  $A(t) = t^{2n+1} \tilde{A}$  and  $B(t) = (1-t)^{2n+1} \tilde{B}$ , we see from the product rule that  $A'(t) = t^{2n} \alpha(t)$  and  $B'(t) = (1-t)^{2n} \beta(t)$  for some polynomials  $\alpha, \beta$  of degree  $n$ . For consistency of notation, we also let  $\gamma(t) = C'(t)$ , so that

$$t^{2n} \alpha(t) + (1-t)^{2n} \beta(t) = \gamma(t).$$

As before, this identity gives us a system of linear equations

$$\begin{aligned} \frac{x_2}{x_1^{2n}} \cdot x_1^{2n} + \frac{y_2}{y_1^{2n}} \cdot y_1^{2n} &= 1 \\ \alpha(x_1) \cdot x_1^{2n} + \beta(x_1) \cdot y_1^{2n} &= \gamma(x_1). \end{aligned}$$

The determinant of this system is

$$\Delta' = \frac{x_2}{x_1^{2n}} \beta(x_1) - \frac{y_2}{y_1^{2n}} \alpha(x_1)$$

and therefore  $\Delta' \neq 0$ , since

$$x_1^{2n} y_1^{2n} \Delta' = x_2 B'(x_1) - y_2 A'(x_1) = h'(x_1) \neq 0.$$

We can now proceed as before, applying Lemma 5. To do this, we need an estimate for the  $L^1$ -norms of  $\alpha, \beta, \gamma$ . Without working out a precise estimate, we note that by Lemma 6 and the relationships

$$A(t) = \pm B(1-t), \quad C(t) = t^n \tilde{A}\left(\frac{1}{t}\right) = \pm t^n \tilde{B}(1-t),$$

it is not hard to see that

$$2 \max\{|\alpha|_{L^1}, |\beta|_{L^1}, |\gamma|_{L^1}\} \leq M_2^n$$

for some constant  $M_2 > 0$ . (Exercise.)

Therefore Lemma 5 yields

$$\begin{aligned} 2nh(P_1) &\leq h(P_2P_1^{-2n}) + h(\alpha(x_1) : \beta(x_1) : \gamma(x_1)) + \log 2 \\ &\leq h(P_2P_1^{-2n}) + nh(P_1) + \log \max\{|\alpha|_{L^1}, |\beta|_{L^1}, |\gamma|_{L^1}\} + \log 2 \\ &\leq h(P_2P_1^{-2n}) + nh(P_1) + n \log M_2. \end{aligned}$$

Therefore

$$\begin{aligned} h(P_1) &\leq \frac{1}{n}h(P_2P_1^{-2n}) + \log M_2 \\ &\leq \frac{1}{n-1}h(P_2P_1^{-2n}) + \log M_2. \end{aligned}$$

□

Using the bounds

$$\frac{1}{2}\hat{h}(P) \leq h(P) \leq \hat{h}(P),$$

Proposition 4 implies the following geometric constraint on  $Z$ :

**Corollary 3.** *If  $u, v \in Z$ , then for all  $n \geq 2$  we have*

$$u \leq \frac{2}{n-1}\|v - 2nu\| + 2M.$$

To conclude the proof that  $Z$  is finite, we now prove the cap principle in the following explicit form:

**Proposition 5.** *If  $u, v$  are elements of  $Z$  with  $\text{angle}(u, v) \leq \frac{1}{10}$  and  $\|u\| > 4M$ , then  $\|v\| \leq 90\|u\|$ .*

*Proof.* Let  $n$  be the greatest integer less than or equal to  $\frac{\|v\|}{2\|u\|}$ , so that  $n \geq 2$  and  $0 \leq \|v\| - 2n\|u\| \leq 2\|u\|$ . Then Corollary 3 and Lemma 4 yield

$$\begin{aligned} \|u\| &\leq \frac{2}{n-1}\|v - 2nu\| + 2M \\ &\leq \frac{2}{n-1}(\|v\| - 2n\|u\| + 2n\|u\| \text{angle}(u, v)) + 2M \\ &\leq \frac{2}{n-1}(2\|u\| + \frac{2n}{10}\|u\|) + 2M \\ &= \frac{4}{n-1}(1 + \frac{n}{10})\|u\| + 2M. \end{aligned}$$

If  $n \geq 45$  then it is easy to verify that

$$\frac{4}{n-1}(1 + \frac{n}{10}) \leq \frac{1}{2},$$

so that  $\|u\| \leq 4M$ , a contradiction.

Therefore  $n \leq 44$ . Since  $\|v\| - 2n\|u\| < 2\|u\|$ , we get the desired inequality

$$\|v\| \leq 90\|u\|.$$

