Number of the solutions of S-unit equation in two variables

Noriko Hirata-Kohno
(joint work with Makoto Kawashima, Anthony Poëls and Yukiko Washio)

Nihon University, Tokyo

June 7, 2024

Summary

1. K : number field of degree $m < \infty$, λ, μ : non-zero elements of K
2. S : finite set of places of K containing all ∞ ones
3. $s = \mathrm{Card}\, S$, U_S : the S-units in K
4. The S-unit equation :

$$\lambda x + \mu y = 1 \text{ in unknowns } x, y \in U_S \tag{1}$$

   Our main result : (1) has solutions at most

$$(3.1 + 68m \log m(1.5)^m)45^s.$$

5. The result due to J. -H. Evertse in 1984 : (1) has solutions at most

$$3 \times 7^{m+2s}.$$

6. Explicit Padé approximation for Binomial function + Loher-Masser bound

Solutions of the S-unit equation

Theorem (Siegel-Mahler)

The S-unit equation

$$\lambda x + \mu y = 1$$

has only finitely many solutions in $x, y \in U_S$.

Remark

1. Finiteness follows by Thue-Siegel-Roth theorem + p-adic version by Mahler (1932), Parry (1950).
2. The 2 variables' case: Linear forms in logarithms : Effective bounds and explicit estimates by Győry-Yu, J. -H. Evertse, Bombieri-Gubler, Bugeaud-Győry. Confer excellent books by Evertse-Győry.
3. More than 3 variables' case, linear forms in logs give no upper bound for height of solutions.

The number of the solutions in 2 variables' case

Theorem (J. -H. Evertse (1984))

$$\lambda x + \mu y = 1$$

has solutions at most

$$3 \times 7^{m+2s}.$$

Remark

1. The bound should depend on s (since $s \geq \frac{m}{2}$ we may omit m).
2. Quantitative Roth + Padé for cubic fct + Counting for bounded height
3. Folklore conjecture (Bombieri, 2000) $\gg \exp\left(s^{1-\varepsilon}\right)$

Remark

Evertse indeed established a better bound (p. 583 of the article)

$$\left(2 + 5 \cdot \left(2e^{24/49}\right)^{m}\right) \cdot 49^{s}. \qquad (2)$$

1st ingredient : Explicit Padé approximation

Define Padé approximants at $z = 0$ (although we perform ours at ∞).

Definition (Padé Approximants of Type I)

For $f_1(z), \ldots, f_m(z) \in K[[z]]$, $0 \leq n_1, \ldots, n_m \in \mathbb{Z}$, $\exists \mathcal{P}_1(z), \ldots, \mathcal{P}_m(z) \in K[z]$ satisfying (i) (ii) (iii). These polynomials $(\mathcal{P}_1(z), \ldots, \mathcal{P}_m(z)) \in K[z]$ are called weight $(n_1, \ldots, n_m)$ Padé approximants of Type I at $z = 0$ $(N = n_1 + \cdots + n_m)$.

(i) $\mathcal{P}_1(z), \ldots, \mathcal{P}_m(z) \not\equiv 0$,

(ii) $\deg \mathcal{P}_i(z) \leq n_i \ (1 \leq i \leq m)$,

(iii) $\mathrm{ord}_{z=0}(\mathcal{P}_1(z) f_1(z) + \cdots + \mathcal{P}_m(z) f_m(z)) \geq N + m - 1.$

1st ingredient : Explicit Padé approximation

Define Padé approximants at $z = 0$ (although we perform ours at ∞).

Definition (Padé Approximants of Type I)

For $f_1(z), \ldots, f_m(z) \in K[[z]]$, $0 \le n_1, \ldots, n_m \in \mathbb{Z}$, $\exists \mathcal{P}_1(z), \ldots, \mathcal{P}_m(z) \in K[z]$ satisfying (i) (ii) (iii). These polynomials $(\mathcal{P}_1(z), \ldots, \mathcal{P}_m(z)) \in K[z]$ are called weight $(n_1, \ldots, n_m)$ Padé approximants of Type I at $z = 0$ ($N = n_1 + \cdots + n_m$).

(i) $\mathcal{P}_1(z), \ldots, \mathcal{P}_m(z) \not\equiv 0$,

(ii) $\deg \mathcal{P}_i(z) \le n_i \ (1 \le i \le m)$,

(iii) $\mathrm{ord}_{z=0}(\mathcal{P}_1(z) f_1(z) + \cdots + \mathcal{P}_m(z) f_m(z)) \ge N + m - 1$.

Definition (Padé Approximants of Type II)

For the $f(z)$ above, the polynomials $(\mathcal{P}_1(z), \ldots, \mathcal{P}_m(z))$ satisfying (iv) (v) (vi) are called weight $(n_1, \ldots, n_m)$ Padé approximants of Type II at $z = 0$.

(iv) $\mathcal{P}_1(z), \ldots, \mathcal{P}_m(z) \not\equiv 0$,

(v) $\deg \mathcal{P}_i(z) \le N - n_i \, (1 \le i \le m)$,

(vi) $\mathrm{ord}_{z=0}(\mathcal{P}_i(z) f_j(z) - \mathcal{P}_j(z) f_i(z)) \ge N + 1 \ (1 \le i < j \le m)$.

Polynomials $\mathcal{P}(z)$ exist by linear algebra, but it is difficult to have in explicit form.

Explicit Padé approximation for Hypergeometric Function

For $k \in \mathbb{Z}_{\geq 0}$, let $(x)_k = x(x+1)(x+2)\cdots(x+k-1), (x)_0 = 1$.
We suppose $a, b, c \in \mathbb{Q}$, $c \notin \mathbb{Z}_{\leq 0}$ throughout the talk.

Definition (Gauss Hypergeometric Function)

$$F = {}_2F_1\left(\begin{matrix} a, \ b \\ c \end{matrix} \middle| \, z\right) = \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{(c)_k} \cdot \frac{z^k}{k!}$$

1. The function converges in $|z| < 1$, $z \in \mathbb{C}$.
   (when $a \in \mathbb{Z}_{\leq 0}$ or $b \in \mathbb{Z}_{\leq 0}$, the function is just a polynomial)
2. The function satisfies a linear differential equation of the shape

$$z(1-z)F'' - ((a+b+1)z - c)F' - abF = 0$$

3. The function $f(z) = \frac{1}{z}\left(1 - \frac{1}{z}\right)^{\omega} = \sum_{k=0}^{\infty} \frac{(-\omega)_k}{k!} \frac{1}{z^{k+1}}$

   is binomial function with exponent ω, a hypergeometric fct in the next slide.

Generalized Hypergeometric Function

Suppose $2 \leq r \in \mathbb{Z}$ and
$\boldsymbol{a} = (a_1, \ldots, a_r) \in (\mathbb{Q} \setminus \mathbb{Z}_{\leq 0})^r, \boldsymbol{b} = (b_1, \ldots, b_{r-1}) \in (\mathbb{Q} \setminus \mathbb{Z}_{\leq 0})^{r-1}$.

Definition (Generalized Hypergeometric Function)

$$ {}_rF_{r-1}\left(\begin{matrix} a_1, a_2, \cdots, a_r \\ b_1, \cdots, b_{r-1} \end{matrix} \middle| z \right) = \sum_{k=0}^{\infty} \frac{(a_1)_k (a_2)_k \cdots (a_r)_k}{(b_1)_k \cdots (b_{r-1})_k} \cdot \frac{z^k}{k!} $$

1. ${}_rF_{r-1}$ converges in $|z| < 1$ and is a G-function for $a_i, b_j \in \mathbb{Q}$ (Siegel).
2. For $x \in \mathbb{Q}$ with $0 \leq x < 1$, Lerch function $(s \in \mathbb{Z}_{\geq 1})$ is defined by

$$ \Phi_s(x, z) = \sum_{k=0}^{\infty} \frac{z^{k+1}}{(k + x + 1)^s} \quad \text{(polylogarithm when } x = 0) $$

$$ = \frac{z}{(x+1)^s} \cdot {}_{s+1}F_s\left(\begin{matrix} 1, x+1, \cdots, x+1 \\ x+2, \cdots, x+2 \end{matrix} \middle| z \right) $$

Linear Independence of Hypergeometric Values

Let K be a number field of any degree over $\mathbb{Q}$. Let $2 \le r \in \mathbb{Z}$.

Theorem (Sinnou David, Makoto Kawashima & NHK, 2024)

Let $\boldsymbol{a} = (a_1, \dots, a_r) \in (\mathbb{Q} \setminus \mathbb{Z}_{\le 0})^r, \boldsymbol{b} = (b_1, \dots, b_{r-1}) \in (\mathbb{Q} \setminus \mathbb{Z}_{\le 0})^{r-1}$, *where* $a_k \notin \mathbb{Z}_{\ge 1}$ *and* $a_k + 1 - b_j \notin \mathbb{Z}_{\ge 1}$ $(1 \le k \le r, 1 \le j \le r-1)$.
Let $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_m) \in (K \setminus \{0\})^m$ *with* α_i *pairwise distinct.*
For v_0 *a place of* K, $B \in K \setminus 0$, *define* $V_{v_0} = V_{v_0}(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{\alpha}, B)$ $\in \mathbb{R}$ *(precised later).*

Assume $V_{v_0} > 0$ $\Longrightarrow$ *Then the* $rm+1$ *values* $(1 \le s \le r-1,\ 1 \le i \le m)$

$${}_rF_{r-1}\left(\begin{matrix} a_1, \dots, a_r \\ b_1 \dots, b_{r-1} \end{matrix} \,\middle|\, \frac{\alpha_i}{B}\right), \dots,$$

$${}_rF_{r-1}\left(\begin{matrix} a_1+1, \cdots, a_r+1 \\ b_1+1, \cdots, b_{r-s}+1, b_{r-s+1}, \dots, b_{r-1} \end{matrix} \,\middle|\, \frac{\alpha_i}{B}\right)$$

and 1 *are linearly independent* $/K$.

These r functions are all linearly independent $/\mathbb{C}(z)$ by Nesterenko (1995).
The theorem is valid (not only in G-function, but) in arithmetic Gevrey series.

$V_{v_0} = V_{v_0}(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{\alpha}, B)$, a special definition precised

Being v_0 a place of K and $\mu(x) = \prod_{\substack{q:\text{prime} \\ q | \mathrm{den}(x)}} q^{q/(q-1)}$, for $B \in K \setminus 0$, define

$$V_{v_0} = V_{v_0}(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{\alpha}, B) = \log |B|_{v_0} + rm \log \|(\boldsymbol{\alpha}, B)\|_{v_0} - rm\mathrm{h}(\boldsymbol{\alpha}, B) - (rm+1) \log \|\boldsymbol{\alpha}\|_{v_0}$$
$$- \left(rm \log(2) + r \left(\log(rm+1) + rm \log \left(\frac{rm+1}{rm} \right) \right) \right)$$
$$- \sum_{j=1}^{r} \left(\log \mu(a_j) + 2 \log \mu(b_j) + \frac{\mathrm{den}(a_j)\mathrm{den}(b_j)}{\varphi(\mathrm{den}(a_j))\varphi(\mathrm{den}(b_j))} \right) .$$

Assume $V_{v_0} > 0 \implies$ Then the $rm+1$ numbers ($1 \le s \le r-1$, $1 \le i \le m$)

$${}_rF_{r-1}\left(\begin{matrix} a_1, \ldots, a_r \\ b_1 \ldots, b_{r-1} \end{matrix} \,\middle|\, \frac{\alpha_i}{B} \right), \cdots,$$

$${}_rF_{r-1}\left(\begin{matrix} a_1+1, \cdots, a_r+1 \\ b_1+1, \cdots, b_{r-s}+1, b_{r-s+1}, \ldots, b_{r-1} \end{matrix} \,\middle|\, \frac{\alpha_i}{B} \right)$$

and 1 are linearly independent $/K$.

- V_{v_0} depends on K and $V_{v_0} > 0$ means B large (such B $\exists$infinitely many).
- Note that V_{v_0} depends on K, and $V_{v_0} > 0$ means B large (α_i / B small).
- By observation on the rationality of values by F. Beukers, B must be large.

Application to a binomial function

Put $\nu_n = 3^{n+\lceil n/2 \rceil}$ and $\nu(x) = \mathrm{den}(x) \cdot \prod_{\substack{q:\mathrm{prime} \\ q|\mathrm{den}(x)}} q^{1/(q-1)}$ for $x \in \mathbb{Q}$.

For $n \in \mathbb{Z}_{\geq 0}$, $G_n := \mathrm{GCD}\left(\nu_n \binom{n+k-1}{k}\binom{n-4/3}{n-k},\ \nu_n \binom{n+k'}{k'}\binom{n+1/3}{n-1-k'}\right)_{\substack{0\leq k\leq n \\ 0\leq k'\leq n-1}}$.

Let $\rho_1 \leq \rho_2$ be the moduli of the roots $2\beta - 1 \pm 2\sqrt{\beta^2 - \beta}$ of the polynomial $P(X) = X^2 - 2(2\beta - 1)X + 1$.

Theorem (Anthony Poëls and M. Kawashima, 2023)

Let $\beta \in \mathbb{Q}$ with $|\beta| > 1$. Put

$$\Delta = 3^{3/2} \cdot \mathrm{den}(\beta) \cdot \limsup_{n\to\infty} G_n^{-1/n},$$

$$Q = \rho_2 \cdot \Delta, \qquad E = \rho_2 \cdot \Delta^{-1}.$$

Assume $E > 1$. Then the irrationality exponent μ of the irrational value $(1 - 1/\beta)^{\frac{1}{3}} \notin \mathbb{Q}$ satisfies $\mu((1 - 1/\beta)^{\frac{1}{3}}) \leq 1 + \dfrac{\log(Q)}{\log(E)}$.

Due to the effective Poincaré-Perron Theorem (slide page 17).

1st ingredient : explicit form of Padé polynomials

Binomial Cubic function is one kind of generalized hypergeometric fcts :

$$f(z) = \frac{1}{z}\left(1 - \frac{1}{z}\right)^{1/3} = \sum_{k=0}^{\infty} \frac{(-1/3)_k}{k!} \frac{1}{z^{k+1}} \in (1/z) \cdot \mathbb{Q}[[1/z]].$$

Lemma (Poëls-Kawashima (2023))

$$\text{Let} \quad Q_n(z) = \sum_{k=0}^{n} (-1)^{n-k} \binom{n+k-1}{k} \binom{n-4/3}{n-k} z^k,$$

$$P_n(z) = \sum_{k=0}^{n-1} (-1)^{n-1-k} \binom{n+k}{k} \binom{n+1/3}{n-1-k} z^k.$$

Then the pair of polynomials (Q_n, P_n) forms Padé approximation of Type II, and $R_n(z) = Q_n(z)f(z) - P_n(z)$ is a Padé approximation of $f(z)$. Moreover,

$$R_n(z) = \frac{(4/3)_n(-1/3)_n}{(2n)!z^{n+1}} \cdot {}_2F_1\left(\begin{matrix} n+1, n-1/3 \\ 2n+1 \end{matrix} \,\middle|\, \frac{1}{z}\right) \in (1/z^{n+1}). \qquad (3)$$

Simpler form of (Q_n, P_n)

Proposition

Define the polynomials $A_n(1-z) = z^{n-1}Q_n(1/z)$ *and* $B_n(1-z) = z^n P_n(1/z)$. *Then, we have*

$$A_n(1-z) = \sum_{\ell=0}^{n-1} \binom{n+1/3}{\ell}\binom{n-4/3}{n-1-\ell}(1-z)^{\ell},$$

$$B_n(1-z) = \sum_{\ell=0}^{n} \binom{n-4/3}{\ell}\binom{n+1/3}{n-\ell}(1-z)^{\ell}.$$

We prove all the statements by the Chu-Vandermonde identity : a simple argument related to binomials.

Better bounds for $A_n(z), B_n(z)$ by simpler forms

The *length* $L(P)$ of P is the sum of the moduli of its coefficients. Define

$$|p|_v = p^{-\frac{[K_v:\mathbb{Q}_p]}{m}} \text{ if } v|p \ , \ |x|_v = |\sigma_v(x)|^{\frac{[K_v:\mathbb{R}]}{m}} \text{ if } v|\infty, \quad \mathrm{H}_v(\beta) = \max\{1, |\beta|_v\}\,,$$

$$s(v) := \begin{cases} 1/m & \text{if } v \ \text{ archimedean real}\,, \\ 2/m & \text{if } v \ \text{ archimedean complex}\,, \\ 0 & \text{if } v \ \text{ non-archimedean}\,. \end{cases}$$

Lemma

Let $z \in \mathbb{C}$. Then we have

$$\max\left(|A_n(z)|, |B_n(z)|\right) \le 4^n \max(1, |z|)^n \quad (n \ge 1),$$

$$L(A_n(z)) + L(B_n(z)) \le 4^n/2 \quad (n \ge 2).$$

Lemma

Let v be an archimedean place and $\alpha \in K$. For $n \ge 1$ we have

$$\max\left(|A_n(\alpha^3)|_v, |B_n(\alpha^3)|_v\right) \le 4^{ns(v)} \mathrm{H}_v(\alpha)^{3n}.$$

2nd ingredient : the Loher-Masser bound

We use the next result due to T. Loher & D. Masser, a uniform bound as below.

Lemma (Loher and Masser, 2004)

Let $\theta \neq 0$ be an algebraic number, not necessarily in K. Let $c \geq 1$ be a constant. Let $m = [K : \mathbb{Q}] \geq 2$. Then the number of $z \in K \setminus \{0\}$ with $H(\theta z) \leq c$ is at most

$$68\, m \log m \cdot c^{2m}.$$

Theorem (Poëls, Kawashima, Washio & Hirata-Kohno, 2023 (IJNT))

The number of the solutions $(x, y) \in U_S^2$ is at most

$$(3.1 + 68\, m \log m \cdot (1.5)^m) \cdot 45^s$$

which is smaller than Evertse' precise bound (2) *for $m \geq 6$, $s \geq 1$. Precisely, the number of the solutions of the unit equation* (1) *is at most, for $m \geq 2$ and $s \geq 1$,*

$$\min\Big\{\Big(2.81864 \cdot (46.8312)^s + \min\Big(5 \cdot (3.22803)^m \cdot 47^s, 68\, m \log m \cdot (1.37597)^m \cdot 47^s\Big),$$

$$\Big(3.06759 \cdot (44.9866)^s + \min\Big(5 \cdot (3.36406)^m \cdot 45^s, 68\, m \log m \cdot (1.41436)^m \cdot 45^s\Big)\Big\}.$$

More application of the explicit Padé approximation

Let β be algebraic of degree d over $\mathbb{Q}$ with $|\beta| > 1$ and $K = \mathbb{Q}(\beta)$. Define $\delta = [K : \mathbb{Q}]/[K_\infty : \mathbb{R}]$ for $K_\infty = \mathbb{R}$ if $K \subset \mathbb{R}$ and $K_\infty = \mathbb{C}$ otherwise. For each conjugate map σ_k $(1 \le k \le d)/\mathbb{Q}$, suppose $P(X) = X^2 - 2(2\sigma_k(\beta) - 1)X + 1$ has the roots whose moduli are distinct with the notation $\rho_1(\sigma_k(\beta)) < \rho_2(\sigma_k(\beta))$. Let

$\Delta = \mathrm{den}(\beta) \cdot \exp\left(\mathrm{den}(\mathrm{x})/\varphi(\mathrm{den}(\mathrm{x}))\right) \cdot \nu(x)$, ($\varphi$ is Euler's fct, $\nu(x)$ is in slide 11),

$Q = \Delta \cdot \prod_{1\le k \le d}\Big(\rho_2(\sigma_k(\beta))\Big)$, $\quad E = \rho_2(\beta)/\rho_1(\beta) = \big(\rho_2(\beta)\big)^2$.

Theorem (R. Muroi, Y. Washio and NHK, 2024)

Let $x \in \mathbb{Q} \cap [0,1)$. Put $\lambda = \dfrac{1}{\delta} - \dfrac{\log Q}{d \log E}$. Consider the shifted logarithmic function $f(z) = (1+x)\Phi_1(x, 1/z) = (1+x)\sum_{k=0}^{\infty} 1/\Big((k+x+1)\cdot z^{k+1}\Big)$.

Whenever $\lambda > 0$, then $f(\beta) = (1+x)\Phi_1(x, 1/\beta) \notin K$, and its effective K-approximation measure satisfies

$$\mu(f(\beta), K) \le \frac{1}{\lambda}.$$

By effective Poincaré-Perron (slide p. 17). This refines previous measures.

What is effective Poincaré-Perron Theorem?

Definition

Let $s_j(n),\ 1 \le j \le \ell$, be functions from $n \in \mathbb{Z}_{\ge 0}$ to $\mathbb{C}$ with $s_\ell(n) \ne 0$ for all n. Let ℓ-th order linear difference equation with unknown functions $x(n)$ on $n \in \mathbb{Z}_{\ge 0}$:

$$x(n+\ell) + s_1(n)x(n+\ell-1) + \cdots + s_\ell(n)x(n) = 0 \tag{5}$$

where the limit $t_j := \lim_{n\to\infty} s_j(n)$ exists in $\mathbb{C}$ for each $1 \le j \le \ell$. Let us write the characteristic equation

$$\lambda^\ell + t_1\lambda^{\ell-1} + \cdots + t_\ell = 0, \tag{6}$$

and denote by λ_j the roots of the equation (6).

The next theorem is useful to have asymptotic behavior of the function $x(n)$.

Theorem (Perron's 2nd theorem, effectively proven by M. Pituk, 2002)

The equation (5) *has either* ℓ *linearly independent solutions* $x_1(n), \ldots, x_\ell(n)$, *or* $x(n) = 0$ *for all large* n, *and in the former case, for each* $1 \le j \le \ell$, *we have:*

$$\limsup_{n\to\infty} \frac{1}{n} \log |x_j(n)| = \log |\lambda_j|.$$

Thank you very much for your cordial invitation.