Note on a theorem of Professor X.

Levent Alpöge

Abstract.

Between his arrival in Frankfurt in 1922 and and his proof of his famous finiteness theorem for integral points in 1929, Siegel had no publications. He did, however, write a letter to Mordell in 1926 in which he explained a proof of the finiteness of integral points on hyperelliptic curves. Recognizing the importance of this argument (and Siegel's views on publication), Mordell sent the relevant extract to be published under the pseudonym "X".

The purpose of this note is to explain how to optimize Siegel's 1926 technique to obtain the following bound. Let K be a number field, S a finite set of places of K, and $f \in \mathfrak{o}_{K,S}[t]$ monic of degree $d \geq 5$ with discriminant $\Delta_f \in \mathfrak{o}_{K,S}^\times$. Then:

$$\#|\{(x,y) : x,y \in \mathfrak{o}_{K,S}, y^2 = f(x)\}| \leq 2^{\operatorname{rank} \operatorname{Jac}(C_f)(K)} \cdot O(1)^{d^3 \cdot ([K:\mathbb{Q}] + \#|S|)}.$$

This improves bounds of Evertse-Silverman and Bombieri-Gubler from 1986 and 2006, respectively.

The main point underlying our improvement is that, informally speaking, we insist on "executing the descents in the presence of only one root (and not three) until the last possible moment".

1 Introduction.

The technique introduced in Siegel's 1926 letter to Mordell [8] to prove the finiteness of integral points on hyperelliptic curves[1] can be summarized as: 2-descent on the curve, then 3-descent on $\mathbb{G}_m$. From this summary the intuitive guess for the bound that the argument "should" produce is of course something of shape $2^{\operatorname{rank}} \cdot 3^{\# \text{ of prime factors of the discriminant}}$. However a consultation of the literature yields worse estimates.[2]

In this note we rectify the situation.

---

[1] By this common and incorrect abbreviation we really mean integral solutions of $y^2 = f(x)$. Siegel's 1926 proof does not control integral solutions of e.g. $y^3 + x \cdot y = x^4$, or, said another way, integral points with respect to an effective divisor containing no nonzero divisor symmetric under the hyperelliptic involution. Baker effectivized Siegel's 1926 argument and thus gave an effective finiteness proof in the $y^2 = f(x)$ cases, but he did not effectivize Siegel's 1929 finiteness theorem in the case of hyperelliptic curves.

[2] To our knowledge the best bounds in the literature are the 2006 estimate of Bombieri-Gubler [4, Theorem 5.3.5] and the 1986 estimate of Evertse-Silverman [6]. In both bounds the "expected" 2^{rank} factor is replaced by a power of the size of the 2-part of the class group of an extension gotten by adjoining three Weierstrass points.

Theorem 1.1. *Let $K/\mathbb{Q}$ be a number field. Let S be a finite set of places of K. Let $f \in \mathfrak{o}_{K,S}[t]$ be monic of degree $d \geq 5$ with discriminant $\Delta_f \in \mathfrak{o}_{K,S}^\times$. Let $\mathcal{C}_f^{\mathrm{aff.}} : y^2 = f(x)$ be the Weierstrass model of the hyperelliptic curve C_f corresponding to f. Then:*

$$\#|\mathcal{C}_f^{\mathrm{aff.}}(\mathfrak{o}_{K,S})| \leq 2^{\mathrm{rank\,Jac}(C_f)(K)} \cdot O(1)^{d^3 \cdot ([K:\mathbb{Q}]+\#|S|)}.$$

Also, for all K-irreducible $g_i \in \mathfrak{o}_{K,S}[t]$ with $\sum_i \deg g_i \geq 3$ and $\prod_i g_i \big| f$,

$$\#|\mathcal{C}_f^{\mathrm{aff.}}(\mathfrak{o}_{K,S})| \leq \left(\prod_i \#|\mathrm{Cl}(\mathfrak{o}_{K_{g_i},S})[2]| \right) \cdot O(1)^{d^3 \cdot ([K:\mathbb{Q}]+\#|S|)},$$

where $K_{g_i} := K[t]/(g_i)$ and $\mathfrak{o}_{K_{g_i},S} := \mathfrak{o}_{K_{g_i}} \otimes_{\mathfrak{o}_K} \mathfrak{o}_{K,S}$.

We include the second statement in case e.g. the product of the linear and quadratic irreducible factors of f is of degree at least 3.

In the elliptic curve case an even stronger bound is available because one can execute the whole descent over K (following Mordell) [1][3]. In the superelliptic ($y^m = f(x), m > 2$) case as usual one does not need to execute a 3-descent on $\mathbb{G}_m$ so the bound also improves.

Let us explain how the argument goes. First, the 2-descent. À la Fermat, given $y^2 = f(x)$, adjoin a root ρ and conclude that $x - \rho$ is almost a square, say $x - \rho = \alpha \cdot \beta^2$ with $\alpha, \beta \in K[\rho] := K[t]/(f)$.

Now adjoin three roots. Let L/K be an extension containing three roots, say ρ_1, ρ_2, ρ_3. Then, for all K-maps $\sigma : K[\rho] \to L$, we find $x - \sigma(\rho) = \sigma(\alpha) \cdot \sigma(\beta)^2$. Thus (obvious notation) $\rho_j - \rho_i = \alpha_i \cdot \beta_i^2 - \alpha_j \cdot \beta_j^2$.

Now pass to $L' := L(\sqrt{\alpha_1}, \sqrt{\alpha_2}, \sqrt{\alpha_3})$. We obtain the six elements $\gamma_{ij,\pm} := \beta_i \sqrt{\alpha_i} \pm \beta_j \sqrt{\alpha_j}$, all divisors of $\rho_j - \rho_i$. By hypothesis they are therefore all S'-units, with S' the set of places of L' above a place in S.

Now for the 3-descent. We may then write $\gamma_{ij,\pm} = \delta_{ij,\pm} \cdot \varepsilon_{ij,\pm}^3$ with $\delta_{ij,\pm}$ representatives for $\mathfrak{o}_{L',S'}^\times/3$. The relation $\gamma_{12,+} - \gamma_{23,+} = \gamma_{13,-}$ becomes the cubic Thue equation $\frac{\delta_{12,+}}{\delta_{13,-}} \cdot \left(\frac{\varepsilon_{12,+}}{\varepsilon_{13,-}} \right)^3 - \frac{\delta_{23,+}}{\delta_{13,-}} \cdot \left(\frac{\varepsilon_{23,+}}{\varepsilon_{13,-}} \right)^3 = 1$.

We conclude with a bound of Evertse.

As a last remark: note that $2^{\mathrm{rank\,Jac}(C_f)(K)}$ is $\ll 1$ on average (ordering as usual by height) [3, 7]. This statement is why we tried for the stronger bound [1] in the elliptic curve case, after which we questioned why a similar bound did not exist in the hyperelliptic curve case.

2 Acknowledgments.

---

[3]See also Chapter 2 of [2], which is based on the same work.

3 Proof of Theorem 1.1.

Proof of Theorem 1.1. Let $g \in \mathfrak{o}_{K,S}[t]$ with $g|f$ be a monic divisor of f with $\deg g \geq 3$ (for the first part of the theorem we will take $g = f$). Write $f =: gh$ with $h \in \mathfrak{o}_{K,S}[t]$. Let L/K be an extension of minimal degree containing at least three roots of g.

Factorize $g =: \prod_i g_i$ into K-irreducible factors $g_i \in \mathfrak{o}_{K,S}[t]$. Note that $(g_i, g_j) = (1)$ for each $i \neq j$ because $\Delta_f \in \mathfrak{o}_{K,S}^{\times}$. Thus $K_g \simeq \bigoplus_i K_{g_i}$, with $K_{g_i} := K[t]/(g_i)$ and $K_g := K[t]/(g)$. We will repeatedly write $\mathfrak{o}_{K_{g_i},S}$ etc. for the evident localizations, e.g. $\mathfrak{o}_{K_{g_i},S} := \mathfrak{o}_{K_{g_i}} \otimes_{\mathfrak{o}_K} \mathfrak{o}_{K,S}$. Write $\rho^{(i)}$ for the image of t in $K_{g_i} = K[t]/(g_i)$.

By the Chebotarev density theorem (and its explicit error term) applied to the Hilbert class field of K_{g_i}, each ideal class of K_{g_i} contains a prime of norm $\ll |\Delta_{K_{g_i}}|^{O(1)}$. (We will only use existence.) Let then $P^{(i)}$ be a minimal set of prime representatives of $\mathrm{Cl}(\mathfrak{o}_{K_{g_i},S})[2]$ in K_{g_i} of norm $\ll |\Delta_{K_{g_i}}|^{O(1)}$.

Now let us begin the argument. For $(x, y) \in \mathcal{C}_f^{\mathrm{aff}} \cdot (\mathfrak{o}_{K,S})$, $(y)^2 = (x - \rho^{(i)}) \cdot \left(\frac{g_i(x)}{x - \rho^{(i)}} \right) \cdot \prod_{j \neq i} (g_j(x)) \cdot (h(x))$ as ideals in $\mathfrak{o}_{K_{g_i},S}$. Without loss of generality $f(x) \neq 0$. Since $\Delta_f \in \mathfrak{o}_{K,S}^{\times}$, it follows that there is an ideal $\mathfrak{a}_i \subseteq \mathfrak{o}_{K_{g_i},S}$ with $(x - \rho^{(i)}) = \mathfrak{a}_i^2$. Thus there is a $\mathfrak{p}_i \in P^{(i)}$ with $\mathfrak{p}_i \equiv \mathfrak{a}_i$ modulo principal ideals of K_{g_i}. Hence $\mathfrak{a}_i$ is principal in $\mathfrak{o}_{K_{g_i},S \cup \{\mathfrak{p}_i\}}$ (obvious meaning).

Let $\alpha_i \in \mathfrak{o}_{K_{g_i},S \cup \{\mathfrak{p}_i\}}$ be such that $\mathfrak{a}_i = (\alpha_i)$ as ideals of $\mathfrak{o}_{K_{g_i},S \cup \{\mathfrak{p}_i\}}$. Thus $x - \rho^{(i)} = \alpha_i^2 \cdot (\in \mathfrak{o}_{K_{g_i},S \cup \{\mathfrak{p}_i\}}^{\times})$, where $(\in \mathfrak{o}_{K_{g_i},S \cup \{\mathfrak{p}_i\}}^{\times})$ denotes an element of the $(S \cup \{\mathfrak{p}_i\})$-units of K_{g_i}.

Let $U^{(i)}$ be a minimal set of representatives of $\mathfrak{o}_{K_{g_i},S \cup \{\mathfrak{p}_i\}}^{\times}/2$ (that is, modulo squares). It follows that there are $\gamma_i \in U^{(i)}$ and $u^{(i)} \in \mathfrak{o}_{K_{g_i},S \cup \{\mathfrak{p}_i\}}^{\times}$ such that $x - \rho^{(i)} = \gamma_i \cdot (\alpha_i \cdot u^{(i)})^2$. Let $\eta_i := \alpha_i \cdot u^{(i)}$. Thus $x - \rho^{(i)} = \gamma_i \cdot \eta_i^2$.

Therefore we find that, for each K-embedding $\sigma : K_{g_i} \hookrightarrow L$, we have $x - \sigma(\rho^{(i)}) = \sigma(\gamma_i) \cdot \sigma(\eta_i)^2$.

By definition of L, for each $1 \leq k \leq 3$ there is an i_k and a K-embedding $\tau_k : K_{g_{i_k}} \hookrightarrow L$ such that the corresponding roots $\tau_k(\rho^{(i_k)})$ are pairwise distinct. Let then:

$$\kappa_k := \tau_k(\rho^{(i_k)}),$$
$$\lambda_k := \tau_k(\gamma_{i_k}),$$
$$\mu_k := \tau_k(\eta_{i_k}).$$

Hence $x - \kappa_k = \lambda_k \cdot \mu_k^2$ as elements of L. Hence for $k \neq \ell$ we find that $\kappa_\ell - \kappa_k = \lambda_k \cdot \mu_k^2 - \lambda_\ell \cdot \mu_\ell^2$.

We are done with the 2-descent on C_f. Now it is time for the 3-descent on $\mathbb{G}_m$. Let $L_{k\ell} := L(\sqrt{\lambda_k}, \sqrt{\lambda_\ell})$. Thus $(1) = (\kappa_\ell - \kappa_k) = (\sqrt{\lambda_k}\mu_k - \sqrt{\lambda_\ell}\mu_\ell) \cdot (\sqrt{\lambda_k}\mu_k + \sqrt{\lambda_\ell}\mu_\ell)$ as ideals of $\mathfrak{o}_{L_{k\ell},S \cup \{\mathfrak{p}_k,\mathfrak{p}_\ell\}}$ (obvious meaning), the first equality following from $\Delta_f \in \mathfrak{o}_{K,S}^{\times}$.

Let $V_{k\ell}$ be a minimal set of representatives of $\mathfrak{o}_{L_{k\ell},S\cup\{\mathfrak{p}_k,\mathfrak{p}_\ell\}}^{\times}/3$ (that is, modulo cubes). It follows that there are $v_{k\ell,\pm}\in V_{k\ell},\zeta_{k\ell,\pm}\in\mathfrak{o}_{L_{k\ell},S\cup\{\mathfrak{p}_k,\mathfrak{p}_\ell\}}^{\times}$ for which $\sqrt{\lambda_k}\mu_k\pm\sqrt{\lambda_\ell}\mu_\ell=v_{k\ell,\pm}\cdot\zeta_{k\ell,\pm}^3$. Hence:

$$0=(\sqrt{\lambda_1}\mu_1\pm\sqrt{\lambda_2}\mu_2)\mp(\sqrt{\lambda_2}\mu_2\pm\sqrt{\lambda_3}\mu_3)-(\sqrt{\lambda_1}\mu_1-\sqrt{\lambda_3}\mu_3)$$
$$=v_{12,\pm}\cdot\zeta_{12,\pm}^3\mp v_{23,\pm}\cdot\zeta_{23,\pm}^3-v_{13,-}\cdot\zeta_{13,-}^3.$$

We may of course rearrange this as $\frac{v_{12,\pm}}{v_{13,-}}\cdot\left(\frac{\zeta_{12,\pm}}{\zeta_{13,-}}\right)^3\mp\frac{v_{23,\pm}}{v_{13,-}}\cdot\left(\frac{\zeta_{23,\pm}}{\zeta_{13,-}}\right)^3=1$.

Now for Evertse's bound. Let $F_\pm(X,Y):=\frac{v_{12,\pm}}{v_{13,-}}\cdot X^3\mp\frac{v_{23,\pm}}{v_{13,-}}\cdot Y^3\in\mathfrak{o}_{L_{123}}[X,Y]$, with $L_{123}:=L(\sqrt{\lambda_1},\sqrt{\lambda_2},\sqrt{\lambda_3})$. Then $F_\pm$ is a cubic form with nonzero discriminant. Moreover if $F_\pm(X,Y)=F_\pm(tX,tY)=1$ then $t^3=1$. Therefore by [5] the number of solutions of $F_\pm(X,Y)=1$ with $(X,Y)\in\mathfrak{o}_{L_{123},S\cup\{\mathfrak{p}_1,\mathfrak{p}_2,\mathfrak{p}_3\}}$ is $\ll O(1)^{[L:\mathbb{Q}]+[L:K]\cdot\#|S|}$.

We are done with the argument. Now let us examine the tally. Let us first show that our original point (x,y) can be recovered from the pairs $\left(\frac{\zeta_{12,\pm}}{\zeta_{13,-}},\frac{\zeta_{23,\pm}}{\zeta_{13,-}}\right)$ up to $O(1)$ many possibilities. To see this first multiply the two X-coordinates together to form $\frac{\zeta_{12,+}\cdot\zeta_{12,-}}{\zeta_{13,-}^2}$. Note that:

$$v_{12,+}\cdot v_{12,-}\cdot\zeta_{12,+}^3\cdot\zeta_{12,-}^3=(\sqrt{\lambda_1}\mu_1+\sqrt{\lambda_2}\mu_2)\cdot(\sqrt{\lambda_1}\mu_1-\sqrt{\lambda_2}\mu_2)$$
$$=\lambda_1\mu_1^2-\lambda_2\mu_2^2$$
$$=\kappa_2-\kappa_1.$$

Thus it follows that $\zeta_{12,+}^3\cdot\zeta_{12,-}^3=\frac{\kappa_2-\kappa_1}{v_{12,+}\cdot v_{12,-}}$. Hence the cube of the product of the two X-coordinates is $\zeta_{13,-}^{-6}\cdot\left(\frac{\kappa_2-\kappa_1}{v_{12,+}\cdot v_{12,-}}\right)$. We note that the term in parentheses is fixed (in terms of our choices up til now). Thus we may recover $\zeta_{13,-}^6$, and hence $\zeta_{13,-}$ up to at most six choices. Having done so we return to the X-coordinates of both solutions and recover $\zeta_{12,+}$ and $\zeta_{12,-}$. Then the equality

$$2\sqrt{\lambda_1}\mu_1=(\sqrt{\lambda_1}\mu_1+\sqrt{\lambda_2}\mu_2)+(\sqrt{\lambda_1}\mu_1-\sqrt{\lambda_2}\mu_2)$$
$$=(v_{12,+}\cdot\zeta_{12,+}^3)+(v_{12,-}\cdot\zeta_{12,-}^3)$$

implies that we can recover $2\sqrt{\lambda_1}\mu_1$. Squaring this we find that we can recover $4\lambda_1\mu_1^2=4(x-\kappa_1)$. Since κ_1 is fixed we can recover x, and then there are at most two choices for y given x, so we can indeed recover the point up to $O(1)$ many choices.

So we see that a point $(x,y)\in\mathcal{C}_f^{\mathrm{aff}\cdot}(\mathfrak{o}_{K,S})$ is determined up to $O(d^3)$ (arising from the choice of three roots in L) many choices by the data

$$(\mathfrak{p}_1,\mathfrak{p}_2,\mathfrak{p}_3,\gamma_1,\gamma_2,\gamma_3,v_{12,+},v_{12,-},v_{23,+},v_{23,-},v_{31,-}).$$

The number of choices for each $\mathfrak{p}_i$ is $\#|P^{(i)}|=\#|\mathrm{Cl}(\mathfrak{o}_{K_{g_i},S})[2]|$. The number of choices for each $\gamma_i\in U^{(i)}$ is $\#|U^{(i)}|\ll 2^{\deg g_i\cdot(\#|S|+[K:\mathbb{Q}])}$ by Dirichlet. Similarly

4

the number of choices for each $v_{k\ell,\pm} \in V_{k\ell}$ is $\#|V_{k\ell}| \ll 3^{4\cdot[L:\mathbb{Q}]+4\cdot[L:K]\cdot\#|S|}$. Of course $[L:K] \le d^3$.

Therefore the total number of tuples $(\mathfrak{p}_1, \mathfrak{p}_2, \mathfrak{p}_3, \gamma_1, \gamma_2, \gamma_3, v_{12,+}, v_{12,-}, v_{23,+}, v_{23,-}, v_{31,-})$ is

$$\ll \left(\prod_i \#|\mathrm{Cl}(\mathfrak{o}_{K_{g_i},S})[2]|\right) \cdot O(1)^{d^3\cdot([K:\mathbb{Q}]+\#|S|)}.$$

We are done with the second part of the theorem.

So take $g = f$, and let us count the ideal classes of $\bigoplus_i \mathfrak{o}_{K_{g_i},S} \simeq \mathfrak{o}_{K_f,S}$ that could possibly arise in the 2-descent step (we have bounded said count by $\le \prod_i \#|\mathrm{Cl}(\mathfrak{o}_{K_{g_i},S})[2]|$ and we claim it is also $\ll 2^{\mathrm{rank\,Jac\,}(C_f)(K)} \cdot O(1)^d$).

Let $J_f := \mathrm{Jac}\,C_f$. Let $W \subseteq C_f(\overline{\mathbb{Q}})$ be the set of Weierstrass points of C_f. Let $\infty \in W$ be a point at infinity. We embed $C_f \hookrightarrow J_f = \mathrm{Pic}^0(C_f)$ via $P \mapsto P - \infty$. Thus as $\mathrm{Gal}(\overline{\mathbb{Q}}/K)$-modules $J_f[2] \simeq \mathbb{F}_2[W - \{\infty\}]/\mathbb{F}_2 \cdot \left(\sum_{P\in W-\{\infty\}} P\right)$. In other words, $\mathrm{Ind}_K^{K_f}\mathbb{F}_2 \simeq \mathbb{F}_2 \oplus J_f[2]$, where we have written $\mathrm{Ind}_K^{K_f}(\bullet) := \bigoplus_i \mathrm{Ind}_K^{K_{g_i}}(\bullet)$. Thus $H^1(K_f, \mathbb{F}_2) \simeq H^1(K, \mathrm{Ind}_K^{K_f}\mathbb{F}_2) \simeq H^1(K, \mathbb{F}_2) \oplus H^1(K, J_f[2])$, where the first isomorphism follows by Shapiro's lemma (and $H^1(K_f, \bullet) := \bigoplus_i H^1(K_{g_i}, \bullet)$). By Kummer it follows that $H^1(K, J_f[2]) \simeq (K_f^\times/2)_{\mathrm{Nm}=\square}$.

Thus by taking invariants of $0 \to J_f[2] \to J_f \to J_f \to 0$ we obtain $J_f(K)/2 \hookrightarrow H^1(K, J_f[2]) \simeq (K_f^\times/2)_{\mathrm{Nm}=\square}$. Write $G \subseteq H^1(K, J_f[2])$ for the image of this map. Note that the restriction of this map to $C_f(K) - W$ is simply $(x, y) \mapsto x - \rho$, so similarly write $G' \subseteq G$ for the image of $\mathcal{C}_f^{\mathrm{aff.}}(\mathfrak{o}_{K,S}) - W$.

We have already seen that each $g \in G' \subseteq G \hookrightarrow (K_f^\times/2)_{\mathrm{Nm}=\square}$ gives rise to a class represented by an $\alpha \in K_f^\times$ for which $v_\mathfrak{p}(\alpha)$ is even for all primes $\mathfrak{p} \subseteq \mathfrak{o}_{K_f,S}$ (because $\Delta_f \in \mathfrak{o}_{K,S}^\times$). From such a class we may produce an element of $\mathrm{Cl}(\mathfrak{o}_{K_f,S})[2]$ via $\alpha \mapsto \mathfrak{a}$ such that $(\alpha) = \mathfrak{a}^2$ as ideals of $\mathfrak{o}_{K_f,S}$.

The corresponding map $\mathcal{C}_f^{\mathrm{aff.}}(\mathfrak{o}_{K,S}) - W \to \mathrm{Cl}(\mathfrak{o}_{K_f,S})[2]$ is precisely (up to our choice of representatives) our map $(x, y) \mapsto (\mathfrak{p}_i)_i$. It therefore suffices to show that $\#|G'| \ll 2^{\mathrm{rank\,}J_f(K)} \cdot O(1)^d$. However $\#|G| = \#|(J_f[2])(K)| \cdot 2^{\mathrm{rank\,}J_f(K)}$ and $G' \subseteq G$. $\square$

References.

[1] Levent Alpöge and Wei Ho, *The second moment of the number of integral points on elliptic curves is bounded*, (2018), arXiv:1807.03761.

[2] Levent Hasan Ali Alpöge, *Points on Curves*, ProQuest LLC, Ann Arbor, MI, 2020, Thesis (Ph.D.)–Princeton University. MR 4209988

[3] Manjul Bhargava and Benedict H. Gross, *The average size of the 2-Selmer group of Jacobians of hyperelliptic curves having a rational Weierstrass point*, Automorphic representations and *L*-functions, Tata Inst. Fundam. Res. Stud. Math., vol. 22, Tata Inst. Fund. Res., Mumbai, 2013, pp. 23–91. MR 3156850

[4] Enrico Bombieri and Walter Gubler, *Heights in Diophantine geometry*, New Mathematical Monographs, vol. 4, Cambridge University Press, Cambridge, 2006. MR 2216774

[5] J.-H. Evertse, *On equations in S-units and the Thue-Mahler equation*, Invent. Math. **75** (1984), no. 3, 561–584. MR 735341

[6] J.-H. Evertse and J. H. Silverman, *Uniform bounds for the number of solutions to $Y^n = f(X)$*, Math. Proc. Cambridge Philos. Soc. **100** (1986), no. 2, 237–248. MR 848850

[7] Arul Shankar and Xiaoheng Wang, *Rational points on hyperelliptic curves having a marked non-Weierstrass point*, Compos. Math. **154** (2018), no. 1, 188–222. MR 3719247

[8] X, *The integer solutions of the equation $y^2 = ax^n + bx^{n-1} + \cdots + k$*, Journal of the London Mathematical Society **s1-1** (1926), no. 2, 66–68, (Extract from a letter from C. L. Siegel to L. J. Mordell.).