ACTA ARITHMETICA
149.2 (2011)

Effective irrationality measures and approximation by algebraic conjugates

by

PAUL M. VOUTIER (London)

1. Introduction. In a recent article [3], we investigated Thue's Fundamentaltheorem [2], showing when it can be used and how to use it in these cases. Using the notation of Theorems 2.1 and 2.4 of [3], we also showed that the case when $[\mathbb{K}(\beta_1) : \mathbb{K}] = 1$ is equivalent to the "usual" hypergeometric method (see Corollary 2.6 of [3]), where, here and in what follows, $\mathbb{K}$ is either $\mathbb{Q}$ or an imaginary quadratic field.

We also considered the case of $[\mathbb{K}(\beta_1) : \mathbb{K}] = 2$ in [3]. The approximants $P_r(x)$ and $Q_r(x)$ that we defined in Lemma 3.3 of [3] have a particularly nice form: an algebraic number plus or minus its algebraic conjugate. This raises the intriguing question of why.

We address that question here and show that the form of $P_r(x)$ and $Q_r(x)$ arises from the fact that Thue's Fundamentaltheorem is a special case of the application to hypergeometric polynomials of a new observation regarding diophantine approximations.

We present this observation here along with a generalisation and extension of Thue's Fundamentaltheorem. In the notation of [3], we are now able to consider more general expressions in place of $W(x)$ (see also Remark 3.3 below) as well as more general expressions for the denominator of $\mathcal{A}(x)$. There are also further improvements such as the consideration of powers m/n rather than just $1/n$, simplification of the numerator of $\mathcal{A}(x), \ldots$

The cost of these improvements is merely in the constant c that appears in our results below. The irrationality measure, κ, itself remains unchanged.

2. Notation. For positive integers m and n with $0 < m < n$, $(m, n) = 1$ and for a non-negative integer r, we put

2010 *Mathematics Subject Classification*: Primary 11J82, 11J68.
Key words and phrases: diophantine approximation, effective irrationality measures, hypergeometric functions.

DOI: 10.4064/aa149-2-3

$$X_{m,n,r}(x) = {}_2F_1(-r, -r - m/n; 1 - m/n; x),$$

where ${}_2F_1$ denotes the classical hypergeometric function.

We use $X^*_{m,n,r}$ to denote the homogeneous polynomials derived from these polynomials, so that

$$X^*_{m,n,r}(x, y) = y^r X_{m,n,r}(x/y).$$

We let $D_{m,n,r}$ denote the smallest positive integer such that the polynomial $D_{m,n,r}X_{m,n,r}(x)$ has rational integer coefficients.

For a positive integer d, we define $N_{d,n,r}$ to be the greatest common divisor of the numerators of the coefficients of $X_{m,n,r}(1 - dx)$.

We will use $v_p(x)$ to denote the largest power of a prime p which divides into the rational number x. With this notation, for positive integers d and n, we put

$$\mathcal{N}_{d,n} = \prod_{p|n} p^{\min(v_p(d), v_p(n)+1/(p-1))}. \tag{2.1}$$

For any complex number w, we can write $w = |w|e^{i\varphi}$, where $|w| \geq 0$ and $-\pi < \varphi \leq \pi$ (with $\varphi = 0$ if $w = 0$). With such a representation, unless otherwise stated, $w^{m/n}$ will signify $(|w|^{1/n})^m e^{im\varphi/n}$ for positive integers m and n, where $|w|^{1/n}$ is the unique non-negative nth root of $|w|$.

Lastly, following the function name in PARI, we define core(n) to be the unique squarefree divisor, n_1, of n such that n/n_1 is a perfect square.

3. Results

Proposition 3.1. *Let $\mathbb{K}$ be either $\mathbb{Q}$ or an imaginary quadratic field. Let $s \geq 2$ be a positive integer and $\mathbb{L}$ be a number field with $[\mathbb{L} : \mathbb{K}] = s$.*

Let $\theta_1 = 1, \theta_2, \ldots, \theta_s \in \mathbb{C}$ be linearly independent over $\mathbb{K}$ and let $\sigma_1 = \mathrm{id}, \sigma_2, \ldots, \sigma_s$ be the s embeddings of $\mathbb{L}$ into $\mathbb{C}$ that fix $\mathbb{K}$.

Suppose that there exist real numbers $k_0, l_0 > 0$ and $E, Q > 1$ such that for all non-negative integers r, there are algebraic integers $p_r \in \mathbb{L}$ with $\max_{1\leq i\leq s} |\sigma_i(p_r)| < k_0Q^r$.

Let β and γ be algebraic integers in $\mathbb{L}$.

(i) *Assume that $\sum_{1\leq i,j\leq s} \{\sigma_i(\beta)\sigma_j(\gamma) - \sigma_j(\beta)\sigma_i(\gamma)\}\, \sigma_i(p_r)\sigma_j(p_{r+1}) \neq 0$ and $\max_{2\leq i\leq s} |p_r\theta_i - \sigma_i(p_r)| < l_0E^{-r}$. Put*

$$\alpha = \frac{\sum_{i=1}^{s} \sigma_i(\beta)\theta_i}{\sum_{i=1}^{s} \sigma_i(\gamma)\theta_i}.$$

For any algebraic integers p and q in $\mathbb{K}$ with $q \neq 0$, we have

$$\left|\alpha - \frac{p}{q}\right| > \frac{1}{c|q|^{\kappa+1}},$$

where

$$c = 2\Big(\sum_{i=1}^{s} |\sigma_i(\gamma)|\Big) k_0 Q \max\Big\{E, 2\Big(\sum_{i=2}^{s} |\sigma_i(\beta) - \alpha\sigma_i(\gamma)|\Big) l_0 E\Big\}^{\kappa},$$
$$\kappa = \frac{\log Q}{\log E}.$$

(ii) *For $s = 2$, assume that $\beta/\gamma, p_r/p_{r+1} \notin \mathbb{K}$, and either $|p_r\theta_2 - \sigma_2(p_r)| < l_0 E^{-r}$ or $|-p_r\theta_2 - \sigma_2(p_r)| < l_0 E^{-r}$. Put*

$$\alpha = \frac{\sigma_2(\beta)\theta_2 \pm \beta}{\sigma_2(\gamma)\theta_2 \pm \gamma},$$

where the operation in the numerator matches the operation in the denominator. If $\mathbb{K} = \mathbb{Q}$, then let $\tau = 1$, else let τ be an algebraic integer in $\mathbb{K}$ such that $\mathbb{L} = \mathbb{K}(\sqrt{\tau})$. For any algebraic integers p and q in $\mathbb{K}$ with $q \neq 0$, we have

$$\left|\alpha - \frac{p}{q}\right| > \frac{1}{c|q|^{\kappa+1}},$$

where

$$c = 2|\sqrt{\tau}|(|\gamma| + |\sigma_2(\gamma)|) k_0 Q \max\{E, 2|\sqrt{\tau}|\,|\sigma_2(\beta) - \alpha\sigma_2(\gamma)| l_0 E\}^{\kappa},$$
$$\kappa = \frac{\log Q}{\log E}.$$

We will use part (ii) of this proposition to prove the following theorems.

THEOREM 3.2. ([1]) *Let $\mathbb{K}$ be either $\mathbb{Q}$ or an imaginary quadratic field. Let $\mathbb{L}$ be a number field with $[\mathbb{L} : \mathbb{K}] = 2$ and let σ be the non-trivial element of* Gal$(\mathbb{L}/\mathbb{K})$. *If $\mathbb{K} = \mathbb{Q}$, then let $\tau = 1$, else let τ be an algebraic integer in $\mathbb{K}$ such that $\mathbb{L} = \mathbb{K}(\sqrt{\tau})$. Let β, γ, η be algebraic integers in $\mathbb{L}$.*

Let g be an algebraic number such that η/g and $\sigma(\eta)/g$ are algebraic integers (not necessarily in $\mathbb{L}$). For each non-negative integer r, let h_r be a non-zero algebraic integer with $h_r/g^r \in \mathbb{K}$ and $|h_r| \le h$ for some fixed positive real number h. Let d be the largest positive rational integer such that $(\sigma(\eta) - \eta)/(dg)$ is an algebraic integer and let $\mathcal{C}_n$ and $\mathcal{D}_n$ be positive real numbers such that

$$\max\left(1, \frac{\Gamma(1 - m/n) r!}{\Gamma(r + 1 - m/n)}, \frac{n\Gamma(r + 1 + m/n)}{m\Gamma(m/n) r!}\right) \frac{D_{m,n,r}}{N_{d,n,r}} < \mathcal{C}_n \left(\frac{\mathcal{D}_n}{\mathcal{N}_{d,n}}\right)^{r} \tag{3.1}$$

for all non-negative integers r.

([1]) Note that our theorems and corollary here correct a small error in Theorems 2.1, 2.4 and Corollary 2.7 of [3], where "max$(1, \ldots$" in the expressions for c should read "max$(E, \ldots$".

Put

$$\begin{aligned}\alpha &= \frac{\beta(\eta/\sigma(\eta))^{m/n} \pm \sigma(\beta)}{\gamma(\eta/\sigma(\eta))^{m/n} \pm \sigma(\gamma)},\\ E &= \left\{ \frac{\mathcal{D}_n}{|g|\mathcal{N}_{d,n}} \min(|\sqrt{\eta} - \sqrt{\sigma(\eta)}|^2, |\sqrt{\eta} + \sqrt{\sigma(\eta)}|^2) \right\}^{-1},\\ Q &= \frac{\mathcal{D}_n}{|g|\mathcal{N}_{d,n}} \max(|\sqrt{\eta} - \sqrt{\sigma(\eta)}|^2, |\sqrt{\eta} + \sqrt{\sigma(\eta)}|^2),\\ \kappa &= \frac{\log Q}{\log E},\\ c &= 4h|\sqrt{\tau}|(|\gamma| + |\sigma(\gamma)|)\mathcal{C}_n Q\\ &\quad \times \max\{E, 5h|\sqrt{\tau}|\,|1 - (\eta/\sigma(\eta))^{m/n}|\,|\beta - \alpha\gamma|\mathcal{C}_n E\}^{\kappa},\end{aligned}$$

where the operation in the numerator of the definition of α matches the operation in its denominator.

If $E > 1$ and either $0 < \eta/\sigma(\eta) < 1$ or $|\eta/\sigma(\eta)| = 1$ with $\eta/\sigma(\eta) \neq -1$, then

$$|\alpha - p/q| > \frac{1}{c|q|^{\kappa+1}} \tag{3.2}$$

for all algebraic integers p and q in $\mathbb{K}$ with $q \neq 0$.

Remark 3.3. Observe that in our definition of α, we take the nth root of $\eta/\sigma(\eta)$. However, this is more general than it may first appear. It can be applied to any quantity $\mu\eta/\sigma(\eta)$ where $\mu \in \mathbb{L}$ and $\mu = \nu/\sigma(\nu)$ for some $\nu \in \mathbb{L}$.

For example, although in Thue's Fundamentaltheorem we take the nth root of $-\eta/\sigma(\eta)$, it, and its generalisations, still follows from our results. Suppose $\mathbb{L} = \mathbb{K}(\sqrt{\tau})$ and put $\eta' = \sqrt{\tau}\eta$; then $-\eta/\sigma(\eta) = \eta'/\sigma(\eta')$, so we can express $-\eta/\sigma(\eta)$ in the form here (i.e., take $\mu = -1$ and $\nu = \sqrt{\tau}$ in the above notation). There appears to be an extra factor of $\sqrt{\tau}$ that will arise in our expressions for E and Q, but these are in fact cancelled out since g also increases by a factor of $\sqrt{\tau}$, so κ is unaffected.

Similarly, if $\mathbb{K} \neq \mathbb{Q}(i)$ and $\mathbb{L} = \mathbb{K}(i)$, then $i\eta/\sigma(\eta) = \eta'/\sigma(\eta')$, where $\eta' = (1+i)\eta$.

Also, if $\mathbb{K} \neq \mathbb{Q}(\sqrt{-3})$ and $\mathbb{L} = \mathbb{K}(\sqrt{-3})$, then $\zeta_3\eta/\sigma(\eta) = \eta'/\sigma(\eta')$, where $\eta' = (1-\sqrt{-3})\eta/2$. And $\zeta_6\eta/\sigma(\eta) = \eta'/\sigma(\eta')$, where $\eta' = (3+\sqrt{-3})\eta$.

As for the other roots of unity of degree at most 4 over $\mathbb{Q}$, it can be shown, via algebraic manipulation, that this is not possible for ζ_8 and ζ_{12}. And since $\mathbb{Q}(\zeta_5)$ contains no subfields besides $\mathbb{Q}$ and $\mathbb{Q}(\sqrt{5})$, we cannot consider $\zeta_5\eta/\sigma(\eta)$.

Remark 3.4. From Lemma 7.4 of [3], the inequality (3.1) holds for $\mathcal{C}_n$ and $\mathcal{D}_n$ as in [3] and hence it does not impose any constraint.

THEOREM 3.5. *Let* $\mathbb{K}$ *be an imaginary quadratic field and* α, β, γ, η, σ, τ, d, g, h, n, $\mathcal{C}_n$, $\mathcal{D}_n$, $\mathcal{N}_{d,n}$ *be as in Theorem* 3.2. *Put*

$$\begin{aligned} E &= \frac{4|g|\mathcal{N}_{d,n}}{\mathcal{D}_n}\frac{(|\eta|-|\sigma(\eta)-\eta|)}{|\sigma(\eta)-\eta|^2}, \\ Q &= \frac{2\mathcal{D}_n}{|g|\mathcal{N}_{d,n}}(|\eta|+|\sigma(\eta)|), \\ \kappa &= \frac{\log Q}{\log E}, \\ c &= 4h|\sqrt{\tau}|(|\gamma|+|\sigma(\gamma)|)\mathcal{C}_n Q \\ &\quad \times \max\{E, 2h|\sqrt{\tau}|\,|1-(\eta/\sigma(\eta))^{m/n}|\,|\beta-\alpha\gamma|\mathcal{C}_n E\}^{\kappa}. \end{aligned}$$

If $E > 1$ *and* $\max(|1-\eta/\sigma(\eta)|, |1-\sigma(\eta)/\eta|) < 1$, *then*

$$|\alpha - p/q| > \frac{1}{c|q|^{\kappa+1}} \tag{3.3}$$

for all algebraic integers p *and* q *in* $\mathbb{K}$ *with* $q \neq 0$.

REMARK 3.6. The condition that $\mathbb{K}$ be an imaginary quadratic field is no restriction since the case of $\mathbb{K} = \mathbb{Q}$ is completely covered by Theorem 3.2.

We now present a corollary of Theorem 3.2 when $\mathbb{K} = \mathbb{Q}$.

COROLLARY 3.7. *Let* $\mathbb{K} = \mathbb{Q}$ *and* $\alpha, \beta, \gamma, \eta$, σ, n, $\mathcal{C}_n$, $\mathcal{D}_n, \mathcal{N}_{d,n}$ *be as in Theorem* 3.2. *Suppose that* $\eta = (u_1 + u_2\sqrt{t})/2$ *where* $t, u_1, u_2 \in \mathbb{Z}$ *and* $t \neq 0$. *Put*

$$\begin{aligned} g_1 &= \gcd(u_1, u_2), \\ g_2 &= \gcd(u_1/g_1, t), \\ g_3 &= \begin{cases} 1 & \textit{if } t \equiv 1 \bmod 4 \textit{ and } (u_1-u_2)/g_1 \equiv 0 \bmod 2, \\ 2 & \textit{if } t \equiv 3 \bmod 4 \textit{ and } (u_1-u_2)/g_1 \equiv 0 \bmod 2, \\ 4 & \textit{otherwise}, \end{cases} \\ g_4 &= \gcd\left(\operatorname{core}(tg_2g_3), \frac{n}{\gcd((u_2/g_1)\sqrt{tg_3/(g_2\operatorname{core}(tg_2g_3))}, n)}\right), \\ g_5 &= \begin{cases} 2 & \textit{if } 2\,|\,n \textit{ and } v_2(u_2^2tg_3/(g_1^2g_2)) = v_2(2n^2), \\ 1 & \textit{otherwise}, \end{cases} \\ g &= \frac{g_1\sqrt{g_2}}{\sqrt{g_3g_4g_5}}, \\ E &= \frac{|g|\mathcal{N}_{d,n}}{\mathcal{D}_n \min(|u_1 \pm \sqrt{u_1^2-u_2^2t}|)}, \\ Q &= \frac{\mathcal{D}_n \max(|u_1 \pm \sqrt{u_1^2-u_2^2t}|)}{|g|\mathcal{N}_{d,n}}, \end{aligned}$$

$$\kappa = \frac{\log Q}{\log E},$$
$$c = 4\sqrt{|2t|}(|\gamma| + |\sigma(\gamma)|)\mathcal{C}_n Q$$
$$\times(\max(E, 5\sqrt{|2t|}\,|1-(\eta/\sigma(\eta))^{m/n}|\,|\beta-\alpha\gamma|\mathcal{C}_n E))^{\kappa},$$

where d *is the largest positive rational integer such that* $u_2\sqrt{t}/(dg)$ *is an algebraic integer. If* $E > 1$ *and either* $0 < \eta/\sigma(\eta) < 1$ *or* $|\eta/\sigma(\eta)| = 1$ *with* $\eta/\sigma(\eta) \neq -1$, *then*

$$|\alpha - p/q| > \frac{1}{c|q|^{\kappa+1}} \tag{3.4}$$

for all rational integers p *and* q *with* $q \neq 0$.

Remark 3.8. The factors g_i used to construct g each arise in natural and distinct ways. Namely, g_1 through g_3 provide ways to remove common factors from η and $\sigma(\eta)$. In turn, g_4 and g_5 arise from the interplay of d and g: under some circumstances (captured by g_4 and g_5), decreasing g can increase d and hence $\mathcal{N}_{d,n}$ by more to provide a net benefit.

Remark 3.9. Using the same argument as in the proof of Corollary 3.7, we can also improve Corollary 2.7 of [3], replacing g_4 there by

$$\gcd\left(\mathrm{core}(g_2g_3), \frac{n}{\gcd((u_1/g_1)\sqrt{g_3/(g_2\mathrm{core}(g_2g_3))}, n)}\right)$$

and adding an appropriate version of the g_5 above by setting $g_5 = 2$ if $2\,|\,n$ and $v_2(u_1^2g_3/(g_1^2g_2)) = v_2(2n^2)$ and setting $g_5 = 1$ otherwise, since the definition of d in Corollary 2.7 of [3] uses $u_1/(dg)$ rather than $u_2\sqrt{t}/(dg)$ as here.

This improved version of Corollary 2.7 of [3] will yield the same results as in Corollary 3.7 together with Remark 3.3.

4. Preliminary lemmas. The next lemma contains the relationship that allows the hypergeometic method to provide good sequences of rational approximations.

Lemma 4.1. *For any positive integers* m *and* n *with* $(m, n) = 1$, *any non-negative integer* r *and any complex number* z *that is not a negative number and not zero,*

$$z^{m/n}z^r X_{m,n,r}(z^{-1}) - X_{m,n,r}(z) = (z-1)^{2r+1}R_{m,n,r}(z), \tag{4.1}$$

where

$$(z-1)^{2r+1}R_{m,n,r}(z) = \frac{\Gamma(r+1+m/n)}{r!\,\Gamma(m/n)}\int_1^z (1-t)^r(t-z)^r t^{m/n-r-1}\,dt.$$

REMARK 4.2. Note that the expression $(z-1)^{2r+1}R_{m,n,r}(z)$ here is the same as the $R_{m,n,r}(z)$ defined in Lemma 7.1 of [3].

Proof of Lemma 4.1. This is shown in the case of $m=1$ in the proof of Lemma 2.3 of [1]. The proof for arbitrary m is identical. ∎

LEMMA 4.3. *Let $\theta \in \mathbb{C}$ and let $\mathbb{K}$ be either $\mathbb{Q}$ or an imaginary quadratic field. Suppose that there exist real numbers $k_0, l_0 > 0$ and $E, Q > 1$ such that for all non-negative integers r, there are algebraic integers p_r and q_r in $\mathbb{K}$ with $|q_r| < k_0 Q^r$ and $|q_r\theta - p_r| \le l_0 E^{-r}$ satisfying $p_r q_{r+1} \ne p_{r+1} q_r$. Then for any algebraic integers p and q in $\mathbb{K}$ with $q \ne 0$, we have*

$$\left|\theta - \frac{p}{q}\right| > \frac{1}{c|q|^{\kappa+1}}, \quad \text{where} \quad c = 2k_0 Q(\max(1, 2l_0)E)^{\kappa}, \quad \kappa = \frac{\log Q}{\log E}.$$

Moreover, if $p/q \ne p_i/q_i$ for any non-negative integer i, then we can put $c = 2k_0(\max(1, 2l_0)E)^{\kappa}$.

Proof. This follows from Lemma 6.1 of [3]. There we proved a similar result for $|q| \ge 1/(2l_0)$ and $c = 2k_0 Q(2l_0 E)^{\kappa}$. Here we merely observe that if we replace l_0 with $\max(0.5, l_0)$, then all the hypotheses of the present lemma still hold. Moreover, $1/(2\max(0.5, l_0)) \le 1$, so the result holds for all non-zero algebraic integers $q \in \mathbb{K}$.

The last statement in the lemma follows since the Q which appears in the expression for c in the statement of Lemma 6.1 of [3] arises only from consideration of the case $p/q = p_i/q_i$ for some positive integer i. ∎

5. Proof of Proposition 3.1. Assume that we have a sequence of p_r's satisfying the hypotheses of Proposition 3.1.

(i) Suppose we have $p_r\theta_i - \sigma_i(p_r) = \delta_{i,r}$ for each $i = 1, \ldots, s$. Then we can write

$$\alpha = \frac{\sum_{i=1}^{s} \sigma_i(\beta)(\delta_{i,r} + \sigma_i(p_r))}{\sum_{i=1}^{s} \sigma_i(\gamma)(\delta_{i,r} + \sigma_i(p_r))}$$

and hence

$$\alpha \sum_{i=1}^{s} \sigma_i(\gamma p_r) - \sum_{i=1}^{s} \sigma_i(\beta p_r) = \sum_{i=2}^{s} (\sigma_i(\beta) - \alpha\sigma_i(\gamma))\delta_{i,r},$$

since $\delta_{1,r} = 0$.

Put $p'_r = \sum_{i=1}^{s} \sigma_i(\beta p_r)$ and $q'_r = \sum_{i=1}^{s} \sigma_i(\gamma p_r)$. Note that both p'_r and q'_r are algebraic integers in $\mathbb{K}$.

Observe that

$$|\alpha q'_r - p'_r| < l_0 \Big(\sum_{i=2}^{s} |\sigma_i(\beta) - \alpha\sigma_i(\gamma)|\Big) E^{-r}$$

and

$$|q_r'| \le k_0 \Big(\sum_{i=1}^{s} |\sigma_i(\gamma)| \Big) Q^r.$$

Since

$$p_r' q_{r+1}' - p_{r+1}' q_r' = \sum_{1 \le i,j \le s} \{\sigma_i(\beta)\sigma_j(\gamma) - \sigma_j(\beta)\sigma_i(\gamma)\} \sigma_i(p_r)\sigma_j(p_{r+1}) \ne 0$$

by our assumption in the statement of the proposition, we can apply Lemma 4.3 with p_r' and q_r' instead of p_r and q_r, respectively, to complete the proof in this case.

(ii) Suppose we have $\zeta_2 p_r \theta_2 - \sigma_2(p_r) = \delta_{2,r}$ for some square root ζ_2 of 1, fixed for a given value of r. As above, we can write

$$\alpha\{\sigma_2(\gamma p_r) \pm \zeta_2 \gamma p_r\} - \{\sigma_2(\beta p_r) \pm \zeta_2 \beta p_r\} = \delta_{2,r}(\sigma_2(\beta) - \alpha\sigma_2(\gamma)).$$

We break the proof into two cases depending on the value of ζ_2.

Case 1: $\pm\zeta_2 = 1$. This case is identical to part (i) with $s = 2$.

Note that in this case ($s = 2$), the condition in part (i) reduces to

$$(\sigma_2(\beta)\gamma - \beta\sigma_2(\gamma))(\sigma_2(p_r)p_{r+1} - p_r\sigma_2(p_{r+1})) \ne 0.$$

This is true under the conditions we have stipulated here, namely $\beta/\gamma \notin \mathbb{K}$ and $p_r/p_{r+1} \notin \mathbb{K}$ (since the fixed field of σ_2 is $\mathbb{K}$).

Also since $|\tau| \ge 1$, our definition of c is valid.

Case 2: $\pm\zeta_2 = -1$. We break this case into two subcases.

Case 2(i): $\pm\zeta_2 = -1$ and $\mathbb{K} = \mathbb{Q}$. If $\mathbb{K} = \mathbb{Q}$, then we can write $\beta p_r = (a + b\sqrt{t})/2$ for some choice of rational integers a, b and t with $t \ne 0$. Hence $\beta p_r - \sigma_2(\beta p_r) = b\sqrt{t}$ and $(\beta p_r - \sigma_2(\beta p_r))/\sqrt{t} \in \mathbb{Z}$. Similarly, $(\gamma p_r - \sigma_2(\gamma p_r))/\sqrt{t} \in \mathbb{Z}$.

In this case, we put $q_r' = (\gamma p_r - \sigma_2(\gamma p_r))/\sqrt{t}$ and $p_r' = (\beta p_r - \sigma_2(\beta p_r))/\sqrt{t}$ and observe that

$$|\alpha q_r' - p_r'| < \frac{l_0|\sigma_2(\beta) - \alpha\sigma_2(\gamma)|}{|\sqrt{t}|} E^{-r} \le l_0|\sqrt{\tau}|\,|\sigma_2(\beta) - \alpha\sigma_2(\gamma)|E^{-r},$$

$$|q_r'| \le \frac{k_0(|\gamma| + |\sigma_2(\gamma)|)}{|\sqrt{t}|} Q^r \le k_0|\sqrt{\tau}|(|\gamma| + |\sigma_2(\gamma)|)Q^r,$$

since $|t| \ge 1$.

Case 2(ii): $\pm\zeta_2 = -1$ and $\mathbb{K}$ is an imaginary quadratic field. If $\mathbb{K}$ is an imaginary quadratic field, then $\beta p_r = a + b\sqrt{\tau}$ for some $a, b \in \mathbb{K}$ and with τ as in the statement of the proposition. Hence $\beta p_r - \sigma_2(\beta p_r) = 2b\sqrt{\tau}$ is an algebraic integer and $(\beta p_r - \sigma_2(\beta p_r))\sqrt{\tau}$ is an algebraic integer in $\mathbb{K}$. Similarly, $(\gamma p_r - \sigma_2(\gamma p_r))\sqrt{\tau}$ is an algebraic integer in $\mathbb{K}$.

In this case, we put $q_r' = (\gamma p_r - \sigma_2(\gamma p_r))\sqrt{\tau}$ and $p_r' = (\beta p_r - \sigma_2(\beta p_r))\sqrt{\tau}$ and observe that

$$|\alpha q_r' - p_r'| < l_0|\sqrt{\tau}|\,|\sigma_2(\beta) - \alpha\sigma_2(\gamma)|E^{-r}, \qquad |q_r'| \le k_0|\sqrt{\tau}|(|\gamma| + |\sigma_2(\gamma)|)Q^r.$$

Note that in both these subcases, we obtain the same upper bound for $|\alpha q_r' - p_r'|$ and for $|q_r'|$.

Here

$$p_r'q_{r+1}' - p_{r+1}'q_r' = \tau(\beta\sigma_2(\gamma) - \sigma_2(\beta)\gamma)(\sigma_2(p_r)p_{r+1} - p_r\sigma_2(p_{r+1})),$$

which we saw in Case 1 can only be zero if $\beta/\gamma \in \mathbb{K}$ or $p_r/p_{r+1} \in \mathbb{K}$.

Therefore, we can apply Lemma 4.3 to find that $\kappa = (\log Q)/(\log E)$ and

$$c = 2k_0|\sqrt{\tau}|(|\gamma| + |\sigma_2(\gamma)|)Q\max\{E, 2l_0|\sqrt{\tau}|\,|\sigma_2(\beta) - \alpha\sigma_2(\gamma)|E\}^{\kappa},$$

concluding the proof of Case 2 and of the proposition.

6. Proof of Theorem 3.2

6.1. Construction of approximations. We construct the approximations under more general conditions. The point is not to generalise for its own sake, but to illustrate the requirements and limitations of our method of proof.

Let ζ_k be a kth root of unity for some k. We apply Lemma 4.1 with $z = \zeta_k\eta/\sigma(\eta)$. Multiplying both sides of (4.1) by $\sigma(\eta)^r$, we obtain

$$\begin{aligned}(\zeta_k\eta/\sigma(\eta))^{m/n}(\zeta_k\eta)^r X_{m,n,r}(\sigma(\eta)/(\zeta_k\eta)) - \sigma(\eta)^r X_{m,n,r}(\zeta_k\eta/\sigma(\eta))\\ = \sigma(\eta)^r(\zeta_k\eta/\sigma(\eta) - 1)^{2r+1}R_{m,n,r}(\zeta_k\eta/\sigma(\eta)),\end{aligned}$$

which we can rewrite as

$$\begin{aligned}(\zeta_k\eta/\sigma(\eta))^{m/n}X^*_{m,n,r}(\sigma(\eta),\zeta_k\eta) - X^*_{m,n,r}(\zeta_k\eta,\sigma(\eta))\\ = \sigma(\eta)^r(\zeta_k\eta/\sigma(\eta) - 1)^{2r+1}R_{m,n,r}(\zeta_k\eta/\sigma(\eta)).\end{aligned}$$

Observe that

$$\begin{aligned}X^*_{m,n,r}(\zeta_k\eta,\sigma(\eta)) &= g^r X^*_{m,n,r}\left(\frac{\zeta_k\eta}{g}, \frac{\sigma(\eta)}{g}\right)\\ &= \left(g\frac{\sigma(\eta)}{g}\right)^r X_{m,n,r}\left(1 - d_k\frac{(\sigma(\eta) - \zeta_k\eta)/g}{d_k\sigma(\eta)/g}\right),\end{aligned}$$

where d_k is the largest positive rational integer such that $(\sigma(\eta) - \zeta_k\eta)/(gd_k)$ is an algebraic integer.

From Lemma 7.4(a) of [3],

$$\frac{D_{m,n,r}}{N_{d_k,n,r}}X_{m,n,r}\left(1 - d_k\frac{(\sigma(\eta) - \zeta_k\eta)/g}{d_k\sigma(\eta)/g}\right) \in \mathbb{Z}\left[\frac{(\sigma(\eta) - \zeta_k\eta)/g}{d_k\sigma(\eta)/g}\right],$$

and, as a consequence,

$$\left(\frac{\sigma(\eta)}{g}\right)^r \frac{D_{m,n,r}}{N_{d_k,n,r}} X_{m,n,r}\left(1 - d_k \frac{(\sigma(\eta) - \zeta_k\eta)/g}{d_k\sigma(\eta)/g}\right)$$

is an algebraic integer by our definition of d_k. Hence

$$p_r = \frac{h_r D_{m,n,r}}{g^r N_{d_k,n,r}} X^*_{m,n,r}(\zeta_k\eta, \sigma(\eta))$$

is an algebraic integer in $\mathbb{L}$.

Similarly,

$$q_r = \frac{h_r D_{m,n,r}}{g^r N_{d_k,n,r}} X^*_{m,n,r}(\sigma(\eta), \zeta_k\eta)$$

is an algebraic integer in $\mathbb{L}$.

Now we want p_r and q_r, or at least numbers obtained from them, to be algebraic conjugates. For this purpose, we must suppose that $1/\zeta_k = \sigma(\zeta_k)$ (note that this implies that $\zeta_k \in \mathbb{L}$).

With this condition, and since $\sigma^2(\cdot)$ is the identity map, we have

$$\begin{aligned}(\zeta_k)^r\sigma(X^*_{m,n,r}(\zeta_k\eta, \sigma(\eta))) &= (\zeta_k)^r\sigma(\sigma(\eta)^r X_{m,n,r}(\zeta_k\eta/\sigma(\eta))) \\ &= (\zeta_k\eta)^r X_{m,n,r}(\sigma(\zeta_k\eta/\sigma(\eta))) \\ &= (\zeta_k\eta)^r X_{m,n,r}(\sigma(\eta)/(\zeta_k\eta)) \\ &= X^*_{m,n,r}(\sigma(\eta), \zeta_k\eta).\end{aligned}$$

Hence, $q_r = \zeta_k^r\sigma(p_r)$ and so q_r and $\sigma(\zeta_k)^r p_r$ are algebraic conjugates over $\mathbb{K}$. Letting $k_1 = k/(2,k)$, we see that $p_{k_1 r}$ and $\pm q_{k_1 r}$ are algebraic conjugates for $k = 1$, 2, 3, 4 and 6, so we could put $p'_r = p_{k_1 r}$ and $q'_r = q_{k_1 r}$.

However here we restrict our attention to $k = 1$ and observe that in this case p_r and q_r are algebraic conjugates (noting that d_1 equals d in the statement of our theorem).

6.2. Estimates. From Lemmas 7.3(a) and 7.4(c) of [3], we have

$$\begin{aligned}|q_r| &\le \frac{2h}{|g|^r} \frac{D_{m,n,r}}{N_{d,n,r}} \frac{\Gamma(1-m/n)r!}{\Gamma(r+1-m/n)} \max(|\sqrt{\eta} + \sqrt{\sigma(\eta)}|, |\sqrt{\eta} - \sqrt{\sigma(\eta)}|)^{2r} \\ &\le 2h\mathcal{C}_n\left(\frac{\mathcal{D}_n}{|g|\mathcal{N}_{d,n}}\right)^r \max(|\sqrt{\eta} + \sqrt{\sigma(\eta)}|, |\sqrt{\eta} - \sqrt{\sigma(\eta)}|)^{2r}.\end{aligned}$$

From Lemma 7.2(a) of [3],

$$\begin{aligned}&|(\sigma(\eta))^r(\eta/\sigma(\eta) - 1)^{2r+1}R_{m,n,r}(\eta/\sigma(\eta))| \\ &\qquad \le 2.38|1 - (\eta/\sigma(\eta))^{m/n}| \frac{n\Gamma(r+1+m/n)}{m\Gamma(m/n)r!} \\ &\qquad\quad \times \min(|\sqrt{\eta} + \sqrt{\sigma(\eta)}|, |\sqrt{\eta} - \sqrt{\sigma(\eta)}|)^{2r}.\end{aligned}$$

Hence

$$\begin{aligned}|q_r(\eta/\sigma(\eta))^{m/n} - p_r| &\le 2.38h\frac{D_{m,n,r}}{|g|^r N_{d,n,r}}|1-(\eta/\sigma(\eta))^{m/n}|\frac{n\Gamma(r+1+m/n)}{m\Gamma(m/n)r!}\\ &\quad\times \min(|\sqrt{\eta}+\sqrt{\sigma(\eta)}|, |\sqrt{\eta}-\sqrt{\sigma(\eta)}|)^{2r}\\ &\le \frac{2.38h}{|g|^r}|1-(\eta/\sigma(\eta))^{m/n}|\mathcal{C}_n\left(\frac{\mathcal{D}_n}{\mathcal{N}_{d,n}}\right)^r\\ &\quad\times \min(|\sqrt{\eta}+\sqrt{\sigma(\eta)}|, |\sqrt{\eta}-\sqrt{\sigma(\eta)}|)^{2r}.\end{aligned}$$

Therefore, in the notation of Proposition 3.1, we have

$$\begin{aligned}k_0 &= 2h\mathcal{C}_n,\\ l_0 &= 2.38h|1-(\eta/\sigma(\eta))^{m/n}|\mathcal{C}_n,\\ E &= \left\{\frac{\mathcal{D}_n}{|g|\mathcal{N}_{d,n}}\min(|\sqrt{\eta}-\sqrt{\sigma(\eta)}|^2, |\sqrt{\eta}+\sqrt{\sigma(\eta)}|^2)\right\}^{-1},\\ Q &= \frac{\mathcal{D}_n}{|g|\mathcal{N}_{d,n}}\max(|\sqrt{\eta}-\sqrt{\sigma(\eta)}|^2, |\sqrt{\eta}+\sqrt{\sigma(\eta)}|^2).\end{aligned}$$

From Proposition 3.1, the expression for κ in the theorem follows immediately, while, upon noting that our β, γ, $\sigma(\beta)$ and $\sigma(\gamma)$ here are $\sigma_2(\beta)$, $\sigma_2(\gamma)$, β and γ respectively in the notation of that proposition,

$$\begin{aligned}c &= 2|\sqrt{\tau}|(|\gamma|+|\sigma(\gamma)|)k_0Q\max\{E, 2|\sqrt{\tau}|(|\beta-\alpha\gamma|)l_0E\}^\kappa\\ &< 4h|\sqrt{\tau}|(|\gamma|+|\sigma(\gamma)|)\mathcal{C}_nQ\\ &\quad\times\max\{E, 5h|\sqrt{\tau}|\,|1-(\eta/\sigma(\eta))^{m/n}|\,|\beta-\alpha\gamma|\mathcal{C}_nE\}^\kappa.\end{aligned}$$

7. Proof of Theorem 3.5. This proof is the same as that of Theorem 3.2, except that we use the upper bounds from parts (b) of Lemmas 7.2 and 7.3 of [3], rather than parts (a). Thus, we find that

$$\begin{aligned}k_0 &= 2h\mathcal{C}_n,\\ l_0 &= h|1-(\eta/\sigma(\eta))^{m/n}|\mathcal{C}_n,\\ E &= \frac{4|g|\mathcal{N}_{d,n}}{\mathcal{D}_n}\,\frac{(|\eta|-|\sigma(\eta)-\eta|)}{|\sigma(\eta)-\eta|^2},\\ Q &= \frac{2\mathcal{D}_n}{|g|\mathcal{N}_{d,n}}\,(|\eta|+|\sigma(\eta)|).\end{aligned}$$

So, from Proposition 3.1, κ is as in the statement of the theorem and, again noting the change of notation mentioned at the end of the proof of Theorem 3.2,

$$\begin{aligned}c &= 2|\sqrt{\tau}|(|\gamma|+|\sigma(\gamma)|)k_0Q\max\{E, 2|\sqrt{\tau}|\,|\beta-\alpha\gamma|l_0E\}^\kappa\\ &= 4h|\sqrt{\tau}|(|\gamma|+|\sigma(\gamma)|)\mathcal{C}_nQ\\ &\quad\times\max\{E, 2h|\sqrt{\tau}|\,|1-(\eta/\sigma(\eta))^{m/n}|\,|\beta-\alpha\gamma|\mathcal{C}_nE\}^\kappa.\end{aligned}$$

8. Proof of Corollary 3.7. This corollary follows from a direct application of Theorem 3.2.

We can write

$$(\sqrt{\eta} \pm \sqrt{\sigma(\eta)})^2 = \eta + \sigma(\eta) \pm 2\sqrt{\eta\sigma(\eta)}. \tag{8.1}$$

The right-hand side of (8.1) is $u_1 \pm \sqrt{u_1^2 - u_2^2 t}$ and $\sigma(\eta) - \eta = -u_2\sqrt{t}$. Hence d is as defined in the corollary.

The analysis of g_1, g_2 and g_3 is identical to that in Section 11 of [3].

As stated in Remark 3.8, g_4 and g_5 arise from the interplay of d and g. Suppose that d_1 is the largest positive rational integer such that the quotient $u_2\sqrt{t}/(d_1 g_1\sqrt{g_2/g_3})$ is an algebraic integer. If there are multiplicative factors of the form $\sqrt{d_2}$ in $u_2\sqrt{t}/(d_1 g_1\sqrt{g_2/g_3})$, then by multiplying η, and hence $u_2\sqrt{t}$, by $\sqrt{d_2}$, we can increase d_1 by a factor of d_2. Under some circumstances, this increases $\mathcal{N}_{d,n}$ by a factor of d_2 while increasing $u_1 \pm \sqrt{u_1^2 - u_2^2 t}$ only by a factor of $\sqrt{d_2}$ for a net reduction in the size of κ. We demonstrate here how g_4 and g_5 capture these circumstances.

Consider the integer $u_2^2 t g_3/(g_1^2 g_2)$ and let d_1^2 be its largest square divisor. Suppose that p is a prime divisor of their quotient. That is, p is a prime divisor of $\mathrm{core}(u_2^2 t g_3/(g_1^2 g_2)) = \mathrm{core}(t g_3/g_2) = \mathrm{core}(t g_2 g_3)$. Note that

$$d_1 = \sqrt{u_2^2 t g_3/(g_1^2 g_2\,\mathrm{core}(t g_2 g_3))} = (u_2/g_1)\sqrt{t g_3/(g_2\,\mathrm{core}(t g_2 g_3))}.$$

First, if $p \nmid n$, then $\mathcal{N}_{pd_1,n} = \mathcal{N}_{d_1,n}$ from the definition of $\mathcal{N}_{d,n}$ in (2.1) and there is no benefit.

Second, if $p \mid n$ and $p \nmid (n/\gcd(d_1,n))$, then $\mathcal{N}_{pd_1,n}$ is at most $\mathcal{N}_{d_1,n} p^{1/(p-1)}$ (again, from (2.1)). That is, we gain at most a factor of $p^{1/(p-1)}$, while increasing the size of $u_1 \pm \sqrt{u_1^2 - u_2^2 t}$ by a factor of $\sqrt{p}$, and hence obtain no benefit for $p > 2$.

Third, if $p \mid n$ and $p \mid (n/\gcd(d_1,n))$, then we gain a factor of p, while we increase the size of $u_1 \pm \sqrt{u_1^2 - u_2^2 t}$ by a factor of $\sqrt{p}$. The product of all such p equals

$$\gcd\left(\mathrm{core}(t g_2 g_3), \frac{n}{\gcd((u_2/g_1)\sqrt{t g_3/(g_2\,\mathrm{core}(t g_2 g_3))}, n)}\right),$$

which is our g_4.

This covers all possible cases except $2 \mid n$ and $2 \nmid (n/\gcd(d_1,n))$. If the power of 2 dividing d equals the power of 2 dividing n, both are positive and $2 \mid \mathrm{core}(t g_2 g_3)$, then we increase $\mathcal{N}_{d_1,n}$ by a factor of 2, while we increase the size of $u_1 \pm \sqrt{u_1^2 - u_2^2 t}$ by a factor of $\sqrt{2}$. Since $u_2^2 t g_3/(g_1^2 g_2) = d_1^2 \mathrm{core}(t g_2 g_3)$, this condition is equivalent to our condition in the definition of g_5.

Lastly, we must consider h_r and h.

Since $g^2 \in \mathbb{Q}$, we can take $h_r = 1$ for r even. As $(g_3 g_4 g_5/g_2)\mathrm{core}(g_2 g_3 g_4 g_5)$ is a perfect square, we can take $h_r = \sqrt{\mathrm{core}(g_2 g_3 g_4 g_5)}$ for r odd. Observe that $g_4 g_5 \mid (2t g_3/g_2)$, $g_2 \mid t$ and $g_3 \mid 4$. Hence $h_r \le \sqrt{|2t|}$ for r odd.

Acknowledgements. The author is grateful to the referees for their very attentive reading of this paper and the corrections and improvements that they suggested. Their diligence has led to a much better paper.

The author also expresses his deep appreciation to the Judith and Lawrence Schulman Foundation for their generous support throughout the entire course of this work.

References

[1] J. H. Chen and P. M. Voutier, *Complete solution of the diophantine equation $X^2 + 1 = dY^4$ and a related family of quartic Thue equations*, J. Number Theory 62 (1997), 71–99.

[2] A. Thue, *Ein Fundamentaltheorem zur Bestimmung von Annäherungswerten aller Wurzeln gewisser ganzer Funktionen*, J. Reine Angew. Math. 138 (1910), 96–108.

[3] P. M. Voutier, *Thue's Fundamentaltheorem, I: The general case*, Acta Arith. 143 (2010), 101–144.

Paul M. Voutier
London, UK
E-mail: Paul.Voutier@gmail.com

Received on 9.9.2009
and in revised form on 7.12.2010 (6145)