ACTA ARITHMETICA
110.4 (2003)

Parametrizing $\mathrm{SL}_2(\mathbb{Z})$ and a question of Skolem

by

UMBERTO ZANNIER (Venezia)

Introduction. This paper is a sequel to [Z], where the *determinantal* equation was considered

$$X_1X_2 - X_3X_4 = 1. \tag{1}$$

We were primarily interested in a question raised by Skolem [S, p. 23], namely: *Can all the integral solutions of* (1) *be obtained from a fixed polynomial solution by letting the variables run through* $\mathbb{Z}$*?*

Skolem expressed his belief in favour of a negative answer. We showed in [Z, Thm. 1] that indeed no suitable polynomial solution may exist depending on three variables at most; actually, we proved an analogous, slightly stronger, result ([Z, Thm. 2]), valid for the integers in an arbitrary number field and for algebraic varieties more general than SL_2. However, we also pointed out that the truth of the Generalized Riemann Hypothesis implies the existence of counterexamples to the analogue for $\mathbb{Z}[\sqrt{2}]$ of Skolem's belief, with a polynomial depending on five variables.

Now, for a diophantine equation it has been proved natural to consider not only the solutions in classical integers (of $\mathbb{Z}$ or a number field), but those in S-integers, where S is a finite set of places; in other words, to allow denominators constructed out only of primes from a given finite set. In the present paper we show unconditionally that the above question has a *positive* answer, contrary to Skolem's expectation, if we replace $\mathbb{Z}$ with the ring of S-integers in $\mathbb{Q}$, for a suitable finite S. Actually, we shall obtain a more explicit result, to be stated in a moment.

First, (as in [Z]) we consider the "general" continued fraction with five partial quotients, namely the expression $Y_0 + \frac{1}{Y_1+}\frac{1}{Y_2+}\frac{1}{Y_3+}\frac{1}{Y_4}$, for variables $Y_0, \ldots, Y_4$. If p_3/q_3 and p_4/q_4 are the last two convergents we see that equation (1) is satisfied if we put $X_1 = p_3$, $X_2 = q_4$, $X_3 = p_4$, $X_4 = q_3$, so we obtain a polynomial solution in five variables, which we denote by $\mathbf{f} = (p_3, q_4, p_4, q_3) \in \mathbb{Z}[Y_0, \ldots, Y_4]^4$; we shall show that it does the job. (As

2000 *Mathematics Subject Classification*: 11D04, 11D09, 11G99.

remarked in [Sz] however, no polynomial solution constructed in this way with any number of partial quotients can give all solutions of (1) over $\mathbb{Z}$ by specializing the Y_i in $\mathbb{Z}$.)

As usual, for a finite set S of prime numbers, we define the ring of *S-integers* (in $\mathbb{Q}$) by

$$\mathcal{O}_S = \Big\{x \in \mathbb{Q} : \exists b \in \mathbb{N},\ x\prod_{l\in S} l^b \in \mathbb{Z}\Big\}.$$

With this notation, we shall prove the following

Theorem. *Let $S = \{2, 3, l\}$, for a prime $l \equiv 1 \pmod 4$. Then, given a solution $\mathbf{x} = (x_1, x_2, x_3, x_4) \in \mathcal{O}_S^4$ of* (1), *there exists $\mathbf{y} = (y_0, \ldots, y_4) \in \mathcal{O}_S^5$ such that $\mathbf{f}(\mathbf{y}) = \mathbf{x}$.*

The proofs in [Z] implicitly show that no polynomial solution in three variables may be found with the same property, no matter the choice of the finite set S. It would be interesting to know whether four variables suffice for a similar example. We believe this is not the case, but have no proof. However, it may be shown that the polynomial solution obtained with four partial quotients indeed does not work.

Proofs. In what follows S will denote a set of primes as in the Theorem. We start by inverting the equation $\mathbf{f}(\mathbf{Y}) = \mathbf{X}$, where $\mathbf{Y} = (Y_0, \ldots, Y_4)$ (and $\mathbf{X}$ satisfies (1)). Let p_n/q_n be the convergents to the continued fraction $[Y_0, \ldots, Y_4]$, so $\mathbf{f}(\mathbf{Y}) = (p_3(\mathbf{Y}), q_4(\mathbf{Y}), p_4(\mathbf{Y}), q_3(\mathbf{Y}))$. From the well-known formulas $p_4 = Y_4p_3 + p_2$, $q_4 = Y_4q_3 + q_2$ we first find that $p_2 = X_3 - Y_4X_1$ and $q_2 = X_2 - Y_4X_4$. We use these formulas in $p_3 = Y_3p_2 + p_1$, $q_3 = Y_3q_2 + q_1$ to obtain $p_1 = (1+Y_3Y_4)X_1 - Y_3X_3$, $q_1 = (1+Y_3Y_4)X_4 - Y_3X_2$. But $q_1 = Y_1$, whence

$$Y_1 = (1 + Y_3Y_4)X_4 - Y_3X_2. \tag{2}$$

Also,

$$Y_0Y_1 = p_1 - 1, \qquad q_2 = Y_2q_1 + 1 = Y_2Y_1 + 1. \tag{3}$$

So, given a value $\mathbf{x}$ of $\mathbf{X}$, we may choose $Y_3 = y_3, Y_4 = y_4$ with the only restriction that the specialization y_1 of Y_1, as defined by (2), becomes nonzero. Then we may put, following (3) ([1]),

$$y_0 = \frac{p_1 - 1}{y_1} = \frac{(1 + y_3y_4)x_1 - y_3x_3 - 1}{y_1}, \qquad y_2 = \frac{q_2 - 1}{y_1} = \frac{x_2 - y_4x_4 - 1}{y_1}.$$

These calculations may be reversed: if $\mathbf{x}$ satisfies (1) and $\mathbf{y}$ is given by the above formulas, we shall find that $\mathbf{f}(\mathbf{y}) = \mathbf{x}$. Suppose now that $\mathbf{x} \in \mathcal{O}_S^4$. In

([1]) The formula for y_0 in [Z] is incorrect, but the error does not affect the arguments therein.

order that a vector $\mathbf{y}$ so obtained lies in $\mathcal{O}_S^5$ it will be sufficient that $y_3, y_4 \in \mathcal{O}_S$ and that the value for y_1 found from (2) lies in $\mathcal{O}_S^*$, the multiplicative group of S-units in $\mathbb{Q}$. In other words, to prove the Theorem it suffices to verify the following

Claim. *For all solutions* $\mathbf{x} \in \mathcal{O}_S^4$ *of* (1), *there exist* $y_3, y_4 \in \mathcal{O}_S$ *such that* $(1 + y_3y_4)x_4 - y_3x_2 \in \mathcal{O}_S^*$.

The strategy will be as follows. We need that $x_4 + y_3(-x_2 + y_4x_4) \in \mathcal{O}_S^*$. We then look for an S-integer Q of the form $-x_2 + y_4x_4$ such that the reduction of $\mathcal{O}_S^*$ modulo Q contains x_4. It seems that the simplest way to ensure this is to find Q such that the above-mentioned reduction contains every class coprime to Q.

A crucial point in this program will be the following lemma:

Fundamental Lemma. *Let* $q, r \in \mathbb{Z}$, $r \equiv 1 \pmod 4$, q *coprime to* r *and to all the primes in* S. *There exist prime numbers* $p_1, p_2 \notin S$ *with the following properties*:

(i) $p_1p_2 \equiv r \pmod{4q}$;
(ii) *the reduction of* $\mathcal{O}_S^*$ *modulo* p_1p_2 *equals the whole* $(\mathbb{Z}/(p_1p_2))^*$.

Of course the real restriction is represented by (ii). For this we shall mimic a method used first by Gupta and Ram Murty [G-RM] and later by Heath-Brown [HB], to deal with Artin's conjecture for primitive roots.

To construct the primes p_1, p_2 we shall appeal to a rather deep result from sieve theory, appearing as Lemma 1 in [HB]. We state here just the corollary we need:

Lemma 2. *Let* $u, v \in \mathbb{Z}$, $u \equiv 3 \pmod 4$, $(u, v) = ((u-1)/2, v) = 1$. *There exists* $\alpha > 1/4$ *with the following property*: *Let* $\mathcal{P}$ *be the set of prime numbers* p *such that* $p \equiv u \pmod v$ *and all prime factors of* $(p-1)/2$ *exceed* p^α. *Then the number of primes in* $\mathcal{P}$ *up to* X *is* $\gg X/\log^2 X$.

This result is the special case $K = 2$ of Lemma 1 in [HB], forgetting the further conclusion given there, about the number of prime factors of $p-1$. (The condition "$16 \mid v$" therein is immaterial here.)

We shall use this lemma similarly to the above-mentioned authors, to show the existence of many primes with a primitive root in $\mathcal{O}_S^*$. The next lemma is a first step in this direction; it shows that there are not many primes p such that the reduction of $\mathcal{O}_S^*$ modulo p is small.

Lemma 3. *Let* Σ *be a set of* 3 *distinct prime numbers and let* $0 < \delta < 1$. *Then the number of primes* $p \le X$ *such that the reduction of* $\mathcal{O}_\Sigma^*$ *modulo* p *has less than* p^δ *elements is* $\ll X^{4\delta/3}$.

Proof. Put $\Sigma = \{l_1, l_2, l_3\}$. For a positive integer L we consider the rational number

$$\varrho_L := \prod_{\mathbf{a}\in B_L} (l_1^{a_1} l_2^{a_2} l_3^{a_3} - 1),$$

where $\mathbf{a} = (a_1, a_2, a_3) \in \mathbb{Z}^3$ and where B_L denotes the cube $[-L, L]^3$ deprived of the origin. Plainly, ϱ_L is a nonzero rational number, whose denominator divides $(l_1 l_2 l_3)^{(2L)^4}$.

Let $p \le (L+1)^{3/\delta}$ be a prime as in the statement; then the order of $\mathcal{O}_\Sigma^*$ modulo p is $< p^\delta \le (L+1)^3$. Then the $(L+1)^3$ numbers $l_1^{a_1} l_2^{a_2} l_3^{a_3}$, for $0 \le a_i \le L$, cannot be pairwise incongruent modulo p. We deduce that p divides the numerator of ϱ_L. If $\varphi(X)$ denotes the number of such primes up to X we then find

$$2^{\varphi((L+1)^{3/\delta})} \le (l_1 l_2 l_3)^{(2L)^4} |\varrho_L|.$$

On the other hand, we easily see that $\log|\varrho_L| \ll L^4$, proving that

$$\varphi((L+1)^{3/\delta}) \ll L^4.$$

The conclusion follows at once by choosing L as the largest integer $\le X^{\delta/3}$.

Proof of the Fundamental Lemma. Recall that $(q, 6r) = 1$. There exists a positive integer a such that $a(a-1)(a-r)$ is coprime to q; in fact, by the Chinese Theorem, it suffices to argue assuming that q is a power of a prime > 3, in which case the result is clear. Let then b be an integer $\equiv 3 \pmod 4$ and $\equiv a \pmod q$. In particular, $(b, 4q) = 1$, so we may pick an integer $c \equiv 3 \pmod 4$ such that $bc \equiv r \pmod{4q}$.

We start by constructing p_1. We pick a quadratic nonresidue $\sigma \pmod l$ and we consider the arithmetic progression modulo $4ql$ defined by

$$x \equiv b \pmod{4q}, \qquad x \equiv \sigma \pmod l. \tag{4}$$

(This set is indeed not empty, since q is coprime to the elements of S by assumption.) Write this progression as $A + \mathbb{Z}(4ql)$. Then the GCD's $(A, 4ql) = 1$ and $(A-1, 4ql) = 2$, as follows by considering separately the moduli $4q, l$ (recall that $(b(b-1), 4q) = 2$ by construction).

We apply Lemma 2 with $u = A$, $v = 4ql$; we have just verified that our construction satisfies the assumptions of the lemma. We find that the set of prime numbers with those properties contains $\gg X/\log^2 X$ elements up to X.

We further appeal to Lemma 3, with $\delta = 1 - \alpha$ and $\Sigma = S$. We deduce that the number of primes $p \le X$, such that the reduction of $\mathcal{O}_S^*$ modulo p has order $< p^{1-\alpha}$, is $\ll X^{(1-\alpha)4/3}$.

Therefore, since $1 - \alpha < 3/4$, throwing away these primes from the set provided by Lemma 2, we are left with an infinite set. Let p_1 be a large prime in this set; thus we may assume that p_1 is not in S, has the properties of Lemma 2 (with $u = A$, $v = 4ql$) and the reduction of $\mathcal{O}_S^*$ modulo p_1 contains at least $p_1^{1-\alpha}$ elements. On the other hand, this reduction has order dividing p_1-1, so it is of the form $(p_1-1)/t$, where t is a divisor of p_1-1. Necessarily we must have $t = 1$ or $t = 2$, since $p_1 \equiv 3 \pmod 4$ and since every prime other than 2 dividing $p_1 - 1$ is $\geq p_1^{\alpha}$. But $t = 2$ is impossible, since $\mathcal{O}_S^*$ contains a quadratic nonresidue of p_1 (e.g. -1, or even l). Therefore the reduction of $\mathcal{O}_S^*$ modulo p_1 equals $\mathbb{F}_{p_1}^*$.

To construct p_2 we argue similarly; we now consider the progression defined by

$$x \equiv c \pmod{4q}, \qquad x \equiv -1 \pmod{l}.$$

Writing this progression as $B+\mathbb{Z}(4ql)$, we contend that we may apply Lemma 2 with $u = B$, $v = 4ql$. In fact, $(u, v) = 1$ and also $((u-1)/2, v) = 1$. This is because $u - 1 \equiv c - 1 \equiv (r - b)b^{-1} \pmod{4q}$ and $(r - b)$ is coprime to q; also, $u \equiv c \equiv 3 \pmod 4$. Now, as before, with the aid of Lemma 3 we may find a (large) prime p_2 in the progression such that the reduction of $\mathcal{O}_S^*$ modulo p_2 equals the whole $\mathbb{F}_{p_2}^*$.

Further, by choosing $p_2 > p_1^4$, we may also assume that $(p_1-1, p_2-1) = 2$; in fact, each factor of $p_2 - 1$ larger than 2 is automatically $> p_2^{1/4} > p_1$, by Lemma 2.

Note that $p_1p_2 \equiv bc \equiv r \pmod{4q}$, so (i) of the Fundamental Lemma is verified.

As to (ii), it is "almost" verified, since the reduction of $\mathcal{O}_S^*$ is as big as possible modulo both p_1 and p_2. We show that it is in fact as big as possible modulo p_1p_2.

Let $G \subset (\mathbb{Z}/(p_1p_2))^*$ be the reduction of $\mathcal{O}_S^*$ modulo p_1p_2. There is a homomorphism $\lambda : G \to \{\pm1\}^2$ given by $g \mapsto ((g|p_1), (g|p_2))$ (Legendre symbols).

We have $(p_1|l) = (\sigma|l) = -1$ by construction, whence by quadratic reciprocity (recall $l \equiv 1 \pmod 4$) we also have $(l|p_1) = -1$. Similarly, $(l|p_2) = (p_2|l) = (-1|l) = 1$, so $\lambda(l) = (-1, 1)$. But $p_1 \equiv p_2 \equiv 3 \pmod 4$, so $\lambda(-1) = (-1,-1)$. Therefore λ is surjective, whence the order of G is divisible by 4. But our construction proved that the reduction of G modulo p_i contains $\mathbb{F}_{p_i}^*$ for $i = 1, 2$, so the order of G is divisible by both $p_1 - 1$ and $p_2 - 1$. Since $(p_1 - 1, p_2 - 1) = 2$, we deduce that the order of G is divisible by $(p_1 - 1)(p_2 - 1) = \varphi(p_1p_2)$, concluding the proof of the Fundamental Lemma.

Now we may easily prove the Claim (and hence the Theorem) as follows.

First, if $x_2 = 0$, equation (1) implies that $x_4 \in \mathcal{O}_S^*$ and we may just choose $y_4 = 0$. Similarly if $x_4 = 0$; therefore, we suppose $x_2x_4 \neq 0$.

For $i = 2, 4$, write $x_i = 2^{a_i}(\varDelta_i/\varDelta)z_i$ where $a_i \in \mathbb{Z}$, $\varDelta_i, \varDelta$ are odd integers in $\mathcal{O}_S^* \cap \mathbb{Z}$ and where z_i are positive integers coprime to every prime in S. This is plainly possible. Moreover, by multiplying $\varDelta_2, \varDelta_4$ and $\varDelta$ by 3 if necessary, we may assume that $\varDelta_2 z_2 \equiv 3 \pmod 4$.

Since $\varDelta_2$ is divisible only by primes in S, it is coprime to z_4. We also have $(z_2, z_4) = 1$, because of equation (1) and the fact that no prime dividing z_2 is in S. Hence we may apply the Fundamental Lemma with $r = -\varDelta_2 z_2$, $q = z_4$, obtaining the existence of primes p_1, p_2 satisfying (i) and (ii) of that lemma. By (i) we may write, for some $m \in \mathbb{Z}$,

$$p_1 p_2 = -\varDelta_2 z_2 + 4mz_4. \tag{5}$$

Now note that, again, $\varDelta_4$ is divisible only by primes in S, while $p_1, p_2 \notin S$. Also, z_4 cannot be divisible by p_1 or p_2, for otherwise, by (5), z_4 would not be coprime to $\varDelta_2 z_2$.

Hence, since by (ii) the reduction of $\mathcal{O}_S^*$ modulo $p_1 p_2$ contains every invertible class, there exists $u \in \mathcal{O}_S^*$ such that $u \equiv 2^{a_4}\varDelta_4 z_4 \pmod{p_1 p_2}$ (the congruence holding in $\mathcal{O}_S$). In other words, we may write

$$u = 2^{a_4}\varDelta_4 z_4 + tp_1 p_2 = 2^{a_4}\varDelta_4 z_4 + t(-\varDelta_2 z_2 + 4mz_4), \tag{6}$$

where $t \in \mathcal{O}_S$. Dividing (6) by $\varDelta$ we obtain

$$\frac{u}{\varDelta} = x_4 + t\left(-x_2 2^{-a_2} + 4\,\frac{m}{2^{a_4}\varDelta_4}\,x_4\right). \tag{7}$$

Now put $y_3 = t2^{-a_2} \in \mathcal{O}_S$, $y_4 = 2^{2+a_2-a_4}\frac{m}{\varDelta_4} \in \mathcal{O}_S$. Equation (7) reads

$$\frac{u}{\varDelta} = x_4 + y_3(-x_2 + y_4 x_4) = (1 + y_3 y_4)x_4 - y_3 x_2.$$

Since the left side is in $\mathcal{O}_S^*$, we have the conclusion of the Claim.

Remark. If one has an explicit version of Lemma 2 at one's disposal, it becomes possible to quantify the Theorem; namely, given a solution $\mathbf{x} \in \mathcal{O}_S^4$ of (1), to estimate the height of a suitable vector $\mathbf{y}$ as in the conclusion.

References

[G-RM] R. Gupta and M. Ram Murty, *A remark on Artin's conjecture*, Invent. Math. 78 (1984), 127–130.

[HB] D. R. Heath-Brown, *Artin's conjecture for primitive roots*, Quart. J. Math. Oxford 37 (1986), 27–38.

[S] T. Skolem, *Diophantische Gleichungen*, Springer, 1938; reprinted Chelsea, 1950.

[Sz] K. Szymiczek, *On some diophantine equations connected with triangular numbers*, Zeszyty Nauk. Wyż. Szkoły Ped. w Katowicach Mat. 4 (1964), 17–22 (in Polish).

[Z] U. Zannier, *Remarks on a question of Skolem about the integer solutions of* $x_1x_2 - x_3x_4 = 1$, Acta Arith. 78 (1996), 153–164.

Ist. Univ. Arch. Venezia – D.C.A.
S. Croce, 191
30135 Venezia, Italy
E-mail: zannier@iuav.it

Received on 19.9.2002 (4373)