ACTA ARITHMETICA
106.2 (2003)

On polynomials taking small values at integral arguments II

by

Roberto Dvornicich (Pisa), Shih Ping Tung (Chung Li)
and Umberto Zannier (Venezia)

1. Introduction. In a recent paper, S. P. Tung [T] considers the problem of estimating from below the quantity

$$S_F(T) := \max_{\substack{x\in\mathbb{N}\\ x\le T}} \min_{y\in\mathbb{Z}} |F(x,y)|,$$

where $F \in \mathbb{Q}[X,Y]$ is a given polynomial and $T \in \mathbb{N}$ is a variable growing to infinity. For a fixed integer x_0, the quantity $\min_{y\in\mathbb{Z}} |F(x_0,y)|$ (which was investigated already in [DZ]) gives a measure of the distance of the roots of $F(x_0,Y)=0$ from the integers; the function $S_F(T)$ expresses the behaviour of this distance as the first variable grows.

Actually, $S_F(T)$ implicitly appears in the statement of Hilbert's Irreducibility Theorem; in fact most proofs of it (see e.g. [S]) reduce to showing the following: *If for every integer x_0 the equation $F(x_0,Y)=0$ has an integral solution y, then there exists a polynomial $f(X)\in\mathbb{Q}[X]$ such that $F(X,f(X))=0$ identically.* Note that the assumption of this statement may be reformulated as $S_F(T)=0$ *for all positive T*. Hence, Hilbert's theorem proves that either $F(X,f(X))=0$ for some polynomial $f\in\mathbb{Q}[X]$ or we have a lower bound $S_F(T)\ge c>0$ for all large T.

Note that it may happen that $S_F(T)$ is bounded, e.g. when there exists a polynomial $f(X)\in\mathbb{Q}[X]$, taking integral values on $\mathbb{Z}$, such that $F(X,f(X))$ is a constant. However, Tung proves, among other things, that this is essentially the only case when $S_F(T)$ is bounded. In fact, Tung has a much sharper conclusion. To state it, we first define, for an infinite set $\mathcal{A}\subset\mathbb{N}$, the symbol

$$\mathcal{A}(T) = \mathcal{A}\cap[1,T],$$

and the function

2000 *Mathematics Subject Classification*: Primary 12E05.

$$S_{\mathcal{A},F}(T) = \max_{x\in\mathcal{A}(T)} \min_{y\in\mathbb{Z}} |F(x,y)|.$$

Also, we recall the classical definitions of upper and lower asymptotic densities:

$$\bar{d}(\mathcal{A}) = \limsup_{T\to\infty} \frac{\#\mathcal{A}(T)}{T}, \qquad \underline{d}(\mathcal{A}) = \liminf_{T\to\infty} \frac{\#\mathcal{A}(T)}{T}.$$

When these numbers coincide, their common value is called the *asymptotic density* of $\mathcal{A}$. With this notation, Tung proves [T, Thm. 3.4] the following statement: *There exists a number* $c > 0$, *depending only on* $\deg F$, *with the following property*: *Either there exists a polynomial* $f(X) \in \mathbb{Q}[X]$ *such that* $F(X, f(X))$ *is constant*, *or*, *for all sets* $\mathcal{A}$ *of positive density*, *we have* $S_{\mathcal{A},F}(T) \gg T^c$. (Here the implied constant may depend both on $\mathcal{A}$ and on F.)

In this statement no attention is given to whether or not the polynomial f is integral-valued on $\mathbb{N}$; Tung studies this condition later on in the abovementioned paper (see also Remark (ii) below). Here we are concerned with a question in a different direction: *how large can one choose the exponent* c *in the above statement?*

Although Tung's method yields in principle an effective estimate for c, he does not mention any explicit lower bound. However, he points out that c cannot exceed $1/2$, in view of the data $F(X,Y) = Y^2 - X$, $\mathcal{A} = \mathbb{N}$. Moreover, under the Generalized Riemann Hypothesis, he obtains the inequality $S_{\mathcal{A},F}(T) \gg \sqrt{T}/\log^2 T$, proving in particular that one can choose $c = 1/2 - \varepsilon$ for any $\varepsilon > 0$.

The purpose of the present note is to show, unconditionally, that in fact one can take $c = 1/2$. We state this as the following

THEOREM 1. *Let* $\mathcal{A} \subset \mathbb{N}$ *be a set of positive lower asymptotic density and let* $F(X,Y) \in \mathbb{Q}[X,Y]$. *Then either there exists* $f(X) \in \mathbb{Q}[X]$ *such that* $F(X, f(X))$ *is constant*, *or* $S_{\mathcal{A},F}(T) \gg \sqrt{T}$ *for* $T \to \infty$.

We shall deduce Theorem 1 from a similar statement, namely

THEOREM 2. *Let* $F(X,Y) \in \mathbb{Q}[X,Y]$. *If* $\mathcal{A}$ *is a set of positive upper asymptotic density and* $y(a)$, $a \in \mathcal{A}$, *are integers such that* $|F(a, y(a))| = o(\sqrt{a})$, *then there exists a polynomial* $f(X) \in \mathbb{Q}[X]$ *such that* $F(X, f(X))$ *is constant.*

We remark that e.g. in the case $\mathcal{A} = \mathbb{N}$ the implict constants are effectively computable.

Our method, of completely different nature compared to [T], will make essential use of the previous paper [DZ]. We shall not use Hilbert's theorem (a proof of which is implicitly given in [DZ]) nor other classical diophantine tools.

2. Proofs. For the reader's convenience, we recall the main result of [DZ]:

THEOREM DZ. *Let $F(X,Y) \in \mathbb{R}[X,Y]$. Assume that $\mathcal{A}$ is a set of natural numbers of positive upper density, such that for $a \in \mathcal{A}$ we may find an integer $y(a)$ satisfying*

$$|F(a,y(a))| = o\Big(\sup_{|\xi - y(a)|\le 1}\Big|\frac{\partial F}{\partial Y}(a,\xi)\Big|\Big). \tag{1}$$

Then there exist a polynomial $f \in \mathbb{Q}[X]$ and a set $\mathcal{B} \subset \mathcal{A}$ such that $\mathcal{A}\setminus\mathcal{B}$ has zero density and

$$|F(b,f(b))| \le |F(b,y(b))| \quad \forall b \in \mathcal{B}. \tag{2}$$

Proof of Theorem 2. Assume, as in the statement, that $\mathcal{A} \subset \mathbb{N}$ is an infinite set of positive upper density, such that $|F(a,y(a))| = o(\sqrt{a})$ for $a \in \mathcal{A}$. We start by writing

$$F(X,Y) = \varphi_0(X) + \varphi_1(X)Y + \ldots + \varphi_d(X)Y^d, \quad \varphi_i \in \mathbb{Q}[X],\ \varphi_d(X) \ne 0.$$

We note at once that, if $d = 0$, then the assumption implies that $|\varphi_0(a)| = o(a)$, whence F is constant, and there is nothing to prove. Hence we assume $d \ge 1$.

Suppose that the leading coefficient $\varphi_d(X)$ in Y of $F(X,Y)$ is constant. Then we normalize F as follows. First we choose $h(X) \in \mathbb{Q}[X]$ such that the second coefficient in Y of $F(X, Y + h(X))$ vanishes, i.e. $h(X) = -\varphi_{d-1}(X)/d\varphi_d$. Next, if r is a common denominator for the coefficients of $h(X)$, we replace $F(X,Y)$ with $F(X, Y/r + h(X))$. We note that this polynomial continues to satisfy the assumptions of Theorem 2: we leave the set $\mathcal{A}$ unchanged, while the function $y(a)$ is replaced by $r(y(a) - h(a))$. Moreover, the conclusion of Theorem 2 for the new polynomial implies the same conclusion for the old one.

Summing up, we may assume that either $\varphi_d(X)$ is not constant or $\varphi_{d-1}(X) = 0$.

Before going on, we recall the following simple fact.

LEMMA. *Let $P(Y) \in \mathbb{C}[Y]$. Then*

$$\sup_{0\le y\le 1} |P(y)| \ge c \sum_{j=0}^{\deg P} |P^{(j)}(0)|,$$

where c is a positive number depending only on $\deg P$.

Proof. Write the Taylor expansion

$$P(Y) = P(0) + P'(0)Y + \ldots + \frac{P^{(k)}(0)}{k!}\,Y^k,$$

where $k=\deg P$. Since the Vandermonde determinant $\det((i/k)^j)_{0\le i,j\le k}$ is nonzero, the formulas

$$P\left(\frac{i}{k}\right)=P(0)+P'(0)\left(\frac{i}{k}\right)+\ldots+\frac{P^{(k)}(0)}{k!}\left(\frac{i}{k}\right)^k,\qquad i=0,\ldots,k,$$

imply that the numbers $P^{(j)}(0)$ may be expressed as linear forms in $P(0)$, $P(1/k),\ldots,P(1)$ with coefficients depending only on k. If C is the maximum of the absolute values of these coefficients, we have

$$\begin{aligned}\sum_{j=0}^{k}|P^{(j)}(0)|&\le(k+1)\sup_j|P^{(j)}(0)|\\&\le(k+1)C\left(P(0)+P\left(\frac{1}{k}\right)+\ldots+P(1)\right)\\&\le(k+1)^2C\sup_{0\le y\le 1}|P(y)|.\ \blacksquare\end{aligned}$$

We now put $G(X,Y)=\frac{\partial}{\partial Y}F(X,Y)$ and, for $a\in\mathcal{A}$,

$$\sigma(a):=\sup_{|\xi-y(a)|\le 1}|G(a,\xi)|.$$

Our next aim is to show that either the conclusion of Theorem 2 is true or

$$\sigma(a)\gg\sqrt{a}\quad\text{ for large } a\in\mathcal{A}.\tag{3}$$

By applying the Lemma to the polynomial $P(Y):=G(a,y(a)+Y)$ we find that

$$\sigma(a)\ge c_1\sum_{j\ge 0}|G^{(j)}(a,y(a))|\tag{4}$$

where $G^{(j)}$ denotes the jth derivative with respect to Y and $c_1>0$ depends only on d.

In the preceding notation we have $G(X,Y)=\varphi_1(X)+2\varphi_2(X)Y+\ldots+d\varphi_d(X)Y^{d-1}$.

In what follows, $c_2,c_3,\ldots$ will denote positive numbers depending only on F. We distinguish two cases.

Case 1: *There exists $i\in\{1,\ldots,d\}$ such that $\varphi_i(X)$ has positive degree* (*i.e.* $\deg_X G>0$). In this case, let q be the maximum index i with this property. If $q=d$, then there exists a positive number c_2 such that $\sigma(a)\ge c_2|a|$ for all large $a\in\mathcal{A}$: in fact, by (4), we have $\sigma(a)\ge c_1d!|\varphi_d(a)|$, and (3) follows.

If $q<d$, then $q<d-1$ in view of the opening normalization. Observe that $G^{(q)}(X,Y)$ is a polynomial in Y alone, of degree $d-q-1>0$, whence $|G^{(q)}(a,y(a))|\ge c_2|y(a)|^{d-q-1}-c_3$. In view of (4) we obtain $\sigma(a)\ge$

$c_1c_2|y(a)|^{d-q-1} - c_1c_3$, whence

$$|y(a)| \leq c_4(\sigma(a)+1)^{1/(d-q-1)}. \tag{5}$$

Further,

$$G^{(q-1)}(X,Y) = q!\varphi_q(X) + \sum_{i=1}^{d-q} \frac{(i+q)!}{i!}\,\varphi_{i+q}Y^i.$$

In particular,

$$|G^{(q-1)}(a,y(a))| \geq |q!\varphi_q(a)| - c_5|y(a)|^{d-q}.$$

Since $\varphi_q(X)$ is not constant by assumption, (4) and (5) imply that, for large $a \in \mathcal{A}$,

$$\sigma(a) \geq c_6|a| - c_7(\sigma(a)+1)^{(d-q)/(d-q-1)}.$$

Since $(d-q)/(d-q-1) \leq 2$ for $q < d-1$ we again deduce (3).

Case 2: *$G(X,Y)$ does not depend on X.* In this case we can assume that $F(X,Y) = \varphi_0(X) + \psi(Y)$ for a polynomial $\psi \in \mathbb{Q}[Y]$, so $G(X,Y) = \psi'(Y)$. If $\varphi_0(X)$ is constant, Theorem 2 follows immediately by letting $f(X)$ be any constant polynomial. Similarly if $d = 1$. Also, the case $d = 0$ was previously excluded, and therefore we assume $\deg_X F > 0$ and $d = \deg_Y F > 1$.

By (4) we have $\sigma(a) \geq c_1|\psi'(y(a))| \geq c_8|y(a)|^{d-1} - c_9$, whence

$$|y(a)| \leq c_{10}(\sigma(a)+1)^{1/(d-1)}.$$

Moreover, since φ_0 is not constant, we have

$$|F(a,y(a))| \geq c_{11}|a| - c_{12}(|y(a)|+1)^d \geq c_{11}|a| - c_{13}(\sigma(a)+1)^{d/(d-1)}.$$

On the other hand we have $|F(a,y(a))| = o(\sqrt{a})$ by assumption, whence

$$c_{14}\sqrt{a} \geq c_{11}|a| - c_{13}(\sigma(a)+1)^{d/(d-1)}.$$

Now, as before, (3) follows by noting that $d/(d-1) \leq 2$ for $d > 1$.

By combining (3) with the assumption $|F(a,y(a))| = o(\sqrt{a})$ for $a \in \mathcal{A}$, we find that

$$|F(a,y(a))| = o\Big(\sup_{|\xi-y(a)|\leq 1}\Big|\frac{\partial}{\partial Y}F(a,\xi)\Big|\Big) \quad \text{for } a \in \mathcal{A}.$$

Since the set $\mathcal{A}$ is assumed to be of positive upper density, Theorem DZ then implies the existence of a polynomial f with rational coefficients and of a set $\mathcal{B} \subset \mathcal{A}$ with the same upper density as $\mathcal{A}$, such that $|F(b,f(b))| \leq |F(b,y(b))| = o(\sqrt{b})$ for $b \in \mathcal{B}$. Since $\mathcal{B}$ is infinite, it follows that $F(X,f(X))$ must be constant, concluding the proof of Theorem 2. ∎

Proof of Theorem 1. We let $\mathcal{A}$ be a set as in the statement. In view of the definition of lower density, there exists a positive number c such that $\#\mathcal{A}(T) > cT$ for all $T > T_0$, say.

We shall prove the existence of a polynomial $f(X)$ with the stated property, under the assumption that $S_{\mathcal{A},F}(T) \gg \sqrt{T}$ does not hold true. This means that there exist positive integers $T_1 < T_2 < \ldots$ such that $S_{\mathcal{A},F}(T_n) \le (1/n)\sqrt{T_n}$ for all positive integers n. We may also assume that $T_1 > T_0$.

For $a \in \mathbb{N}$ we define $g(a) := \min_{y\in\mathbb{Z}} |F(a,y)|$. The numbers $|F(a,y)|$ for $a, y \in \mathbb{Z}$ are nonnegative rational numbers with bounded denominators, so the minimum is attained for every $a \in \mathbb{N}$ and we may write $g(a) = |F(a, y(a))|$ for a suitable rational integer $y(a)$.

In view of our definitions we have

$$\max_{a\in\mathcal{A}(T_n)} g(a) \le \frac{1}{n}\sqrt{T_n}.$$

Define the set $\mathcal{A}'$ to be the union of the sets $\mathcal{A}\cap[(c/2)T_n, T_n]$, over all positive integers n. Since $\#\mathcal{A}(T_n) \ge cT_n$ for all $n \in \mathbb{N}$, the interval $[(c/2)T_n, T_n]$ contains at least $(c/2)T_n$ elements of $\mathcal{A}$, so $\mathcal{A}'$ has positive upper density.

We contend that $g(a) = o(\sqrt{a})$ for $a \in \mathcal{A}'$. In fact, if $a \in \mathcal{A}'$ then a lies in some interval $[(c/2)T_n, T_n]$, and $a \in \mathcal{A}$. Therefore

$$g(a) \le \max_{x\in\mathcal{A}(T_n)} g(x) \le \frac{1}{n}\sqrt{T_n} \le \frac{1}{n}\sqrt{2a/c},$$

since $a \ge (c/2)T_n$. This proves our contention.

Finally, recalling that $g(a) = F(a, y(a))$ for $a \in \mathcal{A}'$, we may apply Theorem 2 to get the desired conclusion. ∎

Remarks. (i) We observe that it is not possible to replace *lower density* with *upper density* in the statement of Theorem 1. It suffices to take $\mathcal{A}$ to be any set containing large intervals of integers and then large gaps, to produce a counterexample. Take, e.g., $\mathcal{A}$ to be the union of the intervals $[2^{n!}, 2^{2\cdot n!}]$ for $n \in \mathbb{N}$; this set clearly has upper density equal to 1 and lower density equal to 0. Also, let $F(X,Y)$ be any polynomial. Plainly, for any positive integer a, we have $\min_{y\in\mathbb{Z}} |F(a,y)| \le |F(a,0)| \ll a^h$ (where $h = \deg_X F$). Then, if $T = 2^{(n+1)!} - 1$, we have $\max_{a\in\mathcal{A}(T)} \min_{y\in\mathbb{Z}} |F(a,y)| \ll \max_{a\le 2^{2\cdot n!}} a^h \ll 2^{2hn!} \ll T^{2h/n}$. Hence a lower bound $S_{\mathcal{A},F}(T) \gg T^c$ does not hold, no matter the value of $c > 0$. On the other hand, for suitable F, e.g. $F(X,Y) = Y^2 - X$ there does not exist a polynomial $f(X)$ such that $F(X, f(X))$ is constant.

(ii) By using the full force of the proof of Theorem DZ (as given in [DZ]), both Theorem 1 and Theorem 2 can be sharpened: one can add to the first alternative of the conclusion of Theorem 1 and to the conclusion of Theorem 2 that f is integral-valued on a sequence $\mathcal{B}$ with $\mathcal{A} \setminus \mathcal{B}$ of zero density.

References

[DZ] R. Dvornicich and U. Zannier, *On polynomials taking small values at integral arguments*, Acta Arith. 42 (1983), 189–196.

[S] A. Schinzel, *Polynomials with Special Regard to Reducibility*, Encyclopedia Appl. Math. 77, Cambridge Univ. Press, 2000.

[T] S. P. Tung, *Max-min of polynomials and exponential Diophantine equations*, to appear.

Dipartimento di Matematica
Via Buonarroti, 2
56127 Pisa, Italy
E-mail: dvornic@dm.unipi.it

Department of Mathematics
Chung Yuan Christian University
Chung Li, Taiwan 32023, R.O.C.
E-mail: sptung@cycu.edu.tw

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

Received on 17.5.2001
and in revised form on 21.6.2002 (4032)