International Journal of Number Theory
(2014)

DOI: 10.1142/S1793042114500535

Counting squarefree values of polynomials with error term

M. Ram Murty* and Hector Pasten†
Department of Mathematics and Statistics
Queen's University Jeffery Hall
University Ave., Kingston, ON, Canada, K7L 3N6
**murty@mast.queensu.ca*
†*hpasten@gmail.com*

Received 22 August 2013
Accepted 22 February 2014
Published

It was shown by Granville that the ABC conjecture allows one to prove asymptotic estimates on the number of squarefree values of polynomials. However, his proof gives no information on the error term of the asymptotic formula. On the ABC conjecture, we prove an asymptotic formula with error term using a different technique. From the ABC conjecture we also deduce an asymptotic formula with error term for the number of squarefree values of polynomials on certain sets of integers that are residually well distributed in a suitable sense.

Keywords: Squarefree values; error term; polynomials; ABC conjecture; Belyi's theorem.

Mathematics Subject Classification 2010: 11N32; 11G32, 11N36, 11J71

1. Introduction and Results

Let $r \geq 2$ and let $f \in \mathbb{Z}[X]$ be a polynomial of degree r. Define

$$N_f(x) = \#\{n \leq x : f(n) \text{ is squarefree}\}$$

and write $G_f := \gcd(f(n) : n \geq 1)$. Define $\omega_f(n)$ to be the number of solutions of the congruence $f(x) \equiv 0 \bmod n$; this is a multiplicative function on n. If f has some repeated factor or if G_f is not squarefree, then trivially $N_f(x)$ is bounded, so we will assume that f has no repeated factors and G_f is squarefree.

On the ABC conjecture, Granville [2] showed the asymptotic formula $N_f(x) \sim c_f x$ for certain explicit constant c_f, although it is not clear how to get an error term from his technique. Lee and Murty [7] provided such an error term under the ABC conjecture and the so-called *abscissa conjecture*. Due to the strong evidence and heuristics supporting the ABC conjecture, it is desirable to obtain an error term

assuming the ABC conjecture without the use of the abscissa conjecture. We prove the following.

Theorem 1.1. *Assume the ABC conjecture. Let f be a polynomial with integer coefficients, of degree $r \geq 2$, without repeated factors, and with G_f squarefree. Then*

$$N_f(x) = c_f x + O_f\left(\frac{x}{(\log x)^\gamma}\right),$$

where $\gamma > 0$ is a computable constant that only depends on r (not on the particular f), and

$$c_f = \prod_p \left(1 - \frac{\omega_f(p^2)}{p^2}\right) > 0.$$

The constant c_f is (on the ABC conjecture) the probability of $f(n)$ being squarefree, while the factor $1 - \omega_f(p^2)/p^2$ can be seen as the probability that p^2 does not divide $f(n)$ as we vary n. Thus, the main term basically says that there is a sort of local–global principle in the problem of counting squarefree values of f. This observation about the main term is already made in [2].

As in [2], one can suitably normalize f (provided that it has no repeated factor) in order to get non-trivial counting of squarefree values of f even when G_f is not squarefree; our method can be modified to obtain a result in that case too.

The proof of Theorem 1.1 uses the ABC conjecture in a way which is different to previous applications in the problem of counting squarefree values of polynomials. For this, we will establish the following result, which is of independent interest.

Theorem 1.2. *Assume the ABC conjecture. Given $\epsilon > 0$ and a positive integer r, there is a constant K_ϵ depending only on ϵ and computable constants α, β depending only on r, such that for all polynomials $F \in \mathbb{Z}[X]$ of degree r without repeated factors and for all integers n one has*

$$|n|^{r-1-\epsilon} < K_\epsilon \exp(\alpha H(F)^\beta) \max\{1, \mathrm{rad}(F(n))\}.$$

Here, $H(F)$ is the height of $F \in \mathbb{Z}[X]$, which is defined as the maximum of the absolute value of the coefficients of F, and $\mathrm{rad}(N)$ is the product of the primes dividing N when N is a non-zero integer (and we set $\mathrm{rad}(0) = 0$). This result is an explicit version of the classical result of Langevin [6] that gives $n^{r-1-\epsilon} \ll_{\epsilon,F} \mathrm{rad}(F(n))$ for F without repeated factors, on the ABC conjecture.

Finally, we mention that the technique of this paper actually allows one to count (with error term) squarefree values of polynomials on sets of integers of positive density that are residually well distributed in a suitable sense (see Sec. 7). For instance, we have the following theorem.

Theorem 1.3. *Assume the ABC conjecture. Let f be a polynomial with integer coefficients, of degree $r \geq 2$, without repeated factors, and with G_f squarefree. Let*

March 18, 2014 20:19 WSPC/S1793-0421 203-IJNT 1450053

$\alpha > 1$ be an irrational real number with finite approximation exponent (for example $\alpha = \sqrt{2}$ or $\alpha = \pi$ are allowed). Consider the set of positive integers

$$A = \{\lfloor k\alpha \rfloor : k \in \mathbb{Z}^+\}.$$

Let $N_f^A(x)$ be the number of integers $n \leq x$ such that $f(n)$ is squarefree and $n \in A$. Then

$$N_f^A(x) = \frac{c_f}{\alpha} x + O\left(\frac{x}{(\log x)^\gamma}\right),$$

where $\gamma > 0$ and $c_f > 0$ are as in Theorem 1.1.

This result is proved in Sec. 8, where the notion of "finite approximation exponent" is recalled. Note that the constant c_f/α can be seen as a product of probabilities; as commented before, c_f is a product of local probabilities for $f(n)$ to be squarefree, while $1/\alpha$ is the probability that the argument n belongs to A.

A more general result is given in Sec. 7, from which Theorem 1.3 is deduced (in Sec. 8) by means of the theory of uniform distribution of sequences modulo 1. Since we care about the error term, it will be crucial to have control on the discrepancy of uniformly distributed sequences.

2. Heights

In this section we recall several height estimates that we will later need in our computations.

For $f \in \mathbb{Q}(X)$ we define its height $H(f)$ as follows: up to sign, there are unique $u, v \in \mathbb{Z}[X]$ coprime such that $f = u/v$. Then we define $H(f)$ as the maximal absolute value among the coefficients of u and v. From the definition, one has $H(u), H(v) \leq H(f)$. Also, note that if $f \in \mathbb{Q}[X]$ is a polynomial then v is the least common denominator of the coefficients of f and $H(f)$ is the usual affine height of f, that is, the affine height of the tuple given by the coefficients of f.

We need to recall the notion of height of an algebraic number. Let α be an algebraic number of degree d, and let F be the minimal polynomial of α over $\mathbb{Q}$, normalized so that it has coprime integer coefficients. We define the (absolute multiplicative) height of α as $H(\alpha) = H(F)^{1/d}$. Note that this is not the same as the Weil height defined in terms of valuations, but it is much simpler to define and both heights agree up to a factor bounded in terms of d (cf. [5, Proposition 4, p. 49]). For instance, if ζ is a primitive 105th root of unity then $H(\zeta) = 2^{1/48}$ although the Weil height of any root of unity is 1.

For the next result, see [5, Proposition 3, p. 48].

Proposition 2.1. *Let $f, g \in \mathbb{Q}[X]$ be non-zero polynomials with $\deg f + \deg g < d$. Then*

$$\frac{1}{4^d} H(fg) \leq H(f)H(g) \leq 4^d H(fg).$$

The height of a polynomial admits the following local decomposition.

Proposition 2.2. *Let $f = c_a X^a + \cdots + c_0 \in \mathbb{Q}[X]$. Let $M_{\mathbb{Q}} = \{\infty, 2, 3, 5, 7, \ldots\}$ be the set of places of $\mathbb{Q}$ and for each $w \in M_{\mathbb{Q}}$ let $|-|_w$ be the normalized absolute value. We have*

$$H(f) = \prod_{w \in M_{\mathbb{Q}}} \max\{1, |c_a|_w, \ldots, |c_0|_w\}.$$

Using this local decomposition we can prove the following proposition.

Proposition 2.3. *Let $f, g \in \mathbb{Q}[X]$ be polynomials of degree $a, b \geq 1$ respectively. Then*

$$H(f \circ g) \leq (a+1)(b+1)^a H(f) H(g)^a.$$

Proof. Write $f = u_a X^a + \cdots + u_0$ and $g = v_b X^b + \cdots + v_0$. Expanding

$$f \circ g = u_a (v_b X^b + \cdots + v_0)^a + \cdots + u_0$$

we see that the coefficients of $f \circ g$ have (Archimedean) absolute value bounded by

$$(a+1) \max\{|u_i|\} (b+1)^a \max\{|v_i|\}^a.$$

Similarly, for each prime p we find that the coefficients of $f \circ g$ have p-adic absolute value bounded by

$$\max\{|u_i|_p\} \max\{|v_i|_p\}^a.$$

The result follows from the local decomposition of the height. □

We will also need a bound for the resultant of two polynomials.

Proposition 2.4. *Let $f, g \in \mathbb{Z}[X]$ be coprime polynomials of degrees $a, b \geq 1$ and height $\leq H$. Let $R \in \mathbb{Z}$ be the resultant of f and g. Then*

$$|R| \leq (a+b)^{a+b} H^{a+b}.$$

Proof. Let $M = [m_{i,j}]_{i,j}$ be the Sylvester matrix of f and g, which is of size $(a+b) \times (a+b)$. Expanding $\det M$ we find

$$|R| = |\det M| \leq \sum_{\sigma \in S_{a+b}} \left| \prod_{i=1}^{a+b} m_{i,\sigma(i)} \right| \leq (a+b)^{a+b} H^{a+b}.$$ □

Finally, we state another useful bound (see [3, p. 237]).

Proposition 2.5. *Let $f \in \mathbb{Q}[X]$ and $d \in \mathbb{Q}$. Then*

$$H(f(X+d)) \leq 4^{\deg f} H(f(X)) H(d)^{\deg f}.$$

March 18, 2014 20:19 WSPC/S1793-0421 203-IJNT 1450053

3. Belyi Maps

Let S be a finite set of m algebraic numbers of degree at most r and absolute multiplicative height bounded by B. A well-known theorem of Belyi shows that there is a rational function $\phi \in \mathbb{Q}(X)$ such that ϕ takes all its *ramification points* in $\mathbb{P}^1$ and all the *elements of* S to $\{0, 1, \infty\}$. Moreover, one can find such a function ϕ that also takes the *point at infinity* to $\{0, 1, \infty\}$.

The construction of ϕ is very explicit and it is clear that one should be able to bound the height of ϕ in terms of m, r and B. Keeping track of the heights during the construction, one concludes the following.

Proposition 3.1. *There are computable constants A_1, A_2, A_3, A_4 depending only on the numbers r, m but not on B or the particular set S, such that one can find a rational function $\phi \in \mathbb{Q}(X)$ with*

$$\deg \phi < A_1 B^{A_2}, \quad H(\phi) < \exp(A_3 B^{A_4}),$$

which maps its ramification points, the elements of S and ∞ to $\{0, 1, \infty\}$.

The proof is a simple height computation which we give below for the sake of completeness. Such a bound has also been worked out in [4] with explicit A_i but the computation is longer and more delicate; for our application the simpler Proposition 3.1 will suffice.

Now we construct ϕ keeping track of the heights and degrees of maps. In the remainder of this section, we write $c_1, c_2, \ldots$ for computable constants that depend only on r, m but not on B or the particular S.

Step I (mapping to $\mathbb{Q}$). This step is standard, see for instance [3, Exercise A.4.7]. Here we do not consider the point at infinity; we will work with non-constant polynomials in this step, hence, ∞ gets mapped to ∞.

Inductively, one uses the monic minimal polynomial F_α of an element α in S to map all the elements of the set and hence reducing the degree over $\mathbb{Q}$ of at least α, at the cost of introducing new elements (the critical points of F_α) whose degrees over $\mathbb{Q}$ are smaller than the degree of α. This procedure stops after c_1 steps, and say that $T \subseteq \mathbb{Q}$ is the final set in this procedure. If $0 \notin T$ we will also include it to simplify the notation later; note that $\#T \leq c_2$. Let F be the composition of all the F_α; then the degree of F is c_3 and F maps all its critical points and all the elements of S to $T \subseteq \mathbb{Q}$. Also, note that F is monic.

At each step of this construction the height of the elements of the new set can increase. However, a simultaneous induction with the height of the F_α and the height of the sets at each step shows that the elements in T and all the F_α have height bounded by $c_4 B^{c_5}$. As there are c_1 polynomials F_α, each with height bounded by $c_4 B^{c_5}$, it follows that $H(F) < c_6 B^{c_7}$.

Step II (mapping to $\{0, 1, \infty\}$). In most references, this step is performed using functions of the form $cX^a(1-X)^b$ to inductively move one element of T each time; this is the first proof that Belyi gave. This procedure is completely explicit but

March 18, 2014 20:19 WSPC/S1793-0421 203-IJNT 1450053

unfortunately it is very expensive in terms of heights. Instead, one can *move all the elements of T at the same time* as in Belyi's second proof — see [1] for a more detailed discussion on these two proofs and an explanation of why this second proof is not so widely known. Note, however, that we follow a different approach for the construction, which makes the estimates simpler.

Enumerate the elements of T as follows: $q_0 = 0, q_1, \ldots, q_t$, where $t < c_2$. We claim that there are (economical) *non-zero* integers k_i such that $\sum_i k_i \neq 0$ and the map

$$\psi(X) = \prod_{i=1}^{t} (X - q_i)^{k_i} \in \mathbb{Q}(X)$$

has all its affine critical points (i.e. possibly excluding ∞) in T. Indeed, away from the poles of ψ (which already belong to T), the affine critical points are the solutions of $d\psi(X)/\psi(X) = 0$. Clearing denominators this becomes

$$D(X) := \sum_{i=1}^{t} k_i P_i(X) = 0, \quad \text{where } P_i(X) = \prod_{j \neq i} (X - q_j).$$

The P_i have degree $t-1$ in X and evaluating at the q_i we see that the P_i are linearly independent. Thus, there are integers k_i not all zero such that $D(X) = aX^{t-1}$ with $a = \sum_i k_i \neq 0$, and again evaluating at the q_i we see that for such a tuple $\mathbf{k} = (k_i)_i \neq 0$ one necessarily has that *all* the k_i are non-zero. For such $\mathbf{k}$, the only affine critical point of ψ which is not a pole of $d\psi/\psi$ is 0 which also belongs to T. This proves the claim, and as we will see below, one can control the size of $\mathbf{k}$.

The condition $\sum_i k_i P_i(X) = D(X) = aX^{t-1}$ can be seen as a vanishing condition on the coefficients of $1, X, \ldots, X^{t-2}$. This is the same as requiring $A\mathbf{k} = 0$ where A is a $(t-1) \times t$ matrix whose entries are $(t-1)$-variable elementary symmetric functions of degree $\leq t-1$ evaluated at the q_i. Moreover, observe that if $\mathbf{k} \neq 0$ satisfies this condition then $a \neq 0$ because the P_i are linearly independent. If δ is the product of the denominators of the q_i then $\mathcal{A} = \delta A$ is a matrix with integers coefficients, and all its entries have absolute value bounded by $M^t \cdot 2^{t-1} M^{t-1} < c_8 B^{c_9}$, where M is the maximal height of an element in T. We are now in a position to apply an elementary version of Siegel's Lemma, which we now recall (see, for instance [3, Lemma D.4.1]).

Proposition 3.2 (Siegel's lemma). *Let $\mathcal{A}$ be an $m \times n$ matrix with integer coefficients. Suppose that $m < n$ and that all the entries of $\mathcal{A}$ have absolute value bounded by X. Then there is a non-zero vector $\mathbf{k} \in \mathbb{Z}^n$ in the kernel of $\mathcal{A}$ such that all the coordinates of $\mathbf{k}$ have absolute value bounded by $(nX)^{m/(n-m)}$.*

Applied in our setting, Siegel's lemma gives that there is a $\mathbf{k} \neq 0$ in the kernel of $\mathcal{A}$ such that all its coordinates are integers and have absolute value bounded by

$$(tc_8 B^{c_9})^{\frac{t-1}{t-(t-1)}} < c_{10} B^{c_{11}}.$$

March 18, 2014 20:19 WSPC/S1793-0421 203-IJNT 1450053

With this choice of k_i, the degree and height of ψ are

$$\deg \psi \leq \sum_{i=1}^{t} |k_i| < c_{12} B^{c_{13}},$$

$$H(\psi) < \delta^{\sum_i |k_i|} 4^{t \sum_i |k_i|} \prod_{i=1}^{t} H(q_i)^{|k_i|} < \exp(c_{15} B^{c_{16}}).$$

The height was estimated as follows: first we clear denominators of the q_i (this is the factor δ) so that the numerator and denominator of ψ become polynomials with integer coefficients, and then we use Proposition 2.1 to estimate the height of these two polynomials.

Now observe that $\psi(0) \in \mathbb{Q}^\times$ and let $\Psi(X) = \frac{1}{\psi(0)}\psi(X)$. Note that it has the same degree as ψ, and $H(\Psi) < \exp(c_{17} B^{c_{18}})$. The function Ψ maps all its affine ramification points and the set T to $\{0, 1, \infty\}$ because of our choice of the k_i and because $k_i \neq 0$ for each i. Moreover, since $\sum_i k_i \neq 0$ we see that $\Psi(\infty) \in \{0, \infty\}$.

Finally, use Step I and Step II to conclude that $\phi = \Psi \circ F$ maps all its ramification points in $\mathbb{P}^1$ and all the elements of S to $\{0, 1, \infty\}$. Moreover, since F is a polynomial $F(\infty) = \infty$ and hence $\phi(\infty) \in \{0, \infty\}$. The degree of ϕ can be bounded using $\deg(\phi) = \deg(F)\deg(\Psi)$, and the height of ψ can be estimated using Proposition 2.3. Therefore, Proposition 3.1 follows.

4. Explicit ABC

In this section we prove Theorem 1.2, so we keep the same notation and assumptions from its statement. We remark that the argument below is essentially due to Langevin and our only contribution is to make explicit the dependence on the height of the polynomial. The same argument, as the one below, gives explicit dependence on the degree of F provided that one uses the bounds from [4] instead of Proposition 3.1. We leave this variation as an exercise for the interested reader.

Let us first introduce some notation. In the computations of this section, we write $c_0, c_1, c_2, \ldots$ for computable constants that only depend on r. Given a nonzero $g \in \mathbb{Z}[X]$, we write $\mathrm{rad}_{\mathbb{Z}[X]}(g)$ for the product of all distinct irreducible factors of g in $\mathbb{Z}[X]$ with positive leading coefficient; this is the radical of g in $\mathbb{Z}[X]$. When $g = 0$ we define $\mathrm{rad}_{\mathbb{Z}[X]}(0) = 0$. Note that $\mathrm{rad}_{\mathbb{Z}[X]}$ agrees with our previously defined rad on $\mathbb{Z} \subseteq \mathbb{Z}[X]$.

Let S be the set of roots of F. Note that S has r elements, all of them with degree $\leq r$ and height bounded by $c_0 B$ where $B := H(F)$. Let ϕ be the function provided by Proposition 3.1 for this S; then the corresponding A_i only depend on r, not on B. Put $D = \deg \phi$ and $H = H(\phi)$.

Let $u, v \in \mathbb{Z}[X]$ be coprime with $\phi = u/v$ and let $w = v - u \in \mathbb{Z}[X]$. Then $H(u), H(v), H(w) \leq 2H$ and $\deg w \leq \max\{\deg u, \deg v\} = D$. Using the Riemann–Hurwitz formula and the fact that ϕ is unbranched away from $\{0, 1, \infty\}$ we see

March 18, 2014 20:19 WSPC/S1793-0421 203-IJNT 1450053

that

$$-2 = -2D + (3D - \#\phi^{-1}\{0,1,\infty\}), \quad \text{hence } \#\phi^{-1}\{0,1,\infty\} = D+2.$$

Note that $\alpha \in \mathbb{C}$ is a root of uvw if and only if $\phi(\alpha) \in \{0,1,\infty\}$, and $\phi(\infty) \in \{0,1,\infty\}$ by construction, hence uvw has $D+1$ distinct roots (without counting multiplicities). We also conclude that F divides $\mathrm{rad}_{\mathbb{Z}[X]}(uvw)$ in $\mathbb{Q}[X]$ because F has no repeated roots and $\phi(S) \subseteq \{0,1,\infty\}$. Hence F divides $\delta \mathrm{rad}_{\mathbb{Z}[X]}(uvw)$ in $\mathbb{Z}[X]$ for some integer $0 < \delta \leq B$, by Gauss lemma.

Let $R \in \mathbb{Z}$ be the resultant of u and w; then $R \neq 0$ as u, w are coprime. Moreover, Proposition 2.4 gives

$$\begin{aligned} |R| &\leq (\deg u + \deg w)^{\deg u + \deg w} \max\{H(u), H(w)\}^{\deg u + \deg w} \\ &\leq (2D)^{2D}(2H)^{2D} < c_1(2H)^{c_2 D^2}. \end{aligned}$$

For $n \in \mathbb{Z}$ put $g_n = \gcd(u(n), w(n))$ which is well defined because u, w have no common root, and observe that for every n we have $g_n | R$. We can apply the ABC conjecture to the equation $u(n)/g_n + w(n)/g_n = v(n)/g_n$ provided that n is not a root of uvw. It follows that for any $\epsilon > 0$ there is K_ϵ depending only on ϵ such that for every integer n one has

$$\frac{1}{R} \max\{|u(n)|, |v(n)|, |w(n)|\}^{1-\epsilon} < K_\epsilon \max\{1, \mathrm{rad}(u(n)v(n)w(n))\},$$

where the extra 1 covers of the case when n is a root of uvw (in which case $\max\{|u(n)|, |v(n)|, |w(n)|\} \leq g_n \leq R$).

Let $G \in \mathbb{Z}[x]$ be such that $FG = \delta \mathrm{rad}_{\mathbb{Z}[X]}(uvw)$; then $\deg(G) = \deg \mathrm{rad}(uvw) - r = D+1-r$. Moreover, G divides δuvw in $\mathbb{Z}[X]$, say $\delta uvw = G \cdot G_0$ with $G_0 \in \mathbb{Z}[X]$, and Proposition 2.1 gives

$$\begin{aligned} H(G) &\leq H(G)H(G_0) \leq 4^{3D+1} H(\delta uvw) \\ &\leq 4^{3D+1} 4^{6D+3} B \cdot H(u)H(v)H(w) \leq e^{c_3 D} B H^3. \end{aligned}$$

Hence, for $n \neq 0$ we have

$$\begin{aligned} \mathrm{rad}(u(n)v(n)w(n)) &\leq \mathrm{rad}(F(n)) \cdot \mathrm{rad}(G(n)) \leq |G(n)| \mathrm{rad}(F(n)) \\ &\leq (\deg(G)+1) H(G) |n|^{\deg G} \mathrm{rad}(F(n)) \\ &\leq e^{c_4 D} H^3 B |n|^{D+1-r} \mathrm{rad}(F(n)). \end{aligned}$$

Combining this with the bound obtained from the ABC conjecture, we get that for every integer $n \neq 0$

$$\begin{aligned} \max\{|u(n)|, |v(n)|, |w(n)|\}^{1-\epsilon} &< K_\epsilon R (1 + e^{c_4 D} H^3 B |n|^{D+1-r} \mathrm{rad}(F(n))) \\ &< K_\epsilon c_1 (2H)^{c_2 D^2} (1 + e^{c_4 D} H^3 B |n|^{D+1-r} \mathrm{rad}(F(n))) \\ &< K_\epsilon c_5 (1+H)^{c_6 D^2} B |n|^{D+1-r} \max\{1, \mathrm{rad}(F(n))\}. \end{aligned}$$

Note that $\max\{|u(n)|, |v(n)|, |w(n)|\} \geq 1$. Let x be a polynomial among u, v, w having degree D; then

$$\begin{aligned}\max\{|u(n)|, |v(n)|, |w(n)|\} &\geq |x(n)| \geq |n|^D - D \cdot H(x)|n|^{D-1} \\ &\geq |n|^{D-1}(|n| - 2DH) > \frac{1}{2}|n|^D\end{aligned}$$

provided that $|n| > 4DH$. Therefore, for all n we have

$$\max\{|u(n)|, |v(n)|, |w(n)|\} \geq \frac{1}{(4DH)^D}|n|^D.$$

It follows that

$$\begin{aligned}|n|^{D(1-\epsilon)} &< K_\epsilon (4DH)^D \cdot c_5(1+H)^{c_6 D^2} B|n|^{D+1-r} \max\{1, \mathrm{rad}(F(n))\} \\ &< K_\epsilon c_5 (1+H)^{c_7 D^2} B|n|^{D+1-r} \max\{1, \mathrm{rad}(F(n))\}\end{aligned}$$

and after choosing a different $\epsilon > 0$ we get that for all n

$$\begin{aligned}|n|^{r-1-\epsilon} &< K_\epsilon c_5 (1+H)^{c_7 D^2} B \max\{1, \mathrm{rad}(F(n))\} \\ &< K_\epsilon c_5 (1 + e^{A_3 (c_0 B)^{A_4}})^{c_7 A_1^2 (c_0 B)^{2A_2}} B \max\{1, \mathrm{rad}(F(n))\} \\ &< K_\epsilon \exp(c_8 B^{c_9}) \max\{1, \mathrm{rad}(F(n))\}\end{aligned}$$

as we wanted. This proves Theorem 1.2.

5. Sieve Preliminaries

The proof of Theorem 1.1 starts with some standard sieve manipulations.

Let us set the notation. Let $r \geq 2$ and let $f \in \mathbb{Z}[X]$ be a polynomial of degree r, without repeated factors, and with G_f squarefree. We write a_r for the leading coefficient of f and Δ_f for the discriminant of f (which is non-zero as f has no repeated factors). The symbol p will denote a prime.

Among the several versions of Hensel's lemma available in the literature, let us recall the following one which is obtained by setting $m = 1$ in [10, Theorem 1, p. 14] (note that the conditions $0 \leq j \leq m$ and $0 < 2k < n$ in the cited result should be $1 \leq j \leq m$ and $0 \leq 2k < n$).

Proposition 5.1 (Hensel's lemma). *Let $F \in \mathbb{Z}_p[X]$ and $x \in \mathbb{Z}_p$ where $\mathbb{Z}_p$ is the ring of p-adic integers. Suppose that for some integers n, k with $0 \leq 2k < n$ we have $F(x) \equiv 0 \bmod p^n$ and $v_p(F'(x)) = k$, where v_p is the p-adic valuation and F' is the derivative of F. Then there is $y \in \mathbb{Z}_p$ with $F(y) = 0$ and $y \equiv x \bmod p^{n-k}$.*

Using this, we obtain the follwing lemma.

Lemma 5.2. *If $p \nmid a_r \Delta_f$ then for every $t \geq 1$ the congruence $f(X) \equiv 0 \bmod p^t$ has at most r solutions. Moreover, there is a constant $C = C(f)$ such that for all primes p and all $t \geq 1$ the congruence $f(X) \equiv 0 \bmod p^t$ has $< C$ solutions.*

Proof. For the first part, it suffices to show that each solution to the congruence $f(X) \equiv 0 \bmod p^t$ lifts to a p-adic solution, because f has at most r roots in $\mathbb{Z}_p$ (recall that $\mathbb{Z}_p$ is an integral domain).

Let $x \in \mathbb{Z}$ such that $f(x) \equiv 0 \bmod p^2$. If $p \,|\, f'(x)$ then $p \,|\, \mathrm{Res}(f, f') = \pm a_r \Delta_f$ which is not possible, hence $p \nmid f'(x)$. Hensel's lemma (with $n = t$ and $k = 0$) gives the desired p-adic lift of n.

For the last part of the lemma, we can restrict our attention to primes p dividing $\mathrm{Res}(f, f')$. Let p be a prime divisor of $\mathrm{Res}(f, f')$ and let t_p be such that $p^{t_p} \nmid \mathrm{Res}(f, f')$. Similarly, we apply Hensel's lemma with $k \leq t_p - 1$ and $n = t \geq 2t_p - 1$ to conclude that f has at most r roots modulo p^{t-t_p+1} that are congruent (modulo p^{t-t_p+1}) to roots in $\mathbb{Z}/p^t\mathbb{Z}$. Hence, f has at most rp^{t_p-1} roots modulo p^t for any p dividing $\mathrm{Res}(f, f')$ and $t \geq t_p$. Since the prime p (divisor of $\mathrm{Res}(f, f')$) and t_p are bounded in terms of f, the result follows. □

Let $\epsilon > 0$ to be chosen later. First we note that

$$\#Q \geq N_f(x) \geq \#Q - \#R - \#S - x^{1-\epsilon} \tag{1}$$

where we write

$$Q = \{n \leq x : \forall\, p \leq y, p^2 \nmid f(n)\},$$

$$R = \{n \leq x : \exists\, p \in (y, z], p^2 \,|\, f(n)\},$$

$$S = \{n \in (x^{1-\epsilon}, x] : \exists\, p > z, p^2 \,|\, f(n)\}$$

and $y < z$ are parameters to be chosen later. These sets depend on ϵ, x, y, z although the notation does not reflect this fact.

To simplify the exposition, let us introduce the following notation: if X is a true statement then write $\delta(X) = 1$, and if X is false then $\delta(X) = 0$. For instance $\delta(3|2) = 0$ because 3 does not divide 2 in $\mathbb{Z}$.

Lemma 5.3. *We have*

$$\#Q = c_f x + O_{f,\epsilon}\left(\exp(\epsilon y) + \frac{x}{y^{1-\epsilon}}\right)$$

provided that $y \gg_{f,\epsilon} 1$.

Proof. Set $P = \prod_{p \leq y} p$. We begin by observing that

$$\#Q = \sum_{n \leq x} \prod_{p \leq y} (1 - \delta(p^2 \,|\, f(n))) = \sum_{n \leq x} \sum_{d | P} \mu(d) \delta(d^2 \,|\, f(n))$$

$$= \sum_{d | P} \mu(d) \omega_f(d^2) \left(\frac{x}{d^2} + O(1)\right) = x \prod_{p \leq y} \left(1 - \frac{\omega_f(p^2)}{p^2}\right) + O\left(\sum_{d | P} \omega_f(d^2)\right)$$

$$= x \prod_{p \leq y} \left(1 - \frac{\omega_f(p^2)}{p^2}\right) + O\left(\prod_{p \leq y} (1 + \omega_f(p^2))\right).$$

By Lemma 5.2, for $y \gg_{\epsilon,f} 1$

$$\prod_{p\le y}(1+\omega_f(p^2)) \ll_f \prod_{p\le y}(r+1) \ll_{f,\epsilon} \exp(\epsilon y).$$

Let us analyze the other product. If $y \gg_{f,\epsilon} 1$ then

$$1 < \prod_{p>y}\left(1-\frac{\omega_f(p^2)}{p^2}\right)^{-1} \le \prod_{p>y}\left(1-\frac{r}{p^2}\right)^{-1} < 1+\sum_{n>y}\frac{1}{n^{2-\epsilon/2}} < 1+\frac{1}{y^{1-\epsilon}}$$

and multiplying times c_f we obtain

$$c_f < \prod_{p\le y}\left(1-\frac{\omega_f(p^2)}{p^2}\right) < \left(1+\frac{1}{y^{1-\epsilon}}\right)c_f.$$

AQ: Please check the sentence "... and multiplying times .." for clarity.

The result now follows. □

Lemma 5.4. *We have*

$$\#R \ll_r \frac{x}{y}+\frac{z}{\log z}$$

provided that $z > y \gg_f 1$.

Proof. Indeed, for $z > y \gg_f 1$

$$\#R \le \sum_{n\le x}\sum_{y<p\le z}\delta(p^2 \mid f(n)) \le \sum_{y<p\le z}\omega_f(p^2)\left(\frac{x}{p^2}+1\right)$$

$$\le \sum_{y<p\le z} r\left(\frac{x}{p^2}+1\right) \le \frac{2rx}{y}+\frac{2rz}{\log z}.$$

□

AQ: Please check the sentence "Now choose ...".

Now choose $y = \log x$ and $z = x$, then by inequalities (1) we have the following proposition.

Proposition 5.5. *Let* $\epsilon > 0$. *Then*

$$N_f(x) = c_f x + O_{f,\epsilon}\left(\frac{x}{(\log x)^{1-\epsilon}}\right) + O(\#S)$$

for $x \gg_{f,\epsilon} 1$, *where*

$$S = \{n \in (x^{1-\epsilon}, x] : \exists\, p > x, p^2 \mid f(n)\}.$$

Note that the proof actually shows that the upper bound does not require one to estimate $\#S$. The problem of bounding $\#S$ is only relevant for the lower bound, and it is exactly the point where one needs to invoke the ABC conjecture. We treat this in the next section, in order to conclude the proof of Theorem 1.1.

March 18, 2014 20:19 WSPC/S1793-0421 203-IJNT 1450053

6. Error Term for Counting Squarefree Values

First, observe that the conditions on f imposed by Theorem 1.1 are compatible with the conditions of the previous section.

With the notation of the previous section, Proposition 5.5 shows that in order to prove Theorem 1.1 it suffices to show that $c_f > 0$ when G_f is squarefree, and (on the ABC conjecture) that

$$\#S = \#\{n \in (x^{1-\epsilon}, x] : \exists\, p > x, p^2 \mid f(n)\} \ll_f \frac{x}{(\log x)^\gamma} \tag{2}$$

with $\gamma > 0$ as in the statement of the theorem, for some $1/2 > \epsilon > 0$ say.

It is well known that $c_f > 0$ when G_f is squarefree, but we sketch a proof for the sake of completeness. Since G_f is squarefree, $\omega_f(p^2) < p^2$ for all primes p, hence no factor in the definition of c_f is zero. For large primes, we use the bound $\omega_f(p^2) \leq r$ from the previous section and it follows that the product defining c_f converges absolutely, hence, it is non-zero.

Now we focus on proving the estimate (2) on the ABC conjecture.

We partition $(x^{1-\epsilon}, x]$ into T intervals I_i, each one having length $\leq 2x/T$ (we will later take T equal to x divided by a power of $\log x$, so that $x/T \to \infty$). First we show, on the ABC conjecture, that $I_i \cap S$ contains at most $\ll_f 1$ elements for suitable choice of T. For $d \geq 1$ define $F_d(X) = f(X)f(X+d)$.

Claim 6.1. *There is a constant M_f depending only on f such that if $d \geq M_f$ then the polynomial F_d has no repeated factors.*

Proof. The roots of $f(X)$ have complex modulus bounded in terms of f. Hence, if $d \gg_f 1$ then $f(X)$ and $f(X+d)$ have no common factor. □

Suppose that $I_i \cap S$ contains more than M_f elements. Then we can find $d \geq M_f$ such that $n, n+d$ are in $I_i \cap S$. By the previous claim, we can apply Theorem 1.2 to F_d (on the ABC conjecture) to obtain

$$n^{2r-1-\epsilon} < K_\epsilon \exp(\alpha H(F_d)^\beta) \operatorname{rad}(F_d(n)) \ll_f K_\epsilon \exp(\alpha H(F_d)^\beta) \left(\frac{x^r}{x}\right)^2,$$

where we have used the fact that $n, n+d \in S$ (note that α and β depend only on r). Hence (as $n > x^{1-\epsilon}$ in S)

$$x^{1-2\epsilon r} < x^{1-2\epsilon r+\epsilon^2} \ll_f K_\epsilon \exp(\alpha H(F_d)^\beta).$$

Let us fix $\epsilon = 1/(4r)$ to get

$$x^{1/2} \ll_f \exp(\alpha H(F_d)^\beta).$$

On the other hand, using Propositions 2.1, 2.5, and the fact that $d < 4x/T$, we obtain

$$H(F_d) \leq 4^{2r+1} H(f) H(f(X+d)) \leq 4^{3r+1} H(f)^2 d^r < 4^{4r+1} H(f)^2 \frac{x^r}{T^r}.$$

March 18, 2014 20:19 WSPC/S1793-0421 203-IJNT 1450053

Therefore, if we choose

$$T = \kappa \frac{x}{(\log x)^{1/(r\beta)}}$$

for some $\kappa > 0$ sufficiently large with respect to r and $H(f)$, then we get a contradiction. It follows that with this choice of T and assuming the ABC conjecture, each $I_i \cap S$ contains at most M_f elements.

Finally, since there are T of these intervals I_i, we conclude that

$$\#S \ll_f \frac{x}{(\log x)^{\gamma}},$$

where $\gamma = 1/(r\beta) > 0$ is computable and depends only on r, not on the particular f. This proves the inequality (2), and hence Theorem 1.1.

7. A More General Result

As said in Sec. 1, the method in this paper allows one to give, on the ABC conjecture, asymptotic formulas *with error term* for the problem of counting squarefree values of polynomials when the variable is restricted to suitable subsets of the positive integers. Let us explain this in more detail.

Given a set A of positive integers, we say that A has *density* $\sigma(A)$ if the following limit exists and equals $\sigma(A)$:

$$\lim_{x\to\infty} \frac{\#\{n \le x : n \in A\}}{x}.$$

For instance the primes have density 0 and the multiples of a fixed positive integer k have density $1/k$. Not all sets of positive integers have a density, but we restrict our attention to those with density.

Given A and integers m, a we define $A(m, a) = \{t \in A : t \equiv a \bmod m\}$.

In this section $g(x)$ will always denote a positive real-valued function satisfying $g(x) = o(x)$, while $\lambda(x)$ will denote a function growing to ∞ and A will denote a set of positive integers with density. We say that A is *residually well distributed with level* $\lambda(x)$ *and discrepancy* $g(x)$ if there are constants C, x_0 such that for all $x > x_0$ one has

$$\left|\#\{n \le x : n \in A(m, a)\} - \frac{\sigma(A)x}{m}\right| < Cg(x)$$

for each $m \le \lambda(x)$ and each residue class a modulo m. Observe that if A is residually well distributed with level λ and discrepancy g, then it is residually well distributed with level λ' and discrepancy g' for any functions λ' and g' satisfying $\lambda'(x) < \lambda(x)$ and $g'(x) > g(x)$ for x sufficiently large, and $\lambda' \to \infty, g'(x) = o(x)$.

We warn the reader that the concept of discrepancy just introduced is not the same as the discrepancy that arises in the theory of uniformly distributed sequences. However, as we will see in the next section, there is indeed a connection between both notions of discrepancy.

The relevant result is the following.

Theorem 7.1. *Assume the ABC conjecture. Suppose that A is residually well distributed with level $(\log x)^2$ and discrepancy g. Let f be a polynomial as in Theorem* 1.1. *Let $N_f^A(x)$ be the number of $n \leq x$ such that $n \in A$ and $f(n)$ is squarefree. Then for any given $\epsilon > 0$ we have*

$$N_f^A(x) = \sigma(A)c_f x + O_{f,A,\epsilon}\left((\log x)^{1+\epsilon}g(x) + \frac{x}{(\log x)^{\gamma}}\right),$$

where c_f and γ are as in Theorem 1.1.

Proof. The proof of this result is similar to the proof of Theorem 1.1. First note that $\omega_f(n) \ll_\epsilon n^\epsilon$ because ω_f is multiplicative and bounded on prime powers (see Lemma 5.2).

Define the sets

$$Q_A = \{n \leq x : n \in A, \forall\, p \leq y, p^2 \nmid f(n)\},$$
$$R_A = \{n \leq x : n \in A, \exists\, p \in (y, z], p^2 \mid f(n)\},$$
$$S_A = \{n \in (x^{1-\epsilon}, x] : n \in A, \exists\, p > z, p^2 \mid f(n)\}$$

with ϵ, y, z to be chosen. Then one observes that

$$\#Q_A \geq N_f^A(x) \geq \#Q_A - \#R_A - \#S_A - x^{1-\epsilon}.$$

However, since $R_A \subseteq R$ and $S_A \subseteq S$ (with R, S as in Sec. 5) we obtain from our previous work that, on the ABC conjecture, the following relation holds:

$$N_f^A(x) = \#Q_A + O\left(\frac{x}{(\log x)^{\gamma}}\right)$$

provided that we choose y, z as in Sec. 5 (namely, $y = \log x, z = x$) and ϵ as in Sec. 6 (namely, any fixed $\epsilon \leq 1/(4r)$). Therefore one just needs to prove that

$$\#Q_A = \sigma(A)c_f x + O\left((\log x)^{1+\epsilon}g(x) + \frac{x}{(\log x)^{\gamma}}\right).$$

This formula requires more work than the estimation of $\#Q$ in Sec. 5, since we want to assume that A is residually well distributed with level $(\log x)^2$, which is rather small (see, for instance, Theorem 8.2 below).

To prove the estimate for $\#Q_A$, write $P = \prod_{p \leq y} p$ and observe that

$$\#Q_A = \sum_{\substack{n \leq x \\ n \in A}} \prod_{p \leq y}(1 - \delta(p^2 \mid f(n))) = \sum_{d \mid P} \mu(d) \sum_{\substack{n \leq x \\ n \in A}} \delta(d^2 \mid f(n))$$
$$= \sum_{d \mid P} \mu(d) \sum_{\substack{a \bmod d^2 \\ f(a) \equiv 0 \bmod d^2}} \sum_{\substack{n \leq x \\ n \in A(d^2, a)}} 1.$$

March 18, 2014 20:19 WSPC/S1793-0421 203-IJNT 1450053

Let us split the latter sum as $U + V$ where U takes the summands with $d \mid P$ and $d \le y$, while V takes the summands with $d \mid P$ and $d > y$. For V one finds

$$|V| \le \sum_{\substack{d|P \\ d>y}} \omega_f(d^2)\left(\frac{x}{d^2}+1\right) \le x \sum_{d>y} \frac{1}{d^{2-\epsilon}} + \sum_{d|P} \omega_f(d^2) \le O\left(\frac{x}{y^{1-\epsilon}} + \exp(\epsilon y)\right),$$

where we used $\omega_f(n) \ll n^{\epsilon}$, and the second summand was bounded as in the proof of Lemma 5.3. Hence, V is absorbed by the error term. On the other hand, since A is residually well distributed with level $(\log x)^2$ and discrepancy g, we find

$$U = \sum_{\substack{d|P \\ d\le y}} \mu(d)\omega_f(d^2)\frac{\sigma(A)x}{d^2} + O(y \cdot y^{\epsilon} \cdot g(x)).$$

Indeed, the error term comes from the discrepancy, and we only used moduli $d^2 \le y^2 = (\log x)^2$. A computation as in the bound for V shows that we can include those $d \mid P$ with $d > y$ obtaining

$$\begin{aligned}\#Q_A = U + V &= \sigma(A)x \sum_{d|P} \mu(d)\frac{\omega_f(d^2)}{d^2} + O\left(\frac{x}{y^{1-\epsilon}} + \exp(\epsilon y) + y^{1+\epsilon} g(x)\right) \\ &= \sigma(A)x \prod_{p\le y}\left(1 - \frac{\omega_f(p^2)}{p^2}\right) + O\left((\log x)^{1+\epsilon} g(x) + \frac{x}{(\log x)^{\gamma}}\right).\end{aligned}$$

Here we used $y = \log x$. The product is treated as in the proof of Lemma 5.3, which introduces an error term $O(x/y^{1-\epsilon})$ that has no effect in the previous error term. This concludes the proof. □

We remark that actually, for Theorem 7.1 it is enough to have an averaged version of residually well-distributed sets.

8. Proof of Theorem 1.3

For $q \in \mathbb{Q}$ one defines $H(q) = \max\{|a|, |b|\}$ where a, b are coprime integers satisfying $q = a/b$. This agrees with our previous definition of the height of an algebraic number. Let $\alpha \in \mathbb{R}\backslash\mathbb{Q}$. Recall that the *approximation exponent* of α (also known as *measure of irrationality*) is defined as

$$\tau(\alpha) = \sup\{t \in \mathbb{R} : \text{there are infinitely many } q \in \mathbb{Q} \text{ with } |q - \alpha| < H(q)^{-t}\}.$$

Dirichlet's box principle shows that $\tau(\alpha) \in [2, \infty]$. On the other hand, Liouville showed $\tau(\alpha) \le r$ whenever α is algebraic of degree $r \ge 2$, which allowed him to construct transcendental numbers by showing examples of real numbers with infinite approximation exponent. A celebrated theorem of Roth shows that actually $\tau(\alpha) = 2$ whenever α is algebraic. There are also very familiar transcendental numbers that have finite approximation exponent, such as π (this is a theorem of Mahler, see [8]).

For later reference, let us recall the Erdös–Turán inequality (see, for instance, [9, Sec. 11.4]).

Theorem 8.1 (Erdös–Turán inequality). *Let $\{x_n\}_n$ be a sequence of real numbers. For all integers $M, N \geq 1$ we have*

$$\sup_{0\leq a<b\leq 1}\left|\frac{\#\{n\leq N : a\leq (x_n)<b\}}{N}-(b-a)\right| \leq \frac{1}{M+1}+3\sum_{k=1}^{M}\frac{1}{Nk}\left|\sum_{n=1}^{N}e^{2\pi i k x_n}\right|,$$

where (x_n) denotes the fractional part of x_n.

The next result provides a source of examples where Theorem 7.1 can be applied. In particular we obtain Theorem 1.3.

Theorem 8.2. *Let $\alpha \in \mathbb{R}\backslash\mathbb{Q}$ be an irrational real number with $\alpha > 1$. Assume that α has finite approximation exponent $\tau = \tau(\alpha)$. Then the set $A = \{\lfloor n\alpha\rfloor : n \geq 1\}$ has density $1/\alpha$ and for all $\epsilon > 0$ it satisfies*

$$\max_{r \bmod m}\left|\frac{\#\{n\leq x : n\in A(m,r)\}}{x}-\frac{\sigma(A)}{m}\right| \ll_{\epsilon,\alpha}\left(\frac{m}{x}\right)^{1/(\tau+\epsilon)}$$

whenever $m \leq x/\alpha$. In particular, A is residually well distributed with level $x^{1/2}$ and discrepancy $x^{1-1/(3\tau)}$.

(Observe that, if $0 < \alpha < 1$ then $A = \mathbb{N}$.)

Proof of Theorem 8.2. The fact that $\sigma(A) = 1/\alpha$ is clear. Given a positive integer m, we define

$$\Delta_m(x) = \max_{r \bmod m}\left|\frac{\#\{n\leq x : n\in A(m,r)\}}{x}-\frac{\sigma(A)}{m}\right|.$$

Given a positive real number β, we define

$$D(\beta, x) = \sup_{0\leq a<b\leq 1}\left|\frac{\#\{k\leq x : a\leq (k\beta)<b\}}{x}-(b-a)\right|,$$

where $(k\beta)$ denotes the fractional part of $k\beta$. This quantity $D(\beta, x)$ is called *discrepancy* in the theory of uniformly distributed sequences, and as we will see, it is very much related to our notion of discrepancy for residually well-distributed sets. Taking $\beta = \alpha/m, b = r/m$ and $a = (r-1)/m$ we see that

$$\Delta_m(x) \leq O\left(\frac{1}{x}\right) + \sigma(A)D(\alpha/m, x/\alpha).$$

AQ: Please check the sentence "This quantity...".

March 18, 2014 20:19 WSPC/S1793-0421 203-IJNT 1450053

Therefore, it suffices to show $D(\alpha/m, y) \ll (m/y)^{1/(\tau+\epsilon)}$ for $m < y$. Write $\|\beta\|$ for the distance of β to the nearest integer, then the Erdös–Turán inequality gives (see also [9, Exercise 11.4.10])

$$\begin{aligned} D(\alpha/m, N) &\leq \frac{1}{M+1} + 3\sum_{k=1}^{M} \frac{1}{Nk} \left| \sum_{n=1}^{N} e^{2\pi i k \cdot \frac{n\alpha}{m}} \right| \\ &\leq \frac{1}{M+1} + 3\sum_{k=1}^{M} \frac{1}{Nk} \frac{1}{|\sin(\pi k\alpha/m)|} \\ &\leq \frac{1}{M+1} + \frac{3}{2N} \sum_{k=1}^{M} \frac{1}{k\|k\alpha/m\|} \end{aligned}$$

for all positive integers M, N. Since α has finite approximation exponent τ we have

$$\|k\alpha/m\| = |j - k\alpha/m| = \frac{k}{m}\left|\alpha - \frac{mj}{k}\right| \gg_{\alpha,\epsilon} \frac{1}{mk^{\tau-1+\epsilon}},$$

where $j \in \mathbb{Z}$ is an integer that satisfies $|j - k\alpha/m| = \|k\alpha/m\|$. Hence, using the fact that $\tau \geq 2$ we get

$$D(\alpha/m, N) \ll_{\alpha,\epsilon} \frac{1}{M} + \frac{m}{N}\sum_{k=1}^{M} k^{\tau-2+\epsilon} \leq \frac{1}{M} + \frac{mM^{\tau-1+\epsilon}}{N}.$$

Choose $M = \lfloor (N/m)^{1/(\tau+\epsilon)} \rfloor$ (provided that $N > m$) to get

$$D(\alpha/m, N) \ll_{\alpha,\epsilon} \left(\frac{m}{N}\right)^{1/(\tau+\epsilon)},$$

which proves the result. □

Acknowledgments

It is our pleasure to thank Joseph Oesterlé for useful discussions on the possibility of using an effective version of Belyi's theorem in the context of the ABC conjecture. We also thank the anonymous referee for carefully reviewing this paper, correcting some issues, and suggesting several changes that improved the presentation of this work. The research of the first author was supported by an NSERC Discovery grant. The research of the second author was supported by an Ontario Graduate Scholarship.

References

[1] W. Goldring, Unifying themes suggested by Belyi's theorem, in *Number Theory, Analysis and Geometry* (Springer, New York, 2012), pp. 181–214.
[2] A. Granville, ABC allows us to count squarefrees, *Internat. Math. Res. Notices* **1998**(19) (1998) 991–1009.
[3] M. Hindry and J. Silverman, *Diophantine Geometry. An Introduction*, Graduate Texts in Mathematics, Vol. 201 (Springer, New York, 2000).

March 18, 2014 20:19 WSPC/S1793-0421 203-IJNT 1450053

[4] L. Khadjavi, An effective version of Belyi's theorem, *J. Number Theory* **96**(1) (2002) 22–47.
[5] S. Lang, *Diophantine Geometry*, Interscience Tracts in Pure and Applied Mathematics, No. 11 (Interscience Publishers, New York, 1962).
[6] M. Langevin, Cas d'égalité pour le théorème de Mason et applications de la conjecture (abc), *C. R. Acad. Sci. Paris Sér. I Math.* **317**(5) (1993) 441–444.
[7] J. Lee and M. R. Murty, Dirichlet series and hyperelliptic curves, *Forum Math.* **19**(4) (2007) 677–705.
[8] K. Mahler, On the approximation of π, *Nederl. Akad. Wet., Proc., Ser. A.* **56** (1953) 30–42.
[9] M. R. Murty, *Problems in Analytic Number Theory*, 2nd edn., Graduate Texts in Mathematics, Vol. 206 (Springer, New York, 2008).
[10] J.-P. Serre, *A Course in Arithmetic*, Graduate Texts in Mathematics, Vol. 7 (Springer, New York, 1973).