Finding Small Degree Factors of Multivariate Supersparse (Lacunary) Polynomials Over Algebraic Number Fields*

Erich Kaltofen
Dept. of Mathematics
Massachusetts Institute of Technology
77 Massachusetts Ave.
Cambridge, Massachusetts 02139-4307, USA
kaltofen@math.mit.edu

Pascal Koiran
Laboratoire LIP
École Normale Supérieure de Lyon
46, Allée d'Italie
69364 Lyon Cedex 07, France
Pascal.Koiran@ens-lyon.fr

ABSTRACT

We present algorithms that compute all irreducible factors of degree $\leq d$ of supersparse (lacunary) multivariate polynomials in n variables over an algebraic number field in deterministic polynomial-time in $(l+d)^n$, where l is the size of the input polynomial. In supersparse polynomials, the term degrees enter logarithmically as their numbers of binary digits into the size measure l. The factors are again represented as supersparse polynomials. If the factors are represented as straight-line programs or black box polynomials, we can achieve randomized polynomial-time in $(l+d)^{O(1)}$. Our approach follows that by H. W. Lenstra, Jr., on computing factors of univariate supersparse polynomials over algebraic number fields. We generalize our ISSAC 2005 results for computing linear factors of supersparse bivariate polynomials over the rational numbers by appealing to recent lower bounds on the height of algebraic numbers and to a special case of the former Lang conjecture.

Categories and Subject Descriptors

I.1.2 [**Computing Methodologies**]: Symbolic and Algebraic Manipulation—*Algorithms*; F.2.2 [**Theory of Computation**]: Analysis of Algorithms and Problem Complexity—*Nonnumerical Algorithms and Problems*

General Terms

algorithms, theory

Keywords

sparse polynomials, lacunary polynomials, multivariate polynomials, polynomial factorization, polynomial-time complexity, algebraic numbers, height, Lang conjecture

1. INTRODUCTION

The algorithms in this paper take as inputs "super"sparse (lacunary) polynomials. A *supersparse* polynomial

$$f(X_1,\ldots,X_n)=\sum_{i=0}^{t} a_i\, X_1^{\alpha_{i,1}}\cdots X_n^{\alpha_{i,n}} \qquad (1)$$

is input by a list of its coefficients and corresponding term degree vectors. One defines the size of f as

$$\text{size}(f)=\sum_{i=0}^{t}\Big(\text{size}(a_i)+\lceil\log_2(\alpha_{i,1}\cdots\alpha_{i,n}+2)\rceil\Big), \qquad (2)$$

where size(a_i) is the bit-size of the scalar coefficients. In our case, the coefficients are in an the algebraic number field F that is represented as a ring of (dense) polynomial residues $F=\mathbb{Q}[\zeta]/(\varphi(\zeta))$ with a monic (dense) irreducible minimum polynomial $\varphi(\zeta)\in\mathbb{Z}[\zeta]$. The measure (2) accounts for very high degrees, say with hundreds of digits as binary numbers, in distinction to the usual sparse representation [21, 12]. One cannot evaluate a supersparse polynomial at algebraic numbers in polynomial-time in its size, because the value of the polynomial can have exponential size, say 2^{100} digits. Important exceptions are evaluating at roots of unity. A supersparse polynomial can be represented by a straight-line program [8] of size $O(\text{size} f)$ via evaluating its terms with repeated squaring. It is NP-hard to test if a supersparse polynomial over the rational numbers has a non-trivial content, i.e., factors depending only on $X_1,\ldots,X_{n-1}$, cf. [18].

A breakthrough polynomial-time result is in [2]. Any integral root of a univariate supersparse polynomial with integral coefficients can be found in $(\text{size} f)^{O(1)}$ bit operations. H. W. Lenstra, Jr., [16, 17] has generalized the result to computing factors of small degree over an algebraic extension, in particular to computing rational roots in polynomial-time. In [11] we generalize Lenstra's results to computing linear factors of bivariate supersparse polynomials over the rational numbers in polynomial-time. We also give a Monte Carlo randomized polynomial-time algorithm based on interpolation for computing quadratic rational polynomial factors, that under restrictions on the leading coefficient. For all

[*] This material is based on work supported in part by the National Science Foundation under Grants CCR-0305314 and CCF-0514585 (Kaltofen).
Kaltofen's permanent address: Dept. of Mathematics, North Carolina State University, Raleigh, North Carolina 27695-8205, USA, kaltofen@math.ncsu.edu.

ISSAC'06, July 9–12, 2006, Genova, Italy.
Copyright 2006 ACM 1-59593-276-3/06/0004 ...$5.00.

problems that we consider there are deterministic algorithms whose bit complexity is of order $(\mathrm{size}(f) + \deg(f))^{O(n)}$ [7, 10].

Here we present an algorithm that given a degree bound d can compute all irreducible factors of degree $\leq d$ and their multiplicities in $(\mathrm{size}(f)+d)^{O(n)}$ bit operations. Since the irreducible factors are also represented as supersparse polynomials of the form (1), the output size can be superpolynomial in $\mathrm{size}(f)+d$ [3, Example 5.1]. If the irreducible factors are represented as straight-line programs [9] or black box polynomials [13], our algorithms become Monte Carlo randomized of $(\mathrm{size}(f) + d)^{O(1)}$ bit complexity.

The deterministic algorithms of [2, 16, 11] are based on gap theorems. For instance, in [2, Proposition 2] it is shown that if $\hat{f}(X) = g(X) + X^u h(X) \in \mathbb{Z}[X]$ with $g \neq 0$, $h \neq 0$, $\deg(g) \leq k$ and $u-k \geq \log_2 \|\hat{f}\|_1$ then for an integer $a \neq \pm 1$, we have $\hat{f}(a) = 0 \Longrightarrow g(a) = h(a) = 0$. In [11] we develop a variant of the gap technique for high degree sums of linear forms.

H. W. Lenstra, Jr. has used the gap method via the height of an algebraic number for computing rational roots and low degree factors of supersparse rational polynomials. The algorithm presented in [16] receives as input a supersparse univariate polynomial $\bar{f}(X) = \sum_{i=0}^{t} a_i X^{\alpha_i} \in F[X]$ over the algebraic number field $F = \mathbb{Q}[\zeta]/(\varphi(\zeta))$. Furthermore, a factor degree bound d is input. The algorithm produces a list of all irreducible factors of $\bar{f}$ over F of degree $\leq d$ and their multiplicities in

$$(\mathrm{size}(\bar{f}) + d)^{O(1)} \tag{3}$$

bit operations. Here $\|\varphi\|$ is the (infinity) norm of the coefficient vector of φ and $\|\bar{f}\|$ is the norm of the vector of norms of the coefficients $a_i(\zeta)$. We assume that a common denominator has been multiplied through and all coefficients of the $a_i(\zeta)$ are integers. For example, for $\varphi = \zeta - 1$, that is, $F = \mathbb{Q}$, and $d = 1$, Lenstra's algorithm finds all rational roots of a supersparse integral polynomial $\bar{f}$ in polynomial-time in $\mathrm{size}(\bar{f})$. We note that there are no more than

$$O(t^2 \cdot 2^D \cdot D \cdot \log(2D\, t)), \quad \text{where } D = d \cdot \deg(\varphi) \tag{4}$$

irreducible factors of $\bar{f}$ of degree $\leq d$ [17, Theorem 1], each of which, with the exception of the possible factor X, has multiplicity at most t [16, Proposition 3.2]. The count (4) is independent of $\deg \bar{f}$ and the coefficient size of $\bar{f}$. In addition, by virtue of (3) the number of factors is always polynomial in $d + \mathrm{size} \bar{f}$.

Here we generalize the gap techniques of [11] in three ways. First, we allow $n \geq 2$ variables. Second, we compute all factors of degrees $\leq d$. Third, we allow algebraic numbers as coefficients. In order to obtain a usable gap theorem, we exploit a special case of the Lang conjecture which essentially limits factors on whose surfaces torsion points, i.e., cyclotomic points, are dense to generalized cyclotomic polynomials. The full Lang conjecture was proved by Faltings (see, e.g, [5]). Arguing in reverse, we can prove that those factors that are not generalized cyclotomic polynomials must have suitable non-cyclotomic roots which we can plug in similarly to the integer root above. Lastly, we need a lower bound on the Weil height of the non-cyclotomic root, which we luckily can lift from the recent literature [1]. For finding generalized cyclotomic factors we develop our own techniques.

2. BACKGROUND FROM NUMBER THEORY AND DIOPHANTINE GEOMETRY

A cyclotomic point of $\mathbb{C}^n$ is a point all of whose coordinates are roots of unity. In group-theoretic language, they are the torsion points of the multiplicative group $(\mathbb{C}^*)^n$. The hypersurfaces on which cyclotomic points are (Zariski) dense play a central role in this paper. Cyclotomic points are clearly dense on any hypersurface defined by an equation of the form

$$\prod_{i=1}^{n} X_i^{a_i} = \theta, \tag{5}$$

where θ is a root of unity and the exponents a_i are in $\mathbb{Z}$. For future use (see in particular Lemma 6 in section 5), note that such a hypersurface is irreducible iff the a_i have no nontrivial common divisor. As it turns out, those hypersurfaces (5) are the complete list of irreducible hypersurfaces on which cyclotomic points are dense, which constitutes a well-known special case of the former Lang conjecture. We first state the general Lang conjecture as given in [5]:

Theorem 1. *Let A be an Abelian variety defined over $\mathbb{C}$, let X be a closed subvariety of A, and let Γ be a subgroup of $A(\mathbb{C})$ of finite rank (i.e., such that there exists a free finitely generated subgroup $\Gamma_0 \subseteq \Gamma$ such that for every $x \in \Gamma$ there exists an integer $n \geq 1$ such that $nx \in \Gamma_0$). Then there exist a finite number of points $\gamma_1, \ldots, \gamma_r \in \Gamma$ and a finite number of Abelian subvarieties $B_1, \ldots, B_r$ of A such that $\gamma_i + B_i \subseteq X$ for all $1 \leq i \leq r$ and*

$$X(\mathbb{C}) \cap \Gamma = \bigcup_{1 \leq i \leq r} \gamma_i + (B_i(\mathbb{C}) \cap \Gamma).$$

We will only need to apply this theorem when A is is the multiplicative group $(\mathbb{C}^*)^n$. A proof of the Lang conjecture in this special case was first given by Laurent [15]. For Γ we will take the group of torsion points of A. In light of our choice for A, Γ is the group of cyclotomic points of $(\mathbb{C}^*)^n$. By Theorem 1, an irreducible hypersurface of $(\mathbb{C}^*)^n$ on which Γ is dense must be a translate of an algebraic subgroup B_i of $(\mathbb{C}^*)^n$ by a point $\gamma_i \in \Gamma$. We therefore need to know what the B_i might look like. This is well known from the basic theory of algebraic groups.

Lemma 1. *Any proper algebraic subgroup G of $(\mathbb{C}^*)^n$ is included in a subgroup defined by an equation of the form*

$$\prod_{i=1}^{n} X_i^{a_i} = 1 \tag{6}$$

where the exponents a_i are in $\mathbb{Z}$ and are not all zero.

For completeness, we sketch the proof of this lemma. Let P be a non identically zero polynomial which vanishes on G. The fact that $P \equiv 0$ on G shows that the monomials occuring in P are linearly dependent (on G). However, any monomial is a character of G, and distinct characters are linearly independent (see for instance [6], p. 102). We conclude that there must exist at least one pair (m, m') of distinct monomials of P such that $m \equiv m'$ on G. This equality yields an equation of the required form. □

In fact, by [6, p. 103], any algebraic subgroup of $(\mathbb{C}^*)^n$ is an intersection of groups of the form (6).

It follows from Lemma 1 and the discussion preceding it that if V is an irreducible hypersurface of $(\mathbb{C}^*)^n$ on which cyclotomic points are dense, it must indeed be defined by an equation of the form (5). Note also that the n hypersurfaces of the form $X_i = 0$, which lie in the complement of $(\mathbb{C}^*)^n$, do not contain any cyclotomic point. We therefore arrive at the following conclusion.

Corollary 1. *Let V be an irreducible hypersurface of $\mathbb{C}^n$. The cyclotomic points are Zariski dense on V iff V is defined by an equation of the form*

$$\prod_{i=1}^{n} X_i^{\beta_i} - \theta \prod_{i=1}^{n} X_i^{\gamma_i} = 0$$

where θ is a root of unity and $\beta_i, \gamma_i \in \mathbb{N}$.

We will also need to use a recent estimate on the height of algebraic numbers (see Lemma 2 below). First, we recall the construction of the height. For any prime number p, the p-adic absolute value on $\mathbb{Q}$ is characterized by the following properties: $|p|_p = 1/p$, and $|q|_p = 1$ if q is a prime number different from p. For any $x \in \mathbb{Q} \setminus \{0\}$, $|x|_p$ can be computed as follows: write $x = p^\alpha y$ where p is relatively prime to the numerator and denominator of y, and $\alpha \in \mathbb{Z}$. Then $|x|_p = 1/p^\alpha$ (and of course $|0|_p = 0$). We denote by $M_\mathbb{Q}$ the union of the set of p-adic absolute values and of the usual (archimedean) absolute value on $\mathbb{Q}$.

Let $d, e \in \mathbb{Z}$ be two non-zero relatively prime integers. By definition, the height of the rational number d/e is $\max(|d|, |e|)$. There is an equivalent definition in terms of absolute values: for $x \in \mathbb{Q}$, $H(x) = \prod_{\nu \in M_\mathbb{Q}} \max(1, |x|_\nu)$. Note in particular that $H(0) = 1$.

More generally, let K be a number field (an extension of $\mathbb{Q}$ of finite degree). The set M_K of *normalized absolute values* is the set of absolute values on K which extend an absolute value of $M_\mathbb{Q}$. For $\nu \in M_K$, we write $\nu|\infty$ if ν extends the usual absolute value, and $\nu|p$ if ν extends the p-adic absolute value. One defines a "relative height" H_K on K by the formula

$$H_K(x) = \prod_{\nu \in M_K} \max(1, |x|_\nu)^{d_\nu}. \qquad (7)$$

Here d_ν is the so-called "local degree". For every p (either prime or infinite), $\sum_{\nu|p} d_\nu = [K : \mathbb{Q}]$. Sometimes, instead of (7) one just writes $H_K(x) = \prod_\nu \max(1, |x|_\nu)$ if it is understood that each absolute value may occur several times (in fact, d_ν times) in the product. The absolute height $H(x)$ of x is $H_K(x)^{1/n}$, where $n = [K : \mathbb{Q}]$. It is independent of the choice of K. In Proposition 1 we will also use the product formula:

$$\prod_{\nu \in M_K} |x|_\nu^{d_\nu} = 1 \qquad (8)$$

for any $x \in K \setminus \{0\}$. More details on absolute values and height functions can be found for instance in [14] or [20]. In the following lemma we work with the logarithmic height $h(x)$, which is defined as the logarithm of the absolute height $H(x)$.

Lemma 2. *Let F be an algebraic number field of degree δ over $\mathbb{Q}$, and θ a root of unity. The logarithmic height of any nonzero algebraic number α that is not a root of unity satisfies*

$$h(\alpha) \geq \frac{c}{d\delta} \left(\frac{\ln(2d\delta)}{\ln\ln(5d\delta)} \right)^{-13}, \qquad (9)$$

where d is the degree of α over $F(\theta)$ and $c > 0$ a universal constant.

Proof. Since $F(\theta)$ is of degree at most δ over $\mathbb{Q}(\theta)$ α is of degree at most $d\delta$ over $\mathbb{Q}(\theta)$. The result follows from Theorem 1.1 of Amoroso-Zannier [1] since the cyclotomic extension $L = \mathbb{Q}(\theta)$ is Abelian over $\mathbb{Q}$ (see for instance [19, Section 8.4]). □

3. A MULTIVARIATE GAP THEOREM

In the following, $\overline{X}$ denotes a tuple of variables $(X_1, \ldots, X_n)$, and $V(P)$ denotes the zero set of a polynomial P.

Let K be a number field and $\nu \in M_K$ a normalized absolute value. We extend ν to $K[\overline{X}]$ by setting $|\sum_i a_i \overline{X}^{\alpha_j}|_\nu = \max_i |a_i|_\nu$. We define a height function on $\overline{\mathbb{Q}}[\overline{X}]$ by the formula $H(f) = \prod_{\nu \in M_K} |f|_\nu^{1/[K:\mathbb{Q}]}$, where K is chosen so that $f \in K[X]$. Note that $H(f)$ is independent of the choice of K. These definitions are natural generalizations of those given by Lenstra for univariate polynomials.

Lemma 3. *Let $f \in \overline{Q}[\overline{X}]$ be a polynomial with k monomials, and let $\theta_1, \ldots, \theta_{n-1}$ be roots of unity. The height of the univariate polynomial $p(X) = f(\theta_1, \ldots, \theta_{n-1}, X)$ satisfies the inequality $H(p) \leq (k - l + 1)H(f)$, where l denotes the number of monomials in p. In particular, we always have $H(p) \leq kH(f)$.*

Proof. Choose the number field K so that $\theta_1, \ldots, \theta_{n-1} \in K$ and $f \in K[\overline{X}]$. We have $|\theta_i|_\nu = 1$ for any $\nu \in M_K$. As a consequence we have $|p|_\nu \leq |f|_\nu$ for any ultrametric absolute value in M_K. For an archimedean absolute value, we have $|p|_\nu \leq (k - l + 1)|f|_\nu$ (each monomial of p "comes" from at most $k - l + 1$ monomials of f). The inequality $H(p) \leq (k - l + 1)H(f)$ follows since there are $[K : \mathbb{Q}]$ archimedean absolute values in M_K. Hence $H(p) \leq kH(f)$ if $l \geq 1$. This inequality also holds true for $l = 0$ since $H(0) = 0$. □

Theorem 2 (multivariate gap theorem). *Let F be an algebraic number field of degree δ over $\mathbb{Q}$, and $f \in F[\overline{X}]$ a multivariate polynomial of the form $f(\overline{X}) = \sum_{j=0}^{t} a_j \overline{X}^{\alpha_j}$.*

Let $P \in F[\overline{X}]$ be a multivariate polynomial of degree d, irreducible in $F[\overline{X}]$. Assume moreover that the cyclotomic points are not dense in $V(P)$.

Let β_j be the exponent of variable X_n in the monomial $\overline{X}^{\alpha_j}$. We assume without loss of generality that the sequence (β_j) is nondecreasing, and assume also that there exists l such that

$$\beta_{l+1} - \beta_l > \frac{d\delta}{c} \left(\frac{\ln(2d\delta)}{\ln\ln(5d\delta)} \right)^{13} \log(t(t+1)\,H(f)), \qquad (10)$$

where $c > 0$ is the absolute constant from Lemma 2.

If P is a factor of f, it is also a factor of the two polynomials $g = \sum_{j=0}^{l} a_j \overline{X}^{\alpha_j}$ and $h = \sum_{j=l+1}^{t} a_j \overline{X}^{\alpha_j}$.

Proof. We first consider the case where P does not depend on variable X_n. This case is completely elementary: if P is a factor f, there are polynomials Q_i such that

$$f(X_1, \ldots, X_n) = P(X_1, \ldots, X_{n-1}) \times \Big(\sum_i Q_i(X_1, \ldots, X_{n-1}) X_n^i \Big).$$

Clearly, $g = \sum_{i \leq \beta_l} P(X_1, \dots, X_{n-1}) Q_i(X_1, \dots, X_{n-1}) X_n^i$, $h = \sum_{i > \beta_l} P(X_1, \dots, X_{n-1}) Q_i(X_1, \dots, X_{n-1}) X_n^i$, and P is a factor of both polynomials.

The case where $P = aX_n$ for some constant $a \in F$ is also easy: if X_n is a factor of f this variable must occur in all monomials of f, so P is obviously a factor of g and h.

The remainder of the proof is devoted to the case where P actually depends on X_n, but P is not of the form aX_n. In $\mathbb{C}[\overline{X}]$, P factors as a product of absolutely irreducible polynomials $P_1, \dots, P_s$. Since the cyclotomic points are not dense in $V(P)$ there exists some P_i, for instance P_1, such that the cyclotomic points are not dense in $V(P_i)$[†].

Let E be the subspace of $\mathbb{C}^n$ spanned by the first $n-1$ coordinate vectors, and $\pi : \mathbb{C}^n \to E$ the orthogonal projection on E. Since P depends on X_n the same is true of P_1, and $\pi(V(P_1))$ is therefore dense in E. We may view P_1 as a polynomial in X_n with coefficients in $\mathbb{C}[X_1, \dots, X_{n-1}]$. Let $Q_1[X_1, \dots, X_{n-1}]$ be its leading monomial. Note that Q_1 is not identically zero. Since the cyclotomic points are dense in E but not in $V(P_1)$, there exists in E a dense set of cyclotomic points $(\theta_1, \dots, \theta_{n-1})$ such that $Q_1(\theta_1, \dots, \theta_{n-1}) \neq 0$ and $\pi^{-1}(\theta_1, \dots, \theta_{n-1}, 0)$ contains a non-cyclotomic point $(\theta_1, \dots, \theta_{n-1}, \alpha) \in V(P_1)$ with $\alpha \neq 0$ (the condition $\alpha \neq 0$ can be enforced thanks to the hypothesis that P is not of the form aX_n).

The algebraic number α is of degree at most d over $F(\theta_1, \dots, \theta_{n-1})$ since it is a root of $X \mapsto P_1(\theta_1, \dots, \theta_{n-1}, X)$ (note that this polynomial is not identically zero since $Q_1(\theta_1, \dots, \theta_{n-1}) \neq 0$). Moreover, by construction α is nonzero and is not a root of unity. Its height therefore satisfies inequality (9). We also have $f(\theta_1, \dots, \theta_{n-1}, \alpha) = 0$ since P_1 is a factor of f. Let us now apply Proposition 1 below to the univariate polynomial $p(X) = f(\theta_1, \dots, \theta_{n-1}, X)$. We have $H(p) \leq (t+1)H(f)$ by Lemma 3. In view of (9) and (10), we conclude that $g(\theta_1, \dots, \theta_{n-1}, \alpha) = h(\theta_1, \dots, \theta_{n-1}, \alpha) = 0$.

We have shown that $g(\theta_1, \dots, \theta_{n-1}, \alpha) = 0$ for a set of points $(\theta_1, \dots, \theta_{n-1}, \alpha)$ that is dense in $V(P_1)$. This implies that $V(P_1) \subseteq V(g)$, and that P_1 is a factor of g since P_1 is squarefree. The polynomials g and P therefore have a nontrivial common divisor. Since P is irreducible over F, P must be a factor of g and it is also of course a factor of h. □

Remark 1. In the above theorem the hypothesis that P is irreducible in $F[\overline{X}]$ is stronger than needed: it is sufficient to assume that P is squarefree. □

The proof of the following proposition is essentially the same as the proof of Proposition 2.3 of [16].

Proposition 1. *Let $p \in \overline{\mathbb{Q}}[X]$ be a polynomial with at most $t+1$ non-zero terms. Assume that p can be written as the sum of two polynomials q and r where each monomial of q has degree at most β and each monomial of r has degree at least γ. Let $x \in \overline{\mathbb{Q}}^*$ be a root of p that is not a root of unity. If $\gamma - \beta > \log(t\, H(p)) / \log H(x)$ then x is a common root of q and r.*

Proof. We may assume that each of the two polynomials q and r collects at most t of the $t+1$ terms of p (otherwise, the result is clear). Assume by contradiction that $q(x) \neq 0$. Pick a number field K which contains x as well as the coefficients of p, and let $\nu \in M_K$. If $|x|_\nu \geq 1$, each term $a_j x^{\beta_j}$ of $q(x)$ satisfies $|a_j x^{\beta_j}| \leq |p|_\nu |x|_\nu^\beta$, therefore

$$|q(x)|_\nu \leq \max(1, |t|_\nu) |p|_\nu |x|_\nu^\beta \quad \text{if} \quad |x|_\nu \geq 1.$$

A similar argument shows that

$$|r(x)|_\nu \leq \max(1, |t|_\nu) |p|_\nu |x|_\nu^\gamma \quad \text{if} \quad |x|_\nu \leq 1.$$

We have $|q(x)|_\nu = |r(x)|_\nu$, so we can combine these two statements in

$$\max(1, |x|_\nu|)^{\gamma - \beta} \cdot |q(x)|_\nu \leq \max(1, |t|_\nu) \cdot |p|_\nu \cdot |x|_\nu^\gamma.$$

Raise this to the power $d_\nu / [K : \mathbb{Q}]$ and take the product over $\nu \in M_K$. Using the fact that $H(t) = t$, and applying the product formula to $q(x)$ and x (which are both supposed to be nonzero) one finds that $H(x)^{\gamma - \beta} \leq t \cdot H(p)$. This is in contradiction with the hypothesis on $\gamma - \beta$. □

4. FACTORS FOR WHICH CYCLOTOMIC POINTS ARE NOT DENSE

In this section we describe an algorithm based on the multivariate gap theorem which, given a supersparse polynomial $f \in F[\overline{X}]$ and an integer d, finds (up to a constant factor) all the factors P of f such that P is irreducible in $F[X]$, $\deg(P) \leq d$, and the cyclotomic points are not dense on $V(P)$. The algorithm also finds the multiplicities of these factors and runs in time polynomial in $(d+l)^n$, where l denotes the length of the input data.

In order to actually implement the algorithm, an explicit knowledge of the real number c of our multivariate gap Theorem 2 above is needed. In [1] no explicit estimate is provided, and if an explicit analysis brings c too close to zero our algorithm would be quite impractical. For the case of linear factors of bivariate polynomials with integer coefficients, reasonable explicit estimates are provided in [11].

Like Lenstra [16] we will use an upper bound for $H(f)$ in Theorem 2. The coefficients a_i of the input polynomial

$$f(X_1, \dots, X_n) = \sum_{i=0}^{t} a_i \, X_1^{\alpha_{i,1}} \cdots X_n^{\alpha_{i,n}} \tag{11}$$

are from a number field $F = \mathbb{Q}[\zeta]/(\varphi(\zeta))$. We shall multiply the coefficients of all residue polynomials by the common rational integer denominator and thus assume without loss of generality that all $a_i(\zeta)$ have integral coefficients. Since φ (of degree δ) is assumed monic, any root ζ is an algebraic integer. Therefore all $a_i(\zeta)$, which are now members in the ring of algebraic integers, are themselves algebraic integers in F.

Lemma 4. *Let $a_i = \sum_{j=0}^{\delta-1} a_{i,j} \zeta^j$ for all $0 \leq i \leq t$, where $a_{i,j} \in \mathbb{Z}$ and let B be an upper bound for the absolute value of any complex root of φ. Then*

$$H(f) \leq \max_i \sum_{j=0}^{\delta-1} |a_{i,j}| B^j. \tag{12}$$

Proof. The proof is an immediate multivariate generalization of Proposition 3.6 in [16]. □

The algorithm proceeds in three steps.

1. Compute an integer G that is an upper bound for the right side gap estimate in (10) using (12).

[†] the irreducibility of P in $F[X]$ implies that the cyclotomic points are not dense on any of the varieties $V(P_i)$, but we do not need to explicitly use this fact.

2. Split f into $g^{[1]} + \cdots + g^{[s]}$ such that for all k with $1 \leq k \leq s$ the summand
$$g^{[k]} = \sum_{i=0}^{t^{[k]}} a_{k_i}\, X_1^{\alpha_{k_i,1}} \cdots X_n^{\alpha_{k_i,n}}$$
has the following property for all j with $1 \leq j \leq n$: If the term degrees of variable X_j,
$$\alpha_{k_0,j}, \alpha_{k_1,j}, \alpha_{k_2,j}, \ldots$$
are sorted in ascending order, any two adjacent degrees are apart by no more than the gap estimate G.

   We can determine all $g^{[k]}$ by first splitting at degree gaps $\geq G$ in X_n, and then proceeding iteratively with the remaining variables $X_{n-1}, \ldots, X_1$ on all parts produced.

3. Compute all irreducible factors over F of degree $\leq d$ of
$$\bar{g}^{[k]} = g^{[k]} \big/ \big(X_1^{\min_i(\alpha_{k_i,1})} \cdots X_n^{\min_i(\alpha_{k_i,n})}\big).$$
   Return those irreducible factors that are common to all $\bar{g}^{[k]}$.

Theorem 3. *The above algorithm returns all irreducible factors of f over F that are not divisible by a polynomial of the form $\prod_{i=1}^n X_i^{\beta_i} - \theta \prod_{i=1}^n X_i^{\gamma_i}$, where θ is a root of unity, in $(\mathrm{size}(f)+d)^{O(n)}$ deterministic bit operations. The factors themselves are represented as supersparse polynomials.*

Proof. Referring to Lemma 4, let $\eta = \max(\|f\|_\infty, \|\varphi\|_\infty)$, where $\|f\|_\infty = \max_{i,j} |a_{i,j}|$. Note that $\log(\eta) = O(\mathrm{size}(f))$. We have the rough estimate $B \leq (\delta+1)\eta$, hence by Theorem 2 we can compute a G with $G = (\mathrm{size}(f)+d)^{O(1)}$. Because for all j we have $\deg_{X_j}(\bar{g}^{[k]}) \leq tG$ the factorization Step 3 can be carried out in $(\mathrm{size}(f)+d)^{O(n)}$ bit operations by algorithms for dense polynomials. □

Remark 2. We assume that our algorithm returns the factors in supersparse representation (11). For that representation, it is known that the size of the output cannot be bounded by $(\mathrm{size}(f)+d)^{O(1)}$ [3, Example 5.1]. However, $(\mathrm{size}(f)+d)^{O(1)}$ bit complexity is achievable if the factors are returned in straight-line program [9] or black box representation [13]. The algorithm is then randomized of the Monte Carlo kind.

By contrast, the algorithm of Section 5 has bit complexity $(\mathrm{size}(f)+d)^{O(1)}$ without appealing to randomization, and outputs factors in supersparse representation. □

Remark 3. The multiplicities of all irreducible factors can be determined by computing the factors of generalized partial derivatives as in [16, Proposition 3.2]: for f as in (11), let
$$D_j^{[1]}(f) = \frac{\partial}{\partial X_j}\Big(\frac{f}{X_j^{\min_i \alpha_{i,j}}}\Big), \quad D_j^{[k]}(f) = D_j^{[1]}(D_j^{[k-1]} f).$$
Then an irreducible factor $h \neq X_j$ of multiplicity μ with $\deg_{X_j}(h) \geq 1$ must divide all $D_j^{[k]}(f)$ for $1 \leq k \leq \mu - 1$. Let $t_j \leq t+1$ be the number of distinct term degrees in the list $\alpha_{0,j}, \alpha_{1,j}, \ldots, \alpha_{t,j}$. The polynomial $D_j^{[t_j-1]}(f)$ has a single power of X_j in its terms and is not divisible by h. Therefore we must have $\mu < t_j$, and μ can be computed by factoring the additional supersparse polynomials $D_j^{[k]}(f)$ for $k = 1, \ldots, t_j - 1$. □

5. FACTORS FOR WHICH CYCLOTOMIC POINTS ARE DENSE

In this section we describe an algorithm which, given a supersparse polynomial $f \in F[\overline{X}]$ and an integer d, finds (up to a constant factor) all the factors $P \in F[\overline{X}]$ of f such that P is irreducible in $F[\overline{X}]$, $\deg(P) \leq d$, and the cyclotomic points are dense on $V(P)$. The algorithm also finds the multiplicities of these factors and runs in time polynomial in $d + l$, where l denotes the length of the input data. It proceeds by reduction to the univariate case.

Let $P \in \mathbb{C}[\overline{X}]$ be an absolutely irreducible factor of f such that cyclotomic points are dense on $V(P)$. By Corollary 1, up to a constant multiplicative factor P must be of the form
$$P_{\beta,\gamma,\theta} = P(X_1, \ldots, X_n) = \prod_{i=1}^n X_i^{\beta_i} - \theta \prod_{i=1}^n X_i^{\gamma_i} \tag{13}$$
where θ is a root of unity and $\beta_i, \gamma_i \in \mathbb{N}$. We have that
$$\left.\begin{array}{l} \forall i, 1 \leq i \leq n\colon \beta_i = 0 \vee \gamma_i = 0 \\ \text{and } \mathrm{GCD}_{1\leq i\leq n}(\beta_i - \gamma_i) = 1, \end{array}\right\} \tag{14}$$
for otherwise P would be reducible.

First, we shall assume that $(\beta_n, \gamma_n) \neq (0,0)$, i.e., P actually depends on X_n. The following lemma shows how to determine candidate degree vectors (β, γ).

Lemma 5. *Let P of the form (13) be an absolute irreducible factor of f and let $\overline{\alpha_\tau} = (\alpha_{\tau,0}, \ldots, \alpha_{\tau,n})$ be the degree vector of the τ-th term in f, where $0 \leq \tau \leq t$. Suppose that $(\beta_n, \gamma_n) \neq (0,0)$. Then there must exist an integer index j with $1 \leq j \leq t$ and an integer k with $k = \pm\,\mathrm{GCD}_{1\leq i\leq n}(\alpha_{0,i} - \alpha_{j,i})$ such that $\alpha_{0,n} \neq \alpha_{j,n}$ and $\forall i, 1 \leq i \leq n\colon \gamma_i - \beta_i = (\alpha_{0,i} - \alpha_{j,i})/k$.*

Proof of Lemma 5. Let λ be such that $\lambda^{\beta_n - \gamma_n} = \theta$. From (13) and the assumption that P is a factor of f we obtain
$$X_n = \lambda \Big(\prod_{i=1}^{n-1} X_i^{\gamma_i - \beta_i}\Big)^{\frac{1}{\beta_n - \gamma_n}} \tag{15}$$
as a root of f in the algebraic closure of $F[X_1, \ldots, X_{n-1}]$. Expanding the 0-th term $\prod_{i=1}^n X_i^{\alpha_{0,i}}$ of f in the Puiseux series field in $X_1, \ldots, X_{n-1}$ at that root, we obtain $\alpha_{0,i} + \frac{\gamma_i - \beta_i}{\beta_n - \gamma_n}\alpha_{0,n}$ as the fractional exponent of X_i, where $1 \leq i \leq n-1$. That term must be cancelled by another term, say term j with $1 \leq j \leq t$, since (15) is a root of f. The fractional exponents must agree, so we have for all i with $1 \leq i \leq n-1$
$$\alpha_{0,i} - \alpha_{j,i} = \frac{\beta_i - \gamma_i}{\beta_n - \gamma_n}(\alpha_{0,n} - \alpha_{j,n}). \tag{16}$$
Now $\alpha_{0,n} \neq \alpha_{j,n}$ because $\alpha_{0,i} \neq \alpha_{j,i}$ for some i, for otherwise all degrees of the j-th term would agree with those of term 0. Immediately, we conclude that then we also have $\beta_i = \gamma_i$ if and only if $\alpha_{0,i} = \alpha_{j,i}$. Let $k = (\alpha_{0,n} - \alpha_{j,n})/(\beta_n - \gamma_n)$. We have
$$\forall i, 1 \leq i \leq n\colon k(\beta_i - \gamma_i) = \alpha_{0,i} - \alpha_{j,i}. \tag{17}$$
Identity (17) implies that k is integral, for if k were not, its denominator would be a common divisor of $\beta_i - \gamma_i$, but those

are relatively prime as by (14). Furthermore, k is a divisor of all $\alpha_{0,i} - \alpha_{j,i}$, and again by the relatively primeness of $\beta_i - \gamma_i$ we conclude from (17) that it must be the greatest common divisor. □

Clearly, Lemma 5 yields a straight-forward method to compute candidate factor exponent vectors (β, γ). Factors P with $(\beta_n, \gamma_n) = (0, 0)$, i.e., those that do not depend on X_n, are also covered by Lemma 5. In that case, for a given j there exists an i such that $\alpha_{0,i} \neq \alpha_{j,i}$ and X_i can assume the role of X_n. Overall, there are at most $2t$ candidate vectors, including the case $(\beta_n, \gamma_n) = (0, 0)$. As a side remark, note that if (β, γ) is one of the pairs, the second pair associated to the same $j \in \{1, \ldots, t\}$ is equal to (γ, β). Since $P_{\gamma,\beta,\theta} = -\theta P_{\beta,\gamma,1/\theta}$ the second pair does not contribute a new factor. The $2t$ pairs can be determined in time polynomial in l', where l' denotes the length of the representation of the tuples of exponents $\overline{\alpha_0}, \ldots, \overline{\alpha_t}$ of f. In particular, the algorithm is completely independent of the field of definition of f, or of the actual values of its coefficients $a_0, \ldots, a_t$.

Now, for each pair (β, γ) we would like to obtain more information on the (possibly empty) set of complex numbers θ such that $P_{\beta,\gamma,\theta}$ is a factor f. This can be done by building on the idea of Lemma 5. Namely, perform the same substitution of variables as in the proof of that lemma, and express the fact that the resulting (finite) Puiseux series is identically zero. After substitution, each monomial of f becomes a monomial in the variables X_i $(1 \leq i \leq n-1)$ with coefficient equal to $a_j \lambda^{\alpha_{j,n}}$. We therefore obtain a system of at most t sparse polynomial equations in the indeterminate λ with coefficients in F. Each polynomial in this system is a sum of at most $t+1$ monomials.

Lemma 6. *A complex number $\lambda \neq 0$ is a solution of the above system if and only if $P_{\beta,\gamma,\theta}$ is a factor of f, where $\theta = \lambda^{\beta_n - \gamma_n}$.*

Proof. By construction, λ is a solution iff f vanishes on $V(P_{\beta,\gamma,\theta}) \cap (\mathbb{C}^*)^n$. The hypothesis $\lambda \neq 0$ implies that $(\mathbb{C}^*)^n$ is dense in $V(P_{\beta,\gamma,\theta})$, so in fact λ is a solution iff f vanishes on $V(P_{\beta,\gamma,\theta})$. Since $P_{\beta,\gamma,\theta}$ is irreducible, f vanishes on $V(P_{\beta,\gamma,\theta})$ iff $P_{\beta,\gamma,\theta}$ is a factor of f. □

Now we would like to determine the roots of unity θ such that θ is of degree at most d over F, and $P_{\beta,\gamma,\theta}$ is a factor of f. As usual, we "determine a root" by computing its minimal polynomial over F. Since we are interested only in roots of unity, our minimal polynomials will be cyclotomic polynomials, that is, factors of polynomials of the form $X^r - 1$ that are monic and irreducible in $F[X]$ (this definition of a cyclotomic polynomial, borrowed from [16], agrees with the traditional definition in the case $F = \mathbb{Q}$). This can be done as follows:

1. Construct the sparse system defined before Lemma 6, and for each polynomial in the system find all its cyclotomic factors in $F[X]$ which are of degree at most $d \cdot |\beta_n - \gamma_n|$.

2. From the set of factors computed at step 1, keep only those polynomials that are factors of all polynomials in the system. Call I the set of remaining factors.

3. For each polynomial $m \in I$ compute the minimal polynomial $M \in F[X]$ of $\theta = \lambda^{\beta_n - \gamma_n}$, where λ denotes a root of m. Output M if it is of degree $\leq d$.

We call *weight* of a pair (β, γ), and denote by $w(\beta, \gamma)$, the quantity $\max(\sum_{i=1}^n \beta_i, \sum_{i=1}^n \gamma_i)$. This is nothing but the degree of $P_{\beta,\gamma,\theta}$, for any $\theta \neq 0$.

Proposition 2. *The above algorithm computes the minimal polynomials of all roots of unity θ such that $P_{\beta,\gamma,\theta}$ is a factor of f, and θ is of degree at most d over F. Its running time is polynomial in $d + l + w(\beta, \gamma)$, where l denotes the length of the sparse representation of f.*

Proof. If θ is of degree at most d over F and $\theta = \lambda^{\beta_n - \gamma_n}$, λ is of degree at most $d|\beta_n - \gamma_n|$ over F. The correctness of the algorithm therefore follows from Lemma 6. Step 1 can be performed within the claimed time bound by [16, Proposition 3.5]. In step 2 we simply compute an intersection of sets, and step 3 is standard. □

We will appeal to this proposition only for pairs of weight at most d (otherwise, the resulting factors of f would be of degree higher than d). For such pairs, the algorithm runs in time polynomial in $d + l$.

By appealing to Lenstra's main theorem instead of his Proposition 3.5 [16], we could as easily compute the set of all complex numbers $\theta \neq 0$ such that θ is of degree at most d over F, and $P_{\beta,\gamma,\theta}$ is a factor of f. There would be some overlap with the factors computed in section 4.

Generalized cyclotomic polynomials

In order to fulfill our goal of finding factors for which cyclotomic points are dense, it is useful to know what those factors can possibly look like. We already know what the *absolutely irreducible* factors look like: they are of form (13), and the auxiliary algorithm described above supplies us with a list of candidates for the pair (β, γ). We are, however, looking for factors that are only irreducible over F. Let P be an absolutely irreducible polynomial of form (13), and let m be the minimal polynomial of θ over F. From now on we assume that θ is a root of unity. The polynomial m is therefore a cyclotomic polynomial, i.e., as explained after Lemma 6, a factor of a polynomial of the form $X^r - 1$ that is irreducible in $F[X]$.

We have the following:

Proposition 3.

$$Q_{\beta,\gamma} = \mathrm{Norm}_{F(\theta)/F}(P_{\beta,\gamma,\theta}) = \prod_{\theta_i\,:\, m(\theta_i)=0} P_{\beta,\gamma,\theta_i}$$

is an irreducible factor of f in $F[\overline{X}]$.

Proof. The norm of an irreducible polynomial over an algebraic extension is a pure power of an irreducible polynomial over the ground field. Since the argument is brief, we shall give it. Suppose $Q_{\beta,\gamma} = Q_1 Q_2$ where Q_1 and Q_2 are relatively prime polynomials over F, and suppose $P_{\beta,\gamma,\theta}$ is a factor of Q_1 over $F(\theta)$. There exists a j such that $P_{\beta,\gamma,\theta_j}$ is a factor of Q_2 over $F(\theta_j)$, which is an isomorphic copy of $F(\theta)$, in the latter of which the division of Q_2 by $P_{\beta,\gamma,\theta}$ leaves again no remainder. Thus Q_1 and Q_2 have a common factor $P_{\beta,\gamma,\theta}$ and cannot be relatively prime.

Now suppose without loss of generality that $\beta_n > 0$ and let $\bar{Q}(X_n) = Q_{\beta,\gamma}(1, \ldots, 1, X_n) = \prod_{\theta_i\,:\, m(\theta_i)=0}(X_n^{\beta_n} - \theta_i)$. Since $m(X)$ divides $X^r - 1$, $\bar{Q}(X_n)$ divides $X_n^{\beta_n r} - 1$, which is squarefree, and therefore $Q_{\beta,\gamma}$ cannot have a multiple factor. □

Note that $Q_{\beta,\gamma}$ is of degree $\deg(m) \cdot w(\beta,\gamma)$ and can be computed from m via a substitution. Let $m(z)$ be the minimal polynomial of θ over F. Then

$$Q_{\beta,\gamma} = \Big(\prod_{i=1}^{n} X_i^{\gamma_i}\Big)^{\deg(m)} \cdot m\Big(\prod_{i=1}^{n} X_i^{\beta_i} \Big/ \prod_{i=1}^{n} X_i^{\gamma_i}\Big). \qquad (18)$$

We can finally describe the main algorithm of section 5.

1. Enumerate all candidate pairs (β, γ).

2. For each candidate pair of weight $w(\beta,\gamma) \leq d$, use the algorithm of Proposition 2 to compute the minimal polynomials of all roots of unity θ such that $P_{\beta,\gamma,\theta}$ is a factor of f, and θ is of degree at most d over F. For each such minimal polynomial m, if $\deg(m) \cdot w(\beta,\gamma) \leq d$ output the factor $Q_{\beta,\gamma}$ of f defined by (18).

As explained after Lemma 5, there are at most $2t$ candidate pairs and they can be computed in time polynomial in the input size. The correctness of the algorithm and the running time claimed at the beginning of section 5 then follow from Propositions 2 and 3.

The multiplicities of all generalized cyclotomic factors can again be determined as in Remark 3.

Acknowledgments: We would like to thank Eric Schost for helpful remarks. Martin Sombra has provided us with some references to the literature of the Bogomolov property of algebraic numbers. We also thank the four reviewers for their helpful comments.

Note added February 3, 2006: On January 24, 2006, Teresa Krick and Martin Sombra have sent us a paper of theirs that contains polynomial-time algorithms similar to ours for the case of two variables. In June and September of 2005, we had by email kept Teresa appraised about the results in this paper, which were also mentioned in Kaltofen's talk in July 2005 at ISSAC in Beijing.

6. REFERENCES

[1] Amoroso, F., and Zannier, U. A relative Dobrowolski lower bound over Abelian varieties. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) XXIX* (2000), 711–727.

[2] Cucker, F., Koiran, P., and Smale, S. A polynomial time algorithm for diophantine equations in one variable. *J. Symbolic Comput. 27*, 1 (1999), 21–29.

[3] von zur Gathen, J., and Kaltofen, E. Factoring sparse multivariate polynomials. *J. Comput. System Sci. 31* (1985), 265–287.

[4] Győry, K., Iwaniec, H., and Urbanowicz, J., Eds. *Number Theory in Progress* (1999), vol. 1 Diophantine Problems and Polynomials, Stefan Banach Internat. Center, Walter de Gruyter Berlin/New York. Proc. Internat. Conf. Number Theory in Honor of the 60th Birthday of Andrzej Schinzel, Zakopane, Poland June 30–July 9, 1997.

[5] Hindry, M., and Silverman, J. H. *Diophantine Geometry: An Introduction*. Springer Verlag, Heidelberg, Germany, 2000.

[6] Humphreys, J. E. *Linear Algebraic Groups*. Springer Verlag, New York, 1975.

[7] Kaltofen, E. Polynomial-time reductions from multivariate to bi- and univariate integral polynomial factorization. *SIAM J. Comput. 14*, 2 (1985), 469–489.

[8] Kaltofen, E. Greatest common divisors of polynomials given by straight-line programs. *J. ACM 35*, 1 (1988), 231–264.

[9] Kaltofen, E. Factorization of polynomials given by straight-line programs. In *Randomness and Computation*, S. Micali, Ed., vol. 5 of *Advances in Computing Research*. JAI Press Inc., Greenwhich, Connecticut, 1989, pp. 375–412.

[10] Kaltofen, E. Polynomial factorization 1987-1991. In *Proc. LATIN '92* (Heidelberg, Germany, 1992), I. Simon, Ed., vol. 583 of *Lect. Notes Comput. Sci.*, Springer Verlag, pp. 294–313.

[11] Kaltofen, E., and Koiran, P. On the complexity of factoring bivariate supersparse (lacunary) polynomials. In *ISSAC'05 Proc. 2005 Internat. Symp. Symbolic Algebraic Comput.* (New York, N. Y., 2005), M. Kauers, Ed., ACM Press, pp. 208–215. ACM SIGSAM's ISSAC 2005 Distinguished Paper Award.

[12] Kaltofen, E., and Lee, W. Early termination in sparse interpolation algorithms. *J. Symbolic Comput. 36*, 3–4 (2003), 365–400. Special issue Internat. Symp. Symbolic Algebraic Comput. (ISSAC 2002). Guest editors: M. Giusti & L. M. Pardo.

[13] Kaltofen, E., and Trager, B. Computing with polynomials given by black boxes for their evaluations: Greatest common divisors, factorization, separation of numerators and denominators. *J. Symbolic Comput. 9*, 3 (1990), 301–320.

[14] Lang, S. *Algebra*. Addison-Wesley, 1993.

[15] Laurent, M. Equations diophantiennes exponentielles. *Inventiones Mathematicae 78*, 2 (1984), 299–327.

[16] Lenstra, Jr., H. W. Finding small degree factors of lacunary polynomials. In Győry et al. [4], pp. 267–276.

[17] Lenstra, Jr., H. W. On the factorization of lacunary polynomials. In Győry et al. [4], pp. 277–291.

[18] Plaisted, D. A. New NP-hard and NP-complete polynomial and integer divisibility problems. *Theoretical Comput. Sci. 13* (1984), 125–138.

[19] van der Waerden, B. L. *Moderne Algebra*. Springer Verlag, Berlin, 1940. English transl. publ. under the title "Modern algebra" by F. Ungar Publ. Co., New York, 1953.

[20] Waldschmidt, M. *Diophantine approximation on linear algebraic groups*. Springer Verlag, Heidelberg, Germany, 2000.

[21] Zippel, R. E. *Probabilistic algorithms for sparse polynomials*. PhD thesis, Massachusetts Inst. of Technology, Cambridge, USA, Sept. 1979.