Записки научных
семинаров ПОМИ
Том 258, 1999 г.

André Galligo

REAL FACTORIZATION OF MULTIVARIATE POLYNOMIALS WITH INTEGER COEFFICIENTS

Abstract. In the recent works ([4, 15]), a new efficient probabilistic semi-numerical absolute (i.e. complex) factorization algorithm for multivariate polynomials with integer coefficients is given. It was based on a simple property of the monomials appearing after a generic linear change of coordinates for bivariate polynomials and a deep result of complex algebraic geometry.

Here we consider the a priori simpler problem of factorization over the field of real numbers. We briefly review our algorithm for complex factorization and adapt it to solve the problem on the reals. This allows to spare a significant part of the computations and improve the range of tractability. The method provides factors with approximative coefficients and eventually exact factors in a suitable real algebraic extension of the field **Q**.

1. Introduction

Absolute factorization of polynomials with coefficients in a number field k means factorization over the algebraic closure of k. This is a problem which has attracted much attention in the domain of Computer Algebra as we see from the works e.g. [1, 2, 7, 5, 9, 10, 11].

Here we restrict to the case $k = \mathbf{Q}$ (so absolute means complex) and when the input polynomial P is irreducible over $\mathbf{Q}$. This last condition is not really a restriction since efficient polynomial factorization (over $\mathbf{Z}$) program are available in most computer algebra systems. Also, we further restrict our discussion to the bivariate case because it already contains most of the difficulties, and it is the main step toward the general case, see e.g. [13].

The real factorization can be deduced from the complex one by multiplying conjugate factors. However complex factorization is not an easy process, so we raise the question to compute directly, at a smaller computational cost, a real factorization.

We propose to adapt to the real case, the method of complex factorization developed in [15].

Although there exist algorithms with a polynomial complexity, our method rely on an extensive search with an exponential complexity. This is acceptable since our aim is principally practical: we wish to extend the range for which factorization over $\mathbf{R}$ may be applied. For "medium" degrees, say around 50, our method becomes quite efficient because the calculations are rather straightforward and most of them can be performed using floating point approximations.

Our approach is probabilistic as it requires the choices of a generic linear change of coordinates and of a new origin. We use a geometric point of view on the plane complex curve $\mathcal{C}$ defined by the input polynomial P and try to deduce qualitative results from computations with floating point numbers.

The first step of our method provides a primality test over $\mathbf{R}$, this is a refinement of the absolute primality test we presented at ISSAC'97 [4] .

We note that recently, J. F. Ragot gave in his thesis [8] a fast absolute primality test which is also probabilistic but based on modular techniques, however it seems that it cannot be easily adapted to a primality test over $\mathbf{R}$.

Our factorization process is divided in three steps. First, we find the partition of the (simple) points of a generic fiber of a projection of $\mathcal{C}$ by the real irreducible components. This also determines the number and degrees of possible factors. Second, by Hensel lemma, we construct polynomials P_j with $j = 1, \ldots, s$ with floating point coefficients such that the factorization obtained is valid to some precision. Third, we construct candidate polynomial factors with exact coefficient in an algebraic extension of $\mathbf{Q}$ and test by a division the validity of the factorization.

The remainder of this article is structured as follows: Section 2 describes our basic observation: the presence of factors implies the vanishing of certain sums of coefficients, then we get the sufficiency of this condition in the generic case by an elaborate geometric argument. Section 3 gives a test of real primality based on computations with a limited precision together with an error analysis needed to validate the test. Section 4 presents the process of detection and calculation of approximated factors followed by a construction of candidate exact factors with coefficients in an algebraic extension of $\mathbf{Q}$.

2. Preparation and genericity

Given a polynomial $P \in \mathbf{Q}[X_1, \ldots, X_n]$ by Hibert irreducibility theorem, P is irreducible iff for a generic hyperplane parametrized by

$X_i = a_i x + b_i y + c_i$ for $i = 1$ to n with a_i, b_i, c_i in $\mathbf{Q}$ the bivariate polynomial $p(x, y) = P(\ldots, a_i x + b_i y + c_i, \ldots)$ is irreducible. Moreover if P and p are reducible their irreducible factors say P_j and p_j correspond bijectively and P_j can be recovered from p_j and P by applying Hensel Lemma. We refer to [13] for this procedure and we will assume from now on that P is a bivariate polynomial.

2.1. Basic facts.

We suppose that P is irreducible over $\mathbf{Q}$.

As the ring $\mathbf{C}[x, y]$ is factorial, the polynomial P admits an irreducible decomposition $P = P_1 ... P_s$ monic in y. As P has integer coefficients, let us denote by K the smallest extension of $\mathbf{Q}$ which contains all the coefficients of P_1, it is a finite algebraic extension.

Let G denote the Galois group of K over $\mathbf{Q}$. As P_1 divides P, for any element $\sigma \in G$ then the conjugate σP_1 also divides P, therefore σP_1 is one of the divisors P_j of P with $j = 1$ to s. The minimality of K hence of G, implies that all these conjugates of P_1 are distinct polynomials.

Let R denote the product of all these conjugate, it is a polynomial with rational coefficients which divides P, if P is irreducible over $\mathbf{Q}$, then $P = R$.

So with our hypothesis , all the s complex factors P_j of P have the same degree n/s and are conjugate by the action of a Galois group G whose cardinal is a divisor s of the degree of P.

For any j as P_j divides P, by complex conjugation so does $\bar{P}_j$. Therefore there exists an integer $t \leq s/2$ such that $P = \prod_{j=1}^{t} P_j \bar{P}_j \cdot \prod_{l=2t+1}^{s} P_l$.

Setting $Q_j = P_j \bar{P}_j$, we get the factorization of P over $\mathbf{R}$. It has $s - t$ factors, t of degree $2n/s$ and $s - 2t$ of degree n/s:

$$P = \prod_{j=1}^{t} Q_j \cdot \prod_{l=2t+1}^{s} P_l \,.$$

Now, we recall a basic lemma:

A linear change of coordinates of the form $x := X + hY; y := Y$ transforms a polynomial $P(x, y)$ of total degree n into a new polynomial $Q(X, Y)$ also of total degree n, which can be written

$$Q(X, Y) = A^0(h)Y^n + A^1(h, X)Y^{n-1} + \ldots$$

where A^0 is a non-zero polynomial of degree less than or equal to n.

We suppose that we have performed a linear change of coordinates such that the polynomial Q is monic in Y of degree n, and write $P(x,y)$ instead of $Q(X,Y)$.

Lemma. *If a polynomial $P(x,y)$, monic in y and such that the degree in y (denoted by n) is the total degree, admits a factorization $P = P_1 \ldots P_s$, then each P_j for $j = 1, \ldots, s$ can be taken monic in y and such that its degree in y (denoted by n_j) is its total degree. Hence, it can be written:*

$$P_j = y^{n_j} + A_j^1 y^{n_j - 1} + A_j^2 y^{n_j - 2} + \ldots + A_j^{n_j}$$

where for each $k = 1, \ldots, n_j$, A_j^k is a polynomial in x of degree less than or equal to k.

2.2. A non-vanishing result.

As P is a square-free polynomial, for all values $x_0 \in \mathbf{Q}$ except at most $n(n-1)$ of them, the complex roots $y_1, \ldots, y_n$ of $P(x_0, y)$ are distinct and all the corresponding points (x_0, y_j) of the complex curve $\mathcal{C}$, defined by P, are smooth, for $j = 1, \ldots, n$.

We fix such a point (x_0, y_j) set $x = x_0 + X$ and let $y = \varphi_j(X)$ denote the implicit function defined by P near by (x_0, y_j). Hence, $P(x_0 + X, \varphi_j(X)) = 0$. We compute a Taylor expansion of order 2 near by $X = 0$ of this function:

$$\varphi_j(X) = y_j + a_j X + b_j X^2 + o(X^2)$$

this defines the coefficients a_j, b_j which are rational function of x_0, y_j and therefore belong to the algebraic extension of $\mathbf{Q}$ containing y_j, their algebraic expressions are the following.

$$\begin{aligned} a_j &= -\alpha_j / \beta_j \\ b_j &= -\frac{\gamma_j}{\beta_j} - \delta_j \frac{\alpha_j^2}{\beta_j^3} + \varepsilon_j \frac{\alpha_j}{\beta_j^2}\ , \end{aligned}$$

where

$$\begin{aligned} &\alpha_j = \frac{\partial P}{\partial x}(x_0, y_j), \quad \beta_j = \frac{\partial P}{\partial y}(x_0, y_j) \\ &\gamma_j = \frac{1}{2}\frac{\partial^2 P}{\partial x^2}(x_0, y_j), \quad \delta_j = \frac{1}{2}\frac{\partial^2 P}{\partial y^2}(x_0, y_j), \quad \varepsilon_j = \frac{\partial^2 P}{\partial x \partial y}(x_0, y_j). \end{aligned}$$

All these quantities, can be expressed as polynomials in y_j of degrees $< 3n$, and we recall that y_j is a solution of $P(x_0, y) = 0$.

Now as detection and discard of degree 1 factors of P is an easier task than complete factorization, see [15], we set:

Hypothesis 1. We assume that P is monic in y such that its degree in y is its total degree and P does not admit A factor of degree 1.

Proposition 1. (cf. [15])

Let P be as in Hypothesis 1, then for almost all (rational) values of x_0, all coefficients b_j with $1 \leq j \leq n$ are non-zero.

2.3. Genericity.

We now analyze the generic behavior of the sums $B_I = b_{i_1} + \ldots + b_{i_m}$ for $I = \{i_1, \ldots, i_m\}$.

First, we suppose that the polynomial P satisfies Hypothesis 1 and is irreducible in $\mathbf{C}[x, y]$.

Definition 1. *We denote by $\mathbf{P}^2(C)^*$ the set of all projective lines in $\mathbf{P}^2(C)$, it is called the dual projective plane. Let $\mathcal{C}$ be a projective curve of degree n in $\mathbf{P}^2(C)$. We call incidence curve of $\mathcal{C}$, and denote it by $\mathcal{C}^*$, the closed subset of $\mathbf{P}^2(C)^*$ consisting of all lines either tangent to $\mathcal{C}$ or passing by a singular point of $\mathcal{C}$.*

Therefore any lines in $\mathbf{P}^2(C)^* - \mathcal{C}^*$ intersects $\mathcal{C}$ in distinct simple points. The following result, due to J. Harris (cf [6]), is used to prove our main result.

Theorem 1 (J. Harris)**.**

Let $\mathcal{C}$ be a projective irreducible plane curve and let $\mathcal{C}^*$ denote its incidence curve in the dual plane. Then the first homotopy group $\Pi_1(\mathbf{P}^2(C)^* - \mathcal{C}^*)$ acts as the symmetric group on any smooth linear section of $\mathcal{C}$.

Now we suppose that P is a product of s prime absolute factors of degree n/s. Hence the complex curve $\mathcal{C}$ is the union of s irreducible curves $\mathcal{C}_k$ with $1 \leq k \leq s$. We fix a direction (d_0, x_0) giving a smooth section of the curve $\mathcal{C}$ and denote the corresponding section of $\mathcal{C}_k$ by $F_k = \{y_j \; ; \; N_{k-1} < j \leq N_k\}$, where $N_k = kn/s$.

For each $1 \leq k \leq s$, we consider a strict subset I_k of $\{N_{k-1} + 1, \ldots, N_k\}$. We let

$$\mathcal{B}_{m_1,\ldots,m_s} = \prod_{\mathrm{card}(I_i)=m_i} B_{I_1 \cup \ldots \cup I_s}.$$

We note that this is a rational function of (d_0, x_0). It is proven in [15] that it is not identically zero, this implies the following result.

Theorem 2 ([15]).

For any direction d_0 except at most $3n^2 2^n$ directions and for any rational value x_0 except at most $3n^2 2^n$ values, the sum $B_I = b_{i_1} + \ldots + b_{i_m}$, with $m < n$, vanishes if and only if it corresponds to the union of the roots of a family of complex prime factors of P.

2.4. Real components. Let as above $P = \prod_{k=1}^{t} Q_k \prod_{l=2t+1}^{s} P_l$ denote the real decomposition of P.

We observe that there is at most a finite number of real zeros for Q_k because they correspond to the intersection of the two complex irreducible curves defined by P_k and its conjugate. Therefore, after a generic change of coordinates, the zeros y_j of $Q_k(0, y)$ are not real numbers but always appears as pairs of distinct conjugate complex numbers; so the corresponding numbers b_j also appears as pairs of conjugate complex numbers.

This observation allows us to improve the previous result.

We gather all real roots of $P(0, y)$ in a subset denoted by $\{y_{2m+1}, \cdots, y_n\}$ and the $2m$ others in a subset denoted by $\{y_1, \bar{y_1}, \cdots, y_m, \bar{y_m}\}$.

We denote with the same indices the corresponding b_j. For each $1 \leq k \leq s$, we let e_k denotes the sum $b_k + \bar{b_k}$ which is also twice the real part of b_k.

Then, for any couple of subsets I included in $\{1, \cdots, m\}$ and J included in $\{2m+1, \cdots, n\}$, we consider the sum

$$E_{I \cup J} = \sum_{k \in I} e_k + \sum_{l \in J} b_l.$$

With these notations we deduce the following useful corollary of the previous theorem.

Theorem 3.

For any direction d_0 except at most $3n^2 2^n$ directions and for any rational value x_0 except at most $3n^2 2^n$ values, the sum

$$E_{I \cup J} = \sum_{k \in I} e_k + \sum_{l \in J} b_l$$

vanishes only if it corresponds to the union of the roots of a family of real prime factors of P.

3. A test of real irreducibility

Let $P(x, y)$ be a polynomial of (total) degree n with coefficients in $\mathbf{Q}$ as above. We suppose that we have chosen a random value $x_0 \in \mathbf{Q}$ and

$-1 < x_0 < 1$ far enough from the discriminant locus : if $Q(x)$ denotes the normalized resultant $Q(x) = Res_y(P, \frac{\partial P}{\partial y})$ then $Q(x_0)$ is clearly not zero. So, at the roots y_i of $P(x_0, y)$, the values of $\frac{\partial P}{\partial y}(x_0, y_i)$ are clearly not zero and this allows us to choose an adequate level of precision.

The general philosophy is that the result of a computation with limited precision can certify that some quantity is non–zero, but it can only hint that a quantity might be zero.

We let

$$y = \varphi_j(x - x_0) = y_j + a_j(x - x_0) + b_j(x - x_0)^2 + o(x - x_0)^2$$

be a limited expansion of the implicit solution of P at (x_0, y_j), computed approximately with the obtained value of y_j.

We also suppose that for any $j = 1, \ldots, n$, the complex numbers b_j are clearly non-zero, or we try another value x_0.

Test 1.

If for each strict subsets I included in $\{1, \cdots, m\}$ and J included in $\{2m+1, \cdots, n\}$, the quantity:

$$E_{I \cup J} = \sum_{k \in I} e_k + \sum_{l \in J} b_l$$

is clearly non-zero, then P is irreducible over $\mathbf{R}$.

An error analysis allows us to quantify the previous claim.

In [15], it is shown that the values b_j can be computed with enough precision such that we can certify that all digits of their long precision floating point representation are correct so, $\left|\frac{\Delta b_j}{b_j}\right| \leq 2^{-p}$. *e.g.* for double precision, $p = 32$.

Then we can certify that a sum $|E_{I \cup J}|$, whose approximatively computed value is greater than some ε as above, is indeed non zero.

We notice that in most of the cases where the $|b_j|$ are " not too small and not too large ", then computations of the sums in double precision suffices.

4. Finding factors via approximations

4.1. Detecting a partition and Hensel lifting.

In this section, we continue with the previous notation and assume that the irreducibility test has failed. With fixed small tolerance ϵ, we suppose that we have kept enough precision in our approximate computations in order:

- first, to compute the family $\mathcal{F}$ of all sets I in $\{1, \ldots, n\}$ such that $|E_{I\cup J}| < \epsilon$
- second, to discard in $\mathcal{F}$ the sets $I \cup J$ which contains a strict subset also belonging to $\mathcal{F}$. We denote by $\mathcal{P}$ this restricted family of sets.

Moreover, we expect that $\mathcal{P}$ defines a partition of $\{1, \ldots, n\}$ with $s - t$ elements and that each subset will correspond to a real factor of P. Degrees can be only n/s or $2n/s$.

Letting $X = x - x_0$, we use $X-$adic Hensel lifting to compute the approximated factors.

4.2. An improved Knapsack procedure.

Our practical target is to factorize polynomials of degree n about 50. It relies on an extensive search for detecting vanishing sums among 2^n possibilities. We recall that for $n = 50$, $2^n \approx 10^{15}$. In order to diminish this combinatorial explosion in "most" cases we propose a pattern matching method.

Our algorithm is based on two ingredients.

- the first one is to replace the sums on b_j by sums of type $E_{I\cup J} = \sum_{k\in I} e_k + \sum_{l\in J} b_l$ as explained above.
- The second is to divide in two part each set of indices and use a pattern matching method. Let us assume for simplicity that $m = 2m', n = 2n'$ so $n - 2m = 2(n' - 2m')$.
  We set $\mathbf{E} = \{e_1, ..., e_{2m'}\}$. and divide it into two subsets

  $$\mathbf{E}' = \{e_1, ..., e_{m'}\}, \;\; \mathbf{E}'' = \{e_{m'+1}, ..., e_{2m'}\}$$

  so card$\mathcal{P}(\mathbf{E}') = $ card$\mathcal{P}(\mathbf{E}'') = 2^{m'}$.
  Similarly we divide $\mathbf{B} = \{b_1, ..., b_{2(n-2m')}\}$ into two subsets $\mathbf{E}'$ and $\mathbf{E}''$. Each subset I of $\mathbf{E}$ is the union of a subset I' of $\mathbf{E}'$ and of a subset I'' of $\mathbf{E}''$, idem for a subset J of $\mathbf{B}$. Then
  $E_{I\cup J} = 0$ iff $E_{I'\cup J'} = -E_{I''\cup J''}$.
  We compute then keep and sort in a vector the $2^{n'-m'}$ real numbers equal to the sums $E_{I'\cup J'}$ and in another vector the $2^{n'-m'}$ sums $-E_{I''\cup J''}$. Then we look for (approximate) matching.

Generically, a polynomial with rational coefficients has about $Log(n)$ real zeros. So for $n = 50$, the total size is about 10^5 double precision floating point numbers, which is rather small.

4.3. Approximation. Our treatment relies on good approximations of the n coefficients b_j. Those are obtained by plugging a good approximation of y_j (needed for the last steps of the factorization process) in the formulaægiven above. Therefore it is important to compute precisely the n roots y_j of $P(x_0, y)$, say with e.g. 250 digits, therefore the e_k and b_j are easily certified e.g. double precision floating point numbers. In this way, for a "usual" size of coefficients of the input polynomial, the sums $E_{I\cup J}$ are still precise without requiring big-float computation.

5. Construction of exact factors and validation

The first task is to precise which algebraic presentation of a complex, respectively real, factorization we choose.

5.1. Complex factorization. A complex factorization of P is well defined once we give:

- an irreducible univariate polynomial q in $\mathbf{Q}[z]$, it defines an algebraic extension $K = \mathbf{Q}[z]/q = \mathbf{Q}(\alpha)$,
- the first factor P_1 which is a bivariate polynomial in $K[x, y]$.

So each coefficient of P_1 is expressed as a rational function $a_{k,l}(\alpha)$. The coefficients $a^j_{k,l}$ of the other factors $P_2, P_3, \cdots P_s$ are obtained by replacing α by the other roots of q in this expression. In order to perform it explicitly, this should be calculated in an algebraic extension of possibly higher degree $\mathbf{Q}(\alpha, \alpha_2, \cdots, \alpha_s)$, but we do not do it till we do not need it.

We take for q the polynomial whose roots are the s values of a fixed generic linear combination of the coefficients applied to each real polynomials $P_j(x_0, y)$, $1 \leq j \leq s$.

The computation of this polynomial q requires to recognize rational numbers by big-float approximations.

5.2. Real factorization.

Let us analyze an algebraic representation of a real factorization of P, say $P = \prod_{k=1}^{t} Q_k \cdot \prod_{l=2t+1}^{s} P_l$.

The coefficients of the first factor Q_1 are real rational function of α but also of $\bar{\alpha}$, where α is the previous root of the previous q. So if we let $\alpha = a + ib$, they belongs to the field L:

$$L = \mathbf{Q}(\alpha, \bar{\alpha}) \cap \mathbf{R} = \mathbf{Q}(a, b^2) = \mathbf{Q}(Re(\alpha), |\alpha|).$$

Such an extension is a subfield of the extension $\mathbf{Q}(\alpha)[z]/\frac{q(z)}{z-\alpha}$ which has a degree lower or equal to $s(s-1)$.

We see that it is more complicate to represent algebraically than the previous K.

Moreover the coefficients of the same monomial in the different Q_k are only $t \leq s/2$ of the $s(s-1)$ conjugates of that coefficient of Q_1. So we cannot apply the previous trick for computing their minimal polynomial.

Another possible representation of these coefficients is "a la Thom" as we used e.g. in [3]. But in our setting it is not efficient for recovering these coefficients from a big-float approximation nor for performing division or multiplication to check validity of the estimated factorization.

Because of this lack of simplicity of an algebraic representation of a real factorization, we propose to compute a complete complex factorization.

This is done just by continuing the approximate computation in order to get an approximate complex representation. Then we recognize the coefficients $a^1_{k,l}(\alpha)$ in $K = \mathbf{Q}[z]/q = \mathbf{Q}(\alpha)$, and check validity. Then we can further embed K in $\mathbf{Q}(\alpha, \bar{\alpha})$, form $\bar{P}_1$, multiply P_1 by $\bar{P}_1$ and express it in $L[x, y]$.

6. Conclusion

In this paper, we presented a new algorithm for computing a factorization over $\mathbf{R}$ of a bivariate polynomial with integer coefficient and irreducible on $\mathbf{Q}$. Our approach associates formal and numerical computations, it proceeds from a basic observation and a generic non-vanishing theorem.

In the introduction we asked whether the computation of a real polynomial factorization could be performed at a lower computational cost than passing via a complex polynomial factorization. We saw that this is actually true for our irreducibility test and the computation of an approximate factorization. However, because of the complexity of an algebraic representation of an exact real factorization, we were not able to propose a fast way for the complete problem. So the situation is more subtle that it might appear at a first look.

In a future work, with the same general philosophy, we plan to develop a specific algorithm and analysis adapted to inputs given with a limited precision.

References

1. C. Bajaj, J. Canny, T. Garrity, and J. Warren, *Factoring rational polynomials over the complexes*. Proc. ISSAC'89, ACM Press (1989), 81–90.

2. D. Duval, *Absolute Factorization of Polynomials: a Geometric Approach. SIAM J. of Comp.*, **20**, No. 1 (1991), 1–21.
3. A. Galligo, N. Vorobjov, *Complexity of finding Irreducible Components of a Semi-algebraic set. J. of Complexity*, **11** (1995), 174-193.
4. A. Galligo, S. Watt, *An absolute primality test for bivariate polynomials. ACM Press*, Proc. ISSAC'97 (1997).
5. E. Kaltofen, *Fast Parallel Absolute Irreducibility Testing. J. Symbolic Computation*, **1** (1985), 57–67.
6. J. Harris, *A bound on the geometric genus of projective varieties. Ann. Sc. Norm. Sup. Pisa Cl.Sci.*, **IV**, Ser. 8 (1981), 35–68.
7. J. Heintz, M. Sieveking, *Absolute primality is decidable in random polynomial time in the number of variables.* Springer Verlag, No. 11, LNCS (1981), 16–28.
8. J. F. Ragot, *PhD thesis, Univ. of Limoges.* 1997.
9. T. Sasaki, M. Susuki, M. Kolár, and M. Sasaki, *Approximate factorization of multivariate polynomials and absolute irreducibility testing.* Japan J. Indus. Appl. Math. **8** (1991), 357-375.
10. Barry Trager, *Integeration of algebraic functions.* MIT (1984).
11. Carlo Traverso, *A study on algebraic algorithms: the normalization. Rend. C. Sem. Mat. Torino* (1986), 111–130.
12. David Y. Y. Yun, *On square-free decomposition algorithms.* Proc. SYMSAC'76, ACM Press (1976), 26-35.
13. Richard Zippel, *Effective Polynomial Computation.* Kluwer Acad. (1993).
14. A. L. Chistov, D. Yu. Grigoriev, *Subexponential time solving of systems of algebraic equations I.* Steklov Mathematical Institute, Leningrad, Technical Report, E-9 83 (1983).
15. D. Rupprech, *Semi-numeric Absolute factorization for polynomials with integer coefficients.* Universite de Nice, Preprint (1998).

Laboratoire de Mathématiques
Université de Nice (France)

Поступило 21 мая 1999 г.