Computing the Irreducible Real Factors and Components of an Algebraic Curve*

Erich Kaltofen

Department of Computer Science, Rensselaer Polytechnic Institute
Troy, New York 12180-3590; ARPA-Net: kaltofen@cs.rpi.edu

ABSTRACT. We present algorithms that decompose an algebraic curve with rational coefficients in its defining bivariate equation into its irreducible real factors and its non-empty irreducible real components. We show that our algorithms are of polynomial bit complexity in the degree of the equation and the size of its coefficients. Our construction is based on computing the irreducible complex factors and then investigating high precision complex floating point coefficients of these factors and the complex norms.

Keywords: Polynomial factorization, algebraic curves, real number arithmetic, polynomial-time complexity

Running title: Computing real factors

*This material is based on work supported in part by the National Science Foundation under Grant No. CCR-87-05363 and under Grant No. CDA-88-05910. A preliminary version of this paper appears in the Proceedings of the 5th Annual Symposium on Computational Geometry, ACM Press, pp. 79–87, (1989).

Appears in *Applic. Algebra Engin. Commun. Comput.*, **1**/2, pp. 135–148 (1990).

1. Introduction

An algorithm is presented that allows to decompose an algebraic curve with rational coefficients into its irreducible real factors. A real algebraic coefficient of a factor is represented by its defining minimal integral polynomial as well as by a rational interval such that the only real root of this minimal polynomial within the given interval is that real coefficient. Our algorithm runs in time polynomial in the degree and the coefficient size of the rational bivariate polynomial defining the curve. As a corollary we get, for example, that the problem of deciding whether a multivariate polynomial factors over the reals is in polynomial-time.

From our representation of the irreducible real factors one can, for instance, find arbitrarily precise floating point approximations of their coefficients, or one can by Seidenberg's [20] algorithm determine whether there are real points that lie on the curves defined by these factors. Our algorithm can also be used as a size reduction tool in any of the algorithms for deciding the theory of real closed fields, e.g., Collins's cylindrical algebraic decomposition method (see [1]) very much in the same way as polynomial factorization over the ground field can be used to fork the reductions in a Gröbner basis completion.

The problem of finding the irreducible real factors of a curve is clearly a problem in the existential theory over the reals. However, the number of unknown real variables, the coefficients of the factors, grows with the degree of the polynomial. The existential theory over the reals is easily shown NP-complete and hence any of the general algorithms, such as Canny's [2] P-space solution or Renegar's [18] efficient method, have running time exponential in the number of variables. Therefore, such algorithms do not solve our problem in polynomial-time.

Our polynomial-time solution relies on the theory of factoring polynomials over the complex numbers. One key result in this theory is that one can represent all the absolutely irreducible factors over small distinct extension fields of the coefficient field, even though their least common superfield would in the worst case have exponential extension degree (see §2). Therefore, one can compare the factors pairwise by only squaring the degrees of the extensions. In particular, the coefficients of the complex norm of these factors—which are the real factors—will not lie in asymptotically higher degree extensions.

As in any of the algorithms dealing with algebraic real numbers, we will combine algebraic field theory with high precision numerical approximation techniques. In order to compute separating intervals, we need to bound the minimum distance between certain algebraic reals from below. In particular, for non-real complex numbers we need to bound their imaginary part away from zero. For all problems of separation we will get bounds that require to work with a maximal precision that is polynomial in the degree and the binary length of the coefficients of the input polynomial.

Our methods generalize to curves with totally real algebraic coefficients. We also can factor implicitly given surfaces over the reals, since our methods work in polynomial-time for any fixed dimension. However, for simplicity's sake we shall treat here the case of curves with rational coefficients only.

We will also briefly describe how one can test whether an irreducible real factor in our output representation contains a real point. The irreducible factors with real points define all irreducible components of the input curve. The method we use is essentially Seidenberg's [20]. and it turns out that this adaptation is also polynomial-time in the input parameters.

Notation. By $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ we denote the set of integers, rational, real, and complex numbers, respectively. The symbols φ, ψ, and χ will be used to denote minimal polynomials over $\mathbb{Q}$, the symbols ζ, η, ξ, and ϑ will be used to denote complex algebraic numbers. $\Re(\zeta)$ and $\Im(\zeta)$ denote the real and imaginary part of ζ, respectively, and ζ^* denotes the complex conjugate of ζ, $\zeta^* = \Re(\zeta) - \sqrt{-1}\,\Im(\zeta)$. Furthermore, $\widetilde{\zeta}$, $\widetilde{\vartheta}$, etc., will denote complex floating point approximations of the algebraic numbers ζ, ϑ. The symbols ε_1, ε_2, etc., denote small quantities, usually a guaranteed precision of an approximation. Finally, we use the symbols $:=$ and $=:$ to define new mathematical objects (the new quantities being on the side of the colon), and we use the symbols $\leftarrow$ and $\rightarrow$ as assignment operators in program code.

2. Factoring over the Complex Numbers

The procedure we will give below is an adaptation of the algorithm presented in [10]. Since we are interested in a polynomial-time solution, the input polynomial can be assumed to be irreducible, for otherwise we factor it over the coefficient field by any of the polynomial-time methods. In addition, the input polynomial can be assumed monic in one of the variables and squarefree when evaluating the other variable at zero, which can be enforced by the simple isomorphic transformation

$$\hat{f}(x,y) := b\, f(x, ax + y + c), \quad a, b, c \in \mathsf{K}, b \neq 0,$$

where a is chosen such that $\deg_x(f) = \deg(\hat{f})$ [11], and c not a root (as a polynomial in y) of the resultant [10, §2]

$$\mathrm{Res}_x \left(f(x, ax + y),\ (\partial/\partial x) f(x, ax + y) \right).$$

The correctness of the algorithm follows from a lemma proved by several persons, among them Chistov and Grigoryev [3, Lemma 1], Trager [21, §3.2], Dvornicich and Traverso [7], and the author [10, Theorem 1].

Lemma 1. *Let $f(x,y) \in \mathsf{K}[x,y]$ be irreducible, monic in x, K a field of characteristic 0. Let $g(x,y) \in \overline{\mathsf{K}}[x,y]$ be an absolutely irreducible factor of f. Then there exists a root α of $f(x,0)$ such that the coefficient field of g is isomorphic to a subfield of $\mathsf{K}(\alpha)$.*

This lemma allows to compute all absolutely irreducible factors of f without construction an algebraic extension common to all the factors, which in the worst case can be of degree $\deg_x(f)$!.

Algorithm *Factorization over the Algebraic Closure*

Input: $f(x,y) \in \mathsf{K}[x,y]$ irreducible and monic in x, $f(x,0)$ squarefree, K a field of characteristic 0. Furthermore, we are given the factorization of $f(x,0)$ into irreducible polynomials over K, $f(x,0) = \varphi_1(x) \cdots \varphi_r(x)$. Notice that this factorization can be found in polynomial-time for the usual representations of K.

Output: Either f will be certified to be absolutely irreducible; or for all $1 \leq i \leq r$ the algorithm returns polynomials

$$f_i(x,y) \in \mathsf{K}_i[x,y] \text{ with } \mathsf{K}_i := \mathsf{K}[z]/(\varphi_i(z))$$

with the following property. For any irreducible factor $g(x,y) \in \overline{\mathsf{K}}(x,y)$ there exists an index j, $1 \le j \le r$, and there exists an embedding $\iota_j \colon \mathsf{K}_j \to \overline{\mathsf{K}}$ that fixes K such that $\iota_j(f_j) = g$. Notice, however, that a factor g may arise as the image of several f_i's, even as the image of several conjugates of a single f_i (see remark following the algorithm).

The idea of the algorithm is to compute the approximation of a root of $f(x,y)$ in any $\mathsf{K}_i[[y]]$, and then find the corresponding minimal polynomial.

Set the order of the approximation

$$\ell_{\max} \leftarrow 2(\deg_x(f) - 1)\deg_y(f).$$

For $i \leftarrow 1, \ldots, r$ Do Steps N and L.

Step N: *Set the initial points for the Newton iteration*

$$\alpha_{i,0} \leftarrow z \bmod \varphi_i(z) \in \mathsf{K}_i, \quad \beta_{i,0} \leftarrow \frac{1}{(\partial f/\partial x)\,(\alpha_{i,0}, 0)} \in \mathsf{K}_i.$$

Notice that $(\partial f/\partial x)(\alpha_{i,0}, 0) \neq 0$ because f was assumed irreducible, thus squarefree.
We now perform Newton iteration with quadratic convergence (see [13, §3.3]).
For $j \leftarrow 0, \ldots, \lfloor \log_2(\ell_{\max}) \rfloor$ Do:

$$\alpha_{i,j+1} \leftarrow \Big(\alpha_{i,j} - \beta_{i,j} f(\alpha_{i,j}, y)\Big) \bmod y^{2^{j+1}};$$

$$\beta_{i,j+1} \leftarrow \left(2\beta_{i,j} - \frac{\partial f}{\partial x}(\alpha_{i,j+1}, y)\,\beta_{i,j}^2\right) \bmod y^{2^{j+1}}.$$

Notice that $\alpha_{i,j+1}$ and $\beta_{i,j+1}$ are polynomials in $\mathsf{K}_i[y]$ with

$$f(\alpha_{i,j+1}, y) \equiv 0 \pmod{y^{2^{j+1}}}, \quad \beta_{i,j+1}\frac{\partial f}{\partial x}(\alpha_{i,j+1}, y) \equiv 1 \pmod{y^{2^{j+1}}}.$$

Set the approximated root

$$\alpha_i \leftarrow \alpha_{i,\lfloor \log_2(\ell_{\max})\rfloor + 1} \bmod y^{k+1} \in \mathsf{K}_i[y].$$

Step L: We now find the lowest degree polynomial in $\mathsf{K}_i[x,y]$ whose root is α_i.
For $m \leftarrow 1, \ldots, \deg_x(f) - 1$ Do:

Here we try to find a polynomial in $\mathsf{K}_i[x,y]$ of degree m in x that α_i satisfies.

Set the needed order of approximation

$$\ell \leftarrow \deg_y(f)(m + \deg_x(f) - 1).$$

We examine whether the equation

$$\alpha_i^m + \sum_{\mu=0}^{m-1} h_{i,\mu}(y)\alpha_i^\mu \equiv 0 \pmod{y^{\ell+1}},$$

has a solution for $h_{i,\mu}(y) \in \mathsf{K}_i[y]$ with $\deg(h_{i,\mu}) \leq \deg_y(f)$. By choosing an indeterminate 'Ansatz' for the coefficients of $h_{i,\mu}$,

$$h_{i,\mu}(y) =: \sum_{\delta=0}^{\deg_y(f)} u_{i,\mu,\delta}\, y^\delta, \quad u_{i,\mu,\delta} \in \mathsf{K}_i,$$

and defining the coefficients of α_i^μ,

$$\alpha_i^\mu \equiv \sum_{\lambda=0}^{\ell} a_{i,\lambda}^{(\mu)} y^\lambda \pmod{y^{\ell+1}}, \quad a_{i,\lambda}^{(\mu)} \in \mathsf{K}_i,$$

we are led to the following problem.
Solve the linear system over the field K_i

$$a_{i,\lambda}^{(m)} + \sum_{\mu=0}^{m-1} \sum_{\delta=0}^{\deg_y(f)} a_{i,\lambda-\delta}^{(\mu)} u_{i,\mu,\delta} = 0, \quad (a_{i,\nu}^{(\mu)} = 0 \text{ for } \nu < 0) \tag{1}$$

for $0 \leq \lambda \leq \ell$ *in the variables* $u_{i,\mu,\delta}$, $0 \leq \mu \leq m-1$, $0 \leq \delta \leq \deg_y(f)$. Notice that if the system (1) has a solution in K_i, then that solution is unique (see [9, Theorem 1]). *If the system has a solution, then set*

$$f_i(x,y) \leftarrow x^m + \sum_{\mu=0}^{m-1} \sum_{\delta=0}^{\deg_y(f)} u_{i,\mu,\delta} y^\delta x^\mu$$

and exit the loop. If the system has no solution and $i = 1$ *and* $m = \deg_x(f) - 1$, *then designate* f *absolutely irreducible and exit the algorithm.* □

The algorithm is subject to several improvements. For one, one can combine steps N and L so that one computes α_i incrementally to the order ℓ needed for the m considered. Furthermore, the systems (1) are not independent for different m, and therefore the triangularizations can be incrementally computed.

We shall briefly discuss how one can algebraically prove two factors f_i and f_j, $i \neq j$, to be the same. Necessarily, the term structure of these factors has to agree. The factors lie in the fields $\mathsf{K}_i := \mathsf{K}[z_1]/(\varphi_i(z_1))$ and $\mathsf{K}_j := \mathsf{K}[z_2]/(\varphi_j(z_2))$, respectively. We use van der Waerden's method of computing a primitive element for a smallest common superfield of K_i and K_j [14]. Such an element can be chosen of the form $\zeta_0 = \zeta_1 + c\zeta_2$ with minimal polynomial $\psi(z_0) \in \mathsf{K}[z_0]$, where $c \in \mathsf{K}$, ζ_1 is a root of $\varphi_i(z_1)$, and ζ_2 is a root of $\varphi_j(z_2)$. Furthermore, both ζ_1 and ζ_2 are algebraically expressible in ζ_0, i.e., there are polynomials $w_1(z_0)$ and $w_2(z_0)$ in $\mathsf{K}[z_0]$ that with $w_1(\zeta_0) = \zeta_1$ and $w_2(\zeta_0) = \zeta_2$, respectively. Interpreting $f_i \in \mathsf{K}[x,y,z_1]/(\varphi_1(z_1))$ and $f_j \in \mathsf{K}[x,y,z_2]/(\varphi_2(z_2))$ we test whether

$$f_i(x,y,w_1(z_0)) \equiv f_j(x,y,w_2(z_0)) \pmod{\psi(z_0)}.$$

If the test fails, f_i and f_j are distinct factors of f.

We shall also briefly discuss how to count the number of distinct conjugates of f_i that are factors of f. Let us first give an example. Consider $f(x,y) = x^4 - 2(y+1)^2 \in \mathbb{Q}[x,y]$. Clearly,

$r = 1$ and $\varphi_1(z) = z^4 - 2$. The factors of f over K_1 are $x^2 + \zeta^2(y+1)$ and $x^2 - \zeta^2(y+1)$, with $\varphi_1(\zeta) = 0$. In other words, only two of the four conjugates of $\varphi(z)$ lead to distinct factors. What one might want to do is find the defining equation of the smallest subfield of K_i that contains all the coefficients of f_i. For the example this is $\mathsf{K}[v]/(v^2-2)$. One way to construct this subfield is as follows. Let $u_{i,1}(z), \ldots, u_{i,t}(z) \in \mathsf{K}[z]/(\varphi_i(z))$ be the coefficients of $f_i(x, y)$. We first find the minimal polynomials $\psi_{i,j}(v_{i,j}) \in \mathsf{K}[v_{i,j}]$ for $v_{i,j} = u_{i,j}(z)$, $1 \le j \le t$, by computing the minimum linear dependence over K of

$$1, u_{i,j}(z) \bmod \varphi_i(z), u_{i,j}(z)^2 \bmod \varphi_i(z), \cdots.$$

Then we compute a primitive element ζ_0 and its defining polynomial $\psi_0(z)$ for the smallest superfield of the fields

$$\mathsf{K}(v_{i,1})/(\psi_{i,1}(v_{i,1})), \ldots, \mathsf{K}(v_{i,t})/(\psi_{i,t}(v_{i,t}))$$

by inductively using the van der Waerden procedure discussed before. All conjugates of ζ_0 now will generate distinct conjugate polynomials of f_i.

Both problems, that of identifying the same factors and that of counting the number of distinct conjugates, require a factorization procedure for $\mathsf{K}[z]$ and are therefore quite costly. For the computation of real factors we will adopt a different strategy, applicable if K is a finite algebraic extension of $\mathbb{Q}$. We will compute a high precision complex floating point approximation of the factors. Using a separation lemma for the coefficients of distinct factors, we then can guarantee that we have approximated distinct factors or identified identical ones.

3. Separation Lemmas

In section 4 we will identify complex and real factors by computing a complex floating point approximation to their coefficients. In order to decide at what precision ("fuzz") a coefficient can be declared real, or two coefficients in two factors distinct, we need so-called separation lemmas. We will formulate these inequalities in term of certain norms of the defining polynomials. We shall first define these norms.

Let

$$\begin{aligned} f(z) &= a_n(z - \zeta_1) \cdots (z - \zeta_n) \\ &= a_n z^n + a_{n-1} z^{n-1} + \cdots + a_0 \in \mathbb{Z}[z],\ \zeta_\nu \in \mathbb{C}. \end{aligned}$$

The *p-norm* of f, $0 < p \le \infty$, is defined as

$$\|f\|_p := \sqrt[p]{|a_0|^p + |a_1|^p + \cdots + |a_n|^p}.$$

The cases that are used most often are for $p = 1, 2$, and $p = \infty$. The ∞-norm of f is referred to as the *height* of f,

$$\|f\|_\infty = \max(|a_0|, |a_1|, \ldots, |a_n|).$$

Clearly, $\|f\|_\infty \le \|f\|_2 \le \|f\|_1$, but we also have

$$\|f\|_2 \le \sqrt{n+1}\, \|f\|_\infty \text{ and } \|f\|_1 \le \sqrt{n+1}\, \|f\|_2.$$

The *measure* of f is defined as

$$M(f) := |a_n| \prod_{\nu=1}^{n} \max(1, |\zeta_\nu|).$$

The *discriminant* of f is

$$\mathrm{Disc}(f) := a_n^{2n-2} \prod_{1 \le \mu < \nu \le n} (\zeta_\mu - \zeta_\nu)^2 = \frac{1}{a_n} \mathrm{Res}\Big(f, \frac{\partial f}{\partial x}\Big),$$

which is always an integer.

The first lemma is a root separation lemma for polynomials with integer coefficients, and is due to Mignotte [16] based on inequalities by Landau [12] and Mahler [15]. Independently, a similar inequality for Gaussian polynomials was proven by Collins and Horowitz [6].

Lemma 2. *Let $n \ge 2$ and let $f(z) = a_n(z - \zeta_1) \cdots (z - \zeta_n) \in \mathbb{Z}[z]$ be squarefree, $\zeta_\nu \in \mathbb{C}$ for $1 \le \nu \le n$. Then $M(f) \le \|f\|_2$ and*

$$\begin{aligned} \forall\, \mu \ne \nu \colon |\zeta_\mu - \zeta_\nu| &> \frac{\sqrt{3\,|\mathrm{Disc}(f)|}}{n^{(n+2)/2} M(f)^{n-1}} \\ &\ge \frac{\sqrt{3}}{n^{(n+2)/2} \|f\|_2^{n-1}} =: \varepsilon_1(n, \|f\|_2). \quad \boxtimes \end{aligned}$$

In section 4 we will need to distinguish real from complex roots. Since the complex roots of an integer polynomial lie symmetric about the real axis, we get the following corollary.

Corollary 1. *Let n, f, and ζ_ν be as in lemma 2. Furthermore, let $\Re(\zeta_\nu)$ be the real and $\Im(\zeta_\nu)$ be the imaginary part of ζ_ν, $1 \le \nu \le n$. Then*

$$\forall \nu \colon \Im(\zeta_\nu) \ne 0 \implies |\Im(\zeta_\nu)| > \frac{1}{2} \varepsilon_1(n, M(f)). \quad \boxtimes$$

The preceding lemma allows us to determine root isolations for f, which will be a crucial tool for our algorithm. Consider approximations $\widetilde{\zeta_\nu}$ to ζ_ν given as complex rational numbers, perhaps in floating point format

$$\mathit{mantissa} \times 2^{\mathit{exponent}}.$$

Such approximations are said to *isolate the roots of $f(z)$* if

$$\forall \nu \colon |\zeta_\nu - \widetilde{\zeta_\nu}| < \varepsilon_0, \quad \text{where } \varepsilon_0 \le \min_{\nu \ne \mu} \{|\zeta_\nu - \zeta_\mu|/4\}.$$

The precision ε_0 is chosen such that a circle of radius ε_0 around each approximate root $\widetilde{\zeta_\nu}$ contains exactly one of the actual roots of f. Furthermore, the distance between the approximations is larger than $2\varepsilon_0$. Now the separation lemma allows us to terminate a root approximation algorithm as soon as $\varepsilon_0 \le \varepsilon_1/4$. We also remark that such root approximations can be computed quite efficiently. E.g., one of Schönhage's procedures [19] has bit complexity

$$O(n^3 \log(\|f\|_\infty/\varepsilon_0)^3).$$

Older algorithms based on the Cauchy principle of argument and the Routh-Hurwitz theorem [17], [23], or real root approximation [5] also have polynomial running time.

We finally give a separation lemma for the roots of relatively prime polynomials.

Lemma 3. *Let*

$$f(z) = a_n(z-\zeta_1)\cdots(z-\zeta_n) \in \mathbb{Z}[z],$$
$$g(z) = b_m(z-\eta_1)\cdots(z-\eta_m) \in \mathbb{Z}[z], \quad \zeta_\nu, \eta_\mu \in \mathbb{C}$$

be two squarefree, relatively prime polynomials. Then

$$\begin{aligned}\forall \nu,\mu\colon |\zeta_\nu - \eta_\mu| &\geq \varepsilon_1(n+m, \sqrt{(n+1)(m+1)}\,\|f\|_2\|g\|_2) \\ &=: \varepsilon_2(n, m, \|f\|_2, \|g\|_2).\end{aligned}$$

Proof. Consider $h(z) := f(z)\,g(z)$. Since f and g are squarefree and relatively prime to one another, h is squarefree. In order to apply lemma 2 to h, we need an upper bound for $\|h\|_2$. A simple argument shows that

$$\|h\|_2 \leq \|fg\|_1 \leq \|f\|_1\|g\|_1 \leq \sqrt{(n+1)(m+1)}\,\|f\|_2\|g\|_2.$$

Hence

$$\begin{aligned}|\zeta_\nu - \eta_\mu| &> \varepsilon_1(n+m, \|h\|_2) \\ &\geq \varepsilon_1(n+m, \sqrt{(n+1)(m+1)}\,\|f\|_2\|g\|_2). \quad \boxtimes\end{aligned}$$

4. Complex Conjugation

The problem addressed in this section is the following. Given be an irreducible polynomial $\varphi(z) \in \mathbb{Z}[z]$ of degree n and one of its non-real roots ζ. Let ζ^* be its complex conjugate. We want to find an algebraic number ϑ and its minimal polynomial $\widehat{\varphi}(t) \in \mathbb{Z}[t]$ such that

$$\zeta = p_1(\vartheta), \quad \zeta^* = p_2(\vartheta), \quad p_1(t), p_2(t) \in \mathbb{Q}[t].$$

The rationale behind this to our problem is the following. If we have found a complex non-real factor $g(x, y, \zeta) \in \mathbb{Q}(\zeta)[x, y]$ of $f(x, y) \in \mathbb{Q}[x, y]$, then

$$g(x, y, \zeta)\, g(x, y, \zeta^*) \in \mathbb{Q}(\vartheta)(x, y)$$

is an irreducible real factor of $f(x, y)$.

The process of the construction of $\widehat{\varphi}$ and p_1, p_2 is again based on the determination of a primitive element. We know that $\vartheta = \zeta + c\zeta^*$ is a primitive element for c chosen appropriately. The polynomials p_1, p_2 are then ζ, ζ^* expressed in terms of ϑ. However, one needs to identify the appropriate minimal polynomial $\widehat{\varphi}$ for ϑ by complex floating point approximations. The precise algorithm follows.

Algorithm *Complex Conjugation*

Input: Given is an irreducible polynomial $\varphi(z) \in \mathbb{Z}[z]$ and a complex floating point number $\widetilde{\zeta}$ of sufficient mantissa length that isolates a non-real root ζ of φ.

Output: An irreducible polynomial $\widehat{\varphi}(t) \in \mathbb{Z}[t]$, polynomials $p_1(t), p_2(t) \in \mathbb{Q}[t]$ and a complex floating point approximation $\widetilde{\vartheta}$ that isolates a root ϑ of $\widehat{\varphi}$ such that $p_1(\vartheta) = \zeta$ and $p_2(\vartheta) = \zeta^*$.

Step 1: *Pick an integer c and compute the resultant*

$$\Phi(z) \leftarrow \mathrm{Res}_y\Big(\varphi(y), c^n\varphi(\frac{z+y}{c})\Big) = a_n^{2n} \prod_{1\le i,j\le n} (z - (\zeta_i - c\zeta_j)),$$

where

$$\varphi(z) =: a_n(z-\zeta_1)\cdots(z-\zeta_n)$$

(see [14]).

Step 2: *Factor $\Phi(z)$ over the integers and identify that irreducible factor $\widehat{\varphi}(z)$ which has $\vartheta := \zeta - c\zeta^*$ as its root.* For that we may have to compute ζ to higher precision than that of $\widetilde{\zeta}$ in order to separate $\widetilde{\vartheta} \leftarrow \widetilde{\zeta} - c(\widetilde{\zeta})^*$ from all other roots of Φ. Automatically, the approximation $\widetilde{\vartheta}$ will isolate a root ϑ of the factor $\widehat{\varphi}$. Notice that $\mathbb{Q}(\vartheta)$ is now isomorphic to $\mathbb{Q}[t]/(\widehat{\varphi}(t))$.

Step 3: *Compute the GCD of $\varphi(z)$ and $\varphi(t+cz)$ over $(\mathbb{Q}[t]/(\widehat{\varphi}(t)))[z]$.* Clearly, for $t = \vartheta$, ζ^* is a common root of both polynomials. By selecting

$$c \neq \frac{\zeta_1 - \zeta_i}{\zeta_1 - \zeta_k}, \quad 1 \le i \le n,\ 2 \le k \le n,$$

this will be the only common root of the two polynomials [22, §40].

Test whether the computed GCD is a linear polynomial $z - p_2(t)$. If not, go back to step 1 and start with a new c. Otherwise, compute $p_1(t) \leftarrow t + cp_2(t)$. $\square$

5. Factoring over the Real Numbers

We now discuss how to factor a polynomial $f(x,y) \in \mathbb{Z}[x,y]$ into its real factors. One of the problems we encounter is how to represent the real coefficients. A standard representation is the one chosen by Collins's [4] in his cylindrical algebraic decomposition algorithm. For a real algebraic number ξ one gives an irreducible integer polynomial $\chi(v)$ that has ξ as a root, and one gives a rational interval that isolates ξ among the real roots of χ. Our output representation is different, but can be easily converted to Collins's representation. The main reason for not performing the conversion inside the algorithm is that Seidenberg's method for testing which of the irreducible real factors has a real point is more efficiently performed with our representation (see §6).

Algorithm *Factorization over the Real Numbers*

Input: $f(x,y) \in \mathbb{Z}[x,y]$ irreducible over $\mathbb{Q}$, monic in x.

Output: A list of distinct monic factors

$$g_l(x,y) = \sum_{j,k} \xi_{l,j,k} x^j y^k \in \mathbb{R}[x,y], \quad 1 \le l \le s,$$

that are irreducible over $\mathbb{R}$. Each factor g_l is represented by an irreducible polynomial $\widehat{\varphi}_l(t) \in \mathbb{Z}[t]$ together with complex floating point number $\widetilde{\vartheta}_l$ that isolates a root $\vartheta_l \in \mathbb{C}$ of $\widehat{\varphi}$, and by coefficient polynomials $\widehat{u}_{l,j,k}(t) \in \mathbb{Q}[t]$ such that

$$\forall j,k\colon\ \xi_{l,j,k} = \widehat{u}_{l,j,k}(\vartheta_l).$$

Step 1: *Factor $f(z,0)$ over $\mathbb{Q}$ into*

$$f(z,0) = \varphi_1(z) \cdots \varphi_r(z), \quad \varphi_i(z) \in \mathbb{Q}[z]$$

and then call the algorithm Factorization over the Algebraic Closure of Section 2. The algorithm returns for each $\varphi_i(z)$, $1 \le i \le r$, an absolutely irreducible factor

$$f_i(x,y,z) = \sum_{j,\,k} u_{i,j,k}(z)\, x^j y^k \in \Big(\mathbb{Q}[z]/(\varphi_i(z))\Big)[x,y].$$

By multiplying through with the least common integral denominator make the coefficients of $u_{i,j,k}$ integral.
Let $\psi_{i,j,k}(v) \in \mathbb{Z}[v]$ be the minimal polynomial of $v = u_{i,j,k}(z) \in \mathbb{Q}[z]/(\varphi_i(z))$. We will not compute these polynomials. However, in order to distinguish the factors via floating point approximation we will need the maximum norm of the $\psi_{i,j,k}(v)$.
For $i \leftarrow 1, \ldots, r$ Do: Compute an upper bound

$$N_i \ge \max\{\|\psi_{i,j,k}\|_2 \mid 0 \le j \le \deg_x(f_i),\ 0 \le k \le \deg_y(f_i)\}.$$

Such a bound can be deduced from the coefficient sizes of $u_{i,j,k}$ and φ_i (see the proof of theorem 1 below).

Step 2: *Initialize the index of factors $s, t \leftarrow 0$.*
For $i \leftarrow 1, \ldots, r$ Do Step 3 and Step 4.

Step 3: Let $d_i := \deg(\varphi_i)$, $d_{\max} \leftarrow \max_{1 \le \iota \le i}\{d_\iota\}$, and $N_{\max} \leftarrow \max_{1 \le \iota \le i}\{N_\iota\}$.
Compute all complex roots of $\varphi_i(z)$ to floating point precision

$$\varepsilon_3(d_{\max}, d_i, N_{\max}, N_i, \|\varphi_i\|_2, \max_{j,\,k}\{\|u_{i,j,k}\|_\infty\}),$$

obtaining the complex floating point numbers $\widetilde{\zeta}_{i,1}, \ldots, \widetilde{\zeta}_{i,d_i}$. Notice that ε_3 is chosen relative to ε_4 below. A possible value for ε_3 is given in the proof of theorem 1 below. To find roots to this precision, we can use any of the arbitrary precision complex root approximation procedures, e.g., the fast algorithms discussed by Schönhage [19]. These algorithms usually identify the real roots among the $\widetilde{\zeta}_{i,\delta}$ as well as match complex conjugate roots.
Next, substitute all these numbers into the polynomials $u_{i,j,k}(z)$ resulting in an approximate polynomial

$$\widetilde{f}_{i,\delta}(x,y) \leftarrow \sum_{j,\,k} u_{i,j,k}(\widetilde{\zeta}_{i,\delta}) x^j y^k, \quad 1 \le \delta \le d_i.$$

Now the coefficients of $\widetilde{f}_{i,\delta}$ are approximations guaranteed to precision

$$\varepsilon_4(d_{\max}, d_i, N_{\max}, N_i) := \min\{\varepsilon_1(d_i, N_i)/4, \varepsilon_2(d_{\max}, d_i, N_{\max}, N_i)/4\}. \tag{2}$$

By that we mean that the approximate coefficients lie within a circle of radius ε_4 of the actual coefficients. By lemma 3 this precision is sufficient to distinguish new complex factors from previously computed ones. By corollary 1 this precision is also sufficient to identify completely real factors, as well as match up complex conjugated ones (see the following step).

Step 4: *For $\widetilde{h}$ in the set of factors $S = \{\widetilde{f}_{i,1}, \ldots, \widetilde{f}_{i,d_i}\}$ Do:*
At this point we have collected distinct complex factors of f in a set of approximate polynomials $T = \{\widetilde{g}_1, \ldots, \widetilde{g}_t\}$. To each of these factors belongs a factor of $f(x, 0)$, designated by $\varphi_{\iota_1}, \ldots, \varphi_{\iota_t}$. Also, the corresponding real factors of the $\widetilde{g}$'s have been produced.

Step 4.1: First, we check whether $\widetilde{h}$ corresponds to a real or a non-real factor. Let $\zeta_{i,\delta} \in \mathbb{C}$ be the root of φ_i whose approximate root $\widetilde{\zeta}_{i,\delta}$ produced $\widetilde{h}$. If $\zeta_{i,\delta}$ is real (which is usually indicated by the root approximation algorithm used in step 3), so must be $f_i(x, y, \zeta_{i,\delta})$. However, even a non-real $\zeta_{i,\delta}$ may lead to a real $f(x, y, \zeta_{i,\delta})$. *Test if*

$$\|\Im(\widetilde{h})\|_\infty > \varepsilon_1(d_i, N_i)/4.$$

From corollary 1 and (2) we deduce that this condition is necessary and sufficient for $f_i(x, y, \zeta_{i,\delta})$ to be a non-real polynomial. We refer to this outcome of the test as *the non-real case*.
Eliminate $\widetilde{h}$ from the set S. Furthermore, if $\zeta_{i,\delta}$ is non-real, let $\delta^\star$ be the index with $\widetilde{\zeta}_{i,\delta^\star} = (\widetilde{\zeta}_{i,\delta})^*$ (the equality is usually guaranteed by the root approximation algorithm employed in step 3). In that case, also *eliminate $\widetilde{f}_{i,\delta^\star}$ from S.*

Step 4.2: Next, we check whether $\widetilde{h}$ has already been treated earlier. *If*

$$\min_{1 \le l \le t} \|\widetilde{g}_l - \widetilde{h}\|_\infty < 2\,\varepsilon_4(d_{\max}, d_i, N_{\max})$$

then, again by (2), the considered factor is already in the list of factors, therefore *process the next factor $\widetilde{h}$.*

Step 4.3: Now $\widetilde{h}$ is the approximate version of a new factor. Clearly, in the non-real case $\widetilde{h}\,(\widetilde{h})^*$ is the approximate version of the corresponding real norm with respect to complex conjugation. If need be, we can compute the coefficients of this real factor to any precision by increasing the precision of the approximate root $\widetilde{\zeta}_{i,\delta}$ of φ_i.
For future tests in step 4.2, *add $\widetilde{h}$, and in the non-real case $(\widetilde{h})^*$, to the set T of $\widetilde{g}$'s and update t accordingly.*
We now treat the two cases. The real case, in which $f(x, y, \zeta_{i,\delta})$ was determined to be a real factor, is easy: *Set $g_{s+1}(x, y) \leftarrow f_i(x, y, t)$; $\widehat{\varphi}_{s+1}(t) \leftarrow \varphi_i(t)$; and $\widetilde{\vartheta} \leftarrow \widetilde{\zeta}_{i,\delta}$.*
In the non-real case, *first compute, by the Complex Trace and Norm algorithm given in section 4, a minimal polynomial $\widehat{\varphi}_{i,\delta}(t) \in \mathbb{Q}[t]$ one of whose roots, $\vartheta_{i,\delta}$, generates both $\zeta_{i,\delta} + \zeta^*_{i,\delta}$ and $\zeta_{i,\delta}\,\zeta^*_{i,\delta}$. Then express*

$$g_{s+1}(x, y) \leftarrow f_i(x, y, \zeta_{i,\delta})\, f_i(x, y, \zeta^*_{i,\delta})$$

as a polynomial in $\mathbb{Q}[x, y, t]/(\widehat{\varphi}(t))$. Also, return the approximation $\widetilde{\vartheta}_{i,\delta}$ for $\vartheta_{i,\delta}$, which is also produced by the Norm and Trace algorithm.
In conclusion to both cases, *increment s*, since we have added a new real factor, *and proceed to the next $\widetilde{h}$.* □

Even though the precision ε_3 is in the worst case unreasonably large (see the proof of theorem 1 below), the algorithm can be implemented in an adaptive more efficient way.

We can work with a lesser precision and first generate all real factors we have certified to be distinct with that precision. If the degrees of those factors add up to the degree of f, we need not certify other factors to coincide with already computed ones. The same holds for the process of identifying non-real complex factors. If we are left with some factors whose precise form is undecided, we increase the precision and try to distinguish again. Our estimate essentially proves that this process will terminate in polynomial-time. We have the following theorem.

Theorem 1. *The bit complexity of the Algorithm Factorization over the Real Numbers is bounded from above by a polynomial in* $\deg(f)$ *and the binary length of the coefficients of* f. $\boxtimes$

This theorem follows from the lemmas 2 and 3 as well as the fairly elaborate analysis of the coefficient growth in the algorithm Factorization over the Algebraic Closure (see [9, §6]) we will not give an explicit upper bound for the bit complexity of the algorithm, but show that all size bounds are polynomially dependent on $\deg(f)$ and $\log(\|f\|_\infty)$. If one were to implement the algorithm, it is paramount to determine the bounds precisely not only with respect these two parameters, but also with respect to $\|\varphi_i\|_\infty$, $\|u_{i,j,k}\|_\infty$, and $|\zeta_{i,\delta}|$. We begin the analysis by first proving a lemma on the height of a the minimal polynomial of an algebraic number in $\mathbb{Q}[z]/(\varphi(z))$.

Lemma 4. *Let* $u(z) \in \mathbb{Z}[z]/(\varphi(z))$, *where* φ *is a monic integral polynomial of degree* n, *and let* $\psi(v) \in \mathbb{Z}[v]$ *be the minimal polynomial of* $v = u(z)$. *Then*

$$\|\psi\|_\infty \leq n^{5n/2}\|\varphi\|_2^{n^3}\|u\|_1^{n^2} =: B_1(n, \|\varphi\|_2, \|u\|_1). \tag{3}$$

Proof. Let $\psi(v) =: b_m(v^m + b_{m-1}v^{m-1} + \cdots + b_0)$, and let, for all $\mu \geq 0$,

$$u(z)^\mu \pmod{\varphi(z)} =: w_{0,\mu} + w_{1,\mu}z + \cdots + w_{n-1,\mu}z^{n-1}.$$

Now $b_0, \ldots, b_{m-1}$ is the unique solution to the linear system

$$\begin{pmatrix} w_{0,0} & w_{0,1} & \ldots & w_{0,m} \\ w_{1,0} & w_{1,1} & \ldots & w_{1,m} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n-1,0} & w_{n-1,1} & \ldots & w_{n-1,m} \end{pmatrix} \begin{pmatrix} b_0 \\ \vdots \\ b_{m-1} \\ 1 \end{pmatrix} = 0.$$

Hence, using Cramer's rule and Hadamard's determinant inequality we get as the bound for both the denominator b_m and the numerators $b_m b_\mu$, $0 \leq \mu \leq m-1$, in the rational solution of this system

$$\|\psi\|_\infty \leq m^{m/2}W^m \text{ with } W := \max_{\nu,\,\mu}\{|w_{\nu,\mu}|\}. \tag{4}$$

We now need to bound W. We first observe that $\|u^\nu\|_\infty \leq \|u\|_1^\nu$ and that

$$\|z^l \pmod{\varphi}\|_\infty \leq \|\varphi\|_2^{l-n+1}, \quad l \geq n.$$

The latter follows again by using Cramer's rule and Hadamard's determinant inequality, now on the linear system arising from computing the coefficients of the quotient and remainder

in the polynomial division of z^l by φ. Combining these two bounds, we get

$$\begin{aligned}\max_{\mu}\{|w_{\mu,\nu}|\} &= \|u^\nu \pmod{\varphi}\|_\infty \\ &\le \Big(\nu(n-1)-n+2\Big)\|\varphi\|_2^{\nu(n-1)-n+1}\|u\|_1^\nu.\end{aligned}$$

Plugging this bound into (4) and crudely bounding μ by n and ν by $n-1$ we obtain (3). ⊠

Next, we show how close an approximation to a root ζ of φ we need in order to approximate the value $u(\zeta)$ of a polynomial $u(z) \in \mathbb{Z}[z]$ to a given precision.

Lemma 5. *Let $u(z) \in \mathbb{Z}[z]$, $\deg(u) < n$, and let $\zeta, \widetilde{\zeta} \in \mathbb{C}$ such that $|\zeta - \widetilde{\zeta}| < \varepsilon$. Then*

$$|u(\zeta) - u(\widetilde{\zeta})| < \varepsilon n^2(|\zeta|+\varepsilon)^n\|u\|_\infty. \tag{5}$$

Proof. This bound is simply established by expanding, with $u(z) =: \sum_j b_j z^j$,

$$\begin{aligned}|u(\zeta) - u(\widetilde{\zeta})| &= \left|\sum_{j=1}^{n-1} b_j(\zeta^j - \widetilde{\zeta}^j)\right| \\ &\le |\zeta - \widetilde{\zeta}| \sum_{j=1}^{n-1}|b_j| \sum_{l=0}^{j-1}|\zeta^l\widetilde{\zeta}^{j-l}|.\end{aligned}$$

Then (5) follows by crudely estimating the double sum. ⊠

We now are in a position to provide a polynomial estimate for N_i and ε_3.

Proof of Theorem 1. By lemma 4, we can choose

$$N_i := \max_{j,k}\{B_1(d_i, \|\varphi_i\|_2, \|u_{i,j,k}\|_1)\}.$$

Finally, we determine ε_3 such that ε_4 satisfies (2). By lemma 5, with

$$Z_i := \max_{\delta}\{|\zeta_{i,\delta}|\} \le \|\varphi_i\|_2,$$

where the inequality follows from the bound for the measure of φ_i given in lemma 2, it is sufficient to have

$$\varepsilon_3 := \min_{j,k}\Big\{\varepsilon_4 \Big/ \Big(d_i^2(|Z_i|+1)^{d_i}\|u_{i,j,k}\|_\infty\Big)\Big\}.$$

Since all $|u_{i,j,k}|$ are bounded polynomially in size (see [9, §6]), $\log(1/\varepsilon_3)$ must also be bounded by a polynomial in $\deg(f)$ and the coefficient size of f. ⊠

6. Seidenberg's Method

We now describe how one can test whether an irreducible factor produced by our algorithm Factorization over the Reals contains, as a curve, a real point. Of course, those factors with real points constitute the irreducible real components of the curve defined by the input polynomial f. The algorithm is due to Seidenberg [20], and can also be found in Jacobson's [8] book, §5.

Algorithm *Real Point Test for a Curve*

Input: A complex algebraic number ϑ given by its minimal polynomial $\psi(\vartheta)$ and an approximate isolated root $\widetilde{\vartheta}$. Furthermore, a polynomial $f(x,y,t) \in \mathbb{Q}[x,y,t]$ such that $f(x,y,\vartheta)$ is an irreducible real polynomial, monic in x.

Output: True or false, depending whether there exists a real point (x_0, y_0) such that $f(x_0, y_0, \vartheta) = 0$.

Step 1: *For* $a \leftarrow 0, 1, \ldots$ *Do:*

Test whether

$$\mathrm{GCD}\Big(f, (x-a)\frac{\partial f}{\partial y} - y\frac{\partial f}{\partial x}\Big)$$

computed over $(\mathbb{Q}[y,t]/(\varphi(t)))[x]$ *is 1. If that is true, set*

$$g(x,y,t) \leftarrow (x-a)\frac{\partial f}{\partial y}(x,y,t) - y\frac{\partial f}{\partial x}(x,y,t)$$

and go to the next step. Otherwise, try the next a. Notice that since f is irreducible, hence squarefree, there are at most $\deg_x(f)$ values of a for which this test can fail.

Step 2: *Choose* b *an integer such that*

$$b > |\Im(x_0)/\Im(y_0)|$$

for all x_0, y_0 *with* $f(x_0, y_0, \vartheta) = g(x_0, y_0, \vartheta) = 0$ *and* $\Im(y_0) \neq 0$, *and compute the resultant*

$$h(Y,t) \leftarrow \mathrm{Res}_X(f(X - bY, Y, t), g(X - bY, Y, t)).$$

Notice that $h(Y,t)$ is a non-zero polynomial in $\mathbb{Q}[Y,t]/(\varphi(t))$, and that by the input assumption $h(Y,\vartheta)$ has real coefficients.

Step 3: *Test whether* $h(Y,\vartheta)$ *has a real root.* By Seidenberg's argument, this is equivalent to the problem whether f possesses a real point. The test itself can be performed by Sturm's method on $\mathbb{Q}(\vartheta)[Y]$. In order to determine the sign of a polynomial remainder at an integer point c one evaluates that remainder at a sufficiently precise approximation of ϑ, which can be obtained from $\widetilde{\vartheta}$. $\square$

The correctness of this algorithm follows as in Seidenberg [20]. A polynomial size bound for b can be derived from the fact that x_0 and y_0 are roots in the corresponding resultants. Also, the Sturm sequence method can be performed with polynomially bounded approximations. We shall omit the arguments, which are similar to those in §5. However, for the record, we state the following theorem.

Theorem 2. *The bit complexity of the algorithm Real Point Test for a Curve is bounded from above by a polynomial in* $\deg(f)$, $\deg(\psi)$, *and the coefficient length of* f *and* ψ.

7. Literature Cited

1. Arnon, D. S., Collins, G. E., and McCallum, S., "Cylindrical algebraic decomposition I: The basic algorithm," *SIAM J. Comp.* **13**, pp. 865–877 (1984).
2. Canny, J., "Some algebraic and geometric computations in P-space," *Proc. 20th Annual ACM Symp. Theory Comp.*, pp. 460–467 (1988).

3. Chistov, A. L. and Grigoryev, D. Yu., "Subexponential-time solving of systems of algebraic equations I," *Tech. Report* **E-9-83**, Steklov Mathematical Institute, Leningrad, 1983.
4. Collins, G. E., "Quantifier elimination for real closed fields by cylindrical algebraic decomposition," *Proc. 2nd GI Conf. Automata Theory Formal Lang., Springer Lec. Notes Comp. Sci.* **33**, pp. 515–532 (1975).
5. Collins, G. E., "Infallible calculation of polynomial zeros to specified precision," in *Mathematical Software III*, edited by J. R. Rice; Academic Press, New York, pp. 35–68, 1977.
6. Collins, G. E. and Horowitz, E., "The minimum root separation of a polynomial," *Math. Comput.* **28**, pp. 589–597 (1974).
7. Dvornicich, R. and Traverso, C., "Newton symmetric functions and the arithmetic of algebraically closed fields," in *Proc. AAECC-5*, Springer Lect. Notes Comput. Sci. **356**; pp. 216–224, 1987.
8. Jacobson, N., *Basic Algebra I*; W. H. Freeman & Co., San Francisco, 1974.
9. Kaltofen, E., "Polynomial-time reductions from multivariate to bi- and univariate integral polynomial factorization," *SIAM J. Comp.* **14**, pp. 469–489 (1985).
10. Kaltofen, E., "Fast parallel absolute irreducibility testing," *J. Symbolic Computation* **1**, pp. 57–67 (1985).
11. Kaltofen, E., "Deterministic irreducibility testing of polynomials over large finite fields," *J. Symbolic Comp.* **4**, pp. 77–82 (1987).
12. Landau, E., "Sur quelques théorèmes de M. Petrovic relatifs aux zéros des fonctions analytiques," *Bull. Soc. Math. France* **33**, pp. 251–261 (1905).
13. Lipson, J., *Elements of Algebra and Algebraic Computing*; Addison-Wesley Publ., Reading, Mass., 1981.
14. Loos, R., "Computing in algebraic extensions," in *Computer Algebra, 2nd ed.*, edited by B. Buchberger et al; Springer Verlag, Vienna, pp. 173–187, 1982.
15. Mahler, K., "An inequality for the discriminant of a polynomial," *Michigan Math. J.* **11**, pp. 257–262 (1964).
16. Mignotte, M., "Some useful bounds," in *Computer Algebra, 2nd ed.*, edited by B. Buchberger et al; Springer Verlag, Vienna, pp. 259–263, 1982.
17. Pinkert, J. R., "An exact method for finding roots of a complex polynomial," *ACM Trans. Math. Software* **2**/4, pp. 351–363 (1976).
18. Renegar, J., "A faster P-space algorithm for deciding the existential theory of the reals," *Proc. 29th Annual Symp. Foundations of Comp. Sci.*, pp. 291–295 (1988).
19. Schönhage, A., "The fundamental theorem of algebra in terms of computational complexity," *Tech. Report*, Univ. Tübingen, 1982.
20. Seidenberg, A., "A new decision method for elementary algebra," *Annals Math.* **60**, pp. 365–374 (1954).
21. Trager, B. M., "Integration of algebraic functions," *Ph.D. Thesis*, MIT, 1984.
22. van der Waerden, B. L., *Modern Algebra*; F. Ungar Publ. Co., New York, 1953.
23. Wilf, H. S., "A global bisection algorithm for computing the zeros of polynomials in the complex plane," *J. ACM* **25**/3, pp. 415–420 (1978).