The basic form of the structure theorem given above has been extended in a number of ways, see for instance [4]. The essential result is not disturbed if two additional modes of extension are admitted:

$$(i \text{ in } X) \quad Dz_i = f_i$$

with the integral non-elementary over F_{i-1}

$$(i \text{ in } A) \quad z_i \text{ satisfies } p(z_i) = 0,$$

where p is a polynomial with coefficients in F_{i-1}.

The Liouvillian extensions X cause no problems at all provided the non-elementary nature of the integral of f_i can be tested. The algebraic case A does not change the statement of the structure theorem but does make some difference to the mechanics of testing the conditions that it imposes. This means that in general it will not be possible to apply the results when several X and A type extensions coexist.

References

[1] Caviness, B. F.: Methods for Symbolic Computation with Transcendental Functions. MAXIMIN **1977**, 16–43.
[2] Kaplansky, I.: An Introduction to Differential Algebra. Paris: Hermann 1957.
[3] Lang, S.: Transcendental Numbers and Diophantine Approximations. Bull. AMS 77/5, 635–677.
[4] Rothstein, M., Caviness, B. F.: A Structure Theorem for Exponential and Primitive Functions. SIAM J. Computing **8**/3, 357–367 (1979).

Dr. A. C. Norman
Computer Laboratory
University of Cambridge
Corn Exchange Street
Cambridge CB2 3QG
United Kingdom

Computing in Algebraic Extensions

R. Loos, Karlsruhe

Abstract

The aim of this chapter is an introduction to elementary algorithms in algebraic extensions, mainly over $\mathbb{Q}$ and, to some extent, over GF(p). We will talk about arithmetic in $\mathbb{Q}(\alpha)$ and GF(p^n) in Section 1 and some polynomial algorithms with coefficients from these domains in Section 2. Then, we will consider the field K of all algebraic numbers over $\mathbb{Q}$ and show constructively that K indeed is a field, that multiple extensions can be replaced by single ones and that K is algebraically closed, i.e. that zeros of algebraic number polynomials will be elements of K (Section 4–6). For this purpose we develop a simple resultant calculus which reduces all operations on algebraic numbers to polynomial arithmetic on long integers together with some auxiliary arithmetic on rational intervals (Section 3). Finally, we present some auxiliary algebraic number algorithms used in other chapters of this volume (Section 7). This chapter does not include any special algorithms of algebraic number theory. For an introduction and survey with an extensive bibliography the reader is referred to Zimmer [15].

0. Introduction

The problem type of this chapter is the task of extending exact operations in a given domain to new domains generated by adjunction of new elements. To solve this problem by algebraic algorithms means to search for exact operations in the extended domains. In most cases the representation of elements in the extension will have a more algebraic part and a more numerical part. Since the full information is always available, the general strategy is to keep the numerical precision as low as possible – just high enough to avoid ambiguities. The process of extension is a powerful method to generate new algebraic structures from given ones; in particular, the process can be iterated leading to telescopic towers of extensions. We hope that the restriction to $\mathbb{Q}$ and GF(p) will not hide the generality of this problem solving strategy.

1. Algorithms in $\mathbb{Q}(\alpha)$ and GF(p^n)

1.1 Representation

The elements of $\mathbb{Q}$ (see the chapter on arithmetic) are represented as 0 or as pairs of integers (a, b) with $b > 0$ and $\gcd(a, b) = 1$. The algebraic number α is defined as a zero of a rational polynomial $M \in \mathbb{Q}[x]$ of positive degree. We will normalize M to be monic and squarefree so that we deal only with separable extensions. In the following we will also assume M to be minimal in order to have a canonical representation of α. From a mathematical point of view things are only slightly more complicated if the minimality assumption is dropped. It would have the computational advantage that only polynomial time algorithms are needed for

arithmetic in K, the field of all algebraic numbers over $\mathbb{Q}$. In practice, however, complete factorization algorithms are fast enough that their exponential worst case computing time can be ignored*. Also, the advantage of working with a defining polynomial of minimal degree is computationally very attractive, since in many applications the speed of arithmetic in $\mathbb{Q}(\alpha)$ is the limiting factor. Also, practical considerations suggest the use of a rational polynomial instead of the similar primitive integral polynomial defining α. Otherwise, one would need a mixed integral and rational polynomial arithmetic.

Elements β of $\mathbb{Q}(\alpha)$ are represented with rational coefficients b_i in the integral base $1, \alpha, \ldots, \alpha^{m-1}$ of the vector space $\mathbb{Q}(\alpha)$ of dimension m, which is the degree of the minimal polynomial M of α. We use therefore the isomorphism $\mathbb{Q}(\alpha) \equiv \mathbb{Q}[x]/(M(x))$.

Algebraically, α has not to be distinguished from its conjugates. However, if α is real, for example, the sign of α or its size may be of interest. For this reason we indicate which root of $M(x)$ is α by an isolating interval $(r, s]$ such that $r < \alpha \leqslant s$, $r, s \in \mathbb{Q}$ and $\alpha \notin \mathbb{Q}$ or by the point interval $[r, r]$ if $\alpha = r$. Computationally, the interval endpoints can be required to have only denominators of the form 2^k in order to simplify the gcd-calculations greatly. Also, it is wise to let the degree of α be greater than 1 in order to avoid overhead. This can be easily achieved by real root calculations (see the chapter on this topic) or by special algorithms for rational root finding [11]. The algorithms to follow will also work for $m = 1$. In case α is not real, the isolating interval is an isolating rectangle in the complex plane with (binary) rational endpoints.

Elements in GF(q), $q = p^m$, $m \geqslant 1$, p a rational prime, are m-dimensional vectors over GF(p). In fact, the isomorphism $\mathrm{GF}(p^n) \equiv \mathrm{GF}(p)[x]/(M(x))$ is used for the arithmetic in GF(q). Here $M(x)$ is any irreducible polynomial of degree m over GF(p). Such defining polynomials can quickly be found by probabilistic methods [3, 13]. For definiteness we assume M to be monic again.

1.2 Arithmetic

The operations of negation, addition and subtraction in $\mathbb{Q}(\alpha)$ are operations in $\mathbb{Q}[x]$; only after multiplication is a reduction mod $M(x)$ sometimes necessary. Let d be the maximal seminorm of the representing polynomials of $\beta, \gamma \in \mathbb{Q}(\alpha)$ and of $M(x)$. Then, the time for $+, -$ is at most $O(mL(d)^2)$ and for multiplication $O(m^3 L(d)^2)$. Similarly, in GF(q) the times are $O(mL(p)^2)$ and $O(m^2 L(p)^2)$ at most; in the most frequently implemented case, $L(p) = 1$ holds.

The inverse β^{-1} is computed in $\mathbb{Q}(\alpha)$ by the extended Euclidean algorithm for the representing polynomial B and $M \bmod M$.

$$UB + VM = 1.$$

Therefore

$$U \equiv B^{-1} \bmod M.$$

Using the modular gcd-algorithm this can be done in $O(m^3 L(d) + m^2 L(d)^2)$ steps. Here the first time the assumption M being minimal is crucial.

* See "Note added in proof", p. 187.

The same method applied in GF(p)[x] yields a maximal computing time of $O(m^2)$.

1.3 The Sign of a Real Algebraic Number

Given α by the minimal polynomial M and its isolating interval $I = (r, t]$ and given furthermore $\beta \in \mathbb{Q}(\alpha)$ by its representing polynomial B, then the following algorithm computes $s = \mathrm{sign}(\beta)$.

(1) [β rational.] If $B = 0$ then $\{s \leftarrow 0$; return$\}$;
If $\deg B = 0$ then $\{s \leftarrow \mathrm{sign}(\mathrm{lc}(B))$; return$\}$.

(2) Compute B^* and b such that $(1/b)B^* = B$, $B^* \in \mathbb{Z}[x]$ and $b \in \mathbb{Q} - \{0\}$.
Set $\bar{B}$ to the greatest squarefree divisor of B^*.

(3) Obtain from I an isolating interval I^* of α containing no roots of B^* by counting the roots of $\bar{B}$ in I^* and bisection.

repeat $\{n \leftarrow \#$ of roots of $\bar{B}$ in I^*;
if $n = 0$ then $\{s \leftarrow \mathrm{sign}(b) * \mathrm{sign}(B^*(t^*))$; return$\}$;
$w \leftarrow (r^* + t^*)/2$;
if $M(r^*)M(w) < 0$ then $t^* \leftarrow w$ else $r^* \leftarrow w\}$. ■

The correctness of the algorithm depends on the fact that B is reduced mod M and M is minimal. Therefore, B and M are relatively prime and the desired interval I^* for α exists. Then $\mathrm{sign}(\beta) = \mathrm{sign}(B(\alpha)) = \mathrm{sign}(B(t^*))$.

How close can the roots $\alpha_1, \ldots, \alpha_m$ of M and the roots $\gamma_1, \ldots, \gamma_n$, $n < m$, of B lie together? We will see in Section 7 that

$$\min_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}} |\alpha_i - \gamma_j| > \frac{1}{d^m},$$

where d is the maximal seminorm of M and B. Since $\max|\alpha_i| < d$, we have $|I| < d$ at the beginning and the number k of bisections required is such that $d/2^k < d^{-m}$, or $k = O(mL(d))$.

With the results of the chapter on root isolation the analysis can be completed.

2. Polynomial Algorithms

Having available the arithmetic in $\mathbb{Q}(\alpha)$ and GF(q), it is straightforward to realize polynomial arithmetic in $\mathbb{Q}(\alpha)[x_1, \ldots, x_r]$ and $\mathrm{GF}(q)[x_1, \ldots, x_r]$. The gcd-algorithm for univariate polynomials can easily be implemented by Euclid's natural remainder sequence. For efficiency reasons it is advisable to use the monic natural p.r.s. Recently, A. K. Lenstra [10] has given a modular gcd algorithm for monic univariate polynomials in $\mathbb{Q}(\alpha)[x]$. He suggests applying the EZGCD algorithm (see chapter on polynomial remainder sequences in this volume) for $\mathbb{Q}(\alpha)[x]$. After this extension, the modified Uspensky algorithm or any other real root isolation algorithm can be extended to polynomials over $\mathbb{Q}(\alpha)$ for real algebraic numbers α. Root finding over GF(q) can be done by interesting probabilistic methods [13, 2].

3. Resultant Calculus

Let $A = \sum_{i=0}^{m} a_i x^i$ and $B = \sum_{i=0}^{n} b_i x^i$ be two polynomials over a commutative ring R with identity. The Sylvester matrix of A and B is the $m+n$ by $m+n$ matrix

$$M = \begin{pmatrix} a_m & a_{m-1} & \cdots & & a_0 & & \\ & a_m & a_{m-1} & \cdots & & a_0 & \\ & & \cdots\cdots & & & & \\ & & & a_m & a_{m-1} & \cdots & a_0 \\ b_n & b_{n-1} & \cdots & b_1 & b_0 & & \\ & b_n & \cdots & & & b_0 & \\ & & \cdots\cdots & & & & \\ & & & b_n & \cdots & & b_0 \end{pmatrix}.$$

The upper part of M consists of n rows of elements of A, the lower part of m rows of elements of B, where all entries not shown are zero. The *resultant* of A and B is defined by

$$\operatorname{res}(A, B) = \det(M).$$

Clearly the resultant is an element of R and we have

$$\operatorname{res}(A, B) = (-1)^{mn} \operatorname{res}(B, A), \tag{1}$$

$$\operatorname{res}(aA, B) = a^n \operatorname{res}(A, B), \qquad a \in R. \tag{2}$$

By definition

$$\operatorname{res}(a, B) = a^n, \qquad a \in R,$$

$$\operatorname{res}(a, b) = 1, \qquad a, b \in R. \tag{3}$$

With m indeterminates α_i, $1 \leqslant i \leqslant m$, we construct

$$A_m(x) = \prod_{i=1}^{m} (x - \alpha_i) = \sum_{i=0}^{m} a_i^{(m)} x^i.$$

Clearly, we will be interested mainly in the case where the roots of $A_m(x)$ are substituted for the indeterminates α_i. But all resultant relations in this section will be derived without the assumption of the existence of roots, thus with the weaker assumption that the α_i are indeterminates. The coefficients a_i are related to the indeterminates α_i by

$$\begin{aligned} a_m^{(m)} &= s_m = 1, \\ -a_{m-1}^{(m)} &= s_{m-1} = \alpha_1 + \cdots + \alpha_m, \\ a_{m-2}^{(m)} &= s_{m-2} = \alpha_1\alpha_2 + \alpha_1\alpha_3 + \cdots + \alpha_{m-1}\alpha_m, \\ &\cdots \\ (-1)^m a_0^{(m)} &= s_0 = \alpha_1\alpha_2 \cdots \alpha_m, \end{aligned}$$

where the s_i are the elementary symmetrical polynomials.

The coefficients $a_i^{(m)}$ are linear in α_m. Let us define $A_{m-1}(x) = A_m(x)/(x - \alpha_m)$. Between the coefficients of A_m and A_{m-1} considered as polynomials in the α_i's the relation

$$a_{i-1}^{(m-1)}(\alpha_1, \ldots, \alpha_{m-1}) = a_i^{(m)}(\alpha_1, \ldots, \alpha_{m-1}, 0), \qquad 1 \leqslant i \leqslant m \tag{4}$$

holds.

We are now ready to prove the

Lemma. *Let $B(x)$ be a polynomial over an integral domain R, $\deg(B) > 0$, and let $m > 1$ be an integer. With m indeterminates α_i let $A_m(x) = \prod_{i=1}^{m} (x - \alpha_i)$ and $A_{m-1}(x) = A_m(x)/(x - \alpha_m)$. Then*

$$\operatorname{res}(A_m, B) = B(\alpha_m) \operatorname{res}(A_{m-1}, B).$$

Proof. For $1 \leqslant i < m + n$ add to the last column of the Sylvester matrix M of A_m and B α_m^{m+n-i} times the ith column. For the resulting matrix M_1 we have $\det(M_1) = \det(M)$ and the elements of the last column from top to bottom are $\alpha_m^{n-1} A_m(\alpha_m), \ldots, \alpha_m^0 A_m(\alpha_m), \alpha_m^{m-1} B(\alpha_m), \ldots, \alpha_m^0 B(\alpha_m)$. Since $A_m(\alpha_m) = 0$ we take the factor $B(\alpha_m)$ out of the last column resulting in a matrix M_2 with the last column $0, \ldots, 0, \alpha_m^{m-1}, \ldots, \alpha_m^0$ and

$$\operatorname{res}(A_m, B) = \det(M) = \det(M_1) = B(\alpha_m) \det(M_2). \tag{5}$$

Let us consider both sides of (5) as polynomials in α_m. Since M has n rows of coefficients of $A(\alpha_m)$, which are at most linear in α_m, the left-hand side is of degree n or less in α_m. On the right-hand side the factor $B(\alpha_m)$ is already of degree n. Since R is an integral domain $\det(M_2)$ is of degree 0 in α_m. Taking $\det(M_2)$ at $\alpha_m = 0$ the last column becomes now $0, \ldots, 0, 0, \ldots, 1$ and the coefficients of A_m are transformed into the coefficients of A_{m-1} according to (4). Expansion of $\det(M_2)|_{\alpha_m = 0}$ with respect to the last column results in the $m + n - 1$ by $m + n - 1$ matrix with $\det(M_3) = \det(M_2) = \operatorname{res}(A_{m-1}, B)$ which together with (2) proves the lemma. ■

Theorem 1 immediately follows, which represents the resultant as symmetrical polynomial in the indeterminates α_i [14].

Theorem 1. *Let $A(x) = a_m \prod_{i=1}^{m} (x - \alpha_i)$ and $B(x) = b_n \prod_{i=1}^{n} (x - \beta_i)$ be polynomials over an integral domain R with indeterminates $\alpha_1, \ldots, \alpha_m$ and $\beta_1, \ldots, \beta_n$. Then*

$$\operatorname{res}(A, B) = (-1)^{mn} b_n^m \prod_{i=1}^{n} A(\beta_i), \tag{6}$$

$$\operatorname{res}(A, B) = a_m^n \prod_{i=1}^{m} B(\alpha_i), \tag{7}$$

$$\operatorname{res}(A, B) = a_m^n b_n^m \prod_{i=1}^{m} \prod_{j=1}^{n} (\alpha_i - \beta_j). \tag{8}$$

Proof. The theorem holds for $m = 0$ or $n = 0$ with the convention $\prod_{i=k}^{l} f_i = 1$, for $l < k$. Eq. (6) follows from (7) by (1), also (8) follows from (7) immediately. We prove (7) by induction on m. $\operatorname{res}(A_1, B) = B(\alpha_1)$, where $A_1(x) = x - \alpha_1$, follows from the expansion of the determinant with respect to the last row. Now by (2),

$\operatorname{res}(A, B) = a_m^n \operatorname{res}(A_m, B)$ and an inductive application of the Lemma results in (7). ■

We state now some resultant relations in which the indeterminates α_i do not occur.

Theorem 2. *Let $A(x)$ and $B(x)$ be polynomials of positive degree over a commutative ring R with identity. Then there exist polynomials $S(x)$ and $T(x)$ over R with* $\deg(S) < \deg(B)$ *and* $\deg(T) < \deg(A)$ *such that*

$$AS + BT = \operatorname{res}(A, B). \tag{9}$$

Theorem 2 is a special instance of (1) in the chapter on polynomial remainder sequences.

Theorem 3. *Let A, B_1 and B_2 be polynomials over an integral domain. Then*

$$\operatorname{res}(A, B_1 B_2) = \operatorname{res}(A, B_1) \operatorname{res}(A, B_2). \tag{10}$$

Proof.

$$\operatorname{res}(A, B_1 B_2) = a_m^{n_1 + n_2} \prod_{i=1}^{m} (B_1(\alpha_i) B_2(\alpha_i)) = \operatorname{res}(A, B_1) \operatorname{res}(A, B_2). \quad ■$$

Theorem 4. *Let A, B, Q be polynomials over an integral domain and let* $\deg(A) = m$, $\operatorname{lc}(A) = a_m$, $\deg(B) = n$, $\deg(AQ + B) = l$. *Then*

$$\operatorname{res}(A, AQ + B) = a_m^{l-n} \operatorname{res}(A, B). \tag{11}$$

Proof. Again we use (7).

$$\operatorname{res}(A, B) = a_m^n \prod_{i=1}^{m} B(\alpha_i) = a_m^n \prod_{i=1}^{m} (A(\alpha_i) Q(\alpha_i) + B(\alpha_i)) = a_m^{n-l} \operatorname{res}(A, AQ + B). \quad ■$$

Theorem 3 may be used to increase the efficiency of the resultant calculation whenever a factorization of one of the polynomials is known. For example by (10) $\operatorname{res}(A, x^k B) = \operatorname{res}(A, B) \prod_{i=1}^{k} \operatorname{res}(A, x)$ and by (6)

$$\operatorname{res}(A, x) = (-1)^m A(0) = (-1)^m a_0.$$

Therefore

$$\operatorname{res}(A, x^k B) = (-1)^{mk} a_0^k \operatorname{res}(A, B). \tag{12}$$

Let $\deg(A) - \deg(B) = k \geqslant 0$. Then Eq. (12) together with (1) shows also that there is no loss of generality in the assumption that the polynomials of the resultant are of a specific degree as we stated in the chapter on polynomial remainder sequences.

Theorem 4 suggests an alternative way to calculate the value of the resultant. Moreover, it provides a proof of the next theorem, sometimes called the standard theorem on resultants [9], which follows immediately from (8), without any reference to the indeterminates α_i and β_i in (8).

Theorem 5. *Let A and B be non-zero polynomials over an integral domain R. Then* $\operatorname{res}(A, B) = 0$ *if and only if* $\deg(\gcd(A, B)) > 0$.

Proof. The theorem holds if A or B is constant. Assume $\deg(A) \geqslant \deg(B) > 0$. Working over the quotient field Q of R let $A = P_1$, $B = P_2$, $P_i = Q_i P_{i+1} + P_{i+2}$, $1 \leqslant i \leqslant k-2$, $k \geqslant 3$, be a polynomial remainder sequence, thus

$$\deg(P_1) \geqslant \deg(P_2) > \cdots > \deg(P_k) \geqslant 0, \qquad P_{k+1} = 0.$$

Let $n_i = \deg(P_i)$ and set $A = P_{i+1}$, $B = P_{i+2}$ and $Q = Q_i$ in (11). Using also (1) we obtain

$$\operatorname{res}(P_i, P_{i+1}) = (-1)^{n_i n_{i+1}} \operatorname{lc}(P_{i+1})^{n_i - n_{i+2}} \operatorname{res}(P_{i+1}, P_{i+2}), \tag{13}$$

or

$$\operatorname{res}(P_1, P_2) = \operatorname{res}(P_{k-1}, P_k) \prod_{i=i}^{k-2} (-1)^{n_i n_{i+1}} \operatorname{lc}(P_{i+1})^{n_i - n_{i+2}}, \tag{14}$$

where lc denotes the leading coefficient.

If $\deg(P_k) = \deg(\gcd(A, B)) = 0$ then $\operatorname{res}(P_{k-1}, P_k) = \operatorname{lc}(P_k)^{n_{k-1}} \neq 0$ by (3). Otherwise we apply (11) again and since $P_{k+1} = 0$ the resultant vanishes. ■

In [5] efficient algorithms for resultant calculation are given which are finally based on Eq. (14). They are superior to an evaluation of the determinant of the Sylvester matrix. In fact, the maximum computing time to calculate the resultant of two r-variate polynomials of maximal degree n and maximal seminorm d is $O(n^{2r+1} L(d) + n^{2r} L(d)^2)$.

4. Arithmetic in the Field K of All Algebraic Numbers over $\mathbb{Q}$

First we consider arithmetical operations on algebraic numbers. The following theorem gives the arithmetic in the field K of all algebraic numbers over $\mathbb{Q}$:

Theorem 6 (Loos 1973). *Let $A(x) = a_m \prod_{i=1}^{m} (x - \alpha_i)$ and $B(x) = b_n \prod_{j=1}^{n} (x - \beta_j)$ be polynomials of positive degree over an integral domain R with roots $\alpha_1, \ldots, \alpha_m$ and $\beta_1, \ldots, \beta_n$ respectively. Then the polynomial*

$$r(x) = (-1)^{mn} g a_m^n b_n^m \prod_{i=1}^{m} \prod_{j=1}^{n} (x - \gamma_{ij})$$

has the $m \cdot n$ roots, not necessarily distinct, such that

(a) $r(x) = \operatorname{res}(A(x - y), B(y))$, $\gamma_{ij} = \alpha_i + \beta_j$, $g = 1$,

(b) $r(x) = \operatorname{res}(A(x + y), B(y))$, $\gamma_{ij} = \alpha_i - \beta_j$, $g = 1$,

(c) $r(x) = \operatorname{res}(y^m A(x/y), B(y))$, $\gamma_{ij} = \alpha_i \beta_j$, $g = 1$,

(d) $r(x) = \operatorname{res}(A(xy), B(y))$, $\gamma_{ij} = \alpha_i / \beta_j$, $B(0) \neq 0$, $g = (-1)^{mn} B(0)^m / b_n^m$.

Proof. The proof is based on relation (6) in all four cases.

(a)
$$\operatorname{res}(A(x - y), B(y)) = (-1)^{mn} b_n^m \prod_{j=1}^{n} A(x - \beta_j)$$
$$= (-1)^{mn} a_m^n b_n^m \prod_{i=1}^{m} \prod_{j=1}^{n} (x - (\alpha_i + \beta_j)).$$

(b) $\operatorname{res}(A(x+y), B(y)) = (-1)^{mn} b_n^m \prod_{j=1}^{n} A(x+\beta_j)$

$$= (-1)^{mn} a_m^n b_n^m \prod_{i=1}^{m} \prod_{j=1}^{n} (x - (\alpha_i - \beta_j)).$$

(c) $\operatorname{res}(y^m A(x(y), B(y)) = (-1)^{mn} b_n^m \prod_{j=1}^{n} \beta_j^m A(x/\beta_j)$

$$= (-1)^{mn} a_m^n b_n^m \prod_{i=1}^{m} \prod_{j=1}^{n} (x - \alpha_i \beta_j).$$

(d) $\operatorname{res}(A(xy), B(y)) = (-1)^{mn} b_n^m \prod_{j=1}^{n} A(x\beta_j)$

$$= (-1)^{mn} a_m^n b_n^m \prod_{i=1}^{m} \prod_{j=1}^{n} (x\beta_j - \alpha_i)$$

$$= (-1)^{mn} a_m^n \prod_{i=1}^{m} b_n \left(\prod_{j=1}^{n} \beta_j \right) \prod_{j=1}^{n} (x - \alpha_i/\beta_j)$$

$$= (-1)^{mn} a_m^n b_0^m \prod_{i=1}^{m} \prod_{j=1}^{n} (x - \alpha_i/\beta_j) \quad \text{mit} \quad b_0 \neq 0. \ \blacksquare$$

Theorem 6 constructs explicit polynomials and we see that, except in case (d) the polynomial $r(x)$, to within a sign, is monic if A and B are. We have therefore

Corollary 1. *All algebraic integers over R form a ring.*

We denote the ring by R_∞.

Corollary 2. *All algebraic numbers over R form a field.*

We denote the field by K, where Q is the quotient field of R.

Since the degree of $r(x)$ is $m \cdot n$, the resultants are linear in the particular case the given polynomials are. We conclude that the rational numbers over R form a subfield Q of K and that R forms a subring of R_∞. We convince ourselves that R_∞ is an integral domain by considering Theorem 6, case (c) with $A(0) = 0$, $r(0) = \operatorname{res}(y^m A(0), B(y))$. By (1), (2), (3) and (12), we find $r(0) = A(0)^n b_n^m$. Since R has no zero divisors the same holds for R_∞.

Theorem 6 is the base of Algorithm 1. Obviously, it is sufficient to consider only addition and multiplication of algebraic numbers. For if the number α is defined by the polynomial $A(x)$, then the polynomial $A(-x)$ defines $-\alpha$ and $x^m A(1/x)$, $m = \deg(A)$, defines $1/\alpha$ if $\alpha \neq 0$.

Algorithm 1 (Algebraic number arithmetic).

Input: Let R be an Archimedian ordered integral domain. α, β are algebraic numbers represented by two isolating intervals I, J having endpoints in Q and by two defining polynomials A and B of positive degree over R.

Output: An isolating interval K and a defining primitive squarefree polynomial $C(x)$ representing $\gamma = \alpha + \beta$ (or $\gamma = \alpha * \beta$ for multiplication).

(1) [Resultant]
$r(x) = \operatorname{res}(A(x-y), B(y))$
$(r(x) = \operatorname{res}(y^m A(x/y), B(y))$ for multiplication).

(2) [Squarefree factorization] $r(x) = D_1(x) D_2(x)^2 \cdots D_f(x)^f$.

(3) [Root isolation] Generate isolating intervals or rectangles $I_{11}, \ldots, I_{1g_1}, \ldots, I_{fg_f}$ such that every root of D_i is contained in exactly one I_{ij} and $I_{ij} \cap I_{kl} = \emptyset$, $1 \leqslant i, k \leqslant f$, $1 \leqslant j \leqslant g_i$, $1 \leqslant l \leqslant g_k$, $(i,j) \neq (k,l)$.

(4) [Interval arithmetic] Set $K = I + J$ ($I * J$ for multiplication) using exact interval arithmetic over Q.

(5) [Refinement] If there is more than one I_{ij} such that $K \cap I_{ij} \neq 0$, bisect I and J and go back to step (4). Otherwise, return K and $C(x) = D_i(x)$. ■

Note that the computing time of the algorithm is a polynomial function of the degrees and seminorms of α and β. In practical implementation it may be preferable to replace in step 2 the squarefree factorization by a complete factorization, which would give the algorithm an exponential maximum computing time.*

The effectiveness of step (3) depends essentially on a non-algebraic property of the underlying ring R, its Archimedean order. A ring is *Archimedean ordered* if there exists for every element A a natural number (i.e. a multiple of the identity of the ring) N such that $N - A > 0$. Let us for example take the non-Archimedean ordered ring of polynomials over the rationals, $\mathbb{Q}[x]$, where an element is called positive if the leading coefficient is positive. Thus the element x is greater than any rational number and there is no assurance that an interval or rectangle containing x can be made arbitrarily small by bisection. It is still possible to count the zeros in intervals over non-Archimedean rings by Sturm's theorem, but root isolation requires Archimedean order.

The loop in step (4) and step (5) is only executed a finite number of times, since by Theorem 6 exactly one isolating interval contains the sum $\alpha + \beta$ (or the product $\alpha * \beta$) and bisection decreases the length of the interval K, computed by exact interval arithmetic, under any bound. Therefore, the input assumption of the Archimedean order enforces the termination of the algorithm.

The proof of the theorem is based on the relation (6) which in turn follows immediately from Lemma 1. We will show, using the equivalent relation (7), how similar constructions of defining polynomials for algebraic numbers can be established. If α is defined by A, we consider the norm $N_\alpha = \operatorname{res}_\alpha(A(\alpha), \cdot)$ as a polynomial operator with indeterminate α. In order to compute any function $g(\alpha)$ composed finally by ring operations on α only, we have to apply the operator N_α to $x = g(\alpha)$ yielding

$$N_\alpha(x - g(\alpha)) = \operatorname{res}_\alpha(A(\alpha), x - g(\alpha)) = a_m^n \prod_{i=1}^{m} (x - g(\alpha_i))$$

which shows that $N_\alpha(x - g(\alpha))$ is a polynomial having $g(\alpha)$ as root. By iteration, the method can be extended to any polynomial function of several algebraic numbers. Let α, β be defined by A and B respectively. In order to compute, say $f(\alpha, \beta) = \alpha + \beta$, we form

$$\begin{aligned} N_\alpha(N_\beta(x - f(\alpha, \beta))) &= \operatorname{res}_\alpha(A(\alpha), \operatorname{res}_\beta(B(\beta), x - (\alpha + \beta))) \\ &= \operatorname{res}_\alpha\left(A(\alpha), b_n \prod_{i=1}^{n} (x - (\alpha + \beta_j))\right) \\ &= a_m^n b_n^m \prod_{i=1}^{m} \prod_{j=1}^{n} (x - (\alpha_i + \beta_j)) \end{aligned}$$

which is up to a sign the defining polynomial of Theorem 6, (a). In fact, the method can still be further generalized. All that is required is that the relation $x = f(\alpha, \beta)$ may be transformed into a polynomial relation, say $F(x, \alpha, \beta)$. The following theorem gives an application. Let us consider fractional powers of algebraic numbers. Since the reciprocal of an algebraic number can be computed trivially we restrict ourselves to positive exponents.

Theorem 7 (Fractional powers of algebraic numbers). *Let A be a polynomial of positive degree m over an integral domain R with roots $\alpha_1, \ldots, \alpha_m$. Let p, q be positive integers. Then*

$$r(x) = \operatorname{res}(A(y), x^q - y^p)$$

has the roots $\alpha_i^{p/q}$, $i = 1, \ldots, m$.

Proof.

$$r(x) = a_m^p \prod_{i=1}^{m} (x^q - \alpha_i^p). \quad \blacksquare$$

We can base on Theorem 7 an algorithm for the computation of fractional powers of algebraic numbers, which would be quite similar to Algorithm 1.

Another application of our algebraic number calculus allows a transformation of the algebraic number representation in $\mathbb{Q}(\alpha)$ as a polynomial $\beta = B(\alpha)$ to a defining polynomial for β. We get $N_\alpha(x - B(\alpha))$ as defining polynomial for β.

Theorem 8. *Let α be an algebraic number over R and A its defining polynomial of degree $m > 0$. Let $\beta = \sum_{i=0}^{m-1} b_i \alpha^i = B(\alpha)$, where $b_i \in Q$, $\deg(B) = n < m$. Then β is algebraic over R and a root of*

$$r(x) = \operatorname{res}(x - B(y), A(y)).$$

If α is an algebraic integer then so is β, provided $b_i \in R$.

Proof. By Theorem 1, Eq. (6), $r(x) = (-1)^{mn} a_m^n \prod_{i=1}^{m} (x - B(\alpha_i))$. $\blacksquare$

Corollary 3. *If A of Theorem 8 is the minimal polynomial of α then B is uniquely determined.*

Proof. Suppose $\beta = B^*(\alpha)$, $\deg(B^*) < m$ and $B^* \neq B$. $B(\alpha) - B^*(\alpha) = \beta - \beta = 0$.

This is a polynomial, not identically vanishing, of degree $< m$, a contradiction to the minimal property $\deg(A) = m$. $\blacksquare$

Theorem 8 can be used to compute the sign of a real algebraic number differently from the approach in Section 1. Given $\beta = B(\alpha)$, we construct $r(x)$ from the theorem and compute $I_\beta = B(I_\alpha)$ by interval arithmetic such that I_β is isolating with respect to $r(x)$. The position of I_β in relation to 0 gives the sign of β. The position in relation to I_α, made disjoint from I_β, gives an algorithm for real algebraic number comparison.

Historical Note. The resultant $\operatorname{res}_y(A((x - y)/2), A((x + y)/2))$ was considered by Houscholder in [8] and stimulated our interest in resultants.

A special case of Theorem 6 is the resultant $\operatorname{res}_y(A(x + y), A(y))$, a polynomial having the roots $\alpha_i - \alpha_j$. The task of obtaining lower bounds on $\min_{i,j} |\alpha_i - \alpha_j|$ is reduced by it to the problem of a lower bound for the roots of $r(x)/x^m$. By this approach Collins [6] improved Cauchy's [4] lower bound for the minimum root separation.

5. Constructing Primitive Elements

The representation $\beta = B(\alpha)$ allows the construction of extensions of R and Q as shown by the next two theorems.

Theorem 9. *Let α be an algebraic integer over R and A its defining polynomial of degree $m > 1$. Then the set of all algebraic numbers represented by $\beta = \sum_{i=0}^{m-1} b_i \alpha^i = B(\alpha)$, where $b_i \in R$, forms a ring of algebraic integers.*

We call the ring a *simple extension of R* and denote it by $R[\alpha]$.

Proof. Since $C(x) = Q(x)A(x) + B(x)$, where $B = 0$ or $\deg(B) < m$, and $A(\alpha) = 0$ we have $B(\alpha) = C(\alpha)$. Hence, there is an isomorphism between $R[\alpha]$ and $R[x]/(A(x))$, the ring of residue classes of A. $\blacksquare$

Theorem 10. *Let α be an algebraic number over a field Q and A its defining polynomial of degree $m > 0$. Then the set of all algebraic numbers represented by $\beta = \sum_{i=0}^{m-1} b_i \alpha^i = B(\alpha)$, where $b_i \in Q$, forms a field.*

We call the field a *simple algebraic extension of Q* and denote it by $Q(\alpha)$.

Proof. We have to show that every non-zero element β of $Q(\alpha)$ has a multiplicative inverse. First, assume $\deg(\gcd(A(x), B(x))) = 0$. By Theorem 5, $\operatorname{res}(A, B) \neq 0$. Theorem 2 gives, for $x = \alpha$, $B(\alpha)T(\alpha) = \operatorname{res}(A, B)$ with $\deg(T) < m$. Therefore, $T(\alpha)/\operatorname{res}(A, B)$ is the inverse of $B(\alpha)$. Now, let $C(x) = \gcd(A, B)$, $\deg(C) > 0$, and $A = CA^*$. Clearly, $C(\alpha) \neq 0$, otherwise $\beta = B(\alpha) = 0$. Therefore, $A^*(\alpha) = 0$. Replace A by A^* and apply the first argument of the proof, observing that $\deg(\gcd(A^*, B)) = 0$. $\blacksquare$

Extensions $Q(\alpha)$ are called *separable*, if the defining polynomial for α is squarefree.

Clearly, $R \subseteq R[\alpha] \subset R_\infty$. Since R_∞ was shown to be an integral domain the same holds for $R[\alpha]$. Also $Q \subseteq Q(\alpha) \subset K$. All previous results remain valid with $R[\alpha]$ and $Q(\alpha)$ in place of R and Q respectively. In particular $R[\alpha][\beta]$ is a ring and $Q(\alpha)(\beta)$ a

field again. We call $Q(\alpha)(\beta) = Q(\alpha, \beta)$ a double extension of Q. Of central importance is

Theorem 11. *Every separable multiple extension is a simple extension.*

We give the proof constructively by an algorithm:

Algorithm 2 (SIMPLE).

Inputs: $A(x)$ and $B(x)$, two primitive squarefree polynomials of positive degree over R, an Archimedean ordered integral domain, I and J two isolating intervals over Q such that α is represented by I and A and β by J and B.

Outputs: An isolating interval K, a defining primitive squarefree polynomial $C_0(x)$ over R, representing γ, and two polynomials over Q, $C_1(x)$ and $C_2(x)$, such that $\alpha = C_1(\gamma)$, $\beta = C_2(\gamma)$.

(1) [Resultant] $r(x, t) = \mathrm{res}(A(x - ty), B(y))$.
[Here, by (6), $r(x, t)$ has root $\gamma_{ij} = \alpha_i + t\beta_j$.]

(2) [Squarefree] Compute the smallest positive integer t_1 such that $\deg(\gcd(r(x, t_1), r'(x, t_1))) = 0$. Set $C_0(x) = r(x, t_1)$.
[This implies that all γ_{ij}, $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n$, are different, so by (1)

$$\alpha_i - \alpha_k \neq t_1(\beta_j - \beta_l)$$

for all pairs (i, j) and (k, l) with $i \neq k$ and $j \neq l$. Since there are only finitely many pairs but infinitely many positive integers, a t_1 can always be found.]

(3) [Interval arithmetic] By repeated bisection of I and J construct K such that $K = I + t_1 J$ and K is an isolating interval for $C_0(x)$. [Obviously $Q(\gamma) \subseteq Q(\alpha, \beta)$ for $\gamma = \gamma_{11}$.]

(4) [gcd] Using arithmetic in $Q(\gamma)$ compute $B^*(x) = \gcd(A(\gamma - t_1 x), B(x))$. [By the construction β_j is a root of $B^*(x)$ and there is only one such β according to (2). Therefore $\deg(B^*) = 1$ and $B^*(x) = x - \beta$, where B^* is monic by convention.]

(5) [Exit] Set $C_2(x) = -$ trailing coefficient of $B^*(x)$, $C_1(x) = x - t_1 C_2(x)$. [We obtain $\beta = C(Y)$ and $\alpha = \beta - t_1 C(Y)$.] ■

We have used a modification of Theorem 6 (a) in step (1). Similarly, (b)–(d) of Theorem 6 may be used for constructing primitive elements γ. The given algorithm occurs under the name SIMPLE in Collin's quantifier elimination algorithm (see the chapter of this name in this volume). As algorithm 1, this algorithm has also polynomial computing time. If n is the maximal degree of α and β and d is its maximal seminorm then the seminorm of C is $O(d^{2n})$. Empirically, it turns out that the most expensive operation of SIMPLE is the division in step (4) to make B^* monic; up to 80% of the total time can be saved if α and β need not to be represented in $Q(\gamma)$. Note that the degree of γ is in general n^2, which indicates the computational limitations of this approach.

6. The Algorithm NORMAL

A last application of the algebraic number calculus allows to represent roots of algebraic number polynomials as roots of integral polynomials. It gives a constructive proof that K is algebraically closed.

Using Theorem 11 we can show that the roots of a polynomial with algebraic number coefficients are algebraic numbers. Let $B^*(x) = \sum_{j=0}^{n} \beta_j x^j$ be such a polynomial. First we compute $Q(\alpha) = Q(\beta_0, \ldots, \beta_n)$ and express $\beta_j = B_j(\alpha)$, i.e. by polynomials over Q according to the construction of the last proof. By the next theorem we obtain a construction for a polynomial $r(x)$ over R having among its roots the roots of $B^*(x)$.

Theorem 12. *Let* $A(y) = \sum_{i=0}^{m} a_i y^i = a_m \prod_{i=1}^{m} (y - \alpha_i)$ *be a primitive squarefree polynomial over an integral domain* R.
Let $B(y, x) = \sum_{j=0}^{n} B_j(y) x^j$ *be a bivariate polynomial over* R *such that* $\deg(\gcd(A(y), B(y))) = 0$. *Let* $k = \deg(B(y, x))$ *and* $r(x) = \mathrm{res}(A(y), B(y, x))$. *Then* $r(x) = a_k^k \prod_{i=1}^{m} B_n(\alpha_i) \prod_{j=1}^{n} (x - \beta_{ij})$, *where the* β_{ij} *are defined by*

$$B(\alpha_i, x) = B_n(\alpha_i) \prod_{j=1}^{n} (x - \beta_{ij}), \qquad 1 \leqslant i \leqslant m.$$

Proof. By Theorem 1 (7),

$$r(x) = a_m^k \prod_{i=1}^{m} B(\alpha_i, x) = a_m^k \prod_{i=1}^{m} B_n(\alpha_i) \prod_{j=1}^{n} (x - \beta_{ij}). \quad ■$$

Corollary 4. *The roots of a polynomial with algebraic number coefficients are algebraic numbers.*

The algorithm NORMAL that occurs with SIMPLE in Collins' quantifier elimination algorithm is based on this theorem.

Algorithm 3 (NORMAL).

Input: A polynomial $B^*(x)$ of degree n over $Q(\beta_0, \ldots, \beta_n)$, where the β_i are given by defining primitive squarefree polynomials $B_i^*(x)$ over R, an integral Archimedean ordered domain, and isolating intervals $I(\beta_i)$, $\beta \leqslant i \leqslant n$.

Output: A polynomial $C(x)$ over R having among its roots the roots of $B^*(x)$. By an obvious modification, isolating intervals of $B^*(x)$ over Q may also be computed.

(1) [$Q(\alpha) = Q(\beta_0, \ldots, \beta_n)$] By repeated application of Algorithm 2 compute a primitive squarefree polynomial A having the root α such that $Q(\alpha) = Q(\beta_0, \ldots, \beta_n)$ and rational polynomials $\bar{B}_j(y)$ such that $\beta_j = \bar{B}_j(\alpha)$ and $\deg(\bar{B}_j) < m = \deg(A)$.

(2) Compute $d \in R$ such that $B_j = d\bar{B}_j \in R[x]$, $0 \leqslant j \leqslant n$, and set $B(y, x) = \sum_{j=0}^{n} B_j(y) x^j$.

(3) [gcd] Set $D(y) = \gcd(A(y), B_n(y))$. Here $\deg(D) \leqslant \deg(B_n) < m$.

(4) [Reduction?] If $\deg(D) > 0$, go to step (6).

(5) Set $C(x) = \text{res}(A(y), B(y, x))$ and exit. By the preceeding Theorem 12 $C(x) \neq 0$ and every root of B^* is a root of $C(x)$.

(6) [Reduce] Set $\bar{A} = A/D$. Now $\bar{A}(\alpha) = 0$, since $B_n(\alpha) \neq 0$ and hence $D(\alpha) \neq 0$. Set $B_j \equiv b_j \bmod \bar{A}$, for $0 \leqslant j \leqslant n$. Compute $d \in R$ such that $B_j = d\bar{B}_j \in R[x]$, for $0 \leqslant j \leqslant n$. Now $\deg(\gcd(\bar{A}, B_n)) = 0$ since A is squarefree. Go back to step (5). ■

Again, the computing time is a polynomial function.

7. Some Applications

Suppose we know that two polynomials are relatively prime. How close can the roots α_i and β_j be? We look for $\min|\alpha_i - \beta_j|$. Using our approach, we construct a polynomial $C(x)$ having all $\alpha_i - \beta_j$ as roots and determine the radius of the circle around the origin, in which no roots of $C(x)$ are located. C is given by $\text{res}_y(A(x + y), B(y))$ according to Theorem 6 (b) and has seminorm $\leqslant d^{2n}$ by Hadamard's inequality (see the chapter on useful bounds). Then we get $\min|\alpha_i - \beta_j| > d^{-2n}$. We need this result to prove termination of algorithms isolating the roots of two relatively prime polynomials "away" from each other.

Next, suppose that a polynomial $A(x)$ is squarefree. How small can $A(\beta_i)$ be at any zero β_i of the derivative? The answer is given by $\text{res}_\beta(A'(\beta), x - A(\beta)) = C(x)$. The lower root bound of C is of the same order as the previous one.

If we do not require A to be squarefree, then the polynomial $C(x)$ is of order $k > 0$, where k gives the degree of $\gcd(A, A')$. Since $c_0, c_1, \ldots, c_{k-1}$ can be expressed by the coefficients of A, the fact that $\deg\gcd(A, A') = k$ is expressible as $c_0 = 0, \ldots, c_{k-1} = 0$ in the coefficients of A only. This is an alternative to Collins' psc-theorem used for quantifier elimination.

The last two observations give termination bounds in the Collins-Loos real root isolation algorithm (see the chapter on root isolation in this volume).

Let $A(x)$ and $B(x)$ be integral polynomials of the form $A = A_1A_2$ and $B = B_1B_2$ such that there exists $k \in Q$ with $B_1(x) = A_1(x - k)$, which we call shifted factors. In order to detect such shifted factors we use again Theorem 6 (b) and compute $C(x) = \text{res}_y(A(x + y), B(y)) = c_{n+m}\prod_j(x - (\alpha_i - \beta_j))$. If A and B have a shifted factor k then there is some root pair α_i, β_j with $\alpha_i - \beta_j = k$ which can be detected as rational root of the polynomial $C(x)$. This idea is used in step 3 of Gosper's algorithm (see the chapter on summation). In a similar manner "rotated" factors $B_1(x) = A_1(x \cdot k)$ can be detected using part (c) of Theorem 6.

With a rational root finder all pairs $\alpha_i + \alpha_j = c_{ij}$ and $\alpha_i\alpha_j = d_{ij}$ can be formed from the resultants (a) and (c) of Theorem 6 with $B = A$. Then, one can form trial divisors $x^2 - c_{ij}x + d_{kl}$ for finding quadratic factors of A [11].

The construction of powers of algebraic numbers used in root-squaring algorithms can be expressed by Theorem 7. Hence, the polynomials entering Graeffe's method are resultants.

Note added in proof: In the meantime A. K. Lenstra, H. W. Lenstra and L. Lovacs discovered a polynomial time bounded factorization algorithm for integral polynomials. Therefore, there is no computational objection against the use of minimal polynomials anymore.

References

[1] Arnon, D. S.: Algorithms for the Geometry of Semi-Algebraic Sets., Ph.D. Thesis, Univ. of Wisconsin, Madison, 1981.
[2] Calmet, J.: A SAC-2 Implementation of Arithmetic and Root Finding over Large Finite Fields. Interner Bericht, Universität Karlsruhe 1981 (forthcoming).
[3] Calmet, J., Loos, R.: An Improvement of Robin's Probabilistic Algorithm for Generating Irreducible Polynomials over GF(p). Information Processing Letters **11**, 94–95 (1980).
[4] Cauchy, A.: Analyse Algébrique. Oeuvres complètes, II série, tome III, p. 398, formula (48). Paris: 1847.
[5] Collins, G. E.: The Calculation of Multivariate Polynomial Resultants. J. ACM **18**, 515–532 (1971).
[6] Collins, G. E., Horowitz, E.: The Minimum Root Separation of a Polynomial. Math. Comp. **28**, 589–597 (1974).
[7] Heindel, G. E.: Integer Arithmetic Algorithms for Polynomial Real Zero Determination. J. ACM **18**, 533–548 (1971).
[8] Householder, A. S.: Bigradients and the Problem of Routh and Hurwitz. SIAM Review **10**, 57–78 (1968).
[9] Householder, A. S., Stuart III, G. W.: Bigradients, Hankel Determinants, and the Padé Table. Constructive Aspects of the Fundamental Theorem of Algebra (Dejon, B., Henrici, P., eds.). London: 1963.
[10] Lenstra, A. K.: Factorisatie van Polynomen. Studieweek Getaltheorie en Computers, Mathematisch Centrum, Amsterdam, 95–134 (1980).
[11] Loos, R.: Computing Rational Zeros of Integral Polynomials by p-adic Expansion. Universität Karlsruhe (1981) (to appear in SIAM J. Comp.).
[12] Pinkert, J. R.: Algebraic Algorithms for Computing the Complex Zeros of Gaussian Polynomials. Ph.D. Thesis, Comp. Sci. Dept., Univ. of Wisconsin, May 1973.
[13] Rabin, M. O.: Probabilistic Algorithms in Finite Fields. SIAM J. Comp. **9**, 273–280 (1980).
[14] van der Waerden, B. L.: Algebra I. Berlin-Heidelberg-New York: Springer 1971.
[15] Zimmer, H.: Computational Problems, Methods, and Results in Algebraic Number Theory. Lecture Notes in Mathematics, Vol. 262. Berlin-Heidelberg-New York: Springer 1972.

Prof. Dr. R. Loos
Institut für Informatik I
Universität Karlsruhe
Zirkel 2
D-7500 Karlsruhe
Federal Republic of Germany