Homogeneous bivariate decompositions

JOACHIM VON ZUR GATHEN

Department of Computer Science, University of Toronto
Toronto, Ontario M5S 1A4, Canada
gathen@cs.toronto.edu

JÜRGEN WEISS

Institut für Informatik I, Universität Bonn and
GMD, Institut I1, Postfach 1316, 53731 St. Augustin, Germany
weiss@cartan.gmd.de

A *homogeneous bivariate decomposition* of a univariate polynomial f is of the form $f = g(h, k)$ with polynomials g, h, k, where g is bivariate and homogeneous. Such decompositions are of interest in robotics applications. This paper gives a Structure Theorem relating these decompositions to certain block decompositions of the roots of f, decomposition algorithms, and a classification of all constellations of degrees for which "almost all" polynomials f have such a decomposition.

1. Introduction

Let $\mathbb{F}$ be a field, $f \in \mathbb{F}[x]$ have degree n, and $r, s, t \in \mathbb{N}$. An (r, s, t)*-decomposition of* f is a triple (g, h, k) consisting of polynomials $h, k \in \mathbb{F}[x]$ of degrees s, t, respectively, and $g = \sum_{0 \leq i \leq r} g_i y^i z^{r-i} \in \mathbb{F}[y, z]$ such that

$$f = g(h, k) = \sum_{0 \leq i \leq r} g_i h^i k^{r-i}.$$

A general form of polynomial decomposition is

$$f_i = g_i(h_1, \ldots, h_m), \tag{1.1}$$

with $f_i, h_1, \ldots, h_m \in \mathbb{F}[x_1, \ldots, x_l]$ and $g_i \in \mathbb{F}[y_1, \ldots, y_m]$ for $1 \leq i \leq u$: an m-variate decomposition of several l-variate polynomials. Our decompositions are the special case of *homogeneous bivariate decompositions of univariate polynomials*, where $u = l = 1$, $m = 2$, and g is homogeneous.

Our main results are:

- a Structure Theorem characterizing decompositions in terms of certain block decompositions of the roots,
- a decomposition algorithm,
- a complete classification of those degrees where (almost) all f have a decomposition.

The paper is organized as follows. In Section 2, we describe the applications in robotics, suggested by Kovács and Hommel, which provided the original motivation for us. Section 3 defines an operation on decompositions that leads to the notion of a normal decomposition; each homogeneous bivariate decomposition is equivalent to a normal one. In Section 4, we give some rather simple examples of decompositions derived from factorizations; later we show that this is the best one can do in general. We also exhibit polynomials with exponentially many inequivalent decompositions.

Section 5 presents a case study for degree four; all homogeneous bivariate decompositions of quartic polynomials are found. In Section 6, we define block decompositions of the roots of f, generalizing a notion introduced by Kozen & Landau (1989). A Structure Theorem characterizes those block decompositions which correspond to homogeneous bivariate decompositions. An appropriate extension in Section 7 of Capelli's Theorem to our homogeneous bivariate decompositions is used in Section 8 to show that certain decompositions of f lead to factorizations of f; this is sort of a converse to the examples of Section 4.

From the Structure Theorem, we obtain in Section 9 a "decomposition algorithm" to find all homogeneous bivariate decompositions. It examines each block decomposition of the roots, and has exponential running time. Example 4.3 shows that any procedure based on exhaustive search will suffer from this problem. If the polynomial is irreducible of degree n, then there are only $O(n^{\log n})$ block decompositions, and the algorithm is correspondingly faster. In Section 10, we report on implementations of our algorithms; they are considerably slower than algorithms for ordinary decompositions of polynomials or rational functions.

In Section 4, we have exhibited normal $(2, s, t)$-decompositions for arbitrary polynomials f, and in Section 5, we have studied these in detail for degree four. In Section 11, a Classification Theorem states that the examples given cover all possibilities for generic (r, s, t)-decompositions; in particular, we always have $r = 2$.

Section 12 relates our decompositions to a special type of decomposition of a polynomial with rational functions.

Ordinary decompositions of the form $f = g(h)$, with $g, h \in \mathbb{F}[x]$, have several pleasant properties described by Ritt's theorems: (essential) uniqueness, and rationality: when a decomposition exists over an extension field, then one exists already over the ground field. Both properties go awry for our homogeneous decompositions: in Section 4, we exhibit polynomials with exponentially many inequivalent decompositions, and Section 5 shows that field extensions may be necessary.

For possible applications of homogeneous bivariate decompositions, this paper has a mixed message. On the one hand, we present some interesting decompositions, in particular in Section 5. On the other hand, we show that as a general tool, homogeneous bivariate decompositions are not available except via factorizations, as in Section 4; as a tool for specific polynomials, they seem to lead to very complicated equations.

2. Applications to robotics

In this section, we demonstrate the usefulness of bivariate homogeneous decompositions in robotics. Paul (1981) explains the fundamental notions of robotics, and much of this section is based on Kovács & Hommel (1992). In Kovács & Hommel (1990), they introduced bivariate homogeneous decompositions in a special case, and a question posed by Peter Kovács to the first author was a starting point for this investigation.

A non-degenerate manipulator with six joints—typical in industrial robots—is described by six 4×4 *arm matrices* A_i for $1 \leq i \leq 6$, and its *kinematic equations*

$$A_1 \cdot A_2 \cdots A_6 = T.$$

T is the effector matrix, containing position and orientation of the effector. 12 of these 16 equations are nontrivial. Each *arm matrix* A_i is fully determined by a tuple $(\theta_i, d_i, a_i, \alpha_i)$ of *Denavit-Hartenberg parameters.* For manipulators with only revolute joints the angles θ_i describe the joint displacements. The remaining Denavit-Hartenberg parameters are structural constants of the manipulator. The kinematics equations exhibit the functional dependence of the effector matrix on the angles θ_i explicitly. The *inverse kinematics*, that is the dependence of the angles θ_i on the effector matrix, is only implicitly given. In general, the equations are rather difficult to solve.

A computational solution of the inverse kinematics problem may proceed in two phases: a pre-processing phase, done in the design laboratory on large computers and using considerable resources (time etc.), and a production phase, where a small amount of computation must be carried out quickly (in a few microseconds) on the processors installed in the robot.

Ideally, one would like to precompute a closed-form (or, at least an easy-to-evaluate) solution for general (indeterminate) effector parameters, which then simple processors in the robot have to evaluate. Often these processors cannot do more, given the real-time constraint, than evaluate numerically a few square roots and some arithmetic operations; even solving an polynomial equation of degree four may be too costly.

A complementary approach proposed by Manocha (1992) and Ghazvini (1993) uses high performance RISC processors for the production phase. The kinematic equations are preprocessed to yield an eigenvalue problem. During the production phase the symbolic matrix elements are numerically evaluated and the eigenvalues and eigenvectors of the resulting matrix are calculated by standard numerical algorithms.

If we restrict ourselves to simple processors in the robots, the cost of solving the kinematic equations is a limiting factor in the design of robots. Finding new classes of equations that can be solved within the robotics constraints means—at least potentially—that new types of robots can be designed.

Typically, the kinematic equations are triangulated into the form

$$f_1(x_1) = f_2(x_1, x_2) = \cdots = f_n(x_1, \ldots, x_n) = 0,$$

where $f_1, \ldots, f_n$ are polynomials with coefficients which depend on the effector matrix. The parameters x_i are related in a simple way to the original Denavit-Hartenberg parameters θ_i; for example, $x_1 = \tan(\theta_3/2)$, etc.

In general, for manipulators with 6 revolute joints such a transformation is nontrivial to find and would lead to a univariate polynomial f_1 of degree 16, with each other f_i being linear in x_i (Lee & Liang 1988). So only robot designs are considered for which such a transformation is known. In these cases f_1 is typically of degree 2 or 4, and for $i \geq 2$, f_i gives a simple (linear or quadratic) dependence of x_i on the previous parameters $x_1, \ldots, x_{i-1}$. Obviously the existence of such a triangulation with polynomial degrees ≤ 4 places restrictive algebraic constraints on the Denavit-Hartenberg parameters.

Then the "remaining" task is to find manipulators with triangulations containing univariate polynomials f_1 such that their roots are easy to calculate, preferably just by square root extractions. One is particularly interested in the case of quartic polynomials, and then an ordinary *decomposition*

$$f = g(h), \tag{2.1}$$

with g, h quadratic, solves the problem. One only has to find the two roots α_1, α_2 of g, and then the two roots of $h - \alpha_i$, for each $i \in \{1, 2\}$.

We already noted that the coefficients of the polynomials depend on the effector matrix. A successful "modular" approach is to substitute random constants for all these parameters, solve the univariate problem, and "lift" this back to the multivariate problem. This tool for decompositions was introduced in von zur Gathen (1990a), Section 5, and has found its way into robotics under the name of "specialized analysis technique" (see Kovács & Hommel 1992).

Over the real or complex numbers only few polynomials f have an ordinary decomposition (2.1). More generally, when $\mathbb{F}$ is a field of characteristic not equal to two, then the set (in fact, algebraic variety) of decomposable monic quartic polynomials has dimension only three, but the space of all monic quartic polynomials has dimension four. Thus a random polynomial has a decomposition with negligible small probability, and this approach is, in general, not applicable.

However, in Section 5 we prove that the quartic monic polynomials $f \in \mathbb{F}[x]$ have, with few exceptions, a bivariate homogeneous decomposition

$$f = g(h, k), \tag{2.2}$$

with g and h quadratic and k linear. In the exceptional cases, such a decomposition also exists, but with smaller degrees. The coefficients of the components lie in a field extension $\mathbb{K}$ of $\mathbb{F}$ of degree at most three, and if $\mathbb{F} = \mathbb{R}$, then usually $\mathbb{K} = \mathbb{R}$. Just as (2.1), this allows us to find the four roots of f by factoring g as $g = (x - \alpha_1 y)(x - \alpha_2 y)$, and then finding the two roots of $h - \alpha_i k$, for each $i \in \{1, 2\}$. Thus this is a *universally applicable* tool for decomposing quartic polynomials.

For the inverse kinematics, this result allows us to replace the calculation of the roots for one fourth degree polynomial by the calculation of the roots for one polynomial of degree 3 and two polynomials of degree 2. The new polynomials depend on the effector matrix. Therefore they have to be treated during the production phase. So the problem is simplified but not reduced to square root extractions only.

This leads us to consider homogeneous bivariate decompositions $f = g(h, k)$ where all components are defined over the ground field $\mathbb{F}$. This is the more traditional Computer Algebra point of view. Using the "modular" approach discussed above a decomposition over the rationals may potentially be "lifted" back to the original problem which depends on the effector parameters. This "lift" will not introduce algebraic functions in these parameters. Though not a universal tool, it generalizes the ordinary decomposition and makes additional classes of manipulators accessible to algorithms relying on square root extractions only. Indeed, some subclasses have already been discussed in the robotics literature, for example in Smith & Lipkin (1990) and reciprocal polynomials in Mavroidis & Roth (1992).

Instead of constructing a triangulation with ordinary polynomials it may be convenient to use trigonometric polynomials $f(\cos\theta, \sin\theta)$. In this context the decomposition of trigonometric polynomials becomes relevant:

$$f(\cos\theta, \sin\theta) \equiv g(h(\cos\theta, \sin\theta)) \bmod (\cos^2\theta + \sin^2\theta - 1)$$

Though this decomposition may be reduced to the homogeneous bivariate decomposition by introducing $x = \tan(\theta/2)$, there exists a direct, polynomial time decomposition algorithm (Weiß 1994) based on the ideas of Kozen & Landau (1989) for the decomposition of univariate polynomials.

3. Normalization

We want to classify all homogeneous decompositions of a polynomial which are *equivalent* in a certain natural sense described below. Let $\mathbb{F}$ denote an arbitrary field. First we define the topic of this paper.

DEFINITION 3.1. *A* homogeneous bivariate decomposition of a univariate polynomial $f \in \mathbb{F}[x]$ *consists of a a homogeneous bivariate polynomial $g \in \mathbb{F}[y, z]$ and two univariate polynomials $h, k \in \mathbb{F}[x]$, such that*

$$f = g(h, k). \tag{3.1}$$

If $\deg g = r$, $\deg h = s$, *and* $\deg k = t$, *then we call (g, h, k) an (r, s, t)-decomposition of f.*

Next we examine the conditions under which the left factor g is uniquely determined by the right factors h and k. We use the following observation, which is probably well known.

LEMMA 3.2. *Let $h, k \in \mathbb{F}[x]$ be linearly independent over $\mathbb{F}$. Then for every $r \in \mathbb{N}$, the $r+1$ polynomials $h^r, h^{r-1}k, \ldots, k^r$ are linearly independent over $\mathbb{F}$. In particular, given $f \in \mathbb{F}[x]$ there is at most one $g \in \mathbb{F}[y, z]$ with $f = g(h, k)$.*

PROOF. Let $g_0, \ldots, g_r \in \mathbb{F}$ with

$$\sum_{0 \le i \le r} g_i h^i k^{r-i} = 0.$$

Let $q, h', k' \in \mathbb{F}[x]$ such that $q = \gcd(h, k)$, $h = qh'$, and $k = qk'$. Then $\gcd(h', k') = 1$ and

$$\sum_{0 \le i \le r} g_i h'^i k'^{r-i} = 0. \tag{3.2}$$

By the linear independence, we may assume that $\deg h' > 0$. Let $p \in \mathbb{F}[x]$ be an irreducible factor of h'. Then p divides all summands of (3.2) except $g_0 k'^r$, which implies $g_0 = 0$. The claim now follows inductively from

$$\sum_{0 \le i \le r-1} g_{i+1} h^i k^{r-1-i} \cdot h = 0. \quad \Box$$

Certain trivial decompositions always exist, as the following corollaries to Lemma 3.2 indicate.

COROLLARY 3.3. *Let $f \in \mathbb{F}[x]$ have degree n, let $h_0, k_0 \in \mathbb{F}$ be distinct, $h = x + h_0$, and either $k = 1$ or $k = x + k_0$. Then there exist unique $g_0, \ldots, g_n \in \mathbb{F}$ with $f = \sum_{0 \leq i \leq n} g_i h^i k^{n-i}$.*

PROOF. The polynomials h and k are linearly independent over $\mathbb{F}$. The $n+1$ polynomials $h^n, h^{n-1}k, \ldots, k^n$ form a basis of the $(n+1)$-dimensional vector space over $\mathbb{F}$ of polynomials in $\mathbb{F}[x]$ with degree at most n by Theorem 3.2. □

COROLLARY 3.4. *For any $n, s, t \in \mathbb{N}$, every polynomial in $\mathbb{F}[x]$ of degree n has*

(i) an $(n, 1, 1)$-decomposition and an $(n, 1, 0)$-decomposition, and

(ii) an $(1, s, t)$-decomposition if $s \geq n$.

The next two results allow us to relate different decompositions of a polynomial f.

LEMMA 3.5. *Let (g, h, k) be an (r, s, t)-decomposition of $f \in \mathbb{F}[x]$, $u \in \mathbb{F}$ nonzero, and $A \in GL_2(\mathbb{F})$.*

(i) The polynomial uf has the (r, s, t)-decomposition (ug, h, k).

(ii) $(g \circ A^{-1}, A(h, k))$ is a (r, s', t')-decomposition of f, with $\max\{s', t'\} = \max\{s, t\}$.

PROOF. (ii) Let $A = (A_{ij})_{1 \leq i,j \leq 2} \in GL_2(\mathbb{F})$, and $B = A^{-1} = (B_{ij})_{1 \leq i,j \leq 2} \in GL_2(\mathbb{F})$. Then

$$\begin{aligned}(g \circ B)(A(h, k)) &= g(B_{11}x + B_{12}y, B_{21}x + B_{22}y)(A_{11}h + A_{12}k, A_{21}h + A_{22}k) \\ &= g(h, k) = f.\end{aligned}$$

Denoting this operation of $GL_2(\mathbb{F})$ on the homogeneous bivariate decompositions by $A(g, h, k) = (g \circ A^{-1}, A(h, k))$, one checks that for any $A, B \in GL_2(\mathbb{F})$,

$$(B \cdot A)(g, h, k) = B(A(g, h, k)). \quad \square$$

We use this operation of $GL_2(\mathbb{F})$ to normalize homogeneous bivariate decompositions. By (i) we may restrict our considerations to monic polynomials f. From each orbit of the operation of $GL_2(\mathbb{F})$ on the set of decompositions of f we choose a unique representative in the following way: Firstly, we enforce $\deg h > \deg k$ by interchanging them if $\deg h < \deg k$, and if $\deg h = \deg k$, by subtracting an appropriate multiple of h from k. Secondly, we make h and k monic. Thirdly, if $k \neq 0$ and $t = \deg k$, we can enforce that the coefficient h_t of x^t in h is zero, by subtracting an appropriate multiple of k from h. All this can be expressed by applying certain matrices from $GL_2(\mathbb{F})$ to the decompositions.

As a consequence of Theorem 6.3 below, which links homogeneous bivariate decompositions with certain block decompositions, for most monic polynomials the operation of $GL_2(\mathbb{F})$ on the set of decompositions divides this set into only a finite number of orbits. The only exceptions are polynomials f which are the rth power of a univariate polynomial q. Let us have a closer look at this degenerate case.

Lemma 3.6. *Suppose that f has an (r, s, t)-decomposition $f = g(h, k)$ with $g = g_0(g_1 y + g_2 z)^r \in \mathbb{F}[y, z]$ and $g_0, g_1, g_2 \in \mathbb{F}$. Then $f = g_0 q^r$ with $q = g_1 h + g_2 k \in \mathbb{F}[x]$.*

Thus we define equivalence classes on the set of decompositions of a polynomial not only by the operation of $GL_2(\mathbb{F})$, but by Lemma 3.6 also. Furthermore, we rule out the trivial decompositions of Corollaries 3.3 and 3.4.

Definition 3.7. *A homogeneous (r, s, t)-decomposition (g, h, k) of f is* normal *if and only if $r, s \geq 2$, f and h are monic, and either k is monic with $s = \deg h > t = \deg k$ and the coefficient h_t of x^t in h is 0, or $k = 0$. In the case that g is the scalar multiple of the rth power of a linear homogeneous polynomial, g must be $g(y, z) = g_r y^r$ and $k = 0$. We call all decompositions* equivalent, *which reduce (as above) to the same normal form.*

4. Decompositions from factorizations

We exhibit two types of decompositions that correspond to factorizations in an easy way and yield examples of exponentially many inequivalent decompositions. They are polynomial analogues, tuned to our point of view, of the correspondence between factorizations of integers and representations by $x^2 - y^2$ in elementary number theory.

Example 4.1. *Suppose that $f \in \mathbb{F}[x]$ is monic of degree n and the product of the monic polynomials f_1 and f_2 with $s = \deg f_1 \geq \deg f_2 = n - s$.*

(i) If $s \neq n/2$, that is $\deg f_1 > \deg f_2$, then $f = f_1 f_2$ is an $(2, s, n-s)$-decomposition with $g_2 = g_0 = 0$.

(ii) If $s = n/2$, let $h = f_1$, $\tilde{k} = f_2 - f_1$ and g_1 the leading coefficient of $\tilde{k}$. Setting $k = g_1^{-1}\tilde{k}$ and $g_2 = 1$, we have

$$f = g_2 h^2 + g_1 hk.$$

This is an $(2, n/2, t)$-decomposition with $t \leq n/2 - 1$, and "in general" $t = n/2 - 1$.

In both cases the conversion to the equivalent normal form preserves the type of the decomposition.

We now demonstrate that all $(2, s, t)$-decompositions, not only those coming from factorizations, can be brought into the form

$$f = g_2 h^2 - g_0 k^2.$$

Lemma 4.2. *If char $\mathbb{F} \neq 2$ and $f \in \mathbb{F}[x]$ has a $(2, s, t)$-decomposition with $t \leq s$, then it has a $(2, s, t')$-decomposition with $g_1 = 0$ and $t' \leq s$.*

Proof. Given $f = g_2 h^2 + g_1 hk + g_0 k^2$, we first assume that $g_2 \neq 0$, and set

$$\tilde{h} = h + \frac{g_1 k}{2g_2} \in \mathbb{F}[x], \ \tilde{g}_0 = -g_0 + \frac{g_1^2}{4g_2}.$$

Then

$$f = g_2 \tilde{h}^2 - \tilde{g}_0 k^2.$$

If $g_2 = 0$, then $f = k \cdot (g_1 h + g_0 k)$. Set $\tilde{h} = g_1 h + (g_0 + 1)k$ and $\tilde{k} = g_1 h + (g_0 - 1)k$. Then

$$f = \frac{1}{4}\tilde{h}^2 - \frac{1}{4}\tilde{k}^2.$$

If $s = t = n/2$ we may have to swap $\tilde{h}$ and $\tilde{k}$ to satisfy $t' \leq s$. □

We now give an example of polynomials with exponentially many different normal decompositions, using Example 4.1(ii).

Theorem 4.3. *For every even $n \in \mathbb{N}$ with $n \geq 4$ there exists a field $\mathbb{F}$ and $f \in \mathbb{F}[x]$ of degree n with $\binom{n}{n/2}$ many different normal $(2, n/2, n/2 - 1)$-decompositions.*

Proof. Let $\mathbb{K}$ be a field, $A = \{\alpha_1, \ldots, \alpha_n\}$ a set of n indeterminates over $\mathbb{K}$, $\mathbb{F} = \mathbb{K}(A)$, and

$$f = \prod_{\alpha \in A} (x - \alpha).$$

For any $B \subseteq A$ with $\#B = n/2$, Example 4.1(ii) and Theorem 3.5 yield the following normal $(2, n/2, n/2 - 1)$-decomposition:

$$f = h^2 + g_1 hk + g_0 k^2,$$

where

$$\begin{aligned}
B' &= A \setminus B, \\
v &= \sigma_1(B) - \sigma_1(B') \in \mathbb{F}, \\
u &= \sigma_1(B) \in \mathbb{F}, \\
g_1 &= v - 2u \in \mathbb{F}, \\
g_0 &= -uv + u^2 \in \mathbb{F}, \\
k &= v^{-1} \cdot \left(\prod_{\alpha \in B'} (x - \alpha) - \prod_{\alpha \in B} (x - \alpha) \right) \in \mathbb{F}[x], \\
h &= \prod_{\alpha \in B} (x - \alpha) + uk \in \mathbb{F}[x].
\end{aligned}$$

Here $\sigma_1(B) = \sum_{\alpha \in B} \alpha$ is the first elementary symmetric function on B, and $\sigma_1(B')$ is defined correspondingly. In particular, $\sigma_1(B) \neq \sigma_1(B')$. The coefficient $c(B)$ of $x^{n/2-2}$ in h is

$$c(B) = \Big(\sigma_1(B)\sigma_2(B') - \sigma_1(B')\sigma_2(B)\Big)v^{-1},$$

where σ_2 denotes the second elementary symmetric function.

For every $C \subseteq A$ with $\#C = n/2$ and $C \notin \{B, A \setminus B\}$, we claim that

$$c(B) - c(C) \neq 0. \tag{4.1}$$

This claim implies a total of at least $\frac{1}{2}\binom{n}{n/2}$ different decompositions, and thus is a generalization of Example 4.1(ii) which gives "in general" three decompositions for $n = 4$.

To check (4.1), we choose

$$\alpha \in B \cap C',\ \beta \in B \cap C,\ \gamma \in B' \cap C,$$

and set all other indeterminates in A to zero. Then

$$c(B) - c(C) = \alpha\beta\gamma/(\beta + \gamma - \alpha)$$

is nonzero, which proves (4.1). $\square$

Theorem 4.3 also holds for $\mathbb{F} = \mathbb{Q}$ or $\mathbb{F}$ a sufficiently large finite field, since we only need the product p of all polynomials in (4.1) for $C \neq B, A \setminus B$ to be nonzero (after clearing all the denominators v). Since p is a nonzero polynomial of degree $d = 4 \cdot \frac{1}{2} \cdot \binom{n}{n/2}^2$, this will hold with probability at least $1/2$ when the elements of A are randomly chosen from a set with $2d$ elements.

5. Decompositions for degree four

In this section, we show that "almost all" polynomials of degree 4 have a normal $(2, 2, 1)$-decomposition, possibly over a field extension of degree at most 3, and that the polynomials without such a decomposition have a normal $(2, 2, 0)$-decomposition. In general the results of the last section would direct us only to such decomposition over extensions of degree 6.

Let $\mathbb{F}$ be a field, $f = f_0 + f_1x + f_2x^2 + f_3x^3 + x^4 \in \mathbb{F}[x]$. We want to determine for which $g_1, g_0, h_0, k_0 \in \mathbb{F}$ we have

$$f = h^2 + g_1hk + g_0k^2, \quad \text{with } h = h_0 + x^2 \text{ and } k = k_0 + x. \tag{5.1}$$

Let

$$
\begin{aligned}
\Delta_3 &= 4f_2f_3 - f_3^3 - 8f_1 \in \mathbb{F}, \\
\Delta &= (f_0f_3^2 - f_1^2) + (-8f_0f_3 + 4f_1f_2 - f_1f_3^2)z \\
&\quad + (2f_1f_3 + 16f_0 + f_2f_3^2 - 4f_2^2)z^2 + \Delta_3 z^3 \in \mathbb{F}[z], \\
\rho &= 256f_0 - 16f_2f_3^2 + 5f_3^4 \in \mathbb{F},
\end{aligned}
$$

and

$$
\begin{gathered}
\Delta(k_0) = 0,\ g_1 = f_3,\ g_0 = f_2 - 2h_0 - f_3k_0, \\
h_0 = (2f_2k_0 - 2f_3k_0^2 - f_1)/(-f_3 + 4k_0),
\end{gathered} \tag{5.2}
$$

$$
\begin{gathered}
\Delta_3 = 0,\ g_1, g_0 \text{ are as in (5.2), } \rho \text{ is a square in } \mathbb{F}, \\
\text{and } h_0 = -f_3^2 - \sqrt{\rho},
\end{gathered} \tag{5.3}
$$

$$
g_1, g_0 \text{ are as in (5.2), } f_1 = 0, \text{ and } f_0 + h_0^2 + f_2k_0^2 = 0. \tag{5.4}
$$

Theorem 5.1.

(i) *If $4k_0 \neq f_3$, then (5.1) is equivalent to (5.2).*

(ii) *If $4k_0 = f_3$ and char $\mathbb{F} \neq 2$, then (5.1) is equivalent to (5.3).*

(iii) *If $f_3 = 0$ and char $\mathbb{F} = 2$, then (5.1) is equivalent to (5.4).*

Proof. Comparing coefficients of powers of x, one finds that (5.1) is equivalent to

$$
\begin{gathered}
g_1 = f_3,\ g_0 = f_2 - 2h_0 - f_3k_0,\ (-f_3 + 4k_0)h_0 = 2f_2k_0 - 2f_3k_0^2 - f_1, \\
f_0 = h_0^2 + f_3h_0k_0 + (f_2 - 2h_0)k_0^2 - f_3k_0^3.
\end{gathered}
$$

For (i), we substitute for h_0 in the last equation and obtain that $\Delta(k_0) = 0$ is equivalent to the last equation. Similar manipulations lead to (ii) and (iii). $\square$

Corollary 5.2.

(i) *(5.1) has a solution if and only if there is a root k_0 of Δ in $\mathbb{F}$ with $4k_0 \neq f_3$ or ρ is a square in $\mathbb{F}$.*

(ii) *Let $\mathbb{F} = \mathbb{R}$. Then (5.1) has a solution if and only if $\Delta_3 \neq 0$ or $\rho \geq 0$.*

If char $\mathbb{F} \neq 2$, then Δ has the following Taylor expansion around $f_3/4$:

$$
\begin{aligned}
\Delta &= -\Delta_3^2/64 + (3f_3^2/16 - f_2/2)\Delta_3^2(z - f_3/4) \\
&\quad + (4f_3^2f_2 - 4f_1f_3 + 16f_0 - 3f_3^4/4 - 4f_2^2)(z - f_3/4)^2 \\
&\quad + \Delta_3(z - f_3/4)^3.
\end{aligned}
$$

In particular, $\Delta(f_3/4) = 0$ is equivalent to $\Delta_3 = 0$.

Over the real numbers, there are "in general" either one or three solutions.

Not unexpectedly, the exceptional polynomials with $\Delta_3 = 0$, for which over $\mathbb{R}$ Theorem 5.1 does not always give an $(2,2,1)$-decomposition, have an $(2,2,0)$-decomposition, i.e., an ordinary decomposition with degrees $2,2$. Let $f \in \mathbb{F}[x]$ be as above, $g_0, g_1, h_1 \in \mathbb{F}$, and consider the three conditions

$$f = h^2 + g_1hk + g_0k^2, \quad h = x^2 + h_1x, \quad k = 1, \tag{5.5}$$

$$g_0 = f_0, \quad g_1 = f_2 - h_1^2, \quad f_3 - 2h_1 = 0, \quad \Delta_3 = 0, \tag{5.6}$$

$$g_0 \text{ and } g_1 \text{ are as in (5.6)}, \quad f_3 = 0, \quad h_1^3 - f_2h_1 + f_1 = 0. \tag{5.7}$$

THEOREM 5.3.

(i) If char $\mathbb{F} \neq 2$, then (5.5) holds if and only if (5.6) holds.

(ii) If char $\mathbb{F} = 2$, then (5.5) holds if and only if (5.7) holds.

The proof is again a simple calculation.

For an arbitrary monic $f \in \mathbb{F}[x]$ of degree 4, we can find a quadratic monic factor $g \in \mathbb{K}[x]$ of f in the splitting field $\mathbb{K}$ of $\mathbb{F}$ with $[\mathbb{K} : \mathbb{F}] \leq 24$. Example 4.1(ii) says that f has a $(2,2,1)$-decomposition over $\mathbb{K}$. If the Galois group $G = \mathrm{Gal}(\mathbb{K}/\mathbb{F})$ is S_4, then one can choose two different roots $\alpha, \beta \in \mathbb{K}$ of f, and f has a quadratic factor over $\mathbb{L} = \mathbb{F}(\alpha+\beta, \alpha\beta)$. Since $\sigma \in G$ leaves $\mathbb{L}$ invariant if and only if either $\sigma(\alpha) = \alpha$, $\sigma(\beta) = \beta$ or $\sigma(\alpha) = \beta$, $\sigma(\beta) = \alpha$, we have $\#\mathrm{Gal}(\mathbb{K}/\mathbb{L}) = 4$ and $[\mathbb{L} : \mathbb{F}] = 6$.

One checks that also for $G \neq S_4$, there will be a quadratic factor of f in an extension of degree at most 6, and for some polynomials, say over $\mathbb{Q}$, 6 is the smallest such degree. Thus Example 4.1(ii) guarantees a $(2,2,1)$-decomposition over an extension of degree 6. In contrast, Theorems 5.1 and 5.3 provide a $(2,2,1)$-decomposition over an extension of degree at most 3. In particular, it shows that Ritt's Theorem on the rationality of ordinary decompositions does not hold for homogeneous bivariate decompositions.

Although Theorem 11.3 below classifies all possible values of n, r, s, t with generic homogeneous bivariate decompositions, the above example shows that we have not classified all such (individual) decompositions.

6. A Structure Theorem

Kozen & Landau (1989) exhibit a bijection between ordinary decompositions $f = g(h)$ of a polynomial $f \in \mathbb{F}[x]$, with $g, h \in \mathbb{F}[x]$ and a special type of block decompositions of the roots of f. We will examine the structure of block decompositions for polynomials with homogeneous decompositions. We will find that there is a similar bijection between inequivalent homogeneous bivariate decompositions of a polynomial and its block decompositions of a special type. This shows that there are only finitely many inequivalent such decompositions, and we will obtain an algorithm for the homogeneous bivariate decomposition of a polynomial. The running time is exponential in general, and quasi-polynomial if the polynomial is irreducible.

We denote by $A = \{\!\{a_1, \ldots, a_n\}\!\}$ the multiset with elements $a_1, \ldots, a_n$. We use the usual set-theoretic notations also for multisets; the only difference is that an element may occur several times in a multiset. (Formally, the set of multisets with n elements from a set F is the set F^n of sequences modulo the action of the symmetric group S_n.)

The following definition of a block decomposition is a slightly generalized version of the one given by Kozen and Landau.

Definition 6.1. *Let $f \in \mathbb{F}[x]$ be monic, $\mathbb{K} \supseteq \mathbb{F}$ the splitting field of f, and $\mathcal{G}$ the Galois group of $\mathbb{K}$ over $\mathbb{F}$. A* block decomposition *for f is a multiset Δ of multisets of elements of $\mathbb{K}$, such that*

(i) $f = \prod_{A\in\Delta} \prod_{\alpha\in A}(x - \alpha)$,

(ii) if $A \in \Delta$ and $\sigma \in \mathcal{G}$, then $B = \{\!\{\sigma(\gamma) | \gamma \in A\}\!\}$ is in Δ.

*If $r = \#\Delta$ and for all $A \in \Delta$, $s = \#A$, then we call Δ an $r \times s$-*block decomposition.

We dismiss Kozen and Landau's condition that for any $\alpha \in A \in \Delta$, $\beta \in B \in \Delta$ and $\sigma \in \mathcal{G}$ with $\sigma(\alpha) = \beta$, we have

$$B = \{\!\{\sigma(\gamma) | \gamma \in A\}\!\}. \tag{6.1}$$

The following example clarifies this. Let

$$f = (x - 1)^3(x - 2) \in \mathbb{Q}[x]. \tag{6.2}$$

Then $\Delta = \{\!\{\{\!\{1, 1\}\!\}, \{\!\{1, 2\}\!\}\}\!\}$ is a block decomposition for f by our definition. The identity in $\mathcal{G}$ maps $1 \in \{\!\{1, 1\}\!\}$ to $1 \in \{\!\{1, 2\}\!\}$, but does not map the block $\{\!\{1, 1\}\!\}$

to $\{\!\{1,2\}\!\}$. So Δ is not a block decomposition in the sense of Kozen and Landau. In fact f has no nontrivial ordinary decomposition, but we will exhibit a $(2,2,1)$-decomposition below.

Let (g,h,k) be a normal (r,s,t)-decomposition of the monic polynomial $f \in \mathbb{F}[x]$. Let $\rho = \deg g(y,1)$ and $\bar{g}(y,z) = z^\rho g(y/z,1)$. Thus $g(y,z) = z^r g(y/z,1) = z^{r-\rho}\bar{g}(y,z)$ has a factor $z^{r-\rho}$, and

$$f = k^{r-\rho}\bar{g}(h,k).$$

Let $\gamma_1,\ldots,\gamma_\rho$ be the roots of $g(y,1)$ in a suitably chosen field extension of $\mathbb{F}$. Then $\bar{g}(y,z) = \prod_{1\le i\le\rho}(y-\gamma_i z)$, and

$$f = k^{r-\rho}\prod_{1\le i\le\rho}(h-\gamma_i k).$$

Let

$$\begin{aligned} A_i &= \begin{cases} \{\!\{\text{roots of } h-\gamma_i k\}\!\} & \text{for } 1\le i\le\rho, \\ \{\!\{\text{roots of } k\}\!\} & \text{for } \rho < i \le r, \end{cases} \\ \Delta &= \{\!\{A_1,\ldots,A_r\}\!\}. \end{aligned}$$

We show that this partition Δ of the roots of f is actually a block decomposition for f. Let $\mathcal{G}$ be the Galois group of the splitting field $\mathbb{K}$ of f over $\mathbb{F}$, $\sigma \in \mathcal{G}$, and $A_i \in \Delta$. We will show that $\sigma(A_i) \in \Delta$.

If $\rho < i \le r$, then $\sigma(A_i) = \sigma(\{\!\{\text{roots of } k\}\!\}) = \{\!\{\text{roots of } \sigma(k)\}\!\} = \{\!\{\text{roots of } k\}\!\} = A_i \in \Delta$. If $1 \le i \le \rho$ and $\gamma_i \in \mathbb{F}$, then again $\sigma(A_i) = A_i$.

So now assume that $1 \le i \le \rho$ and $\gamma_i \in \mathbb{K}\setminus\mathbb{F}$. Since $A_i = \{\!\{\text{roots of } h-\gamma_i k\}\!\}$ and $\sigma(A_i) = \{\!\{\text{roots of } h-\sigma(\gamma_i)k\}\!\}$, it is sufficient to find a j with $1 \le j \le \rho$ such that $\sigma(\gamma_i) = \gamma_j$, since then $\sigma(A_i) = A_j \in \Delta$.

Let $q = \gcd(h,k) \in \mathbb{F}[x]$ and $h = qh'$ and $k = qk'$, so that h', $k' \in \mathbb{F}[x]$ are relatively prime, and

$$\begin{aligned} h-\gamma_i k &= q(h'-\gamma_i k'), \\ A_i &= \{\!\{\text{roots of } q\}\!\} \cup \{\!\{\text{roots of } h'-\gamma_i k'\}\!\}, \\ \sigma(A_i) &= \{\!\{\text{roots of } q\}\!\} \cup \{\!\{\text{roots of } h'-\sigma(\gamma_i)k'\}\!\}. \end{aligned}$$

Let α be a root of $h'-\gamma_i k'$. Then α is a root of $\prod_{1\le k\le\rho}(h'-\gamma_k k') \in \mathbb{F}[x]$. Thus $\beta = \sigma(\alpha)$ is a root of the same polynomial, and there is a j with $1 \le j \le \rho$ such that β is a root of $h'-\gamma_j k'$. In particular, $\beta \in A_j$, and

$$\begin{aligned} \gamma_j k'(\beta) &= h'(\beta) = \sigma(h'(\alpha)) \\ &= \sigma(\gamma_i k'(\alpha)) = \sigma(\gamma_i)k'(\beta). \end{aligned}$$

So $\gamma_j = \sigma(\gamma_i)$, provided that $k'(\beta) \neq 0$. But $k'(\beta) = 0$ implies $h'(\beta) = 0$, which is inconsistent with h' and k' being relatively prime.

In this way, we have obtained a block decomposition Δ from an (r,s,t)-decomposition of f. This decomposition Δ consists of r blocks A_i and satisfies the following additional properties.

Condition 6.2. *Let* $\rho = (\deg f - rt)/(s-t)$*. There are monic polynomials* $h, k \in \mathbb{F}[x]$ *with* $s = \deg h > t = \deg k$ *and* $h_t = 0$*, and* $\gamma_1, \ldots, \gamma_\rho \in \mathbb{K}$ *such that*

(i) for the $r-\rho$ equal *blocks* $A_{\rho+1}, \ldots, A_r$

- $\#A_i = t$,
- $k = \prod_{\alpha \in A_i}(x-\alpha)$,

(ii) for the ρ *blocks* $A_1, \ldots, A_\rho$

- $\#A_i = s$,
- $h - \gamma_i k = \prod_{\alpha \in A_i}(x-\alpha)$.

If $\deg f = rs$*, then* $\rho = r$ *and all blocks contain* s *roots.*

Given $\Delta = \{\!\{A_1, \ldots, A_r\}\!\}$, this condition may be checked constructively as follows. Suppose that $\#A_i = s$ for $1 \leq i \leq \rho$ and $\#A_i = t$ for $\rho < i \leq r$. Then if $r \neq \rho$, the blocks $A_{\rho+1}, \ldots, A_r$ must be equal and define k. For $1 \leq i < j \leq \rho$, write

$$\prod_{\alpha \in A_i}(x-\alpha) - \prod_{\beta \in A_j}(x-\beta) = \delta_{ij} k,$$

with $k \in \mathbb{F}[x]$ monic (and independent of i, j) and $\delta_{ij} \in \mathbb{K}$. If no such k and δ_{ij} exist, Δ does not satisfy Condition 6.2. If all differences are 0, that is, all blocks are equal, set $k = 0$. Because both products are monic, $t = \deg k < s$. Then h and the γ_i are determined by the Condition.

Next we will derive from a block decomposition of f satisfying Condition 6.2 a homogeneous decomposition of f. In the example (6.2), we gave a block decomposition for f whose corresponding polynomials are $b_1 = (x-1)(x-1)$ and $b_2 = (x-1)(x-2)$. Then $b_1 - b_2 = x - 1$ defines k, and $h = x^2 - 1$. The block polynomials are expressed as $b_1 = h - 2k$ and $b_2 = h - 3k$. Condition 6.2 is satisfied. And with $g(y,z) = (y-2z)(y-3z)$ we get the decomposition

$$f = g(h,k) = (x^2-1)^2 - 5(x^2-1)(x-1) + 6(x-1)^2.$$

Now let Δ be a block decomposition for $f \in \mathbb{F}[x]$ satisfying Condition 6.2, so that $f = \prod_{A \in \Delta} \prod_{\alpha \in A}(x-\alpha)$. Since

$$\prod_{\alpha \in A_i}(x-\alpha) = \begin{cases} h - \gamma_i k & \text{for } 1 \leq i \leq \rho, \\ k & \text{for } \rho < i \leq r, \end{cases}$$

we have $f = k^{r-\rho} \prod_{1 \leq i \leq \rho}(h - \gamma_i k)$. Let $g(y,z) = z^{r-\rho} \prod_{1 \leq i \leq \rho}(y - \gamma_i z) \in \mathbb{F}[y,z]$. Then $f = g(h,k)$ is a normal decomposition. The coefficients of g are in $\mathbb{F}$, because by Theorem 3.2 they are solutions to linear equations in $\mathbb{F}$.

From these two constructions we obtain the following *Structure Theorem.*

THEOREM 6.3. *Let $f \in \mathbb{F}[x]$ be a monic polynomial. The above gives a bijection between the inequivalent homogeneous decompositions of f and the block decompositions for f satisfying Condition 6.2.*

The Structure Theorem for ordinary decompositions of polynomials in Kozen & Landau (1989) is an immediate consequence of this theorem by requiring the "right factor" k to be the constant 1, so that $t = 0$. In particular, Definition 6.1 and Condition 6.2 with $t = 0$ imply (6.1).

7. A generalization of Capelli's Theorem

Homogeneous decompositions of irreducible polynomials and homogeneous decompositions of reducible polynomials with irreducible "left factors" g have additional structure.

Let $f \in \mathbb{F}[x]$ be monic, irreducible, and separable. The Galois group $\mathcal{G}$ of f acts transitively on the roots of f. All roots have multiplicity one. Let Δ be a block decomposition for f, and suppose that there is one block $A \in \Delta$ with $b = \prod_{\alpha \in A}(x - \alpha) = h - \gamma k$, $h, k \in \mathbb{F}[x]$, $\deg k = t$ and $h_t = 0$. Then $b \notin \mathbb{F}[x]$, because it is a proper factor of f and f is irreducible over $\mathbb{F}$. Hence $\gamma \notin \mathbb{F}$, and h and k are uniquely determined.

Let $B \in \Delta$, $\beta \in B$, and $\alpha \in A$. There exists $\sigma \in \mathcal{G}$ with $\sigma(\beta) = \alpha$. Since the roots have multiplicity one, we find $\sigma(B) = A$, and $\mathcal{G}$ acts transitively on the blocks. Now

$$\begin{aligned} \prod_{a \in A}(x - a) &= \prod_{b \in B}(x - \sigma(a)) \\ = \sigma(h(x) - \gamma k(x)) &= h(x) - \sigma(\gamma) k(x). \end{aligned}$$

We conclude that if $\prod_{\alpha \in A}(x - \alpha) = h - \gamma k$ for any block A, then this is true for every block B with appropriate γ, the γ's being conjugates under $\mathcal{G}$.

LEMMA 7.1. *Let $f \in \mathbb{F}[x]$ be monic, irreducible, and separable with $\deg f = n$. Let Δ be a block decomposition for f such that there are $A \in \Delta$ and $h, k \in \mathbb{F}[x]$ with $s = \deg h > t = \deg k$ and*

$$\prod_{\alpha \in A}(x - \alpha) = h - \gamma k.$$

Then f has an $(n/s, s, t)$-decomposition.

Let f be monic and irreducible, and α a root of f. There is a bijection between the intermediate fields between $\mathbb{F}$ and $\mathbb{F}(\alpha)$ and the block decompositions for f (Wieland 1964, van der Waerden 1960). So an (r, s, t)-decomposition $f = g(h, k)$

corresponds to an intermediate field $\mathbb{K}$ with $\mathbb{F} \subset \mathbb{K} \subset \mathbb{F}(\alpha)$. $\mathbb{K}$ is generated by a root γ of $g(y, 1)$, that is $\mathbb{K} = \mathbb{F}(\gamma)$ and $[\mathbb{K} : \mathbb{F}] = r$. Furthermore α is a root of $h - \gamma k$, so that $[\mathbb{F}(\alpha) : \mathbb{K}] = s$. The coefficients of the minimal polynomial $h - \gamma k$ of α over $\mathbb{K}$ lie in the $\mathbb{F}$-vector space of dimension 2 generated by $\{1, \gamma\}$. As can easily be seen, this condition is not only necessary but also sufficient.

Part of these results for irreducible f may be extended to the reducible case. Let f have a homogeneous decomposition $f = g(h, k)$ with irreducible left factor g. Let γ be a root of $g(y, 1)$, and α a root of f. The field $\mathbb{K} = \mathbb{F}(\gamma)$ is an intermediate field between $\mathbb{F}$ and all the fields $\mathbb{F}(\alpha)$. Because f is reducible, the fields $\mathbb{F}(\alpha)$ are not necessarily isomorphic. To get some information about the factors of f, we will extend Capelli's theorem (Schinzel 1982, p. 89) to homogeneous decompositions.

Following Schinzel, for $f \in \mathbb{F}[x]$

$$f =_{\mathbb{F}}^{can} \text{const} \prod_{1 \le i \le l} f_i^{e_i}$$

means that $f_1, \ldots, f_l \in \mathbb{F}[x]$ are irreducible over $\mathbb{F}$ and pairwise relatively prime, and that $e_1, \ldots, e_l \in \mathbb{N}$ are positive. As usual, $N_{\mathbb{K}/\mathbb{F}}$ denotes the norm of $\mathbb{K}$ over $\mathbb{F}$. Capelli's Theorem is the special case $k = 1$ of the following result.

THEOREM 7.2. *Let $g \in \mathbb{F}[y, z]$ be homogeneous and irreducible over $\mathbb{F}$, $g(y, 1)$ separable, γ a root of $g(y, 1) \in \mathbb{F}[y]$ in an extension field of $\mathbb{F}$, and $h, k \in \mathbb{F}[x]$ relatively prime. If*

$$h - \gamma k =_{\mathbb{F}(\gamma)}^{can} \mathit{const} \prod_{i=1}^{l} \Phi_i^{n_i},$$

then

$$g(h, k) =_{\mathbb{F}}^{can} \mathit{const} \prod_{i=1}^{l} N_{\mathbb{F}(\gamma)/\mathbb{F}} \Phi_i^{n_i}.$$

PROOF. We extend Schinzel's proof to the homogeneous case. Denote by $\gamma^{(\nu)}$ be the distinct conjugates of γ over $\mathbb{F}$. Thus for $\nu \neq \mu$, $\gcd(x - \gamma^{(\nu)}, x - \gamma^{(\mu)}) = 1$ and by Schinzel (1982), p. 10, $h - \gamma^{(\nu)} k$ is relative prime to $h - \gamma^{(\mu)} k$. Hence any factor of $h - \gamma^{(\nu)} k$ is relative prime to any factor of $h - \gamma^{(\mu)} k$, provided $\nu \neq \mu$. In particular, the conjugate $\Phi_i^{(\nu)}$ of Φ_i under $\mathrm{Gal}(\mathbb{K}/\mathbb{F})$ which is a factor of $h - \gamma^{(\nu)} k$ is relatively prime to $\Phi_j^{(\mu)}$, where $\mathbb{K} = \mathbb{F}(\gamma)$.

Let $p_i \in \mathbb{F}[x]$ be irreducible over $\mathbb{F}$ and such that $\Phi_i \mid p_i$ in $\mathbb{K}[x]$. The conjugates $\Phi_i^{(\nu)}$ of Φ_i under $\mathrm{Gal}(\mathbb{K}/\mathbb{F})$ divide p_i. But, as shown, Φ_i is relative prime to $\Phi_i^{(\nu)}$. So $N_{\mathbb{K}/\mathbb{F}} \Phi_i = \prod_\nu \Phi_i^{(\nu)} \mid p_i$ in $\mathbb{F}[x]$ and $N_{\mathbb{K}/\mathbb{F}} \Phi_i = p_i$ is irreducible.

Now let us show that the norms are coprime as well. For $i \neq j$ and for all ν, μ, $\gcd(\Phi_i^{(\nu)}, \Phi_j^{(\mu)}) = 1$, so that $\gcd(N_{\mathbb{K}/\mathbb{F}} \Phi_i, N_{\mathbb{K}/\mathbb{F}} \Phi_j) = 1$. □

So given a homogeneous decomposition $f = g(h, k)$ with $g \in \mathbb{F}[y, z]$ irreducible and γ a root of $g(y, 1)$, the factors of f that are irreducible over $\mathbb{F}$ are the norms of the irreducible factors of $h - \gamma k$ over $\mathbb{K}$. This shows that each irreducible factor of f has a block decomposition which corresponds to the intermediate field $\mathbb{K}$.

These block decompositions may be trivial (one block containing all roots or each block containing just one root). Condition 6.2 on the blocks in the Structure Theorem 6.3 may not be satisfied for the blocks of each factor separately, but only for corresponding blocks of the factors of f combined.

EXAMPLE 7.3. *Let $f \in \mathbb{Q}[x]$ be given by $f =_{\mathbb{Q}}^{can} f_1 f_2$ with*

$$\begin{aligned} f_1 &= 8x^9 + 12x^8 + 22x^7 + 29x^6 - 6x^5 + 94x^4 - 25x^3 + 45x^2 + 54, \\ f_2 &= x^6 - 2x^5 + x^4 - x^3 + 3x^2 + x + 1, \end{aligned}$$

Both factors of f do not have a nontrivial homogeneous bivariate decomposition, but f has the following one:

$$\begin{aligned} f &= g(h, k), \\ g &= 8y^3 - 4y^2 z + 14yz^2 + 25z^3, \\ h &= x^5 - \frac{x^3}{2} - \frac{x^2}{4} - \frac{x}{4} - \frac{3}{4}, \\ k &= x^4 - x^3 + \frac{3x^2}{2} + \frac{x}{2} + \frac{3}{2}. \end{aligned}$$

Let γ be a root of the irreducible polynomial $g(y, 1) \in \mathbb{Q}[x]$. Between $\mathbb{Q}$ and each field $\mathbb{Q}[x]/(f_i)$, for $i = 1, 2$, there is the intermediate field $\mathbb{K} = \mathbb{Q}(\gamma)$ which can be deduced from the factorizations of f_1 and f_2 over $\mathbb{Q}(\gamma)$:

$$\begin{aligned} f_1 &= 8\left(x^3 + \frac{x^2}{2} + \left(-\frac{\gamma^2}{2} + \frac{3\gamma}{2} - \frac{1}{8}\right)x + \frac{3\gamma^2}{4} + \frac{21}{16}\right) \cdot \\ &\quad \left(x^6 + x^5 + \left(\frac{\gamma^2}{2} - \frac{3\gamma}{2} + \frac{19}{8}\right)x^4 + \left(-\frac{\gamma^2}{2} - \frac{3\gamma}{4} + \frac{5}{4}\right)x^3\right. \\ &\quad \left. + \left(\frac{15\gamma^2}{8} - \frac{7\gamma}{4} + \frac{61}{32}\right)x^2 + \left(\frac{3\gamma}{4} + \frac{33}{8}\right)x - \frac{9\gamma}{4} + \frac{9}{8}\right), \\ f_2 &= \left(x^2 + \left(-\gamma - \frac{1}{2}\right)x + \frac{\gamma^2}{2} - \frac{1}{8}\right) \cdot \\ &\quad \left(x^4 + \left(\gamma - \frac{3}{2}\right)x^3 + \left(\frac{\gamma^2}{2} - \gamma + \frac{3}{8}\right)x^2 - x + \frac{\gamma^2}{2} - \gamma + \frac{11}{8}\right). \end{aligned}$$

The blocks corresponding to the intermediate field $\mathbb{K}$ do not satisfy Condition 6.2, as can be read off these factorizations. Multiplying the first factors of both f_1 and f_2 together, we get

$$b = x^5 - \gamma x^4 + \left(\gamma - \frac{1}{2}\right)x^3 + \left(-\frac{3\gamma}{2} - \frac{1}{4}\right)x^2 + \left(-\frac{\gamma}{2} - \frac{1}{4}\right)x - \frac{3\gamma}{2} - \frac{3}{4}$$

which defines a block of a block decomposition for f and satisfies Condition 6.2.

8. Factors and decompositions

Collecting our results from the last two sections, we state the following two theorems, which describe all possible origins of factors of a decomposable polynomial.

THEOREM 8.1. *Let (g, h, k) be a normal (r, s, t)-decomposition of $f \in \mathbb{F}[x]$ with $\deg_y g = r$.*

(i) Suppose that g is irreducible, $g(y, 1)$ separable, and $\gcd(h, k) = 1$. Let γ be a root of $g(y, 1) \in \mathbb{F}[y]$ and $h - \gamma k =^{can}_{\mathbb{F}(\gamma)} const \prod_{1 \le i \le l} \Phi_i^{n_i}$. Then

$$f =^{can}_{\mathbb{F}} const \prod_{1 \le i \le l} N_{\mathbb{F}(\gamma)/\mathbb{F}} \Phi_i^{n_i}.$$

That is, for every irreducible factor of f there exists a block decomposition (maybe a trivial one) and the corresponding intermediate field is the splitting field of $g(y, 1)$.

(ii) Suppose that g is irreducible, and that $\gcd(h, k) = q \neq 1$. Write $h = qh'$ and $k = qk'$. Then

$$f = q^r g(h', k').$$

The structure of $g(h', k')$ is determined by case (i). The polynomial q may be reducible over $\mathbb{F}$.

(iii) Suppose that g is reducible, say

$$g(y, z) =^{can}_{\mathbb{F}} const \prod_{1 \le i \le l} g_i(y, z)^{n_i}.$$

Then

$$f = const \prod_{1 \le i \le l} g_i(h, k)^{n_i}.$$

The structure of the $g_i(h, k)$ is determined by case (ii).

THEOREM 8.2. *Let (g, h, k) be an (r, s, t)-decomposition of $f \in \mathbb{F}[x]$, and let $\rho = \deg_y g$. Then*

$$f = k^{r-\rho} \bar{g}(h, k),$$

where $\bar{g}(y, z) = z^\rho g(y/z, 1)$ is ρ-homogeneous. The structure of $\bar{g}(h, k)$ is determined by Theorem 8.1.

Unfortunately in constructing decompositions for a decomposable polynomial these two theorems fail to give the origin of a factor *a priori*.

9. Decomposition Algorithm

The overall structure of this homogeneous decomposition algorithm resembles the algorithms for decomposition of polynomials (Kozen & Landau 1989) or rational functions (Zippel 1991).

(i) Calculation of the "right" candidates h and k,

(ii) determination of the "left factor" g,

(iii) verification that $f = g(h, k)$.

First we consider the last two steps. By Lemma 3.2 the left factor g is uniquely determined by linear algebra, if it exists. We adapt an algorithm from Dickerson (1989) for the calculation of left factors. It solves the following problem:

Given $f \in \mathbb{F}[x]$ with $\deg f = n$ and $h, k \in \mathbb{F}[x]$ with $s = \deg h > t = \deg k$. Does a homogeneous polynomial $g \in \mathbb{F}[y, z]$ of degree r exist such that $f = g(h, k)$?

Let $g(y, z) = \sum_{i=0}^{r} g_i y^i z^{r-i}$. For $1 \le i \le r$, the product $h^i k^{r-i}$ has degree $tr + (s-t)i \le rs$. In order for a decomposition to exist we must have $\deg f \le rs$. Furthermore the coefficients of $x^{rs-(s-t)(j+1)+1}$ up to x^{rs} in f only depend on $\sum_{m-j \le i \le r} g_i h^i k^{r-i}$. Hence we are able to determine the coefficients g_i of g by iteration. Besides we may simultaneously verify the decomposition.

$\pi_i = h^i k^{r-i}$ **for** $i = 0, \ldots, r$
$p = f$
for $i = 0$ **to** r **do**
  if $\deg p > rs - (s-t)i$ **then** fail **end**
  $g_{r-i} = \dfrac{\mathrm{coeff}(p, rs - (s-t)i)}{\mathrm{coeff}(\pi_{r-i}, rs - (s-t)i)}$
  $p = p - g_{r-i}\pi_{r-i}$
end
if $p <> 0$ **then** fail **end**

where $\mathrm{coeff}(p, j)$ is the coefficient of x^j in p. The number of arithmetic operations in $\mathbb{F}$ is polynomially bounded in the degree of f. There are asymptotically faster methods for this problem (von zur Gathen 1990a), but the real bottleneck is the first step.

We will apply the Structure Theorem 6.3 to determine the candidates h and k for the decomposition in step 1. First we restrict ourselves to the decomposition of *irreducible* polynomials. Because of the Structure Theorem a homogeneous decomposition of an irreducible polynomial f exists if and only if there is an $r \times s$-block decomposition for f and an arbitrary block of roots satisfies the hypothesis of Lemma 7.1.

Therefore we compute all block decompositions for f. For each block decomposition we check the conditions of Lemma 7.1. Equivalently we may determine all intermediate fields between $\mathbb{F}$ and $\mathbb{F}(\alpha)$, α being a root of f.

The algorithm **BLOCKS** from Landau & Miller (1985), Kozen & Landau (1989), and Yokoyama *et al.* (1990) provides all block decompositions of f. The algorithm gives a *single* block decomposition in polynomial time. The number of block decompositions for f is bounded by $n^{\log n}$, where $n = \deg f$. Alternatively, the algorithm **EQNFIELD** from Lazard & Valibouze (1993) determines all intermediate fields by symmetric resolvents.

Both algorithms expect an irreducible polynomial $f \in \mathbb{F}[x]$ as input. Let α be a root of f. The result of the algorithms is a finite set of polynomials $B \subset \mathbb{F}(\alpha)[x]$. In the block decomposition picture each polynomial $b \in B$ defines a block decomposition for f. The roots of b form a block of this particular block decomposition, namely the block which contains the root α. Hence the coefficients of b are, up to sign, the elementary symmetric functions of the roots in this block. In the intermediate field picture each polynomial b is the minimal polynomial of α over the intermediate field. In order that there be a homogeneous decomposition, these coefficients of b have to satisfy the condition of Lemma 7.1. This may be checked as follows:

```
s = deg b
for i = s by -1 to 0 do
  if coeff(b, i) ∉ F then
    γ = coeff(b, i)
    break
  end
end
for i = 0 to s do
  h_i = coeff(b, i) mod γ
  k_i = coeff(b, i) div γ
  if h_i ∉ F or k_i ∉ F then
    NO DECOMPOSITION
  end
end
```

We assume that the polynomial $b \in \mathbb{F}(\alpha)[x]$ is given by $b \in \mathbb{F}[y][x]$ with the coefficients from $\mathbb{F}[y]$ being reduced by the polynomial $f(y)$. Because f is assumed to be irreducible, b must have at least one coefficient not in $\mathbb{F}$. The division with remainder which determines h_i and k_i is to be calculated in the polynomial ring $\mathbb{F}[y]$. The calculation of the "left factor" g from f, h and k cannot fail, since we have satisfied sufficient conditions for the existence of a homogeneous decomposition.

Now let us have a look at *reducible* polynomials. Assume that $f =_{\mathbb{F}}^{can} \prod_i f_i$ has a decomposition. Each factor f_i of f may be

- a factor of k (Theorem 8.2),
- a factor of $\gcd(h, k)$ (Theorem 8.1 (ii)),
- the norm of a factor Ψ of $h - \gamma k$ over $\mathbb{F}(\gamma)$ (Theorem 8.1(i)). Let α be a root of f_i. There is an intermediate field (maybe trivial) between $\mathbb{F}$ and $\mathbb{F}(\alpha)$ or, equivalently, f_i has a block decomposition. This block, maybe only when combined with corresponding blocks of other factors of f belonging to the same intermediate field, satisfies the condition of the Structure Theorem.

As already noted, we do not know sufficient *a priori* criteria which help to classify a factor f_i. So we have to consider all possibilities, which will result in an exponential algorithm.

There seems to be some kind of trade-off between "few factors = much structure = few homogeneous decompositions, easy to compute" and "many factors = little structure = many homogeneous decompositions, hard to compute".

10. Implementation of the Algorithm

We discuss our experience gained from implementing the decomposition algorithm for an *irreducible* polynomial f. The implementation of the algorithm for the determination of the "left factor" g is straightforward. We compare different approaches for the determination of all block decompositions or equivalently all intermediate fields which give the "right factors" h and k.

Let α be a root of f. As described in the last section, a block decomposition for f is determined by a polynomial $b \in \mathbb{F}(\alpha)[x]$ whose roots just form this block. This polynomial b is actually a factor of f over $\mathbb{F}(\alpha)$.

The main steps of the algorithm **BLOCKS** (Landau & Miller 1985, Yokoyama *et al.* 1990) are

(i) the complete factorization of f over $\mathbb{F}(\alpha)$,

(ii) the determination which products of those irreducible factors describe a block. This is achieved by the calculation of gcds between those factors over algebraic extensions of $\mathbb{F}(\alpha)$.

In practice, using Trager's algorithm (Trager 1976), the factorization of f over $\mathbb{F}(\alpha)$ is technically feasible approximately up to polynomial degree 10. Trager's algorithm factors the norm $N_{\mathbb{F}(\alpha)/\mathbb{F}}(f) \in \mathbb{F}[x]$, a polynomial of degree $(\deg f)^2$.

Empirically the cost of the algorithm is dominated by the second step, the calculation of gcds over algebraic extensions of $\mathbb{F}(\alpha)$. It is beneficial to avoid this step altogether. This can be achieved in the following way. We consider all possible products of irreducible factors of f over $\mathbb{F}(\alpha)$. Every polynomial b describing a block of a block decomposition is among those products. For each product we check the condition of Lemma 7.1. If this condition is satisfied, we try to calculate the "left factor" g from the polynomials h and k defined by this product. If we succeed, the product defines a block of a block decomposition for f and the polynomials g, h and k form a homogeneous decomposition of f.

So we may exchange the costly gcd calculations for an exponential number of linear algebra problems. But given the limits by the polynomial factorization over $\mathbb{F}(\alpha)$ the number of products poses no problem. This variant of the algorithm has been implemented in Maple and AXIOM. The factorization time dominates at least by an order of magnitude the whole computation. On a typical workstation, Maple calculates the decomposition of a polynomial of degree 12 in about 30 minutes.

As a second approach we may determine all intermediate fields between $\mathbb{F}$ and $\mathbb{F}(\alpha)$. By the algorithm **EQNFIELD** (Lazard & Valibouze 1993) all intermediate fields $\mathbb{M}$ of index k in $\mathbb{F}(\alpha)$ may be found by calculating k-symmetric resolvents of f. In the simplest case, a factor of degree $(\deg f)/k$ of a k-symmetric resolvent is the minimal polynomial for a generator of the intermediate field $\mathbb{M}$. The k-symmetric resolvents are polynomials over $\mathbb{F}$ of degree $\binom{\deg f}{k}$. So in general their degree is higher then $(\deg f)^2$, the degree of the norm in Trager's algorithm. We are interested in factors of low degree only, but without a special factorization algorithm, which generates factors of low degree fast, this approach is more costly. This was verified by translating algorithms for handling symmetric polynomials and calculating resolvents originally implemented in MACSYMA (Valibouze 1989) to AXIOM.

If the aim is to calculate intermediate fields and block decompositions for a polynomial, and not only homogeneous decompositions, the second step of the algorithm **BLOCKS**, that is the gcd calculations over extensions of $\mathbb{F}(\alpha)$, may no longer be avoided. This additional cost may well balance the drawbacks of the algorithm **EQNFIELD**.

The costs of these homogeneous decomposition algorithms are rather high. In contrast our implementation of the algorithm described in Kozen & Landau (1989) indicates that an ordinary decomposition of a polynomial of degree 100 can be calculated in less than a minute. Gutiérrez & Recio (1992) calculated decompositions of rational functions of degree 20 in about 10 minutes. Theorem 12.2 seems to indicate that computing homogeneous decompositions is intrinsically harder than computing decompositions of rational functions, because fewer constraints are given.

11. The classification of generic decompositions

In this section, we consider *generic decompositions* which give nontrivial decompositions for almost all polynomials, and show that none exist besides those in Section 4.

LEMMA 11.1. *Let (g, h, k) be a normal (r, s, t)-decomposition of $f \in \mathbb{F}[x]$, and $n = \deg f$.*

(i) If $rs > n$, then $g_r = 0$.

(ii) If $rs > n + s - t$, then $k \neq 0$, $g_r = g_{r-1} = 0$, and either $k \in \mathbb{F}$ or f is not squarefree.

PROOF. Writing $g = \sum_{0 \le i \le r} g_i y^i z^{r-i} \in \mathbb{F}[y, z]$, we have $f = \sum_{0 \le i \le r} g_i h^i k^{r-i}$, and

$$d_i = \deg(h^i k^{r-i}) = \begin{cases} is + (r-i)t & \text{if } k \neq 0, \\ is & \text{otherwise.} \end{cases}$$

Thus $d_r > d_{r-1} > \cdots > d_0$. In particular, $d_r = rs$, which proves (i). In (ii), we have $k \neq 0$, since otherwise $g = g_r y^r$ and $\deg g(h, k) = rs$. Since $d_r > d_{r-1} > n$, we have $g_r = g_{r-1} = 0$, and thus k^2 divides f. □

We denote by $P_n \subseteq \mathbb{F}[x]$ the $(n+1)$-dimensional vector space of all polynomials of degree at most n, and have $P_n \subseteq P_m$ for $n \le m$. We formalize the notion of "almost all" polynomials having a decomposition as follows.

DEFINITION 11.2. *Let $\mathbb{F}$ be an infinite field. For $n, r, s, t \in \mathbb{N}$, we say that P_n has* a generic (r, s, t)-decomposition *if there is no nonzero polynomial $\tau \in \mathbb{F}[F_0, \ldots, F_n]$ such that for all $f_0 + \cdots + f_n x^n \in P_n$ we have*

$$f \text{ has a normal } (r,s,t)\text{-decomposition} \quad \Longrightarrow \quad \tau(f_0, \ldots, f_n) = 0.$$

To explain this definition, we assume $s \ge t \ge 0$ and $rs \ge n$, and make the natural identification of a homogeneous bivariate polynomial of degree r with its coefficient vector in $\mathbb{F}^{r+1}$, and of a monic polynomial of degree k with its coefficient vector in $\mathbb{F}^k$, leaving out the leading one. We consider the normal composition mapping ψ:

$$\begin{array}{rcl} \psi : \mathbb{F}^{r+1} \times \mathbb{F}^{s-1} \times \mathbb{F}^t & \to & P_{rs} \supseteq P_n \\ (g, h, k) & \mapsto & g(h, k) = \sum_i g_i h^i k^{r-i}. \end{array} \tag{11.1}$$

Here g, h, k are coefficient vectors of polynomials as in Definition 3.7. The image of ψ, intersected with P_n, is the set D of polynomials in P_n that have an (r, s, t)-decomposition. The definition says that P_n has a generic (r, s, t)-decomposition if and only if the image is dense in the Zariski topology.

If this is the case and $\mathbb{F}$ is algebraically closed, then "almost all" polynomials in P_n do indeed have an (r,s,t)-decomposition. This follows since $\mathrm{im}\psi \cap P_n$ is always constructible and dense by assumption, hence there exists a nonzero polynomial $\sigma \in \mathbb{F}[F_0,\ldots,F_n]$ such that

$$\sigma(f_0,\ldots,f_n) \neq 0 \quad \Longrightarrow \quad \sum_{0\leq i\leq n} f_i x^i \in D.$$

When $\mathbb{F}$ is not algebraically closed, the situation may be more complicated. The case of finite fields is discussed at the end of this Section. Over $\mathbb{Q}$ and $\mathbb{R}$, the general situation is well illustrated in Corollary 5.2. The polynomial $\Delta \in \mathbb{F}[x]$ usually has a root k_0 over $\mathbb{R}$, and in general (if $k_0 \neq f_3/4$) this leads to a decomposition. But over $\mathbb{Q}$, Δ will usually be irreducible, and a root will only exist in a (small) algebraic extension of $\mathbb{Q}$. In fact, it is a consequence of the classification in Theorem 11.3 and the examples in Section 4 that whenever a generic decomposition of some format exists and a specific polynomial is given, it is sufficient to factor the polynomial in order to find a decomposition of this format.

Whenever there is no generic (r,s,t)-decomposition, then most polynomials actually do not have a decomposition, in the sense that the coefficients have to satisfy a nontrivial algebraic relation, which we might call a "separating" test polynomial, in order for an (r,s,t)-decomposition to exist.

THEOREM 11.3. *Let $\mathbb{F}$ be an infinite field, and $n,r,s,t \in \mathbb{N}$ with $n,r,s \geq 2$ and $rs \geq n > s > t$. Then P_n has a generic (r,s,t)-decomposition if and only if one of the following holds:*

I. $r=2$, $s>n/2$, $t=n-s$, or

II. $r=2$, $s=n/2$, $t=n/2-1$.

PROOF. If $\mathbb{F}$ is algebraically closed, then Examples 4.1 (i) and (ii) exhibit a homogeneous bivariate decomposition of type I and II, respectively, for every polynomial. For an arbitrary infinite $\mathbb{F}$, this implies that generic homogeneous bivariate decompositions of type I and II exist, since otherwise there would be a nonzero test polynomial τ as in Definition 11.2 whose zero set $\{\tau=0\}$ contains $\mathrm{im}\gamma$ over $\mathbb{F}$, and this would imply that the same holds over an algebraic closure of $\mathbb{F}$.

Assume now that P_n has a generic (r,s,t)-decomposition. First suppose that $rs > n+s-t$. Then, by Lemma 11.1 (ii), each decomposition leads to a polynomial f that either has an ordinary decomposition $f=g(h,k)$ with $k \in \mathbb{F}$, or that is contained in the zero set of the discriminant

$$\tau = \mathrm{resultant}_x(F,F') \in \mathbb{F}[F_0,\ldots,F_n],$$

where $F=\sum_i F_i x^i$. The polynomials with an ordinary decomposition form a proper algebraic subset; allowing all $g(y,1)\in\mathbb{F}[y]$ of degree $\rho \leq r$, the dimension is at most $\rho+1+s-1 < n+1 = \rho s+1$.

For a similar reason, we may assume that $t \geq 0$. According to Lemma 11.1 (i), we have two possibilities:

(a) $rs \le n$, or

(b) $n < rs \le n + s - t$, and $g_r = 0$ in any decomposition $f = g(h, k) \in P_n$.

In case (a), we have $n = rs$, and from the dimensions in (11.1) we obtain that $r + 2s - 1 \ge r + 1 + s - 1 + t \ge n + 1$. If $r = 2$, we find $s = n/2$ and $t = n/2 - 1$, corresponding to II. If $r = 3$, then $s = n/3$ and $2 + 2n/3 \ge n + 1$, hence $n \le 3$ and $s \le 1$. If $r \ge 4$, then $s \le n/4$ and $r + n/2 - 1 \ge n + 1$, which implies $r = n$ since r is a divisor of n, and hence $s = 1$, which is ruled out.

In case (b), ψ as in (11.1) is actually a mapping from $\mathbb{F}^r \times \mathbb{F}^{s-1} \times \mathbb{F}^t$, so that $r + s - 1 + t \ge n + 1$. If $r = 2$, the decomposition has the form

$$f = g_1 hk + g_0 k^2 = (g_1 h + g_0 k)k.$$

If f has degree n, then $\deg(g_1 h + g_0 k) = n - t$. Thus $t < n/2$ implies that $s = n - t$, which is I, and $t \ge n/2$ implies that $s \le t$, which is ruled out. If $r \ge 3$, we find that

$$s(r-2) = rs - 2s \le n - t - s \le r - 2,$$

and hence $s \le 1$. This shows that I and II are the only possibilities. $\square$

We next want to bound the degree of the separating test polynomial τ implicit in Theorem 11.3 when I and II are not satisfied. We start with the special case of ordinary decompositions $f = g(h)$. We may assume that f, g, h are monic, and that $h(0) = 0$.

Lemma 11.4. *Let $r, s \in \mathbb{N}$ with $r, s \ge 2$, $n = rs$, and suppose that char $\mathbb{F}$ does not divide r. Then there exists a nonzero polynomial $\tau \in \mathbb{F}[F_{n-1}, \ldots, F_0]$ of degree at most $s+1$ such that for all polynomials $f = x^n + f_{n-1}x^{n-1} + \cdots + f_0$, and $g, h \in \mathbb{F}[x]$ with g, h monic, $r = \deg g$, $s = \deg h$, $h(0) = 0$, and*

$$f = g(h), \tag{11.2}$$

we have

$$\tau(f_{n-1}, \ldots, f_0) = 0.$$

Proof. The decomposition algorithms in Kozen & Landau (1989) and von zur Gathen (1990a) are based on the fact that $f = g(h)$ implies that $h_{r-1}, \ldots, h_0$ are determined by a triangular system of equations, linear in each "new" variable. One checks that this gives polynomials

$$\eta_{r-1}, \ldots, \eta_1 \in \mathbb{F}[F_{n-1}, \ldots, F_0]$$

such that $h_i = \eta(f_{n-1}, \ldots, f_0)$ and $\deg \eta_i = s - i$. The coefficient at x^{n-s} in (11.2) determines g_{r-1}, and the coefficient at x^{n-s-1} gives a nonzero polynomial, in which F_{n-s-1} occurs linearly, which we can take for τ. Its degree is at most $s + 1$. $\square$

The "wild" case, where char $\mathbb{F}$ divides r, is computationally more difficult (von zur Gathen 1990b), and we have not determined a test polynomial τ for it.

The remaining cases of non-ordinary homogeneous bivariate decompositions are dealt with in the following result.

COROLLARY 11.5. *Let $\mathbb{F}$ be an infinite field, $n, r, s, t \in \mathbb{N}$ with $n, r, s \geq 2$ and $n > s > t \geq 1$ and suppose that P_n does not have a generic (r, s, t)-decomposition. Then there exists a nonzero test polynomial*

$$\tau \in \mathbb{F}[F_0, \ldots, F_n]$$

such that $\tau(f_0, \ldots, f_n) = 0$ for any $f = \sum_{0 \leq i \leq n} f_i x^i \in \mathbb{F}[x]$ that has a normal (r, s, t)-decomposition, and

(i) $\deg_{tot} \tau = 1$ *if* $t \geq 0$, $r < n/s + 1$ *and* $r \neq n/s$,

(ii) $\deg_{tot} \tau = 2n - 1$ *if* $t = -\infty$ *or* $r \geq n/s + 1$,

(iii) $\deg_{tot} \tau = (r+1)^{n+1}$ *if* $r = n/s$.

PROOF. If $r < n/s$, then $\tau = F_n$ is sufficient. If $n/s < r < n/s + 1$, then $g_r = 0$ and $\deg(g(h, k)) < n$ for all (g, h, k), so that we can again take $\tau = F_n$. If $r \geq (n+s-t+1)/s$, then all $f \in \mathrm{im}\psi \cap P_n$ are not squarefree by Lemma 11.1(ii), and we can take the discriminant for τ. Together these include all cases where $r > n/s$, so that now we may assume $r = n/s$.

The homogeneous bivariate decomposition map $\psi : \mathbb{F}^{r+s+t} \to \mathbb{F}^{n+1}$, as in (11.1), is given by polynomials in $g_0, \ldots, g_r, h_0, \ldots, h_{s-1}, k_0, \ldots, k_{t-1}$ of total degree at most $r + 1$, and $m = \dim \overline{\mathrm{im}\psi} \leq n$. Pick some $f \in \mathrm{im}\psi$ such that $d = \dim \psi^{-1}(\{f\}) \geq 0$ is minimal, and an affine linear space $L \subseteq \mathbb{F}^{r+s+t}$ of dimension $r+s+t-d$ such that $L \cap \psi^{-1}(\{f\})$ is finite; such an L always exists. By the theorem on the dimension of fibres (see Shafarevich (1974), e.g.), we have $d + m = r + s + t$. Let $\psi' : L \to \mathbb{F}^{n+1}$ be the restriction of ψ to L. Since ψ' has a finite fibre, namely over f, we have $\dim \overline{\mathrm{im}\psi'} = \dim L = \dim \overline{\mathrm{im}\psi}$, so that $\mathrm{im}\psi'$ is dense in $\mathrm{im}\psi$. Now Lemma 3.3 in von zur Gathen (1985) provides a polynomial τ as desired, containing $\mathrm{im}\psi'$ and hence $\mathrm{im}\psi$ in its zero set, and with $\deg \tau \leq (r+1)^{n+1}$; one can also use the proof of Lemma 1 in Heintz & Sieveking (1980). □

The easy cases (i) and (ii) of Corollary 11.5 cover all situations with $n \leq 7$ and $t \geq 1$ except for $n = 6$, $t = 1$ and (r, s) either $(2, 3)$ or $(3, 2)$. The symbolic resultant τ in the first case is a polynomial of degree 10 with 27 terms in 6 variables, and in the second case its computation strains a computer algebra system like MAPLE. For larger values of n, it seems difficult to determine an explicit "separating" test polynomial τ.

Over a finite field $\mathbb{F}_q$ with q elements, $\tau = F_n^q - F_n$ is satisfied by all polynomials, and so Definition 11.2 would trivially imply that there are no generic decompositions over a finite field. However, Corollary 11.5 states that if P_n does not have a generic decomposition, then we may replace "nonzero τ" by "nonzero τ whose degree is at most $(r+1)^{n+1}$". This notion is nontrivial also over $\mathbb{F}_q$ with q large enough—say, $q > 2(r+1)^{n+1}$—and when I and II of Theorem 11.3 are not satisfied, then with high probability a randomly chosen polynomial does not have a decomposition.

12. Relations to decompositions of rational functions

In this section, we note two relations between homogeneous decompositions of univariate polynomials and decompositions of rational functions.

Zippel (1991) reduces the decomposition of rational functions to the simultaneous homogeneous bivariate decomposition of two homogeneous bivariate polynomials. But the decomposition

$$f(x,y) = g(h(x,y), k(x,y))$$

with $h, k \in \mathbb{F}[x,y]$ homogeneous of the same degree is related to the homogeneous bivariate decomposition of univariate polynomials by substituting 1 for y. Vice versa by appropriate homogenization the latter decomposition leads to the former. At least in theory this results in an algorithm for rational function decomposition.

We find a second relation when we try to decompose univariate polynomials as rational functions. Given a polynomial $f \in \mathbb{F}[x]$, the naive approach of decomposing the rational function $f/1 \in \mathbb{F}(x)$ does not lead to any decompositions which could not be obtained by the ordinary decomposition of f (Schinzel 1982, p. 10). So we will try to decompose $f/q \in \mathbb{F}(x)$ with arbitrary $q \in \mathbb{F}[x]$.

Definition 12.1. *A fractional decomposition of a univariate polynomial $f \in \mathbb{F}[x]$ consists of a univariate polynomial $q \in \mathbb{F}[x]$ and two rational functions $G, H \in \mathbb{F}(x)$ such that*

$$\frac{f}{q} = G \circ H = G(H).$$

The relationship between homogeneous decompositions and fractional decompositions of univariate polynomials is explained in the following theorem.

Theorem 12.2. *Let $f \in \mathbb{F}[x]$. To every fractional decomposition of f with q relatively prime to f corresponds a homogeneous decomposition of f. To every homogeneous decomposition of f corresponds a fractional decomposition of f.*

Proof. Let

$$\frac{f}{q} = G \circ H$$

be a fractional decomposition of f, with $q \in \mathbb{F}[x]$ and $G, H \in \mathbb{F}(x)$. Write $G = g_1/g_2$ and $H = h_1/h_2$ with $g_1, g_2, h_1, h_2 \in \mathbb{F}[x]$ and $\gcd(g_1, g_2) = \gcd(h_1, h_2) = 1$. Let $r = \max\{\deg g_1, \deg g_2\}$, $\hat{g}_i(x, y) = y^r g_i(x/y)$ and $\bar{g}_i(x, y) = y^{\deg g_i} g_i(x/y)$ in $\mathbb{F}[x, y]$ for $i = 1, 2$, so that

$$\frac{f}{q} = \frac{g_1\,(h_1/h_2)}{g_2\,(h_1/h_2)} = \frac{\hat{g}_1(h_1, h_2)}{\hat{g}_2(h_1, h_2)}. \tag{12.1}$$

By Schinzel (1982), p. 10, $\bar{g}_1(h_1, h_2)$, $\bar{g}_2(h_1, h_2)$ and h_2 are pairwise relatively prime, because h_1 and h_2 resp. g_1 and g_2 are relatively prime. Hence $\hat{g}_1(h_1, h_2)$ and $\hat{g}_2(h_1, h_2)$ are relatively prime. Since f and q are relatively prime, both denominators and numerators in (12.1) are equal separately, and f has the homogeneous decomposition

$$f = \hat{g}_1(h_1, h_2).$$

For the other direction, let

$$f = g(h, k)$$

be a homogeneous decomposition of f with $g \in \mathbb{F}[x, y]$ homogeneous of degree r. With $\hat{g}(x) = g(x, 1)$, we get

$$\frac{f}{k^r} = \hat{g}\left(\frac{h}{k}\right). \quad \square$$

Open questions

1. *The structure of homogeneous bivariate decompositions.* Given f, we related the set of all homogeneous bivariate decompositions of f to the set of certain block decompositions for f. Is it possible to describe the structure without recourse to block decompositions? If a decomposition exists over an algebraic closure, what is the smallest field extension necessary? Section 5 discusses an instance of this question.

2. *A fast decomposition algorithm.* In practice the proposed algorithm is quite slow when compared to the algorithms for the ordinary decomposition of polynomials and the decomposition of rational functions. Is there an algorithm which avoids the costly factorizations over algebraic extensions? A new approach is required to handle the case of reducible polynomials more efficiently.

3. *Other generic decompositions.* For other instances of the general problem (1.1), classify those values of the parameters for which generic decompositions exists. One might want to start with multivariate decompositions of univariate polynomials (i.e., $l = 1$), and for decomposing multivariate polynomials, one would begin with our homogeneous bivariate decompositions, say for bivariate polynomials (i.e., $l = 2$).

Acknowledgments

The authors are indebted to Peter Kovács for posing the problem, motivated by the applications in robotics. Many thanks go to Daniel Lazard for pointing out the usefulness of resolvents, and to Jaime Gutiérrez for carefully reading a preliminary version of the manuscript and correcting various errors.

The first author was supported by Natural Sciences and Engineering Research Council of Canada, grant A2514, and gratefully acknowledges the hospitality of the Institute for Scientific Computation at ETH Zürich during a sabbatical visit. The second author was supported by Deutsche Forschungsgemeinschaft, grant Kr 393/4-1 and Kr 393/4-2.

Authors' current addresses:

JOACHIM VON ZUR GATHEN, Fachbereich Mathematik-Informatik, Universität-GH Paderborn, 33098 Paderborn, Germany, `gathen@uni-paderborn.de`.

JÜRGEN WEISS, Zentrum für Datenverarbeitung, Universität Mainz, 55099 Mainz, Germany, `weiss@uni-mainz.de`.

References

MATTHEW T. DICKERSON, *The Functional Decomposition of Polynomials.* PhD thesis, Cornell University, Ithaca, NY, 1989.

JOACHIM VON ZUR GATHEN, Irreducibility of multivariate polynomials. *Journal of Computer and System Sciences* **31** (1985), 225–264.

JOACHIM VON ZUR GATHEN, Functional decomposition of polynomials: the tame case. *Journal of Symbolic Computation* **9** (1990a), 281–299.

JOACHIM VON ZUR GATHEN, Functional decomposition of polynomials: the wild case. *Journal of Symbolic Computation* **10** (1990b), 437–452.

MASOUD GHAZVINI, Reducing the inverse kinematics of manipulators to the solution of a generalized eigenproblem. In *Computational Kinematics*, ed. JORGE ANGELES, GÜNTHER HOMMEL, AND PETER KOVÁCS, vol. 28 of *Solid Mechanics and its Application*, Dordrecht, Boston, London, 1993, Kluwer Academic Publishers, 15–26.

JAIME GUTIÉRREZ AND TOMÁS RECIO, A practical implementation of two rational function decomposition algorithms. In *Proceedings of the International Symposium on Symbolic and Algebraic Computation, ISSAC'92*, ed. PAUL S. WANG, New York, NY, 1992, ACM Press, 152–157.

JOOS HEINTZ AND MALTE SIEVEKING, Lower bounds for polynomials with algebraic coefficients. *Theoretical Computer Science* **11** (1980), 321–330.

PETER KOVÁCS AND GÜNTER HOMMEL, Factorization and decomposition in kinematic equation systems. Forschungsbericht 1990–23, Technische Universität Berlin, Fachbereich 20, 1990.

PETER KOVÁCS AND GÜNTHER HOMMEL, An expert system for the optimal symbolic solution of the inverse kinematics problem. In *Proc. Int. Conf. Autom., Robotics and Computer Vision, Singapore*, 1992.

DEXTER KOZEN AND SUSAN LANDAU, Polynomial decomposition algorithms. *Journal of Symbolic Computation* **7** (1989), 445–456.

SUSAN LANDAU AND GARY LEE MILLER, Solvability by radicals is in polynomial time. *Journal of Computer and System Sciences* **30** (1985), 179–208.

DANIEL LAZARD AND ANNICK VALIBOUZE, Computing subfields: Reverse of the primitive element problem. In *Computational Algebraic Geometry*, ed. FRÉDÉRIC EYSSETTE AND ANDRÉ GALLIGO, vol. 109 of *Progress in Mathematics*, Boston-Basel-Berlin, 1993, Birkhäuser, 163–176.

HONG-YOU LEE AND CHONG-GAO LIANG, Displacement analysis of the general spatial 7-link 7R mechanisms. *Mechanism and Machine Theory* **23** (1988), 219–226.

DINESH MANOCHA, *Algebraic and Numeric Techniques in Modeling and Robotics.* PhD thesis, University of California, Berkeley, 1992.

CONSTANTINOS MAVROIDIS AND BERNHARD ROTH, Structural parameters which reduce the number of manipulator configurations. In *Robotics and Spatial Mechanisms, and Mechanical Systems. Proc. ASME 22nd Biennial Mechanisms Conference*, ed. G. KINZEL ET AL., 1992, 359–366. DE-vol. 45.

RICHARD P. PAUL, *Robot Manipulators: Mathematics, Programming, and Control.* MIT Press, 1981.

ANDRZEJ SCHINZEL, *Selected Topics on Polynomials.* University of Michigan Press, Ann Arbor, MI, 1982.

IGOR R. SHAFAREVICH, *Basic algebraic geometry.* Grundlehren Band 213. Springer-Verlag, 1974.

DAVID R. SMITH AND HARVEY LIPKIN, Analysis of fourth order manipulator kinematics using conic sections. In *Proc. IEEE International Conference on Robotics and Automation*, 1990, 274–278.

BARRY M. TRAGER, Algebraic factoring and rational function integration. In *Proceedings of the 1976 ACM Symposium on Symbolic and Algebraic Computation*, ed. R. D. JENKS, New York, 1976, ACM Press, 219–226.

ANNICK VALIBOUZE, Symbolic computations with symmetric polynomials, an extension to macsyma. In *Computers and Mathematics '89*, ed. ERICH KALTOFEN AND STEPHEN M. WATT, Berlin, Heidelberg, New York, 1989, Springer Verlag, 308–320.

BARTEL L. VAN DER WAERDEN, *Algebra*, vol. I. Springer Verlag, Berlin, Göttingen, Heidelberg, fifth edition, 1960.

JÜRGEN WEISS, *Constructive Algebraic Methods for the Inverse Kinematics Problem.* PhD thesis, Universität Bonn, Germany, 1994.

HELMUT WIELAND, *Finite Permutation Groups.* Academic Press, New York, London, 1964.

KAZUHIRO YOKOYAMA, MASAYUKI NORO, AND TAKU TAKESHIMA, On determining the solvability of polynomials. In *Proceedings of the International Symposium on Symbolic and Algebraic Computation 1990*, ed. SHUNRO WATANABE AND MORIO NAGATA, New York, NY, 1990, ACM Press, 127–134.

RICHARD ZIPPEL, Rational function decomposition. In *Proceedings of the 1991 International Symposium on Symbolic and Algebraic Computation*, ed. S.M. WATT, New York, NY, 1991, ACM Press, 1–6.