ACTA ARITHMETICA
XC.4 (1999)

Quadratic factors of $f(x) - g(y)$

by

Yuri F. Bilu (Basel and Graz)

1. Introduction. In this note we consider the following problem:

Problem 1.1. When does a polynomial of the form $f(x) - g(y)$ have a quadratic factor?

Let K be a field and $f(x), g(x) \in K[x]$. It is trivial that $f(x) - g(y)$ has a linear factor if and only if $f(x) = g(ax + b)$, where $a \in K^*$ and $b \in K$.

The problem when $f(x) - g(y)$ has a quadratic factor is considerably more complicated. If $f = \phi \circ f_1$ and $g = \phi \circ g_1$, where $\phi(x), f_1(x), g_1(x) \in K[x]$ and $\max(\deg f_1, \deg g_1) = 2$ then, trivially, $f(x) - g(y)$ has the quadratic factor $f_1(x) - g_1(y)$. However, there also is a famous series of non-trivial examples, provided by the Chebyshev polynomials: $T_n(x) + T_n(y)$ splits (over an algebraically closed field) into quadratic factors (and one linear factor if n is odd; see Proposition 3.1). Recall that the Chebyshev polynomials are defined from $T_n(\cos x) = \cos nx$, or, alternatively, from $T_n((z + z^{-1})/2) = (z^n + z^{-n})/2$.

In this note we completely solve Problem 1.1 for polynomials over a field of characteristic 0. We start from the case of algebraically close base field, which is technically simpler.

Theorem 1.2. *Let $f(x)$ and $g(x)$ be polynomials over an algebraically closed field K of characteristic 0. Then the following assertions are equivalent*:

(a) *The polynomial $f(x) - g(y)$ has a factor of degree at most 2.*

(b) *$f = \phi \circ f_1$ and $g = \phi \circ g_1$, where $\phi(x), f_1(x), g_1(x) \in K[x]$ and either $\deg f_1, \deg g_1 \le 2$, or $f_1(x) = T_{2^k}(\alpha x + \beta)$ and $g_1(x) = -T_{2^k}(\gamma x + \delta)$, where $k \ge 2$, $\alpha, \gamma \in K^*$ and $\beta, \delta \in K$.*

1991 *Mathematics Subject Classification*: Primary 12E05; Secondary 12E10.
Supported by the Lise Meitner Fellowship (Austria), grant M00421-MAT.

For many applications assuming the base field algebraically closed is too restrictive. Therefore it is desirable to drop this assumption. Also, sometimes it is important to know not only that $f = \phi \circ f_1$ and $g = \phi \circ g_1$ with some very special f_1 and g_1, but also that the quadratic factor divides $f_1(x) - g_1(y)$. All this is achieved in the following refinement of Theorem 1.2, where Dickson polynomials $D_n(x, a)$ (see Section 3) replace Chebyshev polynomials.

THEOREM 1.3. *Let $f(x)$ and $g(x)$ be polynomials over a field K of characteristic 0, and let $q(x, y) \in K[x, y]$ be an irreducible (over K) quadratic factor of $f(x) - g(y)$. Then there exist polynomials $\phi(x), f_1(x), g_1(x) \in K[x]$ such that $f = \phi \circ f_1$, $g = \phi \circ g_1$ and one of the following two options takes place*:

(a) *We have* $\max(\deg f_1, \deg g_1) = 2$ *and* $q(x, y) = f_1(x) - g_1(y)$.

(b) *There exists an integer $n > 2$ with $2\cos(2\pi/n) \in K$ such that for some $\alpha \in K^*$ and $a, \beta, \gamma \in K$ we have*

$$f_1(x) = D_n(x + \beta, a), \qquad g_1(x) = -D_n((\alpha x + \gamma) \cdot 2\cos(\pi/n), a),$$

and $q(x, y)$ is a quadratic factor of $f_1(x) - g_1(y)$. Moreover, $e^{2\pi i/n} \notin K$ when $a = 0$.

Notice that in the second option $g_1(x) \in K[x]$ by (6), and $f_1(x) - g_1(y)$ splits over K into irreducible quadratic factors (and one linear factor if n is odd) by Corollary 3.2. If $a \neq 0$ then the quadratic factors are absolutely irreducible.

Problem 1.1 is motivated by Diophantine applications. By the classical theorem of Siegel [15, 10] the finiteness problem for the Diophantine equation $f(x) = g(y)$ reduces to the question of whether or not the corresponding plain curve has a component of genus 0 and with at most 2 points at infinity. Fried [7, Corollary of Theorem 3] showed that the latter question reduces to two independent problems, one of which is Problem 1.1, and the other is a special version of Ritt's second theorem [12, 14]. In the joint paper [1] with Robert F. Tichy we obtain, using Theorem 1.3, a very explicit finiteness criterion for the equation $f(x) = g(y)$.

To the best of my knowledge, the first one to consider Problem 1.1 was Tverberg. In Chapter 2 of his thesis [18] he proved the following: if $\deg f = \deg g$ and $f(x) - g(y)$ has a quadratic factor, then either $f = \phi \circ f_1$ and $g = \phi \circ g_1$, where $\deg \phi > 1$, or $f = \ell \circ T_4 \circ f_1$ and $g = \ell \circ (-T_4) \circ g_1$, where ℓ is a linear polynomial. Tverberg also obtained a similar result about cubic factors.

Our method is quite different from Tverberg's (though some similarities do exist) and relies on the study of the monodromy group of the polynomials f and g. This approach is inspired by Fried [5, 6] and Turnwald [16]. (See especially Remark 1 in [7, p. 48].)

The general problem of factorization of $f(x) - g(y)$ has a long history, which cannot be presented here. We just mention that among the contributors were Cassels, Davenport, Feit, Fried, Lewis, Schinzel, Tverberg, and many others. Fried [6, Theorem 1 on pp. 141–142] proved that if f is an *indecomposable* ([1]) polynomial of degree n and K contains no complex subfield of $\mathbb{Q}(e^{2\pi i/n})$ (in particular, if $K = \mathbb{Q}$), then $f(x) - g(y)$ is reducible (over $\overline{\mathbb{Q}}$) only in trivial cases. He also showed that the problem with indecomposable f and general K reduces to a certain problem in group theory, studied by Feit [3]. For further advances see [4, 8, 9]. Quite recently, Cassou-Noguès and Couveignes [2], essentially using the previous work of Fried and Feit, and assuming the classification of finite simple groups, completely classified the pairs of polynomials f, g with indecomposable f such that $f(x) - g(y)$ is reducible ([2]). The intersection of this result with ours is that, when f is indecomposable, the difference $f(x) - g(y)$ can have a quadratic factor only in trivial cases. This follows also from Tverberg's result.

It is possible (though not obvious) that Theorems 1.2 and 1.3 extend, with suitable modifications, to characteristic $p > 2$. I did not consider the positive characteristics since I could not imagine any applications of such a result.

In Sections 2 and 3 we collect necessary material about two very classical objects: dihedral groups and Dickson polynomials. The results of these two sections are certainly known, but I could not find them in the standard literature.

Some of the results of Section 3, in particular Theorem 3.8, are inspired by Turnwald [16]. (Most of his paper was incorporated in Chapters 2 and 6 of [11].)

The proof of Theorems 1.2 and 1.3 occupies Section 4.

Acknowledgements. This note is a part of a joint project with Robert F. Tichy, supported by the Austrian Science Foundation (FWF). I am pleased to thank Robert Tichy for his kind permission to publish this note separately.

I also thank him, as well as Norbert A'Campo, Roberto Maria Avanzi, Dani Berend, Frits Beukers, Michael Fried, Andrzej Schinzel, Gerhard Turnwald, Helge Tverberg, Umberto Zannier and the referee for stimulating discussions and helpful suggestions.

I am especially grateful to Andrzej Schinzel, who carefully read the paper and detected various inaccuracies in the text.

([1]) A polynomial is *indecomposable* if it is not a composition of two polynomials of smaller degree.

([2]) Cassou-Noguès and Couveignes assume that both f and g are indecomposable. But [6, assertion (2.38) on p. 142] implies that the assumption about g can be dropped.

Finally, I thank Guillaume Hanrot for performing some helpful computations.

Conventions. All fields in this paper are of characteristic 0 (although some of the results are valid in arbitrary characteristic). The capital letter K always stands for a field. We assume that all fields that occur in the paper are contained in one big algebraically closed (unnamed) field. In particular, any field K has a well-defined algebraic closure $\overline{K}$, any two fields K and K' have well-defined intersection $K \cap K'$ and composite KK', etc.

Throughout the paper

- Z_n stands for the cyclic group of order n,
- $\mathcal{D}_n$ stands for the nth dihedral group, and
- $\mathcal{S}_n$ stands for the nth symmetric group.

We use (a, b) for the greatest common divisor of a and b. When it can be confused with (a, b) as an ordered pair, we write $\gcd(a, b)$.

The groups are written multiplicatively, and the neutral element of a group is denoted by 1 or id (the latter is used mainly when the group is realized as a permutation group).

2. Dihedral groups. Recall that the *dihedral group* $\mathcal{D}_n$ is the group generated by two symbols α, β with the relations $\alpha^2 = (\alpha\beta)^2 = \beta^n = 1$. The group $\mathcal{D}_n$ consists of $2n$ elements, and has a cyclic subgroup of index 2, generated by β. All elements outside this subgroup are of order 2.

The identity $\alpha^{-1}\beta\alpha = \beta^{-1}$ implies the following.

PROPOSITION 2.1. *The conjugacy class of any $\gamma \in \langle\beta\rangle$ is $\{\gamma, \gamma^{-1}\}$.* ∎

Notice that $\mathcal{D}_n$ is generated by two elements of order 2 (which are α and $\alpha\beta$). It is important that this property characterizes dihedral groups.

PROPOSITION 2.2 ([13, p. 51]). *Let G be a finite group generated by two of its elements of order 2. Then $G \cong \mathcal{D}_n$, where n is the order of the product of the generators.* ∎

Dihedral subgroups of the symmetric group. Recall that $\mathcal{S}_n$ denotes the nth symmetric group. In this subsection $n \geq 3$.

PROPOSITION 2.3. *Let G be a subgroup of $\mathcal{S}_n$ isomorphic to $\mathcal{D}_m$ for some m, and containing an n-cycle. Then $m = n$.*

Proof. The only cyclic subgroup of $\mathcal{S}_n$ containing a given n-cycle is the group generated by this cycle. Hence the maximal cyclic subgroup of G is of order n. On the other hand, since $G \cong \mathcal{D}_m$, the maximal cyclic subgroup of G is of order $\max\{m, 2\}$. Since $n \geq 3$, we have $m = n$. ∎

DEFINITION 2.4. A *dihedral subgroup* of $\mathcal{S}_n$ is a subgroup isomorphic to $\mathcal{D}_n$ and containing an n-cycle.

THEOREM 2.5. *Any n-cycle $\sigma \in \mathcal{S}_n$ is contained in exactly one dihedral subgroup. The elements of this subgroup may only have the permutation types*

$$(m, \dots, m), \quad (1, 2, \dots, 2), \quad (1, 1, 2, \dots, 2), \tag{1}$$

where $m \mid n$.

(Of course, the second of the types (1) may only occur for odd n, while the third one only for even n.)

P r o o f. *Existence.* We may assume that $\sigma = (1, \dots, n)$. Consider a regular n-gon with vertices numbered $1, \dots, n$. It is well known that the group of its isometries is $\mathcal{D}_n$. Action of this group on the vertices defines a dihedral subgroup containing σ. Obviously, the elements of this subgroup have only the permutation types (1).

Uniqueness. There are exactly $\frac{1}{2}(n-1)!$ distinct 2-element sets of the form $\{\sigma, \sigma^{-1}\}$, where $\sigma \in \mathcal{S}_n$ is an n-cycle. It follows that the normalizer of any of these sets consists of $n!/\big(\frac{1}{2}(n-1)!\big) = 2n$ elements.

On the other hand, by Proposition 2.1, for any n-cycle σ, the set $\{\sigma, \sigma^{-1}\}$ is normalized by any dihedral subgroup containing σ. Since the normalizer of $\{\sigma, \sigma^{-1}\}$ consists of $2n$ elements, the uniqueness follows. ∎

It follows that all dihedral subgroups of $\mathcal{S}_n$ are conjugate. We shall not use this fact.

Let $\mathcal{S}_n$ be realized as the permutation group of the set $\{0, \dots, n-1\}$, and let $\mathcal{D}_n$ be realized as the dihedral subgroup containing the cycle $\sigma := (0, \dots, n-1)$. Then for any $k \in \{0, \dots, n-1\}$ there exists exactly one $\tau = \tau_k \in \mathcal{D}_n \setminus \{\mathrm{id}\}$ such that $\tau_k(k) = k$. (Indeed, there is exactly one non-trivial isometry of the regular n-gon stabilizing a given vertex.) Obviously, $\tau_k^2 = \mathrm{id}$.

PROPOSITION 2.6. *The subgroup generated by τ_0 and τ_k is of index* $\gcd(n, 2k)$ *in $\mathcal{D}_n$.*

P r o o f. By Proposition 2.2, the subgroup generated by τ_0 and τ_k is isomorphic to $\mathcal{D}_m$, where m is the order of $\tau_0\tau_k$. Since $\tau_0\sigma\tau_0 = \sigma^{-1}$ and $\tau_k = \sigma^k\tau_0\sigma^{-k}$, we have $\tau_0\tau_k = \sigma^{-2k}$. Hence $m = n/(n, 2k)$, whence the result. ∎

3. Dickson polynomials. For $a \in K$, the *nth Dickson polynomial* $D_n(x, a)$ is defined from the relation

$$D_n(z + a/z, a) = z^n + (a/z)^n. \tag{2}$$

Sometimes we write $D_{n,a}(x)$ instead of $D_n(x, a)$.

The following identities (where T_n stands for the nth Chebyshev polynomial) will be used in the paper without special reference:

$$D_n(x,0) = x^n; \quad D_n(x,1) = 2T_n(x/2); \tag{3}$$

$$D_1(x,a) = x; \quad D_2(x,a) = x^2 - 2a; \tag{4}$$

$$D_{mn}(x,a) = D_m(D_n(x,a), a^n); \tag{5}$$

$$b^n D_n(x,a) = D_n(bx, b^2 a). \tag{6}$$

The proofs are immediate, upon substituting $x = z + a/z$ into both sides.

For further facts about Dickson polynomials, including equivalent definitions, differential equations, etc., see [11, Chapter 2].

Factorization. The following is a slight modification of Proposition 1.7 from [16].

Proposition 3.1. *Put*

$$\Phi_n(x,y,a) = \prod_{\substack{1\le k<n\\ k\equiv 1 \bmod 2}} (x^2 - xy\cdot 2\cos(\pi k/n) + y^2 - a\cdot 4\sin^2(\pi k/n)). \tag{7}$$

Then

$$D_n(x,a) + D_n(y,a) = \begin{cases} \Phi_n(x,y,a) & \text{if } n \text{ is even},\\ (x+y)\Phi_n(x,y,a) & \text{if } n \text{ is odd}. \end{cases} \tag{8}$$

In particular,

$$T_n(x) + T_n(y) = \begin{cases} \frac{1}{2}\Phi_n(2x,2y,1) & \text{if } n \text{ is even},\\ (x+y)\Phi_n(2x,2y,1) & \text{if } n \text{ is odd}. \end{cases} \tag{9}$$

Proof. Put

$$\Psi_n(x,y,a) = (x-y) \prod_{1\le k\le (n-1)/2} (x^2 - xy\cdot 2\cos(2\pi k/n) + y^2 - a\cdot 4\sin^2(2\pi k/n)).$$

By [16, Proposition 1.7],

$$D_n(x,a) - D_n(y,a) = \begin{cases} \Psi_n(x,y,a) & \text{if } n \text{ is even},\\ (x+y)\Psi_n(x,y,a) & \text{if } n \text{ is odd}. \end{cases} \tag{10}$$

In particular,

$$D_n(x,a)^2 - D_n(y,a)^2 = D_{2n}(x,a) - D_{2n}(y,a) = \Psi_{2n}(x,y,a). \tag{11}$$

Now (8) follows from (10) and (11) after obvious transformations. ∎

Corollary 3.2. *If $a \in K^*$ and $2\cos(2\pi/n) \in K$ then the polynomial $D_n(x,a) + D_n(y\cdot 2\cos(\pi/n), a)$ splits over K into absolutely irreducible quadratic factors (and a linear factor if n is odd). If $2\cos(2\pi/n) \in K$ but $e^{2\pi i/n} \notin K$ then $D_n(x,0) + D_n(y\cdot 2\cos(\pi/n), 0)$ splits over K into irreducible quadratic factors (and a linear factor if n is odd).* ∎

Extrema. Given a polynomial $f(x)$ having s distinct roots in $\overline{K}$, its *root type* is the array $(\mu_1,\dots,\mu_s)$ formed of the multiplicities of its roots. Obviously, $\mu_1+\dots+\mu_s=\deg f$.

Given $\gamma\in\overline{K}$, put

$$\delta(\gamma)=\delta_f(\gamma)=\sum_{i=1}^{s}(\mu_i-1)=\deg f-s,$$

where $(\mu_1,\dots,\mu_s)$ is the root type of $f(x)-\gamma$. We have

$$\sum_{\gamma\in\overline{K}}\delta_f(\gamma)=\sum_{\gamma\in\overline{K}}\deg\gcd(f(x)-\gamma,f'(x))=\deg f'(x)=\deg f-1. \tag{12}$$

We say that $\gamma\in\overline{K}$ is an *extremum* of $f(x)$ if $f(x)-\gamma$ has a multiple root (equivalently, if $\delta_f(\gamma)>0$). The *type* of an extremum γ is the root type of $f(x)-\gamma$.

Proposition 3.3. (a) *The polynomial $D_n(x,0)$ has exactly one extremum $\gamma=0$, of type (n).*

(b) *If $a\neq 0$ and $n\geq 3$ then $D_n(x,a)$ has exactly two extrema $\pm 2a^{n/2}$. If n is odd then both are of type $(1,2,\dots,2)$. If n is even then $2a^{n/2}$ is of type $(1,1,2,\dots,2)$, and $-2a^{n/2}$ is of type $(2,\dots,2)$.*

Proof. (a) is obvious. To prove (b), observe that ([3])

$$D_n(2\sqrt{a},a)=D_n(\sqrt{a}+a/\sqrt{a},a)=2a^{n/2},$$

and $D_n(-2\sqrt{a},a)=(-1)^n2a^{n/2}$. Substituting $y=\pm 2\sqrt{a}$ into (10), we obtain

$$D_n(x,a)\pm 2a^{n/2}=(x\pm 2\sqrt{a})\Delta_n(x,\pm\sqrt{a})^2 \quad (n \text{ odd}), \tag{13}$$

$$D_n(x,a)-2a^{n/2}=(x^2-4a)\Delta_n(x,\sqrt{a})^2 \quad (n \text{ even}), \tag{14}$$

where $\Delta_n(x,\alpha)=\prod_{1\leq k\leq(n-1)/2}(x-\alpha\cdot 2\cos(2\pi k/n))$. Also, if n is even then

$$D_n(x,a)+2a^{n/2}=D_2(D_{n/2}(x,a),a^{n/2})+2a^{n/2}=D_{n/2}(x,a)^2. \tag{15}$$

Now (b) follows from (13)–(15), which show that $\pm 2a^{n/2}$ are the extrema of the required type, and from (12), which implies that no other extrema exist. ∎

It is of fundamental importance that, basically, the Dickson polynomials are characterized by the property established in Proposition 3.3. We shall use this classical fact in the following form.

Theorem 3.4. *Let $f(x)\in K[x]$ be a polynomial of degree n having extrema only of one of the following types*:

$$(n),\quad (2,\dots,2),\quad (1,2,\dots,2),\quad (1,1,2,\dots,2). \tag{16}$$

([3]) We fix a value of the $\sqrt{a}$ and define $a^{n/2}=(\sqrt{a})^n$.

Then either $\deg f = 4$ *or*

$$f(x) = \alpha D_n(x+\beta, a) + \gamma, \qquad \text{where } \alpha \in K^* \text{ and } a, \beta, \gamma \in K. \tag{17}$$

(*We do not assume that the extrema belong to* K.)

P r o o f. If $f(x)$ has at least 3 extrema, then (12) implies that $\deg f = 4$.

If $f(x)$ has a single extremum γ, then, again by (12), it is of type (n). It is immediate now that (17) with $a = 0$ holds.

From now on, assume that $f(x)$ has exactly two extrema. Using induction on n, we shall prove that in this case (17) holds with $a \neq 0$.

If n is odd then both the extrema are of type $(1, 2, \dots, 2)$. In this case the assertion is a particular case of [16, Lemma 1.11] (reproduced in [11] as Lemma 6.16).

Now assume that n is even, and write $n = 2m$. Since $f(x)$ has two extrema, we have $n \geq 4$. By (12), one of the extrema is of type $(2, \dots, 2)$ and the other is of type $(1, 1, 2, \dots, 2)$. Since the extrema have distinct types, they both belong to K. Without loss of generality, we may assume that the polynomial $f(x)$ is monic and that the extremum of type $(2, \dots, 2)$ is 0. This means that $f(x) = g(x)^2$, where $g(x) \in K[x]$ is a monic polynomial of degree m.

If $m = 2$ then $g(x) = D_2(x+\beta, a)$, where $a, \beta \in K$. Moreover, $a \neq 0$, because $g(x)$ has simple roots.

Now assume that $m = \deg g > 2$. Let $\kappa \neq 0$ be the other extremum of $f(x)$. Then $(g(x) - \sqrt{\kappa})(g(x) + \sqrt{\kappa})$ has 2 simple roots, all the other roots being of order 2. It follows that $\pm\sqrt{\kappa}$ are extrema of $g(x)$, of one of the last three types from (16). Identity (12) applied to the polynomial $g(x)$ yields that it has no other extrema. By induction, $g(x) = \alpha D_m(x+\beta, a) + \gamma$, where $a, \alpha \in K^*$ and $\beta, \gamma \in K$. Since $g(x)$ is monic, $\alpha = 1$. Since its extrema $\pm\sqrt{\kappa}$ are symmetric with respect to 0, we have $\gamma = 0$.

Thus, in either case, $m = 2$ or $m > 2$, we have $g(x) = D_m(x+\beta, a)$, where $a \in K^*$ and $\beta \in K$. It follows that $f(x) = g(x)^2 = D_n(x+\beta, a) + 2a^m$, as wanted. ∎

Monodromy. Given a polynomial $f(x) \in K[x]$, consider $f(x) - t$ as a polynomial in x over $K(t)$, and denote by $\mho_f$ its splitting field over $K(t)$. The Galois group $\mathrm{Gal}(\mho_f / K(t))$ is called the *monodromy group of* f *over* K, and denoted by $\mathrm{Mon}_K f$. The *absolute* monodromy group $\mathrm{Mon}_{\overline{K}} f$ is denoted by $\mathrm{Mon}\, f$. We have the standard exact sequence

$$1 \to \mathrm{Mon}\, f \to \mathrm{Mon}_K f \to \mathrm{Gal}(\widehat{K}/K) \to 1, \tag{18}$$

where $\widehat{K}$ is the constant subfield of $\mho_f$.

It will be convenient to number the roots of $D_n(x, a) - t$ as follows. Let $x^{(0)}$ be one of the roots, and let z be one of the roots of $Z + a/Z = x^{(0)}$.

Then the n roots of $D_n(x,a) - t$ are

$$x^{(k)} = \xi^k z + \xi^{-k} a/z \quad (k = 0, \dots, n-1), \tag{19}$$

where $\xi = e^{2\pi i/n}$. Notice that

$$z = (\xi^k x^{(k)} - \xi^j x^{(j)})/(\xi^{2k} - \xi^{2j}) \quad (2k \not\equiv 2j \bmod n),$$

which implies

$$\overline{K}\mathfrak{V}_{D_{n,a}} = \overline{K}(z) \quad (n \geq 3). \tag{20}$$

Notice also that $[\overline{K}(z) : \overline{K}(x^{(0)})] = 2$ when $a \neq 0$, because $(x^{(0)})^2 - 4a$ (the Z-discriminant of $Z^2 - x^{(0)}Z + a$) is not a square in $\overline{K}(x^{(0)})$. This implies that

$$[\overline{K}\mathfrak{V}_{D_{n,a}} : \overline{K}(t)] = [\overline{K}(z) : \overline{K}(t)] = 2n \quad (a \neq 0,\ n \geq 3). \tag{21}$$

The minimal polynomial of z over $\overline{K}(t)$ is $Z^{2n} - tZ^n + a^n$, and the $2n$ conjugates of z are

$$\xi^k z, \quad \xi^k a/z \quad (k = 0, \dots, n-1).$$

Now it is easy to show that

$$\operatorname{Mon} D_{n,a} \cong \begin{cases} Z_n & \text{if } a = 0, \\ \mathcal{D}_n & \text{if } a \neq 0 \text{ and } n \geq 3. \end{cases} \tag{22}$$

(Recall that Z_n stands for the cyclic group of order n, and $\mathcal{D}_n$ is the nth dihedral group.) Indeed, the case $a = 0$ is obvious. Now assume that $a \neq 0$. By definition,

$$G := \operatorname{Mon} D_{n,a} = \operatorname{Gal}(\overline{K}\mathfrak{V}_{D_{n,a}}/\overline{K}(t)) = \operatorname{Gal}(\overline{K}(z)/\overline{K}(t)).$$

Let α and β be the automorphisms of $\overline{K}(z)$ defined by $\alpha(z) = a/z$ and $\beta(z) = \xi z$. Since $\beta^k(z) = \xi^k z$ and $\alpha\beta^k(z) = \xi^k a/z$, the group G is generated by α and β. Since $\alpha^2 = (\alpha\beta)^2 = \beta^n = 1$, the group G is a quotient of $\mathcal{D}_n$. Since $|G| = 2n$, we conclude that $G \cong \mathcal{D}_n$.

Proposition 3.5. *The constant subfield of $\mathfrak{V}_{D_{n,0}}$ is $K(\xi)$. If $a \in K^*$ then the constant subfield of $\mathfrak{V}_{D_{n,a}}$ is $K(2\cos(2\pi/n))$.*

Proof. The first assertion is obvious. Now assume that $a \neq 0$. We have

$$2\cos(2\pi/n) = (x^{(k)} + x^{(k+2)})/x^{(k+1)} \in \mathfrak{V}_{D_{n,a}} \quad (k = 0, \dots, n-3). \tag{23}$$

It remains to prove that $[\mathfrak{V}_{D_{n,a}} : K(t, 2\cos(2\pi/n))] = [\overline{K}\mathfrak{V}_{D_{n,a}} : \overline{K}(t)] = 2n$. Since, obviously, $[\mathfrak{V}_{D_{n,a}} : K(t, 2\cos(2\pi/n))] \geq [\overline{K}\mathfrak{V}_{D_{n,a}} : \overline{K}(t)]$, it suffices to show that

$$[\mathfrak{V}_{D_{n,a}} : K(t, 2\cos(2\pi/n))] \leq 2n. \tag{24}$$

Rewriting (23) as

$$x^{(k+2)} = x^{(k+1)} \cdot 2\cos(2\pi/n) - x^{(k)} \quad (k = 0, \dots, n-3),$$

we conclude that $\mho_{D_{n,a}} \subseteq K(x^{(0)}, x^{(1)}, 2\cos(2\pi/n))$. Since

$$x^{(1)} + x^{(n-1)} = x^{(0)} \cdot 2\cos(2\pi/n) \in K(x^{(0)}, 2\cos(2\pi/n)),$$

and

$$x^{(1)}x^{(n-1)} = (x^{(0)})^2 - 2a(1 - 2\cos(4\pi/n)) \in K(x^{(0)}, 2\cos(2\pi/n)),$$

we have $[K(x^{(0)}, x^{(1)}, 2\cos(2\pi/n)) : K(x^{(0)}, 2\cos(2\pi/n))] \leq 2$. This implies (24), and the proposition follows. ∎

As usual, the group $\operatorname{Mon}_K f$ acts faithfully on the set of roots of $f(x) - t$. This action defines (up to conjugation in $\mathcal{S}_n$) an embedding $\operatorname{Mon}_K f \hookrightarrow \mathcal{S}_n$, where $n = \deg f$. In the sequel, we shall view $\operatorname{Mon}_K f$ as a subgroup of $\mathcal{S}_n$, and $\operatorname{Mon} f$ as a subgroup of $\operatorname{Mon}_K f$.

In particular, we embed $\operatorname{Mon}_K D_{n,a}$ into S_n through the correspondence $k \mapsto x^{(k)}$.

Proposition 3.6. *Let $a \in K$ and $n \geq 3$. Then the following assertions are equivalent.*

(a) $\operatorname{Mon}_K D_{n,a} \cong \mathcal{D}_n$.

(b) $\operatorname{Mon}_K D_{n,a}$ *is the dihedral subgroup of* $\mathcal{S}_n$ (*see Definition* 2.4), *containing the cycle* $(0, \dots, n-1)$.

(c) $2\cos(2\pi/n) \in K$ *and if* $a = 0$ *then* $\xi \notin K$.

Proof. To begin with, observe that the absolute monodromy group $\operatorname{Mon} D_{n,a}$ contains the cycle $(0, \dots, n-1)$, which is given (through $k \mapsto x^{(k)}$) by the automorphism $z \mapsto \xi z$ of $\overline{K}(z)$. Hence $\operatorname{Mon}_K D_{n,a}$ also contains this cycle, which proves the equivalence (a) $\Leftrightarrow$ (b).

When $a \neq 0$, (a) is equivalent to (c) by (18), (22) and Proposition 3.5.

We are left with $a = 0$. Assume first (c), which means in this case that $2\cos(2\pi/n) \in K$ but $\xi \notin K$. Let α and β be the automorphisms of $\mho_{D_{n,0}} = K(\xi, z)$ defined by $\alpha(\xi) = \xi^{-1}$, $\alpha(z) = z$ and $\beta(\xi) = \xi$, $\beta(z) = \xi z$. Then $\alpha^2 = \beta^n = (\alpha\beta)^2 = \mathrm{id}$, which shows that $\operatorname{Mon}_K D_{n,0}$ is a quotient of $\mathcal{D}_n$. Since $[K(\xi, z) : K(t)] = 2n$, we have $\operatorname{Mon}_K D_{n,0} \cong \mathcal{D}_n$.

Conversely, assume that $\operatorname{Mon}_K D_{n,0} \cong \mathcal{D}_n$. Then $\xi \notin K$ (for otherwise $\operatorname{Mon} f = \operatorname{Mon}_K f$). Let $\gamma \in \operatorname{Mon} f$ be defined by $\gamma(z) = \xi z$. By Proposition 2.1, the conjugacy class of γ in $\operatorname{Mon}_K f$ is $\{\gamma, \gamma^{-1}\}$. It follows that the set $\{\xi, \xi^{-1}\}$ is stable under $\operatorname{Mon}_K f$. Hence $2\cos(2\pi/n) = \xi + \xi^{-1} \in K$.

Thus, (a) is equivalent to (c) also when $a = 0$. The proposition is proved. ∎

Proposition 3.7. *Let $f(x)$ be a polynomial of degree n. Then the group* $\operatorname{Mon} f$ *contains an n-cycle. Also, for any extremum $\gamma \in \overline{K}$ of f of type* $(e_1, \dots, e_s)$, *the group* $\operatorname{Mon} f$ *contains a permutation of type* $(e_1, \dots, e_s)$.

Proof. See [16, Lemmas 3.3 and 3.4] or [11, Theorems 6.12 and 6.13].

THEOREM 3.8. *Let $f(x) \in K[x]$ be a polynomial of degree $n \geq 3$. Assume that $\mathrm{Mon}_K f \cong \mathcal{D}_m$ for some m. Then $m = n$ and $f(x) = \alpha D_n(x + \beta, a) + \gamma$, where $\alpha \in K^*$ and $a, \beta, \gamma \in K$.*

P r o o f. To begin with, notice that $\mathrm{Mon}_K f$ contains an n-cycle by Proposition 3.7. Hence $m = n$ by Proposition 2.3, and $\mathrm{Mon}_K f$ is a dihedral subgroup of $\mathcal{S}_n$ (see Definition 2.4).

Assume first that $n = 4$. Since every 2-element subgroup of $\mathcal{D}_4$ is contained in a 4-element subgroup, there exists an intermediate field between $K(t)$ and $K(x_0)$, where x_0 is a root of $f(x) - t$. It follows that $f(x)$ is a composition of two quadratic polynomials, which can be written as $f(x) = \alpha((x - \beta)^2 - 2a)^2 + \gamma'$. Plainly, $\alpha \in K^*$ and $a, \beta, \gamma' \in K$. Further, $f(x) = \alpha D_4(x + \beta, a) + \gamma$ with $\gamma = \gamma' + \alpha a^2 \in K$.

Now assume that $n \neq 4$. By Proposition 3.7 and Theorem 2.5, the polynomial $f(x)$ may have extrema only of the types (1). Identity (12) implies that in the first of the types only $m = n$ or $m = 2$ (for even n) are possible. In other words, $f(x)$ may have extrema only of the types (16). By Theorem 3.4, we have $f(x) = \alpha D_n(x + \beta, a) + \gamma$ with $\alpha \in K^*$ and $a, \beta, \gamma \in K$. The theorem is proved. ∎

REMARK 3.9. Turnwald [16, Theorem 3.11] proved that a polynomial with a solvable monodromy group is a composition of linear polynomials, Dickson polynomials and polynomials of degree 4. Recently [17] he extended this result (with appropriate modifications) to arbitrary characteristic.

The sum of two roots. Fix $a \in K$ and an integer $n \geq 3$. In this subsection we assume that

$$2\cos(2\pi/n) \in K \text{ and if } a = 0 \text{ then } \xi = e^{2\pi i/n} \notin K. \tag{25}$$

It is easy to see that

$$[(K(x^{(0)}) \cap K(x^{(k)})) : K(t)] = (n, 2k) \tag{26}$$

(we use the notation (19)). Indeed, Proposition 3.6 implies that $\mathrm{Mon}_K D_{n,a}$ is the dihedral subgroup of $\mathcal{S}_n$ containing the cycle $(x^{(0)}, \dots, x^{(n-1)})$. By Proposition 2.6, the subgroup of $\mathrm{Mon}_K D_{n,a}$ stabilizing the field $K(x^{(0)}) \cap K(x^{(k)})$ is of index $(n, 2k)$. This proves (26).

PROPOSITION 3.10. *Let x_0 and x_1 be two roots of $D_n(x, a) - t$ satisfying $x_0 + x_1 \neq 0$. Then $x_0 + x_1$ is a root of $D_n(x/(2\cos(\pi k/n)), a) - (-1)^k t$, where $k \in \{0, \dots, n-1\}$ is distinct from $n/2$. If $n \equiv 0 \bmod 4$ and $[(K(x_0) \cap K(x_1)) : K(t)] = 2$ then k is odd.*

P r o o f. Without loss of generality, $x_0 = x^{(0)}$ and $x_1 = x^{(k)}$, where $k \neq n/2$ because $x_0 + x_1 \neq 0$. Then $x_0 + x_1 = x' \cdot 2\cos(\pi/k)$, where

$$x' = e^{\pi i k/n} z + e^{-\pi i k/n} a/z.$$

Hence $D_n(x', a) = (-1)^k(z^n + (a/z)^n) = (-1)^k t$, which proves the first assertion.

If $n \equiv 0 \bmod 4$ and $[(K(x_0) \cap K(x_1)) : K(t)] = 2$ then k is odd by (26). ∎

4. Proof of Theorems 1.2 and 1.3

Proof of Theorem 1.3. Since the proof is rather long, we divide it into short logically complete steps.

Step 0 (preliminaries). We may assume that

$$\min(\deg f, \deg g) > 1, \tag{27}$$

$$\max(\deg f, \deg g) \geq 3 \tag{28}$$

for otherwise there is nothing to prove.

Let $x_0 \in \overline{K(t)}$ be a root of $f(x) - t$. Then there is a root y_0 of $g(x) - t$ such that $q(x_0, y_0) = 0$. Since $q(x, y)$ is irreducible, for any $\Phi(x, y) \in K[x, y]$ we have

$$\Phi(x_0, y_0) = 0 \Rightarrow q(x, y) \mid \Phi(x, y). \tag{29}$$

Step 1. Assume first that $K(x_0) \cap K(y_0)$ is a proper extension of $K(t)$, and write it as $K(z)$, where z is integral over $K[t]$. Then $z = f_0(x_0) = g_0(y_0)$ and $t = \phi_0(z)$, where f_0, g_0 and ϕ_0 are polynomials over K with $\deg f_0 < \deg f$ and $\deg g_0 < \deg g$. We have $f = \phi_0 \circ f_0$ and $g = \phi_0 \circ g_0$. Since $f_0(x_0) - g_0(y_0) = 0$, the polynomial $q(x, y)$ divides $f_0(x) - g_0(y)$ by (29). Using induction on $\deg f$, we conclude that $f_0 = \phi_1 \circ f_1$ and $g_0 = \phi_1 \circ g_1$, where f_1 and g_1 are as required. Putting $\phi = \phi_0 \circ \phi_1$, we complete the proof in this case.

Step 2. From now on,

$$K(x_0) \cap K(y_0) = K(t). \tag{30}$$

Let Ω be a Galois extension of $K(t)$ containing x_0 and y_0, and G the subgroup of $\mathrm{Gal}(\Omega/K(t))$ stabilizing x_0. Since y_0 is at most quadratic over $K(x_0)$, the field $K(x_0, y_0)$ is G-invariant. Similarly, if H is the subgroup stabilizing y_0 then $K(x_0, y_0)$ is H-invariant.

By (30), the subgroups G and H together generate the whole group $\mathrm{Gal}(\Omega/K(t))$. Hence $K(x_0, y_0)$ is invariant with respect to $\mathrm{Gal}(\Omega/K(t))$, which implies that $K(x_0, y_0)$ is a Galois extension of $K(t)$. Thus,

$$K(x_0) \subseteq \mathfrak{V}_f \subseteq K(x_0, y_0).$$

(Recall that $\mathfrak{V}_f$ denotes the splitting field of $f(x) - t$ over $K(t)$.)

Step 3. Another consequence of (30) is

$$[K(x_0, y_0) : K(x_0)] = [K(x_0, y_0) : K(y_0)] = 2. \tag{31}$$

Indeed, the inequality

$$[K(x_0, y_0) : K(x_0)] \le 2$$

is obvious. Now if $y_0 \in K(x_0)$ then (30) implies $y_0 \in K(t)$, which contradicts (27). Hence $[K(x_0, y_0) : K(x_0)] = 2$, and similarly $[K(x_0, y_0) : K(y_0)] = 2$, proving (31). It follows from (31) that

$$\deg f = \deg g. \tag{32}$$

Step 4. Write $q(x, y) = q_{xx}x^2 + q_{xy}xy + q_{yy}y^2 +$ linear terms. Then $q_{xx}q_{yy} \neq 0$ by (31). It is important that also $q_{xy} \neq 0$.

Indeed, if $q_{xy} = 0$, then $q(x, y) = f_0(x) - g_0(y)$, where f_0 and g_0 are polynomials over K of degree at most 2. Since

$$z := f_0(x_0) = g_0(y_0) \in K(x_0) \cap K(y_0) = K(t),$$

we have

$$\deg f = \deg g = [K(x_0) : K(t)] \le [K(x_0) : K(z)] = \deg f_0 \le 2,$$

contradicting (28). Hence $q_{xy} \neq 0$.

Step 5. Let x_1 be the conjugate to x_0 over $K(y_0)$. Then x_1 is a root of $f(x) - t$, and $q_{xy} \neq 0$ implies that $x_0 + x_1 = \alpha_1 y_0 + \gamma_1$ with $\alpha_1 \in K^*$ and $\gamma_1 \in K$. In particular, $y_0 \in \mathfrak{O}_f$, which implies that $\mathfrak{O}_f = K(x_0, y_0)$.

Let σ (respectively, τ) be the non-trivial automorphism of $\mathfrak{O}_f$ over $K(x_0)$ (respectively, $K(y_0)$). By (30), the automorphisms σ and τ generate the group $\mathrm{Mon}_K f$. By Propositions 2.2 and 2.3, we have $\mathrm{Mon}_K f \cong \mathcal{D}_n$, where $n = \deg f = \deg g$. Theorem 3.8 implies that $f(x) = \kappa D_n(x + \beta, a) + \lambda$ where $\kappa \in K^*$ and $a, \beta, \lambda \in K$. By Proposition 3.6,

$$2\cos(2\pi/n) \in K \text{ and if } a = 0 \text{ then } e^{2\pi i/n} \notin K. \tag{33}$$

Step 6. Thus, $x_0 + \beta$ and $x_1 + \beta$ are two roots of $D_n(x, a) - t'$, where $t' = (t - \lambda)/\kappa$. Proposition 3.10 implies that $x_0 + x_1 + 2\beta = \alpha_1 y_0 + \gamma_1 + 2\beta$ is a root of $D_n(x/(2\cos(\pi k/n)), a) - (-1)^k t'$, where $k \in \{0, \dots, n-1\}$ and $k \neq n/2$.

It follows that the polynomials $g(x) - t$ and

$$D_n\left(\frac{\alpha_1 x + \gamma_1 + 2\beta}{2\cos(\pi k/n)}, a\right) - (-1)^k t'$$

have a common root y_0. Since both the polynomials are irreducible over $K(t)$, we have

$$g(x) - t = c\left(D_n\left(\frac{\alpha_1 x + \gamma_1 + 2\beta}{2\cos(\pi k/n)}, a\right) - (-1)^k (t - \lambda)/\kappa\right)$$

with $c \in K^*$. Comparing the coefficients of t, we find $c = (-1)^k \kappa$. Thus,

$$g(x) = (-1)^k \kappa D_n\left(\frac{\alpha_1 x + \gamma_1 + 2\beta}{2\cos(\pi k/n)}, a\right) + \lambda. \tag{34}$$

Step 7. If

(35) at least one of the numbers k and n is odd,

then $\cos(\pi k/n)\cos(\pi/n) \in K$, and we can rewrite (34) as

$$g(x) = \kappa(-D_n((\alpha x + \gamma)\cdot 2\cos(\pi/n), a)) + \lambda$$

where

$$\alpha = \frac{(-1)^{k+1}\alpha_1}{4\cos(\pi k/n)\cos(\pi/n)} \in K^*, \quad \gamma = \frac{(-1)^{k+1}(\gamma_1 + 2\beta)}{4\cos(\pi k/n)\cos(\pi/n)} \in K.$$

Putting $\varphi(x) = \kappa x + \lambda$, we complete the proof in the case (35).

Step 8. Now assume that

(36) both k and n are even.

The group $G \le \mathrm{Mon}_K f$ stabilizing $K(x_0) \cap K(x_1)$ is generated by σ and $\sigma' = \tau\sigma\tau$. The order of $\sigma\sigma' = (\sigma\tau)^2$ is $m = n/2$. Proposition 2.2 implies that $G \cong \mathcal{D}_m$, whence $[(K(x_0) \cap K(x_1)) : K(t)] = 2$. Therefore $n \equiv 2 \bmod 4$ by the second assertion of Proposition 3.10.

Thus, m is odd. It follows from (33) that $2\cos(\pi/m) \in K$ and if $a = 0$ then $e^{2\pi i/m} \notin K$. Also, since k is even, $2\cos(\pi k/n) \in K$. Hence we can rewrite (34) as $g(x) = \kappa D_n((\alpha x + \gamma)\cdot 2\cos(\pi/m), a) + \lambda$, where

$$\alpha = \varepsilon\frac{\alpha_1}{4\cos(\pi k/n)\cos(\pi/m)} \in K^*, \quad \gamma = \varepsilon\frac{\gamma_1 + 2\beta}{4\cos(\pi k/n)\cos(\pi/m)} \in K,$$

and $\varepsilon \in \{1, -1\}$ is to be defined later. Thus,

$$f(x) - g(y) = \kappa(f_1(x) - g_1(y))(f_1(x) + g_1(y)),$$

where

$$f_1(x) = D_m(x + \beta, a), \quad g_1(x) = -D_m((\alpha x + \gamma)\cdot 2\cos(\pi/m), a).$$

Now we can define ε so that $q(x, y)$ divides $f_1(x) - g_1(x)$. Putting $\varphi(x) = \kappa D_2(x, a^m) + \lambda$, we complete the proof also in the case (36). ∎

Proof of Theorem 1.2. If $f(x) - g(y)$ has a linear factor then there is nothing to prove. Hence we may assume that it has no linear factors, but has an absolutely irreducible quadratic factor. Theorem 1.3 implies that $f = \varphi_0 \circ f_0$ and $g = \varphi_0 \circ g_0$, where either $\max(\deg f_0, \deg g_0) = 2$ or

(37) $$f_0(x) = D_n(\alpha_1 x + \beta_1, a), \quad g_0(x) = -D_n(\gamma_1 x + \delta_1, a).$$

In the former case the proof is complete. Now assume (37). Since $f(x) - g(x)$ has no linear factors, $a \ne 0$. Hence $f_0 = 2b^{-n}T_n(\alpha x + \beta)$ and $g_0 = -2b^{-n}T_n(\gamma x + \delta)$, where $b = \sqrt{a}$ and $\alpha = \alpha_1/(2b)$, $\beta = \beta_1/(2b)$, etc.

Write $n = m \cdot 2^k$, where m is odd. Putting $\varphi(x) = \varphi_0(2b^{-n}T_m(x))$, we complete the proof. ∎

References

[1] Yu. F. Bilu and R. F. Tichy, *The Diophantine equation $f(x) = g(y)$*, submitted.

[2] P. Cassou-Noguès et J.-M. Couveignes, *Factorisations explicites de $g(y) - h(z)$*, Acta Arith. 87 (1999), 291–317.

[3] W. Feit, *On symmetric balanced incomplete block designs with doubly transitive automorphism groups*, J. Combin. Theory Ser. A 14 (1973), 221–247.

[4] —, *Some consequences of the classification of finite simple groups*, in: Proc. Sympos. Pure Math. 37, Amer. Math. Soc., 1980, 175–181.

[5] M. Fried, *On a conjecture of Schur*, Michigan Math. J. 17 (1970), 41–55.

[6] —, *The field of definition of function fields and a problem in the reducibility of polynomials in two variables*, Illinois J. Math. 17 (1973), 128–146.

[7] —, *On a theorem of Ritt and related Diophantine problems*, J. Reine Angew. Math. 264 (1974), 40–55.

[8] —, *Exposition on an arithmetic-group theoretic connection via Riemann's existence theorem*, in: Proc. Sympos. Pure Math. 37, Amer. Math. Soc., 1980, 571–601.

[9] —, *Variables separated polynomials, the genus 0 problem and moduli spaces*, in: Number Theory in Progress (Zakopane, 1997), de Gruyter, 1999, 169–228.

[10] S. Lang, *Fundamentals of Diophantine Geometry*, Springer, 1983.

[11] R. Lidl, G. L. Mullen and G. Turnwald, *Dickson Polynomials*, Pitman Monographs Surveys Pure Math. 65, Longman Sci. Tech., 1993.

[12] J. F. Ritt, *Prime and composite polynomials*, Trans. Amer. Math. Soc. 23 (1922), 51–66.

[13] D. J. S. Robinson, *A Course in the Theory of Groups*, Grad. Texts in Math. 80, Springer, 1982.

[14] A. Schinzel, *Selected Topics on Polynomials*, The Univ. of Michigan Press, Ann Arbor, MI, 1983.

[15] C. L. Siegel, *Über einige Anwendungen Diophantischer Approximationen*, Abh. Preuss. Akad. Wiss. Phys.-Math. Kl. 1929, Nr. 1.

[16] G. Turnwald, *On Schur's conjecture*, J. Austral. Math. Soc. 58 (1995), 312–357.

[17] —, *Some notes on monodromy groups of polynomials*, in: Number Theory in Progress (Zakopane, 1997), de Gruyter, 1999, 539–552.

[18] H. A. Tverberg, *A study in irreducibility of polynomials*, Ph.D. thesis, Department of Mathematics, University of Bergen, 1968.

Mathematisches Institut
Universität Basel
Rheinsprung 21
4051 Basel, Switzerland
E-mail: yuri@math.unibas.ch

Institut für Mathematik (A)
Technische Universität Graz
Steyrergasse 30,
8010 Graz, Austria

Received on 28.8.1998
and in revised form on 12.3.1999 (3454)