A Weil–bound free proof of Schur's conjecture

Peter Müller*
Department of Mathematics
University of Florida
Gainesville, FL 32611
E-mail: *pfm@math.ufl.edu*

Abstract

Let f be a polynomial with coefficients in the ring $\mathcal{O}_K$ of integers of a number field. Suppose that f induces a permutation on the residue fields $\mathcal{O}_K/\mathfrak{p}$ for infinitely many non-zero prime ideals $\mathfrak{p}$ of $\mathcal{O}_K$. Then Schur's conjecture, namely that f is a composition of linear and Dickson polynomials, has been proved by M. Fried. All the present versions of the proof use Weil's bound on the number of points of absolutely irreducible curves over finite fields in order to get a Galois theoretic translation and to finish the proof by means of finite group theory.

This note replaces the use of this deep result by elementary arguments.

Introduction

Let K be a number field, with $\mathcal{O}_K$ its ring of integers. We say that a polynomial $f \in \mathcal{O}_K[X]$ is *exceptional* if the following holds. For infinitely many non–zero prime ideals $\mathfrak{p}$ of $\mathcal{O}_K$, f as a function on $\mathcal{O}_K/\mathfrak{p}$ permutes the elements of this field. The aim of this note is

*The author thanks the Deutsche Forschungsgemeinschaft (DFG) for its support in form of a postdoctoral fellowship.

Theorem 1. *Let $f \in \mathcal{O}_K[X]$ be an exceptional polynomial of degree $n \geq 2$. Let t be a transcendental over K, and let x_i $(i = 1, 2, \ldots, n)$ be the roots of $f(X) - t$ in some algebraic closure of $K(t)$. Then $\sum_{i=1}^{n} \zeta^i x_i = 0$ for a primitive n-th root of unity ζ and a suitable numbering of the x_i.*

I. Schur had this assertion on page 128 in his paper [8] from 1923 for $K = \mathbb{Q}$. Of course there was no Weil bound (or sufficiently strong substitute) available at that time. Schur used a quite complicated series of arguments, involving the Lagrange inversion formula for power series and computations with multinomial coefficients. Further, Schur's method seems to work only for $K = \mathbb{Q}$. Our method works for any number field and is certainly more transparent.

After having proved Theorem 1 and Corollary 3, we sketch how Schur's conjecture follows from that by easy standard arguments. For more details about this consult [1] and [10].

An analogous result about exceptional polynomials over finite fields holds provided that the characteristic p does not divide the degree of f. Namely if f is exceptional under this assumption, then f is a composition of linear and Dickson polynomials. Under this assumption, Theorem 1 holds without change. However, the argument giving Corollary 3 does not work anymore. Instead, it seems that one cannot remove the use of the Weil-bound. The Appendix contains an account of that. M. Fried (who has a different proof) asked the author to supply a proof of Theorem 4.

Proof of Theorem 1.

Let a be the leading coefficient of f. Replacing $f(X)$ by $a^{n-1}f(X/a)$ does neither affect the hypothesis, nor the conclusion of Theorem 1, so henceforth we assume that $f \in \mathcal{O}_K[X]$ is monic. Denote by ζ a primitive n–th root of unity. Let R be the ring extension of $\mathcal{O}_K$ generated by ζ and $\frac{1}{n}$, so $R = \mathcal{O}_K[\zeta, \frac{1}{n}]$. Let z be a variable. We will work in the ring of formal power series $R[[z]]$. If $\mathfrak{p}$ is a prime ideal of $\mathcal{O}_K$, and $S_1, S_2 \in R[[z]]$, then the congruence $S_1 \equiv S_2 \pmod{\mathfrak{p}}$ means that $S_1 - S_2 \in R[[z]]\mathfrak{p}$. (We will use this notion only when $\mathfrak{p}$ does not divide n.) Also, a congruence modulo a power of z has its obvious meaning. Set

$$f_z(X) = z^n f(\frac{X}{z}) - 1.$$

Lemma 2. *For $i = 1, 2, \dots, n$ there are $y_i \in R[[z]]$ with $f_z(y_i) = 0$ and $y_i \equiv \zeta^i \pmod{z}$.*

Proof. We have $f_z(X) \equiv X^n - 1 \pmod{z}$. For $i = 1, 2, \dots, n$ set $y_i^{[0]} = \zeta^i$. The usual proof of Hensel's Lemma (see e.g. [5, XII.7.6]) shows that the sequence $y_i^{[m]}$ $(m = 0, 1, \dots)$ defined by $y_i^{[m+1]} = y_i^{[m]} - f_z(y_i^{[m]})/f_z'(y_i^{[m]})$ converges to $y_i \in R[[z]]$, as $\frac{f_z(X)}{f_z'(X)} = \frac{X^n+\dots}{nX^{n-1}+\dots} \in \frac{1}{n}\mathcal{O}_K[[X]] \subseteq R[[X]]$. Also, $y_i \equiv \zeta^i \pmod{z}$ is a consequence of this proof. □

([1] has a direct proof of this Lemma that goes like this. Substitute a power series in z with unknown coefficients in the polynomial for X. Then, inductively compute the coefficients of this power series.)

Write $f_z(X) = \prod_{i=1}^n (X - y_i)$ according to Lemma 2. Replace X by zY to get

$$z^n f(Y) - 1 = \prod_{i=1}^{n} (zY - y_i).$$

Now let $\mathfrak{p}$ be a non-zero prime ideal of $\mathcal{O}_K$ which does not divide n, such that f is a permutation polynomial on $\mathcal{O}_K/\mathfrak{p}$. Let $F \subset \mathcal{O}_K$ be a set of representatives of $\mathcal{O}_K/\mathfrak{p}$, and set $q = |\mathcal{O}_K/\mathfrak{p}|$. Carry out the following calculations in the ring $R[[z]]$, modulo the ideals specified. We use the easy fact that for $U, V \in R[[z]]$

$$\prod_{a \in F} (Ua - V) \equiv VU^{q-1} - V^q \pmod{\mathfrak{p}}.$$

As f induces a permutation on $\mathcal{O}_K/\mathfrak{p}$, we get

$$\begin{aligned} \prod_{a \in F} (z^n f(a) - 1) &\equiv \prod_{a \in F} (z^n a - 1) \\ &\equiv z^{n(q-1)} - 1 \qquad \pmod{\mathfrak{p}}. \end{aligned}$$

On the other hand,

$$\begin{aligned}\prod_{a\in F}(z^n f(a)-1) &\equiv \prod_{a\in F}\prod_{i=1}^{n}(za-y_i)\\ &= \prod_{i=1}^{n}\prod_{a\in F}(za-y_i)\\ &\equiv \prod_{i=1}^{n}(y_i z^{q-1}-y_i^q) \qquad \pmod{\mathfrak{p}}.\end{aligned}$$

Hence

$$\prod_{i=1}^{n}(y_i z^{q-1}-y_i^q) \equiv z^{n(q-1)}-1 \pmod{\mathfrak{p}}.$$

Multiply by the unit $c = 1/\prod_{i=1}^{n} y_i^q$ to give

$$\prod_{i=1}^{n}(y_i \frac{z^{q-1}}{y_i^q}-1) \equiv c(z^{n(q-1)}-1) \pmod{\mathfrak{p}}. \tag{1}$$

From Lemma 2 we get

$$y_i^q \equiv \zeta^{iq} \pmod{z^q, \mathfrak{p}}.$$

Hence

$$\frac{z^{q-1}}{y_i^q} \equiv z^{q-1}\zeta^{-iq} \pmod{z^{2(q-1)}, \mathfrak{p}}.$$

Substitute this in (1) (using $n \geq 2$):

$$\prod_{i=1}^{n}(y_i\zeta^{-iq}z^{q-1}-1) \equiv -c \pmod{z^{2(q-1)}, \mathfrak{p}}.$$

This yields

$$\sum_{i=1}^{n} y_i\zeta^{-iq} \equiv 0 \pmod{z^{q-1}, \mathfrak{p}}.$$

Now use the assumption that f is a permutation on the field $\mathcal{O}_K/\mathfrak{p}$ of cardinality q for infinitely many non-zero prime ideals $\mathfrak{p}$ of $\mathcal{O}_K$. At least one residue of q modulo n, say r, thus appears infinitely often. So, using $\zeta^q = \zeta^r$ for these prime ideals (which furthermore should not divide n) shows that

$$\sum_{i=1}^{n} y_i \zeta^{-ir} \equiv 0 \pmod{z^{q-1}, \mathfrak{p}}$$

holds for infinitely many $\mathfrak{p}$ with q becoming arbitrarily large. For $k = 0, 1, \dots$ let b_k be the coefficient of the power series expansion of $\sum_{i=1}^{n} y^i \zeta^{-ir}$ with respect to z. We get that the the congruence $b_k \equiv 0 \pmod{\mathfrak{p}}$ holds in $R = \mathcal{O}_K[\zeta, \frac{1}{n}]$ for infinitely many prime ideals which do not divide n. Hence b_k vanishes. So $\sum_{i=1}^{n} y^i \zeta^{-ir} = 0$. Now set $t = 1/z^n$ and $x_i = y_i/z$. Then the x_i are the roots of $f(X) - t$, and the claim follows by replacing ζ^{-r} with ζ (note that r is prime to n).

The proof of Schur's conjecture

The Galois theoretic translation relies on

Corollary 3. *Let $f \in \mathcal{O}_K[X]$ be an exceptional polynomial of degree $n \geq 2$. Let $\overline{K}$ be an algebraic closure of K and t a transcendental over K. Then the Galois group G of $f(X) - t$ over $\overline{K}(t)$ does not act doubly transitively on the roots x_i of $f(X) - t$.*

Proof. We assume that G is doubly transitive, and aim for a contradiction. Let V be the subspace of $\overline{K}^n$ defined by $V = \{(u_1, u_2, \dots, u_n) | \sum u_i = 0\}$. Identify the digits $1, 2, \dots, n$ with $x_1, x_2, \dots, x_n$, and define an action of G on V by $(u_1, u_2, \dots, u_n)^\sigma = (u_{1^\sigma}, u_{2^\sigma}, \dots, u_{n^\sigma})$ for $\sigma \in G$. As G permutes the components of V doubly transitively, the module V is irreducible (see [3, Theorem 4.3.4]; the change from $\mathbb{C}$ to $\overline{K}$ is immediate). Now

$$W = \{(u_1, u_2, \dots, u_n) \in V | \sum u_i x_i = 0\}$$

is a G-invariant subspace of V, as $0 = \sum u_i x_i = (\sum u_i x_i)^\sigma = \sum u_i x_{i^\sigma} = \sum u_{i^{\sigma^{-1}}} x_i$. But $(\zeta^1, \zeta^2, \dots, \zeta^n)$ from Theorem 1 is contained in W, so $W = V$. But this gives $x_1 = x_2 = \cdots = x_n$, which of course is nonsense. □

We are still assuming that f is monic. To prove Schur's conjecture write f, which fulfills the hypothesis of Theorem 1, as a composition of indecomposable polynomials over K. These indecomposable constituents can be chosen with coefficients in $\mathcal{O}_K$ (see [10, 2.3]), they of course fulfill the hypothesis of Theorem 1. So in order to prove Schur's conjecture, one may assume from the beginning that f is indecomposable over K. Then f is also indecomposable over $\overline{K}$, see [1, Lemma 1], [10, 2.2(ii)]. As a consequence of Lüroth's Theorem, the Galois group G of $f(X) - t$ over $\overline{K}(t)$ is primitive, see [1, Lemma 2] or [10, 3.1]. On the other hand, G is not doubly transitive by Corollary 3. But G contains an n-cycle (where n is the degree of f), which by classical theorems of Schur and Burnside then forces n to be a prime and G to be solvable, normalizing the Sylow n-subgroup, see [1] or [10]. The original paper by Schur [8], [1], or [10] determine the shape of f from this. Or see the Appendix. Thus, the main tools for proving the Schur conjecture are the group theoretic theorems of Schur and Burnside for which [6] has short self-contained proofs.

Appendix.

Let F be a finite field of characteristic p. Suppose that $f \in F[X]$ is exceptional in the usual sense; f is a permutation polynomial on infinitely many finite extensions E of F. Suppose that the degree of f is not divisible by p. The method above allows us to draw the same conclusion as in Theorem 1. However, the representation theoretic part in the proof of Corollary 3 fails in general. The module V will be irreducible in general only under additional assumptions like $p > n$, or $n - 1$ is a power of p (see [7]). We do not see how to avoid the use of the Weil–bound in this situation. Anyway, the usual argument (see [2]) gives that we are reduced to the same configuration as in characteristic 0, namely that the Galois group of $f(X) - t$ over $\bar{F}(t)$ has prime degree and is solvable. However, the possibility of wild ramification requires different arguments to actually determine the polynomials. By the following theorem then f is, up to composition with linear polynomials over the algebraic closure of F, either cyclic or a Cebychev polynomial. From [10, 1.9(iii)] it then follows that f is, up to composition with linear polynomials over F, a Dickson polynomial over F.

In the following let K be an algebraically closed field of any characteristic p. We say that two polynomials a and b are *linearly related*, if $a(X) =$

$\lambda(b(\mu(X)))$ with linear polynomials $\lambda, \mu \in K[X]$.

Theorem 4. *Let $f \in K[X]$ be a polynomial of prime degree n, with p not dividing n (if $p > 0$). Suppose that the Galois group G of $f(X) - t$ over $K(t)$ is solvable. Then either $G = C_n$ (the cyclic group of degree n), and f is linearly related to X^n, or $G = D_n$ (the dihedral group of degree n), and f is linearly related to the Cebychev polynomial T_n, which is uniquely defined by $T_n(z + \frac{1}{z}) = z^n + \frac{1}{z^n}$.*

Proof. Let $x = x_1, x_2, \ldots, x_n$ be the solutions of $f(X) = t$, and let $L = K(x_1, x_2, \ldots, x_n)$ be the splitting field of $f(X) - t$ over $K(t)$. So G is the Galois group of $L|K(t)$. We view G as permutation group on the set $\{x_1, \ldots, x_n\}$, and denote by H_j the stabilizer of x_j. As n is prime and G is solvable, we get that $G = N \rtimes H$, where N is transitive of order n, and H is any of the H_j. Set $|H| = d$. So $[L : K(t)] = nd$ and $[L : K(x_j)] = d$.

We use notions and results in valuation theory from [9] to first show that L is a rational field. A different proof for that is contained in [4], and yet another proof has been communicated to the author by M. Fried.

Let $\mathbb{P}$ be the set of places of L which are ramified over $K(t)$. For a subfield E of L let P_E be the restriction of P to E. Denote by I_P the inertia group of a place $P \in \mathbb{P}$ over $P_{K(t)}$. Let g be the genus of L, and $d(P|P_{K(t)})$ (or $d(P|P_{K(x_j)})$) the different exponent of P over $P_{K(t)}$ (or over $P_{K(x_j)}$). The Riemann–Hurwitz genus formula [9, III.4.12] for the extensions $L|K(t)$ and $L|K(x_j)$ gives

$$2(nd - 1 + g) = \sum_{P \in \mathbb{P}} d(P|P_{K(t)}) \tag{2}$$

and

$$2(d - 1 + g) = \sum_{P \in \mathbb{P}} d(P|P_{K(x_j)}) \tag{3}$$

respectively. Subtract (3) for $j = 1, 2, \ldots, n$ from (2):

$$2(n-1)(1-g) = \sum_{P \in \mathbb{P}} (d(P|P_{K(t)}) - \sum_{j=1}^{n} d(P|P_{K(x_j)})). \tag{4}$$

We now compute the contributions in (4) from the places P.

The inertia group of P over $P_{K(x_j)}$ of course is $H_j \cap I_P$.

First suppose that I_P is intransitive. Then there is exactly one index j_0 such that $I_P \leq H_j$. (Note that I_P being intransitive implies that $I_P \cap H = 1$, so I_P maps injectively into the cyclic group HN/N; therefore I_P is generated by a non-zero element which fixes exactly one letter.) If $j = j_0$, then $P_{K(x_j)}|P_{K(t)}$ is unramified, and therefore $d(P|P_{K(x_j)}) = d(P|P_{K(t)})$ (e.g. by [9, III.4.11(b)]). If however $j \neq j_0$, then $I_P \cap H_j = 1$, so $P|P_{K(x_j)}$ is unramified, hence $d(P|P_{K(x_j)}) = 0$. Thus we have no contribution for these places in (4).

Now suppose that I_P is transitive. Then $P_{K(x_j)}|P_{K(t)}$ is totally ramified. As $p \neq n$, the group I_P does not contain a normal Sylow p-subgroup. This implies that the inertia group I_P is cyclic (see [9, III.8.6(e)]), hence $I_P = N$. In particular $I_P \cap H_j = 1$, hence $P|P_{K(x_j)}$ is unramified. Thus (see [9, III.4.11(b), III.5.1(b)]) $d(P|P_{K(x_j)}) = 0$ and (as the ramification is tame) $d(P|P_{K(t)}) = |I_P| - 1 = n - 1$. The places $Q \in \mathbb{P}$ with $Q_{K(t)} = P_{K(t)}$ are conjugate under the action of G, so their number is $[G : I_P] = [G : N] = |H| = d$. Hence these places Q together contribute $d(n-1)$ to (4). We get

$$2(n-1)(1-g) = ld(n-1),$$

where l is the number of places of $K(t)$ which are totally ramified in $K(x)$. (Note that I_P being transitive is equivalent to $P_{K(t)}$ being totally ramified in $K(x)$.)

We get $1 - g = ld/2$. However, $l \geq 1$, as f is a polynomial and so the infinite place of $K(t)$ is totally ramified in $K(x)$. This gives $g = 0$ and $ld = 2$. If $l = 2$, then there is a totally ramified finite place $t \mapsto a$ of $K(t)$. So $f(X) - a$ is the n–th power of a linear polynomial, and the assertion follows in this case.

For the remainder we assume $d = 2$. As L has genus 0, we have $L = K(z)$ for some z. The automorphism group of $K(z)$ is $\mathrm{PGL}_2(K)$. Denote the image of $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}_2(K)$ in $\mathrm{PGL}_2(K)$ by $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$. The action on $K(z)$ is given by sending z to $\frac{az+b}{cz+d}$. It is easy to see that our dihedral group G of order $2n$ is conjugate in $\mathrm{PGL}_2(K)$ to the group generated by $\sigma = \left[\begin{smallmatrix} 1 & 0 \\ 0 & \zeta \end{smallmatrix}\right]$ and $\tau = \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$, where ζ is a primitive n–th root of unity. (To see this, first note that the Sylow n-subgroup of G is diagonalizable. Then compute its normalizer in $\mathrm{PGL}_2(K)$.) The fixed field in $K(z)$ of $<\sigma, \tau>$ is obviously $K(z^n + \frac{1}{z^n})$. So replacing z by a linear fractional change, we may assume that $K(t) = K(z^n + \frac{1}{z^n})$. Further, as the involutions in G are conjugate, we may further assume that the fixed field of τ, namely $K(z+\frac{1}{z})$, equals $K(x)$. Therefore $\lambda(z+\frac{1}{z})=x$ and $z^n+\frac{1}{z^n}=\mu(t)=\mu(f(x))$ for linear fractional functions λ and μ. So

$$\mu(f(\lambda(X)))=T_n(X).$$

It remains to show that λ and μ are indeed polynomials. First suppose that μ is not a polynomial. Then, using the usual rules to compute with ∞, we have $\mu(\infty)=\omega\in K$. As T_n is not linearly related to X^n, there are distinct $\omega_1,\omega_2\in K$ with $T_n(\omega_1)=T_n(\omega_2)=\omega$. For $i=1,2$ we get

$$f(\lambda(\omega_i))=\mu^{-1}(T_n(\omega_i))=\mu^{-1}(\omega)=\infty.$$

Thus, as f is a polynomial,

$$\lambda(\omega_1)=\lambda(\omega_2)=\infty,$$

contrary to $\omega_1\neq\omega_2$. So μ is a polynomial. If λ were not a polynomial, then $\lambda(\infty)\neq\infty$, and setting $X=\infty$ yields a contradiction. □

Remark. The fact that p does not divide the degree n of f does not allow us to argue as in characteristic 0. Indeed, the branch points of the polynomials T_n ($n\geq 3$) are 2 and -2. So in characteristic 2, these two branch points collapse, and give wild ramification over 0.

References

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

[2] M. Fried, R. Guralnick, J. Saxl, Schur covers and Carlitz's conjecture, *Israel J. Math.* **82** (1993), 157–225.

[3] D. Gorenstein, "Finite Groups," Harper and Row, New York–Evanston–London, 1968.

[4] R. Guralnick, D. Wan, Finite covers of curves over finite fields and fixed point free elements in a permutation group, *to appear in Israel J. Math.*.

[5] S. Lang, "Algebra," Addison–Wesley, Menlo Park, 1984.

[6] R. Lidl, G. L. Mullen, G. Turnwald, "Dickson Polynomials," Pitman Monographs and Surveys in Pure and Applied Mathematics **65**, Longman, Essex, 1993.

[7] B. Mortimer, The modular permutation representations of the known doubly transitive groups, *Proc. London Math. Soc. (3)* **41** (1980), 1–20.

[8] I. Schur, Über den Zusammenhang zwischen einem Problem der Zahlentheorie und einem Satz über algebraische Funktionen, *S.–B. Preuss. Akad. Wiss., Phys.–Math. Klasse* (1923), 123–134.

[9] H. Stichtenoth, “Algebraic Function Fields and Codes,” Springer, Berlin Heidelberg, 1993.

[10] G. Turnwald, On Schur’s conjecture, *J. Austr. Math. Soc. (Series A)* **58** (1995), 312–357.