The complex polynomials $P(x)$ with $\mathrm{Gal}(P(x) - t) \cong M_{23}$

Noam D. Elkies

Department of Mathematics, Harvard University, Cambridge, MA 02138
elkies@math.harvard.edu
Supported in part by NSF grants DMS-0501029 and DMS-1100511

1 Introduction

For $P \in \mathbf{C}[x]$ of degree $n > 0$, define G_P to be the Galois group of $P(x) - t$ over $\mathbf{C}(t)$. Since $P(x) - t$ is irreducible, G_P is a transitive subgroup of the symmetric group S_n. Generically[1] G_P is all of S_n, but it can be as small as the cyclic or dihedral group for special choices such as $P = x^n$ or $P = T_n(x)$ (Čebyšev polynomial) respectively. If P decomposes as $P(x) = P_1(P_2(x))$ with each $\deg(P_i) > 1$, then G_P permutes the proper subsets $\{x : P_2(x) = u\}$ of the roots with $P_1(u) = t$, and is therefore imprimitive. The converse implication is shown in [FMcR, Prop. 3.4]. Müller [Mü] determined all G_P that can arise for indecomposable polynomials: they are the symmetric and alternating groups, the cyclic groups of prime order, the dihedral groups of order twice an odd prime, and twelve exceptional permutation groups with $n = 6, 7, \ldots, 23, 31$, the last two for the sporadic Mathieu group M_{23} and the linear group $\mathrm{GL}_5(\mathbf{Z}/2\mathbf{Z})$.

The proof uses covering-space methods and Riemann's existence theorem, and thus does not yield explicit polynomials. But it is still a natural question to exhibit all P that realize each possible group G_P, except for for the cases of A_n and S_n, which occur in "many, not reasonably classifiable types" [Mü]. Say $P, Q \in \mathbf{C}[x]$ are *equivalent* if $Q(x) = L_1(P(L_2(x)))$ for some polynomials L_1, L_2 both of degree 1; then $G_P = G_Q$. Up to this equivalence, the cyclic and dihedral groups occur only for powers and Čebyšev polynomials respectively. Some of the exceptional groups were realized in [Mü], or earlier by Matzat [Ma]; most of the others were realized by Cassou-Noguès and Couveignes [CNC],[2] leaving only M_{23}. Here we find the polynomials P with $G_P \cong M_{23}$.

The main novelty here is not in the computation of P but in the proof that $G_P \cong M_{23}$. The coefficients of P were computed using a known p-adic method

[1] In particular, $G_P = S_n$ if dP/dx has $n - 1$ distinct roots at which P takes distinct values; equivalently, if $\mathrm{disc}_t(\mathrm{disc}_x(P(x) - t)) \neq 0$. This sufficient (but far from necessary) condition was already noted by Hilbert ([Hi], see also [Se, §4.4]); the formulation in terms of the discriminant of the discriminant is attributed to Davenport in [BSD, p.422].

[2] Michael Zieve had already obtained but not published polynomials for a few of these cases, with groups $\mathrm{PGL}_2(\mathbf{Z}/7\mathbf{Z})$ $(n = 8)$, $\mathrm{P\Gamma L}_2(\mathbf{F}_8)$ $(n = 9$, both classes), and M_{11} $(n = 11)$; he also calculated that there are four M_{23} polynomials up to equivalence, but was not able to exhibit such a polynomial.

for finding polynomial identities by solving the equivalent system of nonlinear equations in the coefficients, though here the search for the initial approximation took several CPU days. The difficulty was that these equations cannot distinguish between polynomials with Galois group M_{23} and A_{23}, and there are four M_{23} covers but numerous A_{23} covers with the same cycle structure (with all the A_{23} covers probably defined only over number fields of rather high degree). Once we found P with coefficients in a quartic number field F, we quickly convinced ourselves that G_P must be M_{23} by factoring $P(x) - t_0 \bmod \lambda$ for many primes λ of F and choices of $t_0 \bmod \lambda$ at which $P(x) - t_0$ has distinct roots: in each case the degrees of the factors matched one of the 12 cycle structures of elements of M_{23}, out of the 632 that arise in A_{23}. Moreover, the fraction of t_0 values that yield each cycle structure was quite near to the fraction of elements of M_{23} with that cycle structure, as promised by the Čebotarev density theorem for Galois extensions of function fields. Still this did not amount to a proof that $G_P \cong M_{23}$.

However, if G_P were actually A_{23} then we would observe a very different distribution of cycle structures, which would contradict the Čebotarev theorem once the residue field of λ got large enough. In our function-field setting such a calculation turns out to be feasible thanks to Weil's proof of the Riemann hypothesis for curves over finite fields. We did this for a λ whose residue field is prime of characteristic $l = 10^8 + 7$ (the smallest 9-digit prime, which happens to lie under a degree-1 prime of F). We showed that the resulting distribution of cycle structures implies that G_P is not 5-transitive, which soon yields $G_P \cong M_{23}$ as desired.

The factorization of 10^8 polynomials mod λ was a somewhat extravagant computation (two days of CPU time in `gp` [Pa]). This is not the only way to prove $G_P \cong M_{23}$; for example, one could do it also by numerically lifting monodromy generators to permutations of 23 preimages, as Granboulan did for the 24 preimages of an M_{24} cover [Gr]. Still our technique using Čebotarev plus Weil has some advantages over the monodromy computation: while our computation took rather long to run, it was very easy to code, whereas the monodromy calculation would require some careful estimates to guarantee that the precision was sufficient to obtain the correct permutations; and our technique works also for Galois groups of extensions in positive characteristic. This approach also raises the theoretical question of how large a residue field is necessary: perhaps it can be shown that the counts over a field of size much smaller than 10^8 would have sufficed.

In the next section we exhibit F and $P \in F[x]$ and give some details on its calculation. In the following section we report on the results of our computation mod λ, use them to prove that $G_P \not\cong A_{23}$, and deduce that a polynomial P_1 satisfies $G_{P_1} \cong M_{23}$ if and only if P_1 is equivalent to the image of our P under one of the four embeddings of F into $\mathbf{C}$.

2 Computation of P

Suppose $G_P \cong M_{23}$. By [Mü], the map $P : \mathbf{P}^1 \to \mathbf{P}^1$ is branched above only three points, with orders 23 (at $t = \infty$), 2, and 4. The group M_{23} contains only one conjugacy class of order 2 and one of order 4. The corresponding monodromy generators γ_2 and γ_4 must have $\gamma_2\gamma_4$ of order 23. Up to conjugation in M_{23}, there are four such pairs (γ_2, γ_4), two for each of the two conjugacy classes of elements of order 23 in M_{23}, and in each case γ_2 and γ_4 generate M_{23}. Since M_{23} is its own normalizer in S_{23}, we conclude that there are four equivalence classes of M_{23} polynomials, each defined over a number field F containing $\mathbf{Q}(\sqrt{-23})$ with degree 1 or 2. We eventually found that F is the dihedral quartic field of discriminant $3 \cdot 23^3$ generated by a root of $g^4 + g^3 + 9g^2 - 10g + 8$, which indeed contains the square roots $\pm(2g^3 + 4g^2 + 16g - 7)/3$ of -23.

The permutations γ_2 and γ_4 of 23 objects have cycle structures $1^7 2^8$ and $1^3 2^2 4^4$. Thus P is equivalent to a monic polynomial with two double and four quadruple roots. Then, if τ is the value of P at its finite critical points other than zeros, we can write

$$P = P_2^2 P_3 P_4^4 = P_7 P_8^2 + \tau, \tag{1}$$

where the P_i $(i = 2, 3, 4, 7, 8)$ are pairwise coprime monic polynomials of degree i, and τ is a nonzero constant. It may seem that we have 10 coefficients to determine: the 2+3+4 non-leading coefficients of P_2, P_3, P_4, together with τ. We can reduce this to 8 variables using the remaining equivalences (translate x, and multiply x by some nonzero μ and divide each P_i by μ^i). One further variable is eliminated using a familiar[3] differentiation trick: dP/dx has leading term $23x^{22}$ and is a multiple of $P_2 P_4^3 P_8$, so must equal $23 P_2 P_4^3 P_8$; hence

$$P_8 = \frac{1}{23} \frac{dP/dx}{P_2 P_4^3} = \frac{1}{23} \left(2P_2' P_3 P_4 + P_2 P_3' P_4 + 4 P_2 P_3 P_4' \right). \tag{2}$$

Still the remaining nonlinear equations are too complicated to solve directly by techniques such as Gröbner bases, especially since they do not distinguish between M_{23} and A_{23} covers.

Instead we use the following strategy. Suppose the solution is defined over a number field F with a prime π of small residue field at which the cover $P : \mathbf{P}^1 \to \mathbf{P}^1$ has good reduction. We can then find our cover mod π by exhaustive search. An arbitrary lift to the π-adic numbers is then an approximate solution, which can be improved by a multivariate Newton iteration. Once we have the solution to high enough π-adic precision, we can recognize it as an F-rational point by lattice reduction, and verify that it satisfies the equations exactly.

For a general system of nonlinear equations we could not know in advance which π satisfy the condition of good reduction. In our setting, we are seeking a "Belyi map" (a cover of $\mathbf{P}^1$ ramified only above three points), so Beckmann's

[3] The earliest published references I know of are [El,Bi], but the trick must have been known and used long before that.

theorem [Be] gives a sufficient condition: if the characteristic of the residue field of π does not divide the order of the Galois group then the cover has good reduction at π. But we do not know F in advance, and thus do not know which residue fields arise. We therefore tried small prime fields $\mathbf{Z}/p\mathbf{Z}$ in the hope that one would work. But searches over $(\mathbf{Z}/p\mathbf{Z})^7$ became ever longer without finding the desired cover. For example, a search mod 13 (the smallest prime not dividing $|M_{23}|$) found only

$$P_2 = x^2 - 3x - 6, \quad P_3 = x^4 - 4x - 4, \quad P_4 = x^4 + 5x^2 - 5x - 1 \tag{3}$$

with $\tau = 5$; but the resulting $P = P_2^2 P_3 P_4^4$ cannot have Galois group M_{23} because there are $t_0 \neq 0, 5$ for which the factorization of $P - t_0$ mod 13 has degrees not seen in any of the M_{23} cycle structures — for instance, $P - 1$ has an irreducible factor of degree 19. In retrospect we know there is no M_{23} polynomial over $\mathbf{Z}/13\mathbf{Z}$, because F has no prime of degree 1 above 13 (even though 13 does split in the quadratic subfield $\mathbf{Q}(\sqrt{-23})$).

To bring larger p within reach, we applied the following refinement. For $j \geq 0$ and any $Q \in \mathbf{C}[x]$, denote by $c_j(Q)$ the x^j coefficient of Q; for example $c_i(P_i) = 1$ for each $i = 2, 3, 4, 7, 8$. For any monic P_2, P_3, P_4, let R be the remainder when P_{23} is divided by P_8^2, where P_{23} and P_8 are defined by (1) and (2). Then R has degree $\deg P_8^2 - 1 = 15$ generically, but must vanish at the desired solution. We noticed that if we hold all but $c_0(P_4)$ and $c_1(P_4)$ fixed then $c_{15}(R)$ and $c_{14}(R)$ are polynomials of degree only 2 in $c_0(P_4)$ and of degree 3 in $c_1(P_4)$; in fact, $c_{15}(R)$ and $c_{14}(R)$ have degrees 2 and 3 respectively in $(c_0(P_4), c_1(P_4))$ together. We could have solved the simultaneous equations $c_{15}(R) = c_{14}(R) = 0$ in $(c_0(P_4), c_1(P_4))$, reducing the search from $O(p^7)$ to $O(p^5)$ but with quite a large O-constant. Instead we opted for the following strategy, which is still $O(p^7)$ but with a much smaller constant. Having fixed all but $c_0(P_4)$ and $c_1(P_4)$, compute R at the 12 sample points with $c_0(P_4) = 0, 1, 2$ and $c_1(P_4) = 0, 1, 2, 3$, and then use the fact that both $c_{15}(R)$ and $c_{14}(R)$ are quadratic in $c_0(P_4)$ and cubic in $c_1(P_4)$ to recursively evaluate them at all other choices of $c_0(P_4)$ and $c_1(P_4)$. If both vanish, test whether $\deg(R) = 0$. This way, instead of computing p^2 polynomial remainders we need on average only 13: twelve sample points, and one more for the expected number of solutions of $c_{15}(R) = c_{14}(R) = 0$.

We implemented this search in `gp` (which we used also for the earlier $O(p^7)$ method), and finally succeeded at $p = 29$. We assumed that $c_2(P_3) = 0$, and that $c_0(P_3) = c_1(P_3)$ if both $c_0(P_3)$ and $c_0(P_1)$ are nonzero; every choice of P_2, P_3, P_4 with $c_0(P_3)c_1(P_3) \neq 0$ is equivalent to exactly one satisfying these conditions. (One can also make a unique choice if $c_0(P_3) = 0$ or $c_1(P_3) = 0$, but here this was not necessary.) The search took 46 CPU hours, compressed to less than five hours by running on 10 heads in parallel, which is an order of magnitude smaller than the time to compute some 29^7 polynomial remainders. The resulting list of solutions contained two for which every $P(x) - t_0$ has a factorization consistent with $G_P \cong M_{23}$. One of these was

$$P_2 = x^2 - x - 3, \quad P_3 = x^3 - 3x - 3, \quad P_4 = x^4 - 3x^3 - 11x^2 + 13x + 7 \tag{4}$$

with $\tau = 5$. Lifting to $\mathbf{Z}/p^{128}\mathbf{Z}$ (while retaining the conditions $c_2(P_3) = 0$ and $c_0(P_3) = c_1(P_3)$) gave more than enough precision to identify all the coefficients as elements of the quartic field $F = \mathbf{Q}[g]/(g^4 + g^3 + 9g^2 - 10g + 8)$.

These elements of F are quite complicated because of the normalization $c_0(P_3) = c_1(P_3)$. Once we have found one choice of $P_2, P_3, P_4 \in F[x]$ that works, we can find equivalent but simpler ones by removing this normalization and the spurious bad reduction that it entails. One reasonably simple choice we found (dropping also the condition that the P_i be monic) is as follows:

$$\begin{aligned} P_2 &= (8g^3 + 16g^2 - 20g + 20)x^2 - (7g^3 + 17g^2 - 7g + 76)x \\ &\quad - 13g^3 + 25g^2 - 107g + 596; \end{aligned} \tag{5}$$

$$\begin{aligned} P_3 &= 8(31g^3 + 405g^2 - 459g + 333)x^3 + (941g^3 + 1303g^2 - 1853g + 1772)x \\ &\quad + 85g^3 - 385g^2 + 395g - 220; \end{aligned} \tag{6}$$

$$\begin{aligned} P_4 &= 32(4g^3 - 69g^2 + 74g - 49)x^4 + 32(21g^3 + 53g^2 - 68g + 58)x^3 \\ &\quad - 8(97g^3 + 95g^2 - 145g + 148)x^2 + 8(41g^3 - 89g^2 - g + 140)x \\ &\quad - 123g^3 + 391g^2 - 93g + 3228. \end{aligned} \tag{7}$$

With this choice,

$$\tau = \frac{2^{38}3^{17}}{23^3}(47323g^3 - 1084897g^2 + 7751g - 711002), \tag{8}$$

the last factor having norm $2^{27}3^{23}5^{10}$.

3 Proof of $\mathrm{Gal}(P(x) - t) \cong M_{23}$

We chose the degree-1 prime λ of F above the rational prime $l = 10^8 + 7$ at which $g \equiv 36436770 \bmod l$. We reduced $P \bmod \lambda$ to obtain a polynomial $\overline{P}$ with coefficients in $F_\lambda = \mathbf{Z}/l\mathbf{Z}$, and factored $\overline{P} - t_0$ for each of the $l - 2$ values of $t_0 \bmod l$ for which $\overline{P} - t_0$ has no repeated roots. In each case the degrees of the irreducible factors, and thus the cycle structure of the action of Frobenius at $t = t_0$, agreed with the cycle structure of one or two of the conjugacy classes of M_{23}. Table 1 lists, for each class or pair of classes $c \subset M_{23}$: its ATLAS label [C&, p.71], the cycle structure, the fraction $|c|/|M_{23}|$, the integer nearest to $(|c|/|M_{23}|)(l - 2)$ (which is the expected number of occurrences of this cycle structure), the actual number of times it appeared, and the difference between the actual and expected counts.

$1A$	$2A$	$3A$	$4A$	$5A$	$6A$	$7A, 7B$
1^{23}	$1^7 2^8$	$1^5 3^6$	$1^3 2^2 4^4$	$1^3 5^4$	$1\ 2^2 3^2 6^2$	$1^2 7^3$
$1/\lvert M_{23}\rvert$	1/2688	1/180	1/32	1/15	1/12	2/14
10	37202	555556	3125000	6666667	8333334	14285715
9	37235	556547	3123317	6665816	8329354	14290600
−1	33	991	−1683	−851	−3980	4885

Table 1

$8A$	$11A, 11B$	$14A, 14B$	$15A, 15B$	$23A, 23B$
$1\,2\,4\,8^2$	$1\,11^2$	$2\,7\,14$	$3\,5\,15$	23
$1/8$	$2/11$	$2/14$	$2/15$	$2/23$
12500001	18181819	14285715	13333334	8695653
12493007	18185450	14289505	13331689	8697476
−6994	3631	3790	−1645	1823

Table 1, continued

The agreement is quite close: the discrepancy never exceeds twice the square root of the expected value.

In particular, because each of the M_{23} cycle structures occurs (and $G_P \subseteq A_{23}$ because $\operatorname{disc}_x(P(x) - t)$ is a square) we know that G_P is a transitive subgroup of A_{23} containing elements of order p for each of the prime factors $p = 2, 3, 5, 7, 11, 23$ of $|M_{23}| = 2^7\, 3^2\, 5 \cdot 7 \cdot 11 \cdot 23 = 10200960$. This shows that G_P is either M_{23} or A_{23}.

One could try various strategies for deducing $G_P \not\cong A_{23}$ from the counts in Table 1. The following approach was the one that worked most easily. We shall take C_0 and C_1 to be the projective t- and x-lines in the following general set-up.

Suppose C_1/C_0 is a degree-n covering of curves over some finite field F_λ. Let $\widetilde{C}$ be the Galois closure, with Galois group $G \subseteq S_n$. Assume that G is k-transitive. Let G_k be the stabilizer of a k-element *set*, so the action of G_k on that set gives a surjective homomorphism $G_k \to S_k$ whose kernel is the k-point stabilizer; write $C_k = \widetilde{C}/G_k$, so C_k/C_0 is a cover of degree $\binom{n}{k}$. If the cover C_1/C_0 is given by a polynomial Q of degree n, then with finitely many exceptions a point of C_k corresponds to a degree-k factor of a specialization of Q.

Let N_k be the number of F_λ-rational points of C_k. For an unramified F_λ-rational point t_0 on C_0, let $N_k(t_0)$ be the number of F_λ-rational points of C_k lying over t_0. We next express $N_k(t_0)$ in terms of the Galois structure of the preimage of t_0 in C_1. Let ϕ be the Frobenius permutation of the preimage of t_0 in C_1.

Lemma. *Let $c_1, c_2, \ldots, c_m$ (with $\sum_{i=1}^m c_m = n$) be the cycle lengths of ϕ. Then $N_k(t_0)$ is the X^k coefficient of the polynomial $\prod_{i=1}^m (1 + X^{c_i})$.*

Proof: A k-element subset of the preimage of t_0 yields a rational point of C_k if and only if it is taken to itself by ϕ; equivalently, if and only if it is the union of orbits of ϕ. Since these orbits have sizes c_i, the expansion of $\prod_{i=1}^m (1+X^{c_i})$ yields a sum of 2^m monomials, with each monomial X^k corresponding to a k-element subset, Q.E.D.

We now take C_0 and C_1 to be the t- and x-lines. Then $G = G_P$ by Beckmann's criterion [Be] (since l is too large to be a factor of $|G|$ even if $G = A_{23}$). Using the entries in Table 1, we find for each $k = 1, 2, \ldots, 22$ the sum of $\prod_{i=1}^m (1+X^{c_i})$ over the $l - 2$ unramified points t_0. The sum is invariant under $k \leftrightarrow n - k$, so we need only tabulate up to $k = 11$. In each case we write $\sum_{t_0} N_k(t_0) = Al - B$

with $A \in \mathbf{Z}$ minimizing $|B|$:

k	1	2	3	4	5	6
A	1	1	1	1	2	3
B	10	6592	19784	2326	10892	60120

k	7	8	9	10	11
A	4	5	5	5	7
B	109978	243430	487620	742744	883854

Table 2

In each case $Al - B$ is a lower bound for N_k, with the difference coming from the counts above the three ramified points. If G acts k-transitively then C_k is an irreducible curve, and then the Weil bound gives $|N_k - (l+1)| \leq 2l^{1/2} g(C_k)$. Table 2 suggests that this might happen for $k \leq 4$ but not for $k = 5$ (and indeed C_5 has two components, one for each of the orbits of the action of M_{23} on 5-element subsets). We next prove that G is not 5-transitive by bounding $g(C_5)$. If $G_{\overline{P}} = A_{23}$ then C_k has genus at most

$$1 + \frac{1}{2}\left(1 - \frac{1}{2} - \frac{1}{4} - \frac{1}{23}\right)[C_k : C_0] = 1 + \frac{1}{2}\frac{19}{92}\binom{23}{k}$$

by the Riemann-Hurwitz formula. For $k = 5$ this gives $27805/8$, so $g(C_5) < 3476$. Therefore

$$|N_5 - (l+1)| < 2l^{1/2} \cdot 3476 < 7 \cdot 10^7. \tag{9}$$

But the $k = 5$ column of Table 2 gives

$$N_5 - (l+1) \geq l - 10893 > 9 \cdot 10^7, \tag{10}$$

even without including the preimages of the ramified points. The conflict between (9) and (10) refutes the hypothesis that $G_P = A_{23}$ and completes the proof that $G_P \cong M_{23}$.

Acknowledgements

I thank Michael Zieve for telling me of the remaining M_{23} case in Müller's list, and of the earlier work of himself and of Cassou-Noguès and Couveignes on the other groups. I also thank: the referee, for a careful reading resulting in several local corrections and improvements; John Voight, for the reference [Bi]; and Mark Watkins, as well as Voight and Zieve, for helpful correspondence on this and similar problems.

References

[Be] Beckmann, S.: Ramified primes in the field of moduli of branched coverings of curves, *J. Algebra* **125** (1989) #1, 236–255 (dx.doi.org/10.1016/0021-8693(89)90303-7).

[Bi] Birch, B.J.: Noncongruence subgroups, covers, and drawings, pp.25–46 in *The Grothendieck Theory of Dessins d'Enfants* (L.Schneps, ed.), Cambridge: , Cambridge Univ. Press, 1994.

[BSD] Birch, B.J., and Swinnerton-Dyer, H.P.F.: Note on a problem of Chowla. *Acta Arith.* **5** (1959), 418–423.

[CNC] Cassou-Noguès, P., and Couveignes, J.-M.: Factorisations explicites de $g(y) - h(z)$, *Acta Arith.* **87** (1999) #4, 291–317 (http://www.math.u-bordeaux1.fr/~couveign/publi/CaCou.pdf).

[C&] Conway, J.H., et al.: *ATLAS of Finite Groups: Maximal Subgroups and Ordinary Characters for Simple Groups.* Oxford: Clarendon Press, 1985.

[El] Elkies, N.D.: ABC implies Mordell, *International Math. Research Notices* **1991** #7, 99–109 [bound with *Duke Math. J.* **64** (1991)].

[FMcR] Fried, M.D., and MacRae, R.E.: On the invariance of chains of fields, *Illionis J. Math.* **13** (1969) #1, 165–171.

[Gr] Granboulan, L.: Construction d'une extension régulière de $\mathbf{Q}(T)$ de groupe de Galois M_{24}, *Experimental Math.* **5** (1996), 3–14.

[Hi] Hilbert, D.: Ueber die Irreducibilität ganzen rationaler Functionen mit ganzzahligen Coefficienten, *J. reine angew. Math.* **110**, 104–129 (= *Ges. Abh. II*, 264–286).

[Ma] Matzat, H.B.: Konstruktion von Zahl- und Funktionenkörpern mit vorgegebener Galoisgruppe, *J. reine angew. Math.* **349** (1984), 179–220.

[Mü] Müller, P.: Primitive monodromy group of polynomials, *Contemp. Math.* **186** (1995), 385–401 (http://mathematik.uni-wuerzburg.de/~mueller/Papers/mon.dvi).

[Pa] PARI/GP, version `2.4.3`, Bordeaux, 2011, `http://pari.math.u-bordeaux.fr/`.

[Se] Serre, J.-P.: *Topics in Galois Theory.* Boston: Jones & Bartlett 1992.