PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 122, Number 4, December 1994

SOME MORE MATHIEU GROUP COVERINGS IN CHARACTERISTIC TWO

SHREERAM S. ABHYANKAR AND IKKWON YIE

(Communicated by Ronald M. Solomon)

ABSTRACT. Explicit equations are given for unramified coverings of the affine line in characteristic two with Mathieu groups of degrees 23 and 24 as Galois groups.

1. INTRODUCTION

Let k be a field of characteristic $p \neq 0$, and consider the polynomial $\overline{F}_{23,20,1,1} = Y^{23} + XY^3 + 1$ of degree 23 in Y with coefficients in $k[X]$. Inspired by Serre's *linearization trick* (cf. [6] as reported in Section 1 of [5]), in the case of $p = 2$, in (1.5) of [4] a linearization lemma was proved for this polynomial and, together with the transitivity lemma (1.3) of [4], it showed that the said polynomial gives an unramified covering of the affine line L_k (in characteristic two) having M_{23} (= the Mathieu group of degree 23) as Galois group. In the present paper, by modifying this procedure, we shall prove the following:

First Mathieu Group Theorem (1.1). *If $p = 2$ then, for any $\alpha \in k$, the Galois group $\mathrm{Gal}(Y^{24} + \alpha Y^4 + Y + X, k(X))$ equals the Mathieu group M_{24} of degree 24.*

From (1.1) it follows that, for $p = 2$, the equation $Y^{24} + \alpha Y^4 + Y + X = 0$ gives an unramified covering of the affine line L_k with Galois group M_{24}. It may be noted that this covering is a special case of the family of unramified coverings given in Proposition 2 of the 1957 paper [1]. Moreover, the subcase $\alpha = 0$ is part of the tilde family on pp. 74 and 103–108 of [2] and in (9.5) of [3] it was called a *border value* case giving interesting Galois group. In the subcase $\alpha = 0$, recently McKay and Conway have independently shown the Galois group to be M_{24}.

By "throwing away" a root of the above equation [see (2.1)], by (1.1) we get the following:

Received by the editors March 30,1993.

1991 *Mathematics Subject Classification.* Primary 12F10, 14H30, 20D06, 20E22.

The first author's work was partly supported by NSF grant DMS 91-01424 and NSA grant MDA 904–92–H–3035. The second author's work was partly supported by PRF grant 690-1395-1920.

Second Mathieu Group Theorem (1.2). *If $p = 2$ and α is any element of k then upon letting $\Phi = Y^{-1}[(Y+X)^{24} - X^{24}] + \alpha Y^{-1}[(Y+X)^4 - X^4] + 1 = Y^{23} + X^8Y^{15} + X^{16}Y^7 + \alpha Y^3 + 1$, the equation $\Phi = 0$ gives an unramified covering of the affine line L_k with* $\mathrm{Gal}(\Phi, k(X)) = M_{23}$.

To explain the linearization trick, let us recall that an *additive polynomial* Θ over a domain D of characteristic p is a polynomial of the form $\Theta = Y^{p^M} + \sum_{i=0}^{M-1} a_i Y^{p^i}$ where $a_i \in D$ with $a_0 \neq 0$. Clearly the Y-derivative of Θ equals the nonzero constant a_0 and hence, for any overfield Λ of D, the Galois group $\mathrm{Gal}(\Theta, \Lambda)$ is defined. Moreover, the roots of Θ obviously form an elementary abelian group of order p^M and hence, as a permutation group, $\mathrm{Gal}(\Theta, \Lambda)$ is a subgroup of $\mathrm{GL}(M, p)$. Now let there be given a monic polynomial Γ of degree $N > 0$ in Y with coefficients in D. Assume that the Y-discriminant of Γ is nonzero so that we can talk about the Galois group $\mathrm{Gal}(\Gamma, \Lambda)$. We shall say that Γ *linearizes over D at M* if there exists an additive polynomial Θ over D of degree p^M such that $\Gamma\Gamma^* = \Theta$ for some $\Gamma^* \in D[Y]$. If this is so, then $\mathrm{Gal}(\Gamma, \Lambda)$ is a homomorphic image of $\mathrm{Gal}(\Theta, \Lambda)$ and hence the order $|\mathrm{Gal}(\Gamma, \Lambda)|$ of $\mathrm{Gal}(\Gamma, \Lambda)$ divides the order $|\mathrm{GL}(M, p)|$ of $\mathrm{GL}(M, p)$. It is easily seen that Γ always linearizes over Λ at $M = N$. But if it linearizes at a significantly smaller value of M then we can obtain reasonable bounds for the prime power factors of $|\mathrm{Gal}(\Gamma, \Lambda)|$.

In (1.5) of [4] it was shown that, for $p = 2$, the polynomial $\overline{F}_{23,20,1,1} = Y^{23} + XY + 1$ linearizes over $k[X]$ at 11. In the Linearization Lemma (5.1) of Section 5, by slightly modifying the proof of (1.5) of [4], we shall show that, for $p = 2$ and for any element T in an overfield of $k[X]$ (for instance, T could be transcendental over $k[X]$), the polynomial $F^* = Y\overline{F}_{23,20,1,1} + T$ linearizes over $k[X, T]$ at 12. By taking (α, X) for (X, T) in F^*, it follows that $|\mathrm{Gal}(Y^{24} + \alpha Y^4 + Y + X, k(X))|$ divides $|\mathrm{GL}(12, 2)|$.

For $p = 2$, in (1.3) of [4] it was shown that $\mathrm{Gal}(Y^{23} + XY^3 + 1, k(X))$ is doubly transitive and, as said above, this together with the fact that $Y^{23} + XY^3 + 1$ linearizes at 11 shows that $\mathrm{Gal}(Y^{23} + XY^3 + 1, k(X)) = M_{23}$. This time, in the Transitivity Lemma (4.1) of Section 4, we shall show that for certain monic polynomials F of degree $n = mq$ in Y with coefficients in $k[X]$ where q is the highest power of p which divides n and where p need not be 2, the Galois group $\mathrm{Gal}(F, k(X))$ is doubly transitive and its order is divisible by $n(n-1)(q-1)$. The proof of the Transitivity Lemma (4.1) will be based on some auxilliary lemmas which we shall prove in Section 2 and an irreducibility lemma which we shall prove in Section 3. Thus the Linearization Lemma (5.1) shows that $\mathrm{Gal}(Y^{24} + \alpha Y^4 + Y + 1, k(X))$ is not too big and the Transitivity Lemma (4.1) shows that it is not too small. In Section 6, Theorem (1.1) will be deduced from these two facts.

2. Auxilliary lemmas

Let $F = F(Y) = Y^n + B_{n_1}Y^{n_1} + B_{n_2}Y^{n_2} + \cdots + B_{n_h}Y^{n_h} + X$ where h and $n > n_1 > n_2 > \cdots > n_h = 1$ are positive integers, and $0 \neq B_{n_i} \in k$ for $1 \leq i \leq h$. Assume that n is divisible by p and let m and q be the unique positive integers with $n = mq$ such that m is nondivisible by p and q is a power of p. For $1 \leq i \leq h-1$ assume that n_i is divisible by p and let m_i and q_i be the unique positive integers with $n_i = m_i q_i$ such that m_i is nondivisible by p and q_i is a power of p. [Note that the Y-derivative of F equals the nonzero element B_{n_h} of k and hence the Galois group $\mathrm{Gal}(F, k(X))$ makes sense and the equation $F = 0$ gives an unramified covering of the affine line L_k.]

In the Transitivity Lemma (4.1) of Section 4 we shall show that if certain conditions are satisfied then $\mathrm{Gal}(F, k(X))$ is doubly transitive and its order is divisible by $n(n-1)(q-1)$. To prepare the ground work for this, here we shall prove two auxilliary lemmas.

First let us note that F is irreducible because it is linear in X, and hence $\mathrm{Gal}(F, k(X))$ is transitive. By "throwing away" the root Y of F we get the monic polynomial $\Omega(Y, Z) = Z^{-1}[F(Z+Y) - F(Y)]$ of degree $n-1$ in Z with coefficients in $k(Y) = k(X, Y)$; for the method of "throwing away" roots and its relation to one-point stabilizers, see [2]. Since the Y-derivative of F is a nonzero element of k, it follows that the Z-discriminant of $\Omega(Y, Z)$ is a nonzero element of k and $\mathrm{Gal}(\Omega(Y, Z), k(Y))$ is isomorphic to the one-point stabilizer of $\mathrm{Gal}(F, k(X))$. For every positive integer u let $E_u(Y, Z) = Z^{-1}[(Z+Y)^u - Y^u]$. Then clearly $\Omega(Y, Z) = E_n(Y, Z) + \sum_{i=1}^{h} B_{n_i} E_{n_i}(Y, Z)$. Therefore, by writing X and Y for Y and Z, respectively, we get the following:

Auxilliary Lemma (2.1). *With F as above, let $\Omega(X, Y)$ be the monic polynomial of degree $n-1$ in Y with coefficients in $k[X]$ obtained by putting*

$$\Omega(X, Y) = E_n(X, Y) + \sum_{i=1}^{h} B_{n_i} E_{n_i}(X, Y)$$

where, for every positive integer u, by $E_u(X, Y)$ we are denoting the homogeneous polynomial of degree $u-1$ in (X, Y) with coefficients in k obtained by putting $E_u(X, Y) = Y^{-1}[(Y+X)^u - X^u]$. Then the Y-discriminant of $\Omega(X, Y)$ is a nonzero element of k and hence the equation $\Omega(X, Y) = 0$ gives an unramified covering of the affine line L_k. Moreover, $\mathrm{Gal}(F, k(X))$ *is transitive and its one-point stabilizer is isomorphic to* $\mathrm{Gal}(\Omega(X, Y), k(X))$.

Now the $k(X)$-automorphism $Y \mapsto XY$ of $k(X)[Y]$ sends $X^{1-n}\Omega(X, Y)$ to $\Omega^*(X, Y) = X^{1-n}\Omega(X, XY) = E_n(1, Y) + \sum_{i=1}^{h} B_{n_i} X^{n_i - n} E_{n_i}(1, Y)$ which is a monic polynomial of degree $n-1$ in Y with coefficients in $k(X)$. Likewise the k-automorphism $(X, Y) \mapsto (1/X, Y)$ of $k(X)[Y]$ sends $\Omega^*(X, Y)$ to $\Omega'(X, Y) = \Omega^*(1/X, Y) = E'_n(Y) + \sum_{i=1}^{h} B_{n_i} X^{n-n_i} E'_{n_i}(Y)$ which is a monic polynomial of degree $n-1$ in Y with coefficients in $k[X]$, where for every positive integer u we have put $E'_u(Y) = Y^{-1}[(Y+1)^u - 1]$. It follows that the Y-discriminant of $\Omega'(X, Y)$ is a nonzero element of $k[X]$, and $\mathrm{Gal}(\Omega'(X, Y), k(X))$ is isomorphic to $\mathrm{Gal}(\Omega(X, Y), k(X))$. Therefore, by (2.1) we get the following:

Auxilliary Lemma (2.2). *With F as above, let $\Omega'(X, Y)$ be the monic polynomial of degree $n-1$ in Y with coefficients in $k[X]$ obtained by putting $\Omega'(X, Y) = E'_n(Y) + \sum_{i=1}^{h} B_{n_i} X^{n-n_i} E'_{n_i}(Y)$ where, for every positive integer u, by $E'_u(Y)$ we are denoting the monic polynomial of degree $u-1$ in Y with coefficients in k obtained by putting $E'_u(Y) = Y^{-1}[(Y+1)^u - 1]$. Then the Y-discriminant of $\Omega'(X, Y)$ is a nonzero element of $k[X]$. Moreover,*

$\mathrm{Gal}(F, k(X))$ *is transitive and its one-point stabilizer is isomorphic to* $\mathrm{Gal}(\Omega'(X, Y), k(X))$.

3. Irreducibility

As another step toward the Transitivity Lemma, let us prove the following:

Irreducibility Lemma (3.1). *Let V be a real discrete valuation of a field K (note that then V maps K onto $\mathbb{Z} \cup \{\infty\}$), let $0 \le r < d$ be integers, let $f(Y) = \sum_{j=0}^{d} b_j Y^{d-j}$ be a polynomial of degree d in Y with coefficients b_j in K such that $V(b_j) \ge V(b_0) = V(b_r) = 0 < V(b_d) < \infty$ for $0 < j < r$, and $V(b_j)/V(b_d) > (j-r)/(d-r)$ for $r < j < d$, and let*

$$s = (d-r)/\mathrm{GCD}(V(b_d), d-r).$$

Then we have the following.

(3.1.1) *If y is a root of $f(Y)$ in an overfield of K and W is an extension of V to $K(y)$ with $W(y) > 0$ (where we again assume W to map $K(y)$ onto $\mathbb{Z} \cup \{\infty\}$), then the reduced ramification exponent e of W over V is divisible by s.*

(3.1.2) *If $d = d - r = s$, then $f(Y)$ is irreducible in $K[Y]$.*

(3.1.3) *If $f(Y)$ is irreducible in $K[Y]$ and has no multiple root in any overfield of K, then $|\mathrm{Gal}(f(Y)/b_0, K)|$ is divisible by s.*

For a moment let the situation be as in (3.1.1). Then $W(b_j) = eV(b_j)$ for $0 \le j \le d$, and hence $W(b_j) \ge W(b_0) = W(b_r) = 0 < W(b_d) < \infty$ for $0 < j < r$, and $W(b_j) > (j-r)W(b_d)/(d-r)$ for $r < j < d$. Since $W(y) > 0$, we see that $W(b_j y^{d-j}) > W(b_r y^{d-r})$ for $0 \le j < r$; therefore, since $f(y) = 0$, there must be at least two minimal W-value terms amongst $(b_j y^{d-j})_{r \le j \le d}$. Now if $W(y) > W(b_d)/(d-r)$ then for $r \le j < d$ we would have

$$W(b_j y^{d-j}) = W(b_j) + (d-j)W(y) > [(j-r) + (d-j)]W(b_d)/(d-r) = W(b_d)$$

which would contradict the existence of two minimal value terms. Likewise, if $W(y) < W(b_d)/(d-r)$ then for $r < j \le d$ we would have

$$\begin{aligned} W(b_j y^{d-j}) = W(b_j) + (d-j)W(y) &\ge (j-r)(d-r)^{-1}W(b_d) + (d-j)W(y) \\ &> (j-r)W(y) + (d-j)W(y) \\ &= (d-r)W(y) = W(b_r y^{d-r}) \end{aligned}$$

which would again contradict the existence of two minimal value terms. Consequently we must have $W(y) = W(b_d)/(d-r)$. Therefore $eV(b_d)/(d-r) = W(y) \in \mathbb{Z}$ and hence e is divisible by s. This proves (3.1.1).

Next for a moment let the situation be as in (3.1.2). We can take a root y of $f(y)$ in an overfield of K and we can take an extension W of V to $K(y)$. Since $f(y) = 0$, there must be at least two minimal W-value terms amongst $(b_j y^{d-j})_{0 \le j \le d}$. Since $d = d - r$, we must have $W(b_0) = 0 < W(b_j)$ for $1 \le j \le d$. Since $f(y) = 0$ and $V(b_d) \ne \infty$, we must also have $y \ne 0$. Consequently $W(y) > 0$ because otherwise $b_0 y^d$ would be the only minimal value term. Therefore, since $s = d$, by (3.1.1) we see that the reduced ramification exponent of W over V is divisible by d and hence it must equal d and $f(Y)$ must be irreducible in $K[Y]$. This proves (3.1.2).

Finally let the situation be as in (3.1.3). Let $y_1, y_2, \dots, y_d$ be the distinct roots of $f(Y)$ in a splitting field L of $f(Y)$ over K, and take an extension U of V to L. Since $V(b_0) = 0 \le V(b_j)$ for $1 \le j \le d$, we must have $U(b_i) \ge 0$ for $1 \le i \le d$. Since $y_1 y_2 \dots y_d = (-1)^d b_d / b_0$ and $V(b_0) = 0 < V(b_d)$, we conclude that $U(y_i) > 0$ for some i. Let $y = y_i$ and let W be the extension of V to $K(y)$ such that U is an extension of W to L. Now $W(y) > 0$ and hence by (3.1.1) we see that the reduced ramification exponent of W over V is divisible by s. Therefore $|\mathrm{Gal}(f(Y)/b_0, K)|$ is divisible by s. This proves (3.1.3).

4. Transitivity

Finally let us state and prove the:

Transitivity Lemma (4.1). *Let* $F = F(Y) = Y^n + B_{n_1} Y^{n_1} + B_{n_2} Y^{n_2} + \cdots + B_{n_h} Y^{n_h} + X$ *where* h *and* $n > n_1 > n_2 > \cdots > n_h = 1$ *are positive integers, and* $0 \ne B_{n_i} \in k$ *for* $1 \le i \le h$. *Assume that* n *is divisible by* p *and let* m *and* q *be the unique positive integers with* $n = mq$ *such that* m *is nondivisible by* p *and* q *is a power of* p. *For* $1 \le i \le h-1$ *assume that* n_i *is divisible by* p *and let* m_i *and* q_i *be the unique positive integers with* $n_i = m_i q_i$ *such that* m_i *is nondivisible by* p *and* q_i *is a power of* p. [*Note that the* Y*-derivative of* F *equals the nonzero element* B_{n_h} *of* k *and hence the Galois group* $\mathrm{Gal}(F, k(X))$ *makes sense and the equation* $F = 0$ *gives an unramified covering of the affine line* L_k.] *Now considering the conditions*

(*)

$$\mathrm{GCD}(n-1, q-1) = 1 \quad \textit{and} \quad (q_i - 1)(n-1) > (q-1)(n_i - 1) \quad \textit{for } 1 \le i \le h-1,$$

and

(**)
$$(n - n_i)(q-1) > (n-1)(q-q_i) \quad \textit{for } 1 \le i \le h-1,$$

we have that: (*) $\Rightarrow$ $\mathrm{Gal}(F, k(X))$ *is doubly transitive, and* (*) + (**) $\Rightarrow$ $|\mathrm{Gal}(F, k(X))|$ *is divisible by* $n(n-1)(q-1)$.

To prove (4.1), in view of (2.2), it suffices to show that in the situation of (2.2), (*) $\Rightarrow$ $\Omega'(X, Y)$ is irreducible in $k(X)[Y]$, and (*) + (**) $\Rightarrow$ $|\mathrm{Gal}(\Omega'(X, Y), k(X))|$ is divisible by $q-1$. So let the situation be as in (2.2) and assume (*), let $K = k(X)$ and $d = n-1$, and let V be the order of zero at $X = 0$, i.e., $V(X^t P(X)/Q(X)) = t$ for all integers t and all $P(X)$ and $Q(X)$ in $k[X]$ with $P(0) \ne 0 \ne Q(0)$. Note that now $V(E'_n(X)) = q-1$, $V(E'_{n_i}(X)) = q_i - 1$ for $1 \le i \le h-1$, and $E'_{n_h}(X) = 1$.

For a moment let $r = 0$ and $f(Y) = \Omega'(Y, X) = \sum_{j=0}^{d} b_j Y^{d-j}$ with $b_j \in K$. Then $b_0 = B_{n_h} E'_{n_h}(X) =$ the nonzero element B_{n_h} of k. Also $b_d = E'_n(X)$ and hence $V(b_d) = q-1$. Moreover, for $1 \le i < h$ we have $0 < n_i - 1 < d$ and $b_{n_i - 1} = B_{n_i} E'_{n_i}(X)$ and hence $V(b_{n_i - 1}) = q_i - 1$ and therefore $V(b_{n_i - 1})/V(b_d) > (n_i - 1)/d$. Clearly $b_j = 0$ for all $j \in \{1, 2, \dots, d\} \setminus \{n_1, n_2, \dots, n_h\}$, and hence $V(b_j)/V(b_d) > j/d$ for $0 < j < d$. Since GCD $(n-1, q-1) = 1$, we also get $\mathrm{GCD}(V(b_d), d) = 1$, and hence upon letting $s = (d-r)/\mathrm{GCD}(V(b_d), d-r)$ we have $d = d - r = s$. Consequently by (3.1.2) we conclude that $\Omega'(Y, X)$ is irreducible in $k(X)[Y]$. Therefore $\Omega'(X, Y)$ is irreducible in $k(Y)[X]$, and hence by Gauss's Lemma we see that $\Omega'(X, Y)$ is irreducible in $k(X)[Y]$.

Now assume $(**)$ and let $r = n - q$ and $f(Y) = \Omega'(X, Y) = \sum_{j=0}^{d} b_j Y^{d-j}$ with $b_j \in K$. Then $V(b_j) \geq V(b_0) = V(b_r) = 0 < V(b_d) = n - 1 = d$ for $0 < j < r$, and $V(b_j)/V(b_d) > (j-r)/(d-r)$ for $r < j < d$. Also $(d-r)/\mathrm{GCD}(V(b_d), d-r) = q - 1$. Therefore by (3.1.3) we see that $|\mathrm{Gal}(\Omega'(X, Y), k(X))|$ is divisible by $q - 1$.

5. Linearization

Let us now prove the:

Linearization Lemma (5.1). *If $p = 2$ and T is any element in an overfield of $k(X)$ [for instance, T could be transcendental over $k(X)$], then there exist elements $A_0, A_1, \ldots, A_{12}$ in $k[X, T]$ with $A_0 \neq 0$ and $A_{12} = 1$ such that $\sum_{i=0}^{12} A_i Y^{2^i} = HF^*$ for some $H \in k[X, T][Y]$ where $F^* = Y\overline{F}_{23,20,1,1} + T$.*

The proof of (5.1) is simply obtained by adding obvious terms involving T in the RHS of various equations occurring in the proof of (1.5) of [4] given in Section 5 of [4]. In greater detail: To prove (5.1) assume that $p = 2$. Now

$$F^* = Y^{24} + XY^4 + Y + T$$

and by adding $F^* + Y^{24}$ to both sides of this we get

$$(J'_{24}) \qquad Y^{24} = XY^4 + Y + T + F^*.$$

Let $P \equiv Q$ mean $P - Q = HF^*$ for some $H \in k[X, T][Y]$. Then multiplying (J'_{24}) by Y^{i-24} for $i = 24, 26, 32, 36$ we get:

$$(J_{24}) \qquad Y^{24} \equiv XY^4 + Y + [T],$$

$$(J_{26}) \qquad Y^{26} \equiv XY^6 + Y^3 + [TY^2],$$

$$(J_{32}) \qquad Y^{32} \equiv XY^{12} + Y^9 + [TY^8],$$

$$(J_{36}) \qquad Y^{36} \equiv XY^{16} + Y^{13} + [TY^{12}].$$

Squaring (J_{32}) we get

$$Y^{64} \equiv X^2Y^{24} + Y^{18} + [T^2Y^{16}],$$

and using (J_{24}) we obtain

$$(J_{64}) \qquad Y^{64} \equiv Y^{18} + X^3Y^4 + X^2Y + [T^2Y^{16} + X^2T].$$

Likewise, by squaring (J_{64}) and then using (J_{36}) we obtain

$$(J_{128}) \qquad Y^{128} \equiv XY^{16} + Y^{13} + X^6Y^8 + X^4Y^2 + [T^4Y^{32} + TY^{12} + X^4T^2].$$

Again, by squaring (J_{128}) and then using (J_{24}), (J_{26}), and (J_{32}) we obtain

$$(J_{256}) \qquad \begin{aligned} Y^{256} \equiv{}& X^{12}Y^{16} + X^3Y^{12} + X^2Y^9 + XY^6 + X^8Y^4 + Y^3 \\ &+ [T^8Y^{64} + X^2TY^8 + XT^2Y^4 + TY^2 + T^2Y + X^8T^4 + T^3]. \end{aligned}$$

Similarly, by squaring (J_{256}) and then using (J_{24}) and (J_{32}) we obtain (J_{512})

$$\begin{aligned} Y^{512} \equiv{}& X^4Y^{18} + (X^2 + X^{25})Y^{12} + X^{24}Y^9 + X^{16}Y^8 + Y^6 + X^7Y^4 + X^6Y \\ &+ [T^{16}Y^{128} + X^4T^2Y^{16} + (X^2T^4 + X^{24}T)Y^8 \\ &+ T^2Y^4 + T^4Y^2 + (X^{16}T^8 + T^6 + X^6T)]. \end{aligned}$$

Likewise, by squaring (J_{512}) and then using (J_{24}) and (J_{36}) we obtain

$$(J_{1024})\qquad \begin{aligned} Y^{1024} \equiv\ & X^{48}Y^{18} + (X^9 + X^{32})Y^{16} + X^8Y^{13} + Y^{12} \\ & + X^{14}Y^8 + (X^5 + X^{51})Y^4 + X^{12}Y^2 + (X^4 + X^{50})Y \\ & + [T^{32}Y^{256} + X^8T^2Y^{32} + (X^4T^8 + X^{48}T^2)Y^{16} + X^8TY^{12} \\ & \quad + (T^4 + X^{18}T + X^{64}T)Y^8 + T^8Y^4 \\ & \qquad + (X^{32}T^{16} + T^{12} + X^{12}T^2 + X^{50}T + X^4T)]. \end{aligned}$$

Finally, by squaring (J_{1024}) and then using (J_{24}), (J_{26}), (J_{32}), and (J_{36}) we obtain

$$(J_{2048})$$

$$\begin{aligned} Y^{2048} \equiv\ & (X^{28} + X^{97})Y^{16} + X^{96}Y^{13} + (X^{19} + X^{65})Y^{12} \\ & + (X^{18} + X^{64})Y^9 + (X^{10} + X^{102})Y^8 + X^{17}Y^6 \\ & + (X + X^{24})Y^4 + X^{16}Y^3 + (X^8 + X^{100})Y^2 + Y \\ & + [T^{64}Y^{512} + X^{16}T^4Y^{64} + (X^8T^{16} + X^{96}T^4)Y^{32} \\ & \quad + (T^8 + X^{36}T^2 + X^{128}T^2)Y^{16} + X^{96}TY^{12} \\ & \quad + (T^{16} + X^{18}T + X^{64}T)Y^8 + X^{17}T^2Y^4 + X^{16}TY^2 + X^{16}T^2Y \\ & \qquad + (X^{64}T^{32} + T^{24} + X^{24}T^4 + X^{16}T^3 + X^{100}T^2 + X^8T^2 + T)]. \end{aligned}$$

Since the above formulas (J_{24}), (J_{26}), ..., (J_{2048}) are obtained by adding T-terms in the RHS of the corresponding formulas (I_{24}), (I_{26}), ..., (I_{2048}) of Section 5 of [4], by modifying the last formula of that section to compensate for the T-terms we get

$$\begin{aligned} & Y^{2048} + T^{64}Y^{512} + X^{16}Y^{256} + X^{96}Y^{128} + (X^8T^{16} + X^{64})Y^{32} \\ & \quad + (T^{16} + X^{10})Y^8 + XY^4 + X^8Y^2 + Y + (X^{64}T^{32} + T^{24} + X^8T^2 + T) \equiv 0. \end{aligned}$$

Alternatively the above equation can be proved directly by using (J_{2048}), (J_{512}), (J_{256}), (J_{128}), and (J_{32}). By multiplying the above equation by its constant term $X^{64}T^{32} + T^{24} + X^8T^2 + T$ and then adding the resulting equation to the square of the above equation we get

$$\begin{aligned} & Y^{4096} + (X^{64}T^{32} + T^{24} + X^8T^2 + T)Y^{2048} + T^{128}Y^{1024} \\ & \quad + (X^{64}T^{96} + T^{88} + X^8T^{66} + T^{65} + X^{32})Y^{512} \\ & \quad + (X^{80}T^{32} + X^{16}T^{24} + X^{24}T^2 + X^{16}T + X^{192})Y^{256} \\ & \quad + (X^{160}T^{32} + X^{96}T^{24} + X^{104}T^2 + X^{96}T)Y^{128} + (X^{16}T^{32} + X^{128})Y^{64} \\ & \quad + (X^{72}T^{48} + X^8T^{40} + X^{128}T^{32} + X^{64}T^{24} \\ & \quad + X^{16}T^{18} + X^8T^{17} + X^{72}T^2 + X^{64}T)Y^{32} + (T^{32} + X^{20})Y^{16} \\ & \quad + (X^{64}T^{48} + T^{40} + X^{74}T^{32} + X^{10}T^{24} \\ & \quad + X^8T^{18} + T^{17} + X^{18}T^2 + X^{10}T + X^2)Y^8 \\ & \quad + (X^{65}T^{32} + XT^{24} + X^9T^2 + XT + X^{16})Y^4 \\ & \quad + (X^{72}T^{32} + X^8T^{24} + X^{16}T^2 + X^8T + 1)Y^2 \\ & \quad + (X^{64}T^{32} + T^{24} + X^8T^2 + T)Y \equiv 0, \end{aligned}$$

and this proves (5.1).

6. Mathieu group

To prove (1.1) assume that $p = 2$, let α be any element of k, and let $G = \mathrm{Gal}(Y^{24} + \alpha Y^4 + Y + X, k(X))$. Then by (4.1) we see that G is a doubly transitive permutation group of degree 24 whose order is divisible by 7, and hence by CTT and Special CDT on pp. 86 to 89 of [2], we must have $G = M_{24}$ or A_{24} or S_{24}. As said in the Introduction, by taking (α, X) for (X, T) in (5.1) we see that $|G|$ divides $|\mathrm{GL}(12, 2)|$. Finally, as a factorization of $|\mathrm{GL}(12, 2)|$ into powers of prime numbers we have

$$\begin{aligned}|\mathrm{GL}(12, 2)| &= \prod_{i=0}^{11}(2^{12} - 2^i) \\ &= 2^{66} \times 3^8 \times 5^3 \times 7^4 \times 11 \times 13 \times 17 \times 23 \times 31^2 \times 73 \times 89 \times 127,\end{aligned}$$

but $|A_{24}|$ and $|S_{24}|$ are obviously divisible by 11^2 and hence we must have $G = M_{24}$.

References

1. S. S. Abhyankar, *Coverings of algebraic curves*, Amer. J. Math. **79** (1957), 825–856.
2. ——, *Galois theory on the line in nonzero characteristic*, Bull. Amer. Math. Soc. (N.S.) **27** (1992), 68–133.
3. ——, *Fundamental group of the affine line in positive characteristic*, Proceedings of the 1992 Bombay International Colloquium on Geometry and Analysis held at the Tata Institute of Fundamental Research (to appear).
4. ——, *Mathieu group coverings in characteristic two*, C. R. Acad. Sci. Paris Sér. I Math. **316** (1993), 267–271.
5. S. S. Abhyankar and I. Yie, *Small degree coverings of the affine line in characteristic two*, Discrete Math. (to appear).
6. J.-P. Serre, e-mail to Abhyankar dated October 1991.

Department of Mathematics, Purdue University, West Lafayette, Indiana 47907

E-mail address: `ram@cs.purdue.edu`

E-mail address: `yie@math.purdue.edu`