A Galois Approach to mth Roots of Matrices with Rational Entries

Robert Reams
Programs in Mathematical Sciences
The University of Texas at Dallas
Box 830688, Richardson, Texas 75083-0688

Abstract

Let A be a given $n \times n$ matrix with rational entries and irreducible characteristic polynomial $f(x)$. We investigate the Galois groups of $f(x)$ and $f(x^m)$, to find necessary and sufficient conditions for the existence of a solution B to the matrix equation $A = B^m$, where B is also a matrix with rational entries. We do this by finding necessary and sufficient conditions that $f(x^m)$ has a factor of degree n (with rational coefficients).

Introduction

We concern ourselves with finding matrix solutions B to the equation $A = g(B)$, where A is some given matrix and $g(x)$ is a polynomial. Previous work has been done by other authors (see for instance [1] and [4]) where all the matrices have entries from an arbitrary field, or just complex entries. We look at the situation where all the entries of A are rational i.e. $A \in M_n(\mathbf{Q})$, and the characteristic polynomial of A, namely $f(x)$, is irreducible. Then by using Galois theory and looking at the structure of the Galois groups of $f(x)$ and $f(x^m)$, we find conditions on these groups that the matrix A has an mth root $B \in M_n(\mathbf{Q})$, under certain fairly general restrictions. First we prove a proposition due to T. J. Laffey and B. Cain, previously unpublished, and which provides the motivation for what follows.

Proposition: Let $\mathbf{F}$ be a field and $A \in M_n(\mathbf{F})$ have irreducible characteristic polynomial $f(x)$. Let $g(x) \in \mathbf{F}[x]$. Then the equation $g(B) = A$ is solvable for $B \in M_n(\mathbf{F})$ if and only if $f(g(x))$ has a factor of degree n in $\mathbf{F}[x]$.
Proof: Suppose such a B exists and let $m(x)$ be its minimal polynomial. Since $\mathbf{F}[B]$ contains $\mathbf{F}[A]$, $m(x)$ has degree n. Also, $f(g(B)) = f(A) = 0$, so $m(x)$ divides $f(g(x))$.

Conversely, let $h(x)$ be a factor of $f(g(x))$ of degree n, and let C be the companion matrix of $h(x)$. Then $f(g(C)) = 0$, and since $f(x)$ is irreducible and has degree n, it follows that $g(C)$ is similar to the companion matrix of $f(x)$ and thus $g(C)$ is similar to A, say $T^{-1}g(C)T = A$, where $T \in GL(n, \mathbf{F})$. But then $g(T^{-1}CT) = A$, and so take $B = T^{-1}CT$.

In consequence of this proposition we may (and will) concentrate on the existence of a factor of $f(g(x))$ of degree n. We now prove two theorems, restricted to the case where $g(x) = x^m$, and include some counterexamples to show that there are some directions in which the results cannot be improved. We will use the notation that $|G|$ denotes the order of a group G, and $G(K/k)$ is the Galois group of the extension K over k.

Theorem 1: Let $m,\ n$ be natural numbers, m odd, and $A \in M_n(\mathbf{Q})$ have irreducible characteristic polynomial $f(x)$. Let $\mu_i,\ 1 \le i \le n$, be the roots of $f(x)$, and for some choice $\lambda_1, ..., \lambda_n$ of n roots of $f(x^m)$, where $\lambda_i^m = \mu_i,\ 1 \le i \le n$, suppose that $\mathbf{Q}(\lambda_1, ..., \lambda_n) \cap \mathbf{Q}(\zeta) = \mathbf{Q}$, where $\zeta = e^{\frac{2\pi i}{m}}$. Then the following are equivalent:

(i) the equation $A = B^m$ is solvable with $B \in M_n(\mathbf{Q})$,

(ii) $f(x^m)$ has a factor of degree n in $\mathbf{Q}[x]$,

(iii) $|G(K/\mathbf{Q})| = \phi(m)|G(L/\mathbf{Q})|$,

where $\phi(\cdot)$ is Euler's ϕ-function, K is the splitting field for $f(x^m)$ over $\mathbf{Q}$ and L is the splitting field for $f(x)$ over $\mathbf{Q}$.

Proof: That (i) is equivalent to (ii) follows from the proposition, with $g(x) = x^m$.

To prove that (ii) implies (iii), let $h(x) \in \mathbf{Q}[x]$ be a factor of degree n of $f(x^m)$ and let us say $h(x) = (x - \nu_1)(x - \nu_2) \cdots (x - \nu_n)$, where $\nu_i \in \overline{\mathbf{Q}}$, $1 \le i \le n$.

Then $f(\nu_1^m) = 0$, so that $\nu_1^m = \mu_i$, for some $i \in \{1, 2, ..., n\}$. But since $[\mathbf{Q}(\nu_1) : \mathbf{Q}] = [\mathbf{Q}(\nu_1) : \mathbf{Q}(\mu_i)][\mathbf{Q}(\mu_i) : \mathbf{Q}]$ and $[\mathbf{Q}(\mu_i) : \mathbf{Q}] = n$, this implies $[\mathbf{Q}(\nu_1) : \mathbf{Q}] = n$, so $h(x) \in \mathbf{Q}[x]$ must be irreducible and so the roots $\nu_i,\ 1 \le i \le n$ must be distinct. We also have that $\nu_i^m,\ 1 \le i \le n$ must be distinct, since suppose not, then $\nu_i^m = \nu_j^m$, for some $i \ne j$, which implies $\nu_i = \zeta^r \nu_j$, for some $r \in \{0, 1, 2, ..., m-1\}$. Now $\nu_i^m = \mu_{k_i}$, $\nu_j^m = \mu_{k_j}$ for some $\mu_{k_i},\ \mu_{k_j} \in \{\mu_1, ..., \mu_n\}$, and $[\mathbf{Q}(\nu_i) : \mathbf{Q}] = [\mathbf{Q}(\nu_i) : \mathbf{Q}(\mu_{k_i})][\mathbf{Q}(\mu_{k_i}) : \mathbf{Q}]$ implies $\mathbf{Q}(\nu_i) = \mathbf{Q}(\mu_{k_i})$ so $\nu_i \in \mathbf{Q}(\mu_{k_i})$ and similarly $\nu_j \in \mathbf{Q}(\mu_{k_j})$. But $\zeta^r \nu_j,\ \nu_j \in \mathbf{Q}(\mu_{k_i}, \mu_{k_j}) \subset \mathbf{Q}(\mu_1, ..., \mu_n) \subset \mathbf{Q}(\lambda_1, ..., \lambda_n)$ where $\lambda_1, ..., \lambda_n$ are as in the hypotheses of the theorem. Thus $\mathbf{Q}(\lambda_1, ..., \lambda_n) \cap \mathbf{Q}(\zeta) = \mathbf{Q}$ and therefore $\zeta^r = 1$, then $\nu_i = \nu_j$, contradiction.

Thus
$$
\begin{aligned}
f(x^m) = (x^m - \nu_1^m) \cdots (x^m - \nu_n^m) = &(x - \nu_1)(x - \zeta\nu_1) \cdots (x - \zeta^{m-1}\nu_1) \\
&(x - \nu_2)(x - \zeta\nu_2) \cdots (x - \zeta^{m-1}\nu_2) \\
&\quad\vdots \\
&(x - \nu_n)(x - \zeta\nu_n) \cdots (x - \zeta^{m-1}\nu_n).
\end{aligned}
$$

By definition $\mathbf{Q}(\nu_1, ..., \nu_n)$ is the splitting field for $h(x)$, and is therefore a Galois extension. Similarly, $\mathbf{Q}(\nu_1, ..., \nu_n, \zeta)$ is the splitting field for $f(x^m)$ and also a Galois extension. (Note that $\mathbf{Q}(\nu_1, ..., \nu_n, \zeta) = K = \mathbf{Q}(\lambda_1, ..., \lambda_n, \zeta)$ by unique factorization of $f(x^m)$ in $\overline{\mathbf{Q}}[x]$). Thus we have the tower of fields:

$$\begin{array}{c} \mathbf{Q}(\nu_1, ..., \nu_n, \zeta) \\ | \\ \mathbf{Q}(\nu_1, ..., \nu_n) \\ | \\ \mathbf{Q} \end{array}$$

and

$$|G(K/\mathbf{Q})| = [\mathbf{Q}(\nu_1, ..., \nu_n, \zeta) : \mathbf{Q}] = [\mathbf{Q}(\nu_1, ..., \nu_n, \zeta) : \mathbf{Q}(\nu_1, ..., \nu_n)][\mathbf{Q}(\nu_1, ..., \nu_n) : \mathbf{Q}] \ (*).$$

Since (again) $[\mathbf{Q}(\nu_i) : \mathbf{Q}] = [\mathbf{Q}(\nu_i) : \mathbf{Q}(\mu_i)][\mathbf{Q}(\mu_i) : \mathbf{Q}]$, where $\mu_i = \nu_i^m$, $1 \leq i \leq n$, we deduce as before that $\mathbf{Q}(\nu_i) = \mathbf{Q}(\mu_i)$, for each i, $1 \leq i \leq n$. Thus $\mathbf{Q}(\nu_1, ..., \nu_n) = \mathbf{Q}(\mu_1, ..., \mu_n) = L$, and so

$$[\mathbf{Q}(\nu_1, ..., \nu_n) : \mathbf{Q}] = [\mathbf{Q}(\mu_1, ..., \mu_n) : \mathbf{Q}] = |G(L/\mathbf{Q})|.$$

We assumed $\mathbf{Q}(\lambda_1, ..., \lambda_n) \cap \mathbf{Q}(\zeta) = \mathbf{Q}$, where $\lambda_1, ..., \lambda_n$ are as in the statement of the theorem, and we know $\mathbf{Q}(\lambda_1, ..., \lambda_n) \supset \mathbf{Q}(\mu_1, ..., \mu_n) = \mathbf{Q}(\nu_1, ..., \nu_n)$ so $\mathbf{Q}(\nu_1, ..., \nu_n) \cap \mathbf{Q}(\zeta) = \mathbf{Q}$, giving $G(\mathbf{Q}(\nu_1, ..., \nu_n, \zeta)/\mathbf{Q}(\nu_1, ..., \nu_n)) \cong G(\mathbf{Q}(\zeta)/\mathbf{Q})$ [3, p.305]. Then from $(*)$ we get

$$|G(K/\mathbf{Q})| = \phi(m)|G(L/\mathbf{Q})| \text{ , which is (iii).}$$

Conversely, to prove that (iii) implies (ii), we have $f(x) = (x - \mu_1) \cdots (x - \mu_n)$, so $f(x^m) = (x^m - \mu_1) \cdots (x^m - \mu_n) = \Pi_{j=1}^n (x - \lambda_j)(x - \zeta\lambda_j) \cdots (x - \zeta^{m-1}\lambda_j)$, where $\lambda_j^m = \mu_j$, $1 \leq j \leq n$, and $\mathbf{Q}(\lambda_1, ..., \lambda_n) \cap \mathbf{Q}(\zeta) = \mathbf{Q}$.

Now consider the tower of fields:

$$\begin{array}{c} \mathbf{Q}(\lambda_1, ..., \lambda_n, \zeta) \\ | \\ \mathbf{Q}(\lambda_1, ..., \lambda_n) \\ | \\ \mathbf{Q}(\mu_1, ..., \mu_n) \\ | \\ \mathbf{Q} \end{array}$$

We know $[\mathbf{Q}(\lambda_1, ..., \lambda_n, \zeta) : \mathbf{Q}(\lambda_1, ..., \lambda_n)] = \phi(m)$, $[\mathbf{Q}(\mu_1, ..., \mu_n) : \mathbf{Q}] = |G(L/\mathbf{Q})|$, and $[\mathbf{Q}(\lambda_1, ..., \lambda_n, \zeta) : \mathbf{Q}] = |G(K/\mathbf{Q})|$, and since we're given $|G(K/\mathbf{Q})| = \phi(m)|G(L/\mathbf{Q})|$, we must have that $\mathbf{Q}(\lambda_1, ..., \lambda_n) = \mathbf{Q}(\mu_1, ..., \mu_n)$.

Therefore $\mathbf{Q}(\lambda_1, ..., \lambda_n)$ is a Galois extension of $\mathbf{Q}$, and $\tau(\mathbf{Q}(\lambda_1, ..., \lambda_n)) = \mathbf{Q}(\lambda_1, ..., \lambda_n)$, for all $\tau \in G(\mathbf{Q}(\lambda_1, ..., \lambda_n, \zeta)/\mathbf{Q})$. We know $f^\tau(x^m) = f(x^m)$, since all the coefficients of $f(x)$ are in $\mathbf{Q}$, then by unique factorization in $\overline{\mathbf{Q}}[x]$ we know τ just permutes the roots of

$f(x^m)$. But τ must also just permute $\lambda_1, ..., \lambda_n$ since if $\tau(\lambda_i) = \zeta^s \lambda_j \in \mathbf{Q}(\lambda_1, ..., \lambda_n)$ then $\zeta^s \in \mathbf{Q}(\lambda_1, ..., \lambda_n)$, but $\mathbf{Q}(\lambda_1, ..., \lambda_n) \cap \mathbf{Q}(\zeta) = \mathbf{Q}$, so we must have that $\zeta^s = 1$ (as m is odd). Let $h(x) = (x-\lambda_1) \cdots (x-\lambda_n)$, then $h^\tau(x) = h(x)$, for all $\tau \in G(\mathbf{Q}(\lambda_1, ..., \lambda_n, \zeta)/\mathbf{Q})$, so $h(x) \in \mathbf{Q}[x]$ and we have the desired factor.

Discussion of Theorem 1: Notice that the fact that (i) is equivalent to (ii) did not require that $\mathbf{Q}(\lambda_1, ..., \lambda_n) \cap \mathbf{Q}(\zeta) = \mathbf{Q}$ for some choice of $\lambda_1, ..., \lambda_n$, as stated in the theorem. Also, Theorem 1 does not hold when $m = 2$, since consider $f(x) = x^3 + 3$ then it is easy to check that $|G(K/\mathbf{Q})| = |G(L/\mathbf{Q})|$ (here $\phi(2) = 1$) and $f(x^2)$ has no factor of degree 3 in $\mathbf{Q}[x]$, (see [2] for a consideration of the Galois group of a polynomial of form $f(x^2)$).

It is not difficult to see that to prove (ii) implies (iii), it would have been sufficient to assume in the hypotheses of the theorem that $\mathbf{Q}(\mu_1, ..., \mu_n) \cap \mathbf{Q}(\zeta) = \mathbf{Q}$.

To prove (iii) implies (ii) in the special case of $m = p$ an odd prime, it again is sufficient to assume $\mathbf{Q}(\mu_1, ..., \mu_n) \cap \mathbf{Q}(\zeta) = \mathbf{Q}$ in the statement of the theorem, though it is necessary to change the argument as follows: we know

$$[\mathbf{Q}(\lambda_1, ..., \lambda_n, \zeta) : \mathbf{Q}] = [\mathbf{Q}(\lambda_1, ..., \lambda_n, \zeta) : \mathbf{Q}(\mu_1, ..., \mu_n, \zeta)]$$
$$\times [\mathbf{Q}(\mu_1, ..., \mu_n, \zeta) : \mathbf{Q}(\mu_1, ..., \mu_n)][\mathbf{Q}(\mu_1, ..., \mu_n) : \mathbf{Q}],$$

where $\lambda_i^p = \mu_i,\ 1 \leq i \leq n$, and λ_i are *any* p^{th} roots of μ_i.

But $|G(K/\mathbf{Q})| = [\mathbf{Q}(\lambda_1, ..., \lambda_n, \zeta) : \mathbf{Q}]$, $\phi(p) = [\mathbf{Q}(\mu_1, ..., \mu_n, \zeta) : \mathbf{Q}(\mu_1, ..., \mu_n)]$, and $|G(L/\mathbf{Q})| = [\mathbf{Q}(\mu_1, ..., \mu_n) : \mathbf{Q}]$, so that $|G(K/\mathbf{Q})| = \phi(p)|G(L/\mathbf{Q})|$ implies that $\mathbf{Q}(\lambda_1, ..., \lambda_n, \zeta) = \mathbf{Q}(\mu_1, ..., \mu_n, \zeta)$. Thus

$$G = G(\mathbf{Q}(\lambda_1, ..., \lambda_n, \zeta)/\mathbf{Q}(\mu_1, ..., \mu_n)) = G(\mathbf{Q}(\mu_1, ..., \mu_n, \zeta)/\mathbf{Q}(\mu_1, ..., \mu_n)) \cong G(\frac{\mathbf{Q}(\zeta)}{\mathbf{Q}}) ,$$

so G is isomorphic to R_p, the multiplicative group of residue classes modulo p. Moreover G is cyclic, and let us say is generated by σ, an element of order $\phi(p) = p - 1$. Note that σ is determined by its action $\sigma(\zeta) = \zeta^i$, say, and the fact that it fixes all the $\mu_j,\ 1 \leq j \leq n$.

Let λ_j be a root of the equation $\lambda_j^p = \mu_j,\ (j = 1, 2, ..., n)$.

Claim: σ fixes $\lambda_j\zeta^l$, for some $l = l(j)$ for each $j = 1, 2, ..., n$.

Proof: First, we know $\sigma(\lambda_j) = \lambda_j\zeta^t$, for some $t = t(j),\ j = 1, 2, ..., n$, since $\lambda_j^p = \mu_j$. Let l be the solution of the congruence $(i-1)l \equiv -t \bmod p$. Then

$$\sigma(\lambda_j\zeta^l) = \sigma(\lambda_j)\sigma(\zeta^l) = \lambda_j\zeta^t\zeta^{il} = \lambda_j\zeta^{t+il} = \lambda_j\zeta^l,$$

and we have the desired l, proving the claim.

Since σ generates G we must have that $\lambda_j\zeta^l \in \mathbf{Q}(\mu_1, ..., \mu_n),\ j = 1, 2, ..., n$. Let us denote $\lambda_j\zeta^l$ by λ'_j $(j = 1, 2, ..., n)$. Notice that if $\tau \in G(\mathbf{Q}(\lambda_1, ..., \lambda_n, \zeta)/\mathbf{Q})$, then

$$[\tau(\lambda'_j)]^p = \tau((\lambda'_j)^p) = \tau(\mu_j) = \mu_k,\ \text{for some } k \in \{1, ..., n\}.$$

So $\tau(\lambda'_j) = \lambda'_k\zeta^s$, for some s. Therefore $\lambda'_k\zeta^s \in \mathbf{Q}(\mu_1, ..., \mu_n)$, but $\lambda'_k \in \mathbf{Q}(\mu_1, ..., \mu_n)$, so $\zeta^s \in \mathbf{Q}(\mu_1, ..., \mu_n)$. Hence, $\tau(\lambda'_j) = \lambda'_k$, and we conclude τ just permutes $\lambda'_1, ..., \lambda'_n$. Then if $h(x) = (x - \lambda'_1)\cdots(x - \lambda'_n)$ we must have $h(x) \in \mathbf{Q}[x]$ as before, and $f(x^p)$ has a factor of degree n.

For the following result, where $m = p$ is an odd prime, we retain all the notation and hypotheses from Theorem 1, i.e. $f(x)$ has roots $\mu_i,\ \lambda_i^p = \mu_i,\ 1 \le i \le n$, and ζ is a p^{th} root of unity, $\zeta \ne 1$.

Theorem 2: Let $f(x)$ be an irreducible polynomial of degree n in $\mathbf{Q}[x]$, let p be an odd prime, and suppose $\mathbf{Q}(\lambda_1, ..., \lambda_n) \cap \mathbf{Q}(\zeta) = \mathbf{Q}$ for some choice $\lambda_1, ..., \lambda_n$ of n roots of $f(x^p)$, where $\lambda_i^p = \mu_i,\ 1 \le i \le n$. Then $f(x^p)$ has a factor of degree n in $\mathbf{Q}[x]$ if and only if $G = G(\mathbf{Q}(\lambda_1, ..., \lambda_n, \zeta)/\mathbf{Q}(\mu_1, ..., \mu_n))$ is abelian.

Proof: We already saw in the first part of the proof of Theorem 1 that if $f(x^p)$ has a factor of degree n, with roots $\nu_1, ..., \nu_n$ then $\mathbf{Q}(\nu_1, ..., \nu_n) = \mathbf{Q}(\mu_1, ..., \mu_n)$. We also saw there that $\mathbf{Q}(\nu_1, ..., \nu_n, \zeta) = \mathbf{Q}(\lambda_1, ..., \lambda_n, \zeta)$. Then we have $G = G(\mathbf{Q}(\lambda_1, ..., \lambda_n, \zeta)/\mathbf{Q}(\mu_1, ..., \mu_n)) = G(\mathbf{Q}(\nu_1, ..., \nu_n, \zeta)/\mathbf{Q}(\nu_1, ..., \nu_n)) \cong R_p$, and therefore G is abelian.

Conversely, assume that G is abelian, and take $\lambda_1, ..., \lambda_n$ as stated in the hypotheses of the theorem.

We know that $\mathbf{Q}(\mu_1, ..., \mu_n) \subset \mathbf{Q}(\lambda_1, ..., \lambda_n)$. If $\mathbf{Q}(\mu_1, ..., \mu_n) \neq \mathbf{Q}(\lambda_1, ..., \lambda_n)$ then we also know there exists $\sigma \in G$, and λ_i for some i, $1 \leq i \leq n$, such that $\sigma(\lambda_i) = \zeta^s \lambda_i$, where p does not divide s (since λ_i^p is left fixed by σ).

Let $\tau \in G(\mathbf{Q}(\lambda_1, ..., \lambda_n, \zeta)/\mathbf{Q}(\lambda_1, ..., \lambda_n))$ be such that $\tau(\zeta) = \zeta^t$, where p does not divide $t-1$, then

$$\sigma\tau(\lambda_i) = \sigma(\lambda_i) = \zeta^s \lambda_i$$

$$\text{and} \quad \tau\sigma(\lambda_i) = \tau(\zeta^s \lambda_i) = \tau(\zeta)^s \lambda_i = \zeta^{st} \lambda_i \ .$$

If now $\zeta^s \lambda_i = \zeta^{st} \lambda_i$ then $\zeta^{s(t-1)} = 1$, but then p divides $s(t-1)$. This contradiction would imply $\sigma\tau \neq \tau\sigma$, for some σ, $\tau \in G$, and then G would be non-abelian. So we must have that $\mathbf{Q}(\lambda_1, ..., \lambda_n) = \mathbf{Q}(\mu_1, ..., \mu_n)$.

Now we can proceed as in Theorem 1, however we give another argument:

$$\begin{array}{c} \mathbf{Q}(\mu_1, ..., \mu_n) = \mathbf{Q}(\lambda_1, ..., \lambda_n) \\ | \\ \mathbf{Q}(\lambda_i) \\ | \\ \mathbf{Q}(\mu_i) \\ | \\ \mathbf{Q} \end{array}$$

In the tower of fields above we know $\mathbf{Q}(\mu_i) \subset \mathbf{Q}(\lambda_i)$. If $\mathbf{Q}(\mu_i) \neq \mathbf{Q}(\lambda_i)$, then there exists $\rho \in G(\mathbf{Q}(\mu_1, ..., \mu_n)/\mathbf{Q}(\mu_i))$ such that $\rho(\lambda_i) = \zeta^r \lambda_i$, where p does not divide r. But $\lambda_i \in \mathbf{Q}(\mu_1, ..., \mu_n)$ implies $\rho(\lambda_i) \in \rho(\mathbf{Q}(\mu_1, ..., \mu_n)) = \mathbf{Q}(\mu_1, ..., \mu_n)$, then $\zeta^r \lambda_i \in \mathbf{Q}(\mu_1, ..., \mu_n) = \mathbf{Q}(\lambda_1, ..., \lambda_n)$ so $\zeta^r \in \mathbf{Q}(\lambda_1, ..., \lambda_n)$. But we assumed $\mathbf{Q}(\lambda_1, ..., \lambda_n) \cap \mathbf{Q}(\zeta) = \mathbf{Q}$. So we must have that $\mathbf{Q}(\lambda_i) = \mathbf{Q}(\mu_i)$. Also, $[\mathbf{Q}(\lambda_i) : \mathbf{Q}] = [\mathbf{Q}(\lambda_i) : \mathbf{Q}(\mu_i)][\mathbf{Q}(\mu_i) : \mathbf{Q}]$, so $[\mathbf{Q}(\lambda_i) : \mathbf{Q}] = n$, the degree of $\mathrm{Irr}(\lambda_i, \mathbf{Q}, x)$ is n and $\mathrm{Irr}(\lambda_i, \mathbf{Q}, x)$ divides $f(x^p)$ so we have proved the theorem.

We would like to know if the last theorem can be improved upon by allowing $p = 2$ or $\mathbf{Q}(\lambda_1, ..., \lambda_n) \cap \mathbf{Q}(\zeta) \neq \mathbf{Q}$. Thus we ask whether $G(\mathbf{Q}(\lambda_1, ..., \lambda_n, \zeta)/\mathbf{Q}(\mu_1, ..., \mu_n))$ being abelian forces $f(x^p)$ to have a factor of degree n in $\mathbf{Q}[x]$.

The answer is seen to be no, by considering the following two counterexamples:

For $p = 2$ take $f(x) = x^2 - 2$ with $\zeta = -1$ then we find that $G(\mathbf{Q}(\lambda_1, \lambda_2, \zeta)/\mathbf{Q}(\mu_1, \mu_2))$ $= G(\mathbf{Q}(\sqrt[4]{2}, i)/\mathbf{Q}(\sqrt{2}))$ is abelian, but $x^4 - 2$ has no factor of degree 2 in $\mathbf{Q}[x]$.

Take $f(x) = x^2 + 3$ where $m = p = 3$ and $\zeta = \frac{-1+\sqrt{-3}}{2}$ then $G(\mathbf{Q}(\lambda_1, \lambda_2, \zeta)/\mathbf{Q}(\mu_1, \mu_2))$ $= G(\mathbf{Q}(\sqrt[6]{-3}, \zeta)/\mathbf{Q}(\sqrt{-3}))$ is abelian, but $x^6 + 3$ has no factor of degree 2 in $\mathbf{Q}[x]$, and note that $\mathbf{Q}(\lambda_1, \lambda_2) \cap \mathbf{Q}(\zeta) \neq \mathbf{Q}$.

Theorem 2 also does not extend to the non-prime case. We see this when we take $m = 4$ and $f(x) = x^2 - 14x + 1$. Then $x^8 - 14x^4 + 1$ has no factor of degree 2 although $G(\mathbf{Q}(\lambda_1, \lambda_2, \zeta)/\mathbf{Q}(\mu_1, \mu_2)) = G(\mathbf{Q}(\sqrt[4]{7 + 4\sqrt{3}}, \sqrt[4]{7 - 4\sqrt{3}}, i)/\mathbf{Q}(\sqrt{3}))$ is abelian.

Acknowledgements

These results are taken from the author's Ph.D. thesis at University College Dublin, Ireland, where his advisor Thomas J. Laffey contributed the proposition at the beginning.

References

[1] E. C. Evard and F. Uhlig, On the Matrix Equation $f(X) = A$, *Linear Algebra Appl.,* 162\164:447–517 (1992).

[2] R. Gow, Construction of Some Wreath Products as Galois Groups of Normal Real Extensions of the Rationals, *Journal of Number Theory,* 24:360–372 (1986).

[3] S. Lang, *Algebra,* Addison-Wesley, Menlo Park, California, 1984.

[4] D. E. Otero, Extraction of mth Roots in Matrix Rings over Fields, *Linear Algebra Appl.,* 128:1–26 (1990).