A Galois Approach to $m$th 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. $A \in M_n(\mathbb{Q})$, 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 $m$th root $B \in M_n(\mathbb{Q})$, 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 First we prove a proposition due to T. J. Laffey and B. Cain, previously unpublished, and which provides the motivation for what follows.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 $F$ be a field and $A \in M_n(F)$ have irreducible characteristic polynomial $f(x)$. Let $g(x) \in F[x]$. Then the equation $g(B) = A$ is solvable for $B \in M_n(F)$ if and only if $f(g(x))$ has a factor of degree $n$ in $F[x]$.

Proof: Suppose such a $B$ exists and let $m(x)$ be its minimal polynomial. Since $F[B]$ contains $F[A]$, $m(x)$ 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 $T \in GL(n, F)$. But then $g(T^{-1}CT) = A$, and so take $B = T^{-1}CT$. 

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

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

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

(iii) \( |G(K/Q)| = \phi(m)|G(L/Q)| \),

where \( \phi(\cdot) \) is Euler’s \( \phi \)-function, \( K \) is the splitting field for \( f(x^m) \) over \( Q \) and \( L \) is the splitting field for \( f(x) \) over \( Q \).

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

To prove that (ii) implies (iii), let \( h(x) \in Q[x] \) 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 \( \nu_i \in Q \), \( 1 \leq i \leq n \).

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

Thus \( f(x^m) = (x^m - \nu_1^m) \cdots (x^m - \nu_n^m) = (x - \nu_1)(x - \zeta \nu_1) \cdots (x - \zeta^{m-1} \nu_1) \) \( (x - \nu_2)(x - \zeta \nu_2) \cdots (x - \zeta^{m-1} \nu_2) \) \( \vdots \) \( (x - \nu_n)(x - \zeta \nu_n) \cdots (x - \zeta^{m-1} \nu_n) \).

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

and

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

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

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

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

\[ |G(K/\mathbb{Q})| = \phi(m)|G(L/\mathbb{Q})| \], which is (iii).

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

Now consider the tower of fields:

\[ \mathbb{Q}(\lambda_1, \ldots, \lambda_n, \zeta) \]
\[ \mathbb{Q}(\lambda_1, \ldots, \lambda_n) \]
\[ \mathbb{Q}(\mu_1, \ldots, \mu_n) \]
\[ \mathbb{Q} \]

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

Therefore \( \mathbb{Q}(\lambda_1, \ldots, \lambda_n) \) is a Galois extension of \( \mathbb{Q} \), and \( \tau(\mathbb{Q}(\lambda_1, \ldots, \lambda_n)) = \mathbb{Q}(\lambda_1, \ldots, \lambda_n) \), for all \( \tau \in G(\mathbb{Q}(\lambda_1, \ldots, \lambda_n, \zeta)/\mathbb{Q}) \). We know \( f^\tau(x^m) = f(x^m) \), since all the coefficients of \( f(x) \) are in \( \mathbb{Q} \), then by unique factorization in \( \mathbb{Q}[x] \) we know \( \tau \) just permutes the roots of
$f(x^m)$. But $\tau$ must also just permute $\lambda_1, ..., \lambda_n$ since if $\tau(\lambda_i) = \zeta^s \lambda_j \in \mathbb{Q}(\lambda_1, ..., \lambda_n)$ then $\zeta^s \in \mathbb{Q}(\lambda_1, ..., \lambda_n)$, but $\mathbb{Q}(\lambda_1, ..., \lambda_n) \cap \mathbb{Q}(\zeta) = \mathbb{Q}$, 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 $\tau \in G(\mathbb{Q}(\lambda_1, ..., \lambda_n, \zeta)/\mathbb{Q})$, so $h(x) \in \mathbb{Q}[x]$ and we have the desired factor.

**Discussion of Theorem 1:** Notice that the fact that (i) is equivalent to (ii) did not require that $\mathbb{Q}(\lambda_1, ..., \lambda_n) \cap \mathbb{Q}(\zeta) = \mathbb{Q}$ 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 $|G(K/\mathbb{Q})| = |G(L/\mathbb{Q})|$ (here $\phi(2) = 1$) and $f(x^2)$ has no factor of degree 3 in $\mathbb{Q}[x]$, (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 $\mathbb{Q}(\mu_1, ..., \mu_n) \cap \mathbb{Q}(\zeta) = \mathbb{Q}$.

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

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

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

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

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

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 $\sigma$, an element of order $\phi(p) = p - 1$. Note that $\sigma$ 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 $\lambda_j$ be a root of the equation $\lambda_j^p = \mu_j$, $(j = 1, 2, ..., n)$.
Claim: $\sigma$ fixes $\lambda_j \zeta^t$, for some $l = l(j)$ for each $j = 1, 2, \ldots, n$.

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

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

and we have the desired $l$, proving the claim.

Since $\sigma$ generates $G$ we must have that $\lambda_j \zeta^t \in \mathbb{Q}(\mu_1, \ldots, \mu_n)$, $j = 1, 2, \ldots, n$. Let us denote $\lambda_j \zeta^t$ by $\lambda'_j$ ($j = 1, 2, \ldots, n$). Notice that if $\tau \in G(\mathbb{Q}(\lambda_1, \ldots, \lambda_n, \zeta)/\mathbb{Q})$, then

$$[\tau(\lambda'_j)]^p = \tau((\lambda'_j)^p) = \tau(\mu_j) = \mu_k,$$

for some $k \in \{1, \ldots, n\}$. So $\tau(\lambda'_j) = \lambda'_k \zeta^s$, for some $s$. Therefore $\lambda'_k \zeta^s \in \mathbb{Q}(\mu_1, \ldots, \mu_n)$, but $\lambda'_k \in \mathbb{Q}(\mu_1, \ldots, \mu_n)$, so $\zeta^s \in \mathbb{Q}(\mu_1, \ldots, \mu_n)$. Hence, $\tau(\lambda'_j) = \lambda'_k$, and we conclude $\tau$ just permutes $\lambda'_1, \ldots, \lambda'_n$. Then if $h(x) = (x - \lambda'_1) \cdots (x - \lambda'_n)$ we must have $h(x) \in \mathbb{Q}[x]$ 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 $\mu_i$, $\lambda'_i = \mu_i$, $1 \leq i \leq n$, and $\zeta$ is a $p^{th}$ root of unity, $\zeta \neq 1$.

Theorem 2: Let $f(x)$ be an irreducible polynomial of degree $n$ in $\mathbb{Q}[x]$, let $p$ be an odd prime, and suppose $\mathbb{Q}(\lambda_1, \ldots, \lambda_n) \cap \mathbb{Q}(\zeta) = \mathbb{Q}$ for some choice $\lambda_1, \ldots, \lambda_n$ of $n$ roots of $f(x^p)$, where $\lambda'_i = \mu_i$, $1 \leq i \leq n$. Then $f(x^p)$ has a factor of degree $n$ in $\mathbb{Q}[x]$ if and only if $G = G(\mathbb{Q}(\lambda_1, \ldots, \lambda_n, \zeta)/\mathbb{Q}(\mu_1, \ldots, \mu_n))$ 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 $\nu_1, \ldots, \nu_n$ then $\mathbb{Q}(\nu_1, \ldots, \nu_n) = \mathbb{Q}(\mu_1, \ldots, \mu_n)$. We also saw there that $\mathbb{Q}(\nu_1, \ldots, \nu_n, \zeta) = \mathbb{Q}(\lambda_1, \ldots, \lambda_n, \zeta)$. Then we have $G = G(\mathbb{Q}^\ast(\lambda_1, \ldots, \lambda_n, \zeta)/\mathbb{Q}^\ast(\mu_1, \ldots, \mu_n)) = G(\mathbb{Q}(\nu_1, \ldots, \nu_n, \zeta)/\mathbb{Q}(\nu_1, \ldots, \nu_n)) \cong R_p$, and therefore $G$ is abelian.

Conversely, assume that $G$ is abelian, and take $\lambda_1, \ldots, \lambda_n$ as stated in the hypotheses of the theorem.
We know that $\mathbb{Q}(\mu_1, \ldots, \mu_n) \subset \mathbb{Q}(\lambda_1, \ldots, \lambda_n)$. If $\mathbb{Q}(\mu_1, \ldots, \mu_n) \neq \mathbb{Q}(\lambda_1, \ldots, \lambda_n)$ then we also know there exists $\sigma \in G$, and $\lambda_i$ for some $i$, $1 \leq i \leq n$, such that $\sigma(\lambda_i) = \zeta^s \lambda_i$, where $p$ does not divide $s$ (since $\lambda_i^p$ is left fixed by $\sigma$).

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

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

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

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 $\sigma$, $\tau \in G$, and then $G$ would be non-abelian. So we must have that $\mathbb{Q}(\lambda_1, \ldots, \lambda_n) = \mathbb{Q}(\mu_1, \ldots, \mu_n)$.

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

$$\mathbb{Q}(\mu_1, \ldots, \mu_n) = \mathbb{Q}(\lambda_1, \ldots, \lambda_n)$$

$$\mathbb{Q}(\mu_i) \quad \mathbb{Q}(\lambda_i)$$

$$\mathbb{Q}$$

In the tower of fields above we know $\mathbb{Q}(\mu_i) \subset \mathbb{Q}(\lambda_i)$. If $\mathbb{Q}(\mu_i) \neq \mathbb{Q}(\lambda_i)$, then there exists $\rho \in G(\mathbb{Q}(\mu_1, \ldots, \mu_n)/\mathbb{Q}(\mu_i))$ such that $\rho(\lambda_i) = \zeta^r \lambda_i$, where $p$ does not divide $r$. But $\lambda_i \in \mathbb{Q}(\mu_1, \ldots, \mu_n)$ implies $\rho(\lambda_i) \in \rho(\mathbb{Q}(\mu_1, \ldots, \mu_n)) = \mathbb{Q}(\mu_1, \ldots, \mu_n)$, then $\zeta^r \lambda_i \in \mathbb{Q}(\mu_1, \ldots, \mu_n) = \mathbb{Q}(\lambda_1, \ldots, \lambda_n)$ so $\zeta^r \in \mathbb{Q}(\lambda_1, \ldots, \lambda_n)$. But we assumed $\mathbb{Q}(\lambda_1, \ldots, \lambda_n) \cap \mathbb{Q}(\zeta) = \mathbb{Q}$. So we must have that $\mathbb{Q}(\lambda_i) = \mathbb{Q}(\mu_i)$. Also, $[\mathbb{Q}(\lambda_i) : \mathbb{Q}] = [\mathbb{Q}(\lambda_i) : \mathbb{Q}(\mu_i)][\mathbb{Q}(\mu_i) : \mathbb{Q}]$, so $[\mathbb{Q}(\lambda_i) : \mathbb{Q}] = n$, the degree of $\text{Irr}(\lambda_i, \mathbb{Q}, x)$ is $n$ and $\text{Irr}(\lambda_i, \mathbb{Q}, x)$ 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 $\mathbb{Q}(\lambda_1, \ldots, \lambda_n) \cap \mathbb{Q}(\zeta) \neq \mathbb{Q}$. Thus we ask whether $G(\mathbb{Q}(\lambda_1, \ldots, \lambda_n, \zeta)/\mathbb{Q}(\mu_1, \ldots, \mu_n))$ being abelian forces $f(x^p)$ to have a factor of degree $n$ in $\mathbb{Q}[x]$.

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 $G(\mathbb{Q}(\lambda_1, \lambda_2, \zeta)/\mathbb{Q}(\mu_1, \mu_2)) = G(\mathbb{Q}(\sqrt{2}, i)/\mathbb{Q}(\sqrt{2}))$ is abelian, but $x^4 - 2$ has no factor of degree 2 in $\mathbb{Q}[x]$.

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

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 $G(\mathbb{Q}(\lambda_1, \lambda_2, \zeta)/\mathbb{Q}(\mu_1, \mu_2)) = G(\mathbb{Q}(\sqrt[4]{7} + 4\sqrt[4]{3}, \sqrt[4]{7} - 4\sqrt[4]{3}, i)/\mathbb{Q}(\sqrt{3}))$ 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


