Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 40, pp. 564-573, August 2024.

RATIONAL SOLUTIONS OF THE MATRIX EQUATION $p(X) = A^*$

G.J. GROENEWALD[†], G. GOOSEN[†], D.B. JANSE VAN RENSBURG[†],
A.C.M. RAN[‡], AND M. THIERSEN[†]

Abstract. We extend Theorem 1 of R. Reams, A Galois approach to m-th roots of matrices with rational entries, *LAA*, 258:187–194, 1997. Let $p(\lambda)$ be any polynomial over $\mathbb{Q}$, and let $A \in M_n(\mathbb{Q})$ have irreducible characteristic polynomial $f(\lambda)$ with degree n. We provide necessary and sufficient conditions for the existence of a solution $X \in M_n(\mathbb{Q})$ of the polynomial matrix equation $p(X) = A$. Specifically, we find necessary and sufficient conditions for $f(p(\lambda))$ to have a factor of degree n over $\mathbb{Q}$.

Key words. Matrix polynomial equation, Rational solution, Nonderogatory matrix.

AMS subject classifications. 15A20, 15A24, 15B33, 12D05, 12F10.

1. Introduction.

Let A be an $n \times n$ matrix with rational entries, let $f(\lambda)$ denote the characteristic polynomial of A, and suppose further that $f(\lambda)$ is irreducible. Let $p(\lambda)$ be a polynomial of degree m with rational coefficients. We are interested in finding rational solutions X of the matrix equation $p(X) = A$.

This problem has been considered in the literature before. In [15], a purely algebraic approach is taken for the special case $p(\lambda) = \lambda^m$ with m odd. A more constructive approach for general $p(\lambda)$ was taken in [3] in case $f(\lambda)$ is irreducible. In [9], the condition that $f(\lambda)$ is irreducible is relaxed to A being nonderogatory, following a constructive approach very much related to the one in [3]. Another, more numerical approach can be found in [5, 6, 7]. Allowing for any complex solution, the problem of finding solutions to $p(X) = A$, with $p(\lambda)$ being a holomorphic function, was considered in [4]; see also [18, 19].

The special case where $p(\lambda) = \lambda^m$, that is, the case where X is an mth root of A, has been studied in detail in several papers; see for example [10, 12, 14, 21, 22]. The case where the additional symmetry of H-selfadjointness is involved is treated in [8].

The goal of the paper is as follows: given a polynomial $p(\lambda)$ with rational coefficients, and given an $n \times n$ rational matrix A, find conditions in terms of the characteristic polynomial $f(\lambda)$ of A and the polynomial $p(\lambda)$ for the existence of a rational solution X. One such a condition is mentioned in a paper by Robert Reams [15]; he attributes this result to previously unpublished work by Tom Laffey and Bryan Cain. They showed that existence of a rational solution of $p(X) = A$ is equivalent to the polynomial $f(p(\lambda))$ having a factor $h(\lambda)$ of degree n in $\mathbb{Q}[\lambda]$, where $\mathbb{Q}[\lambda]$ denotes the ring of polynomials in the indeterminate λ with rational coefficients.

Reams considers the special case $p(\lambda) = \lambda^m$, where m is an odd integer greater than 2, subject to the following extra condition. Let $\mu_1, \ldots, \mu_n$ be the eigenvalues of A (or, equivalently, the roots of $f(\lambda)$).

*Received by the editors on May 9, 2024. Accepted for publication on July 30, 2024. Handling Editor: Heike Fassbender. Corresponding Author: Gerrit Goosen.

[†]School of Mathematical and Statistical Sciences, North-West University, Research Focus: Pure and Applied Analytics, Private Bag X6001, Potchefstroom 2520, South Africa (gilbert.groenewald@nwu.ac.za, gerrit.goosen@mandela.ac.za, dawie.jansevanrensburg@nwu.ac.za, madelein.thiersen@nwu.ac.za).

[‡]Department of Mathematics, Faculty of Science, VU Amsterdam, De Boelelaan 1111, 1081 HV Amsterdam, The Netherlands and Research Focus: Pure and Applied Analytics, North-West University, Potchefstroom, South Africa (a.c.m.ran@vu.nl)

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 40, pp. 564-573, August 2024.

Introduce the splitting fields K of $f(\lambda^m)$ and L of $f(\lambda)$ over $\mathbb{Q}$. For some choice of $\gamma_1, \ldots, \gamma_n$, where $\gamma_i^m = \mu_i$, suppose that $\mathbb{Q}(\gamma_1, \ldots, \gamma_n) \cap \mathbb{Q}(e^{2\pi i/m}) = \mathbb{Q}$, where $\mathbb{Q}(\gamma_1, \ldots, \gamma_n)$ denotes the field extension of $\mathbb{Q}$ obtained by adjoining $\gamma_1, \ldots, \gamma_n$ to $\mathbb{Q}$. Under these conditions, Reams shows that there is a rational matrix X such that $X^m = A$ if and only if there is the following relation between the orders of the two related Galois groups: $|Gal(K : \mathbb{Q})| = \phi(m)|Gal(L : \mathbb{Q})|$, where $\phi(\cdot)$ is Euler's ϕ-function.

In Section 3, we expand on the condition of Laffey and Cain and connect it to a condition found in the paper [3] by Michael P. Drazin, which results in an explicit construction of the factor $h(\lambda)$. It will be shown that for a matrix A with irreducible characteristic polynomial and for general $p(\lambda)$, there is an additional condition on solvability of a scalar polynomial equation which is equivalent to the solvability of $p(X) = A$ with a rational solution X. This replaces the condition of Reams on the orders of the Galois groups and holds in more generality. Section 4 extends the result to the case where A is a simple matrix, that is, the case where A has n distinct eigenvalues.

In [9, Theorem 7.1], the equation $p(X) = A$ is considered using a predominantly linear algebraic approach, leading to a result for nonderogatory matrices A. The starting point for this approach is a construction given in [3], and if a rational solution exists it can be constructed explicitly following the approach in [9]. Finally, we restate the existence result in a purely algebraic way in Section 5.

2. Preliminaries. We shall use the following notation: the eigenvalues of A (and hence the roots of $f(\lambda)$) are denoted by μ_i, $i = 1, \ldots, n$, and in case we have $h(\lambda)$ given, then the roots of $h(\lambda)$ are denoted by γ_i, $i = 1, \ldots, n$. Here and in the sequel, n is a natural number.

We recall in this section several definitions and results for the convenience of the reader. Notation will conform with usual practices in abstract algebra, see, for example, [13]. In particular, we will adopt the convention of referring to roots (or zeroes) of a polynomial $f(\lambda)$, which will also be taken to mean solutions of the polynomial equation $f(\lambda) = 0$.

Let $f(\lambda)$ be a monic polynomial of degree n over $\mathbb{Q}$. Write

$$f(\lambda) = (\lambda - \mu_1)(\lambda - \mu_2) \cdots (\lambda - \mu_n),$$

where $\mu_j \in \overline{\mathbb{Q}}$ for each j, with $\overline{\mathbb{Q}}$ denoting the algebraic closure of $\mathbb{Q}$. The *Galois group of* $f(\lambda)$ over $\mathbb{Q}$ is defined to be the group of automorphisms of $\mathbb{Q}(\mu_1, \cdots, \mu_n)$ which fix $\mathbb{Q}$. We denote this group by $Gal(\mathbb{Q}(\mu_1, \cdots, \mu_n) : \mathbb{Q})$ or simply by G when the context is clear.

In the sequel, we will require our Galois group G to act on a certain set of roots. We therefore state the necessary definitions. Let H be a group and X a set. A *group action* of H on X is a map:

$$\cdot : H \times X \to X,$$

where we write $h \cdot x$ for the value of the map on the pair (h, x), satisfying the following properties:

(i) $e \cdot x = x$ for all $x \in X$, where e is the identity element of H.
(ii) $h_1 \cdot (h_2 \cdot x) = (h_1 h_2) \cdot x$ for all $h_1, h_2 \in H$ and all $x \in X$.

A group H is said to act *transitively* on X if for all $x, y \in X$, there exists an element $h \in H$ such that $h \cdot x = y$.

The following result can be found in [20, Proposition 22.3] or [2, Section 10.10].

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 40, pp. 564-573, August 2024.

THEOREM 2.1. [20, Proposition 22.3] *Let $f(\lambda) \in \mathbb{Q}[\lambda]$ be irreducible and monic. Let $X = \{\mu_1, \ldots, \mu_n\}$ be the roots of $f(\lambda)$, which lie in some fixed algebraic closure of $\mathbb{Q}$. Let G be the Galois group of $f(\lambda)$. Then G acts transitively on X via the action:*

$$g \cdot \mu = g(\mu),$$

where $g \in G$, $\mu \in X$.

We recall here the spectral mapping theorem for the finite dimensional case. For a proof, see, for example, [17, Theorem 8.3].

THEOREM 2.2. [17, Theorem 8.3] *Let V be a finite dimensional vector space over an algebraically closed field $\mathbb{F}$. Let $T : V \to V$ be a linear map with spectrum $\sigma(T)$, and let $p(\lambda) \in \mathbb{F}[\lambda]$. Then*

$$\sigma(p(T)) = p(\sigma(T)) = \{p(\mu) \mid \mu \in \sigma(T)\}.$$

The following proposition is a special case of [3, Proposition 2.3]. The proposition is stated in [3] for nonderogatory matrices. Since f being irreducible implies that A is nonderogatory, the version we state here is a special case.

PROPOSITION 2.3. [3, Proposition 2.3] *Let $A, X \in M_n(\mathbb{Q})$ and $p(\lambda) \in \mathbb{Q}[\lambda]$, and suppose that the characteristic polynomial of A, denoted by $f(\lambda)$, is irreducible and that $A = p(X)$. Then, for each eigenvalue μ_j of A, $j = 1, \ldots, n$, the equation $p(\gamma) = \mu_j$ has at least one solution $\gamma = \gamma_j \in \mathbb{Q}(\mu_j)$.*

3. Main results. The main theorem that we prove in this article is an extension of Theorem 1 of [15] and is stated as follows. In [15], only the case $p(\lambda) = \lambda^m$ with odd m is considered.

THEOREM 3.1. *Let $p(\lambda)$ be any polynomial over $\mathbb{Q}$ and let $A \in M_n(\mathbb{Q})$ have irreducible characteristic polynomial $f(\lambda)$ with degree n. Let $\mu_i, 1 \leq i \leq n$, denote the roots of $f(\lambda)$. Then the following are equivalent:*

(i) $A = p(X)$ has a solution over $\mathbb{Q}$;
(ii) $f(p(\lambda))$ has a factor $h(\lambda)$ of degree n over $\mathbb{Q}$;
(iii) There exist an eigenvalue $\mu \in \sigma(A)$ and an element $\gamma \in \mathbb{Q}(\mu)$ such that $p(\gamma) = \mu$.

It will follow from the proof that the third statement above is also equivalent to: for every eigenvalue $\mu \in \sigma(A)$, there is an element $\gamma \in \mathbb{Q}(\mu)$ such that $p(\gamma) = \mu$.

The implication (i) implies (iii) follows from [3, Proposition 2.3]. We will provide an alternative independent argument in the proof below.

Proof. The equivalence of (i) and (ii) is already stated as [15, Proposition 1], which attributes the result to T.J. Laffey and B. Cain. For completeness' sake, we provide the main ideas of the proof. Assuming (i) holds, let X be a rational solution of $p(X) = A$, let $h(\lambda)$ be the minimal polynomial of X, and let $\gamma_i, i = 1, \ldots, n$ be the eigenvalues of X. By the spectral mapping theorem $p(\gamma_i) = \mu_i$ (after possibly reordering), and since $f(\lambda)$ is irreducible, this means that the γ_i are all different, as are the μ_i. Hence, $h(\lambda)$ is a polynomial of degree n in $\mathbb{Q}[\lambda]$. Moreover, by the Cayley–Hamilton theorem $f(A) = f(p(X)) = 0$, and hence $h(\lambda)$ divides $f(p(\lambda))$.

Conversely, suppose $h(\lambda)$ is a polynomial of degree n in $\mathbb{Q}[\lambda]$ which divides $f(p(\lambda))$. Let C_h denote the companion matrix of $h(\lambda)$. Then $p(C_h)$ is similar to A, because $f(p(C_h)) = 0$ and $f(\lambda)$ is irreducible.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 40, pp. 564-573, August 2024.

Hence, there is an invertible rational matrix S such that $S^{-1}p(C_h)S = A$. Take $X = S^{-1}C_hS$, then $p(X) = p(S^{-1}C_hS) = S^{-1}p(C_h)S = A$.

(ii) $\Rightarrow$ (iii). Let $h(\lambda) \in \mathbb{Q}[\lambda]$ be a factor of degree n of $f(p(\lambda))$, and let us say $h(\lambda) = (\lambda-\gamma_1)\cdots(\lambda-\gamma_n)$, where $\gamma_i \in \overline{\mathbb{Q}}$, for $1 \leq i \leq n$.

Then $f(p(\gamma_i)) = 0$, so that $p(\gamma_i) = \mu_{j_i}$ for some $j_i \in \{1, 2, \ldots, n\}$. Hence, $\mu_{j_i} \in \mathbb{Q}(\gamma_i)$.

Denote by $[\mathbb{Q}(\gamma_i) : \mathbb{Q}]$ the degree of the field extension $\mathbb{Q}(\gamma_i)$ over $\mathbb{Q}$. Then, since $n \geq [\mathbb{Q}(\gamma_i) : \mathbb{Q}] = [\mathbb{Q}(\gamma_i) : \mathbb{Q}(\mu_{j_i})] \cdot [\mathbb{Q}(\mu_{j_i}) : \mathbb{Q}]$ and $[\mathbb{Q}(\mu_{j_i}) : \mathbb{Q}] = n$, it follows that $[\mathbb{Q}(\gamma_i) : \mathbb{Q}] = n$, so $h(\lambda) \in \mathbb{Q}[\lambda]$ must be irreducible. By the same argument, $[\mathbb{Q}(\gamma_i) : \mathbb{Q}(\mu_{j_i})] = 1$, and hence $\gamma_i \in \mathbb{Q}(\mu_{j_i})$, so that in fact $\mathbb{Q}(\gamma_i) = \mathbb{Q}(\mu_{j_i})$. This shows part (iii).

(iii) $\Rightarrow$ (ii). Let μ be an eigenvalue of A such that there is a $\gamma \in \mathbb{Q}(\mu)$ with $p(\gamma) = \mu$. Without loss of generality, possibly after renumbering the μ_i, we may assume that this holds for μ_1, and let us denote γ by γ_1.

Since $f(\lambda)$ is irreducible, the Galois group G of $f(\lambda)$ acts transitively on $\{\mu_1, \cdots, \mu_n\}$. That is, by Theorem 2.1, for each $\mu_j \in \sigma(A)$ there is an automorphism g_j in G such that $g_j(\mu_1) = \mu_j$. Define $\gamma_j = g_j(\gamma_1)$, and take

$$h(\lambda) = (\lambda - \gamma_1)(\lambda - \gamma_2)\cdots(\lambda - \gamma_n).$$

Then $p(\gamma_j) = p(g_j(\gamma_1)) = g_j(p(\gamma_1)) = g_j(\mu_1) = \mu_j$, because $p(\lambda)$ has rational coefficients, and g_j fixes $\mathbb{Q}$. Hence $f(p(\gamma_j)) = f(\mu_j) = 0$, and since the μ_j are all different, so are the γ_j. Thus, $h(\lambda)$ is a divisor of $f(p(\lambda))$ of degree n.

It remains to show that $h(\lambda) \in \mathbb{Q}[\lambda]$. We shall show this in a direct manner based on the use of elementary symmetric polynomials, see, for example, [2]. The proof is based on the fact that

$$h(\lambda) = \sum_{j=0}^{n} \lambda^{n-j}(-1)^j e_j(\gamma_1, \ldots, \gamma_n),$$

where $e_j(\gamma_1, \ldots, \gamma_n)$ is the elementary symmetric polynomial of degree j, that is

$$e_j(\gamma_1, \ldots, \gamma_n) = \sum_{1 \leq i_1 < i_2 < \cdots < i_j \leq n} \gamma_{i_1}\gamma_{i_2}\cdots\gamma_{i_j}.$$

In a similar manner,

$$f(\lambda) = (\lambda - \mu_1)\cdots(\lambda - \mu_n) = \sum_{j=0}^{n} \lambda^{n-j}(-1)^j e_j(\mu_1, \ldots, \mu_n),$$

and since we know that $f(\lambda) \in \mathbb{Q}[\lambda]$, we have that $e_j(\mu_1, \ldots, \mu_n) \in \mathbb{Q}$ for $j = 1, \ldots, n$.

Now, since $\gamma_1 \in \mathbb{Q}(\mu_1)$, and $f(\lambda)$ is the minimal polynomial of μ_1 over $\mathbb{Q}$, there are rational numbers $\alpha_0, \ldots, \alpha_{n-1}$ such that

$$\gamma_1 = \alpha_0 + \alpha_1\mu_1 + \alpha_2\mu_1^2 + \cdots + \alpha_{n-1}\mu_1^{n-1}.$$

Applying g_i left and right we obtain

$$\gamma_i = \alpha_0 + \alpha_1\mu_i + \alpha_2\mu_i^2 + \cdots + \alpha_{n-1}\mu_i^{n-1},$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 40, pp. 564-573, August 2024.

for the same $\alpha_0, \ldots, \alpha_{n-1}$. We have to show that $e_j(\gamma_1, \ldots, \gamma_n) \in \mathbb{Q}$ for $j = 0, 1, \ldots, n$. To see this, note that by inserting the formulas for γ_i in terms of μ_i, we have

$$e_j(\gamma_1, \gamma_2 \ldots, \gamma_n) = e_j\left(\sum_{j=0}^{n-1} \alpha_j \mu_1^j, \sum_{j=0}^{n-1} \alpha_j \mu_2^j, \ldots, \sum_{j=0}^{n-1} \alpha_j \mu_n^j\right).$$

This is a symmetric polynomial in $\mu_1, \ldots, \mu_n$ with rational coefficients. By the *fundamental theorem of symmetric polynomials* (see [2, Theorem 5.1]), any symmetric polynomial in $\mu_1, \ldots, \mu_n$ with rational coefficients has a unique representation as a polynomial in $e_1(\mu_1, \ldots, \mu_n), \ldots, e_n(\mu_1, \ldots, \mu_n)$ with rational coefficients. Since the numbers $e_j(\mu_1, \ldots, \mu_n)$ are rational as well, it follows that also the numbers $e_j(\gamma_1, \ldots, \gamma_n)$ are rational. ∎

In Theorem 1 of [15], for the special case $p(\lambda) = \lambda^m$ with $m > 2$, a connection is made between the solvability of $X^m = A$ and the orders of the Galois groups of $f(\lambda^m)$ and $f(\lambda)$. As stated in [15], the equivalence of (i) and (ii) in Theorem 3.1 does not require the condition $\mathbb{Q}(\lambda_1, \ldots, \lambda_n) \cap \mathbb{Q}(\zeta) = \mathbb{Q}$ for some choice of roots $\lambda_1, \ldots, \lambda_n$ of $f(\lambda^m)$, as given in Theorem 1 of [15]. Moreover, Theorem 3.1 does hold when $m = 2$. This is in contrast with the main theorem in [15], which does not hold for $m = 2$. To illustrate this point, consider $f(\lambda) = \lambda^3 + 3$, then the following example shows that $f(\lambda^2) = \lambda^6 + 3$ has no factor $h(\lambda)$ of degree 3 in $\mathbb{Q}[\lambda]$.

Example 3.2. Consider

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -3 & 0 & 0 \end{bmatrix}.$$

The matrix A has characteristic polynomial $f(\lambda) = \lambda^3 + 3$. We take $p(\lambda) = \lambda^2$. The roots of $f(p(\lambda)) = \lambda^6 + 3$ are given by $\gamma_k = \sqrt[6]{3} e^{i(\pi/6 + k\pi/3)}$ for $k = 0, 1, \ldots, 5$. The minimum polynomial of each of these roots is of degree 6. Since $f(p(\lambda))$ is irreducible over $\mathbb{Q}$, this is the minimum polynomial for each of these roots. Hence, for each of these roots we have $[\mathbb{Q}(\gamma_k) : \mathbb{Q}] = 6$. From the proof of Theorem 3.1, this would have to be 3 for the existence of a rational matrix X such that $X^2 = A$. None of the γ_k are in $\mathbb{Q}(\sqrt[3]{3})$ or in $\mathbb{Q}(\sqrt[3]{3} e^{i\pi/3})$ because that would imply that the minimal polynomial of such a γ_k would have degree 3 rather than 6. □

The previous example is a special case of the following proposition.

Proposition 3.3. *Suppose $p(\lambda) \in \mathbb{Q}[\lambda]$ is a monic quadratic polynomial, and that $f(\lambda) \in \mathbb{Q}[\lambda]$ is irreducible and of degree n. Furthermore, assume that for every root μ of $f(\lambda) = 0$ there is a $\gamma \in \mathbb{Q}(\mu)$ such that $p(\gamma) = \mu$. Then the splitting field of $f(p(\lambda))$ over $\mathbb{Q}$ is equal to the splitting field of $f(\lambda)$ over $\mathbb{Q}$.*

Proof. Notice that for every complex number μ there are two solutions γ_1, γ_2 of $p(\lambda) = \mu$. Let $p(\lambda) = \lambda^2 + p_1\lambda + p_0$. Then $\gamma_1\gamma_2 = p_0 - \mu$. Hence, if $\gamma_1 \in \mathbb{Q}(\mu)$, then also $\gamma_2 \in \mathbb{Q}(\mu)$.

Let $\mathbb{Q}(\mu_1, \ldots, \mu_n)$ be the splitting field of $f(\lambda)$ over $\mathbb{Q}$. By assumption, for each μ_i there is at least one $\gamma_{i1} \in \mathbb{Q}(\mu_i)$ such that $p(\gamma_{i1}) = \mu_i$. As argued above, it follows that also the other solution γ_{i2} must be in $\mathbb{Q}(\mu_i)$, so both solutions are in $\mathbb{Q}(\mu_1, \ldots, \mu_n)$. The roots of $f(p(\lambda)) = 0$ are given by the $2n$ solutions of $p(\lambda) = \mu_i$, for $i = 1, \ldots, n$, and therefore $\mathbb{Q}(\gamma_{11}, \gamma_{12}, \gamma_{21}, \gamma_{22}, \ldots, \gamma_{n1}, \gamma_{n2}) \subset \mathbb{Q}(\mu_1, \mu_2, \ldots, \mu_n)$.

Since $\mu_i = p(\gamma_{i1})$ for all $i = 1, \ldots, n$ the other inclusion is evident. ∎

It is clear that a similar argument will fail when $p(\lambda)$ is of degree larger than 2. This is illustrated in the following example.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 40, pp. 564-573, August 2024.

EXAMPLE 3.4. Let $A = \begin{bmatrix} 1 & -2 \\ -4 & 1 \end{bmatrix}$. The characteristic polynomial of A is $f(\lambda) = (\lambda - 1)^2 - 8$, which is irreducible. The splitting field of $f(\lambda)$ is $\mathbb{Q}(\sqrt{2})$; the eigenvalues of A are $1 \pm 2\sqrt{2}$. One checks directly that with $p(\lambda) = \lambda^3 - 4\lambda + 1$ and $X = \begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}$, we have $p(X) = A$. Then $f(p(\lambda)) = (\lambda^3 - 4\lambda)^2 - 8 = \lambda^6 - 8\lambda^4 + 16\lambda^2 - 8$, which factorizes as:

$$f(p(\lambda)) = (\lambda^2 - 2)(\lambda^4 - 6\lambda^2 + 4) = (\lambda^2 - 2)((\lambda^2 - 3)^2 - 5).$$

Hence, the six roots of $f(p(\lambda)) = 0$ are given by $\pm\sqrt{2}, \pm\sqrt{3 \pm \sqrt{5}}$. Thus, the splitting field of $f(p(\lambda))$ is $\mathbb{Q}(\sqrt{2}, \sqrt{3 + \sqrt{5}})$. □

To illustrate the fact that the connection made in [15] between the solvability of $X^m = A$ and the orders of the Galois groups of $f(\lambda^m)$ and $f(\lambda)$ is very specific for the case $p(\lambda) = \lambda^m$ with $m > 2$, we consider in the next example these two Galois groups for a quadratic polynomial $p(\lambda)$.

EXAMPLE 3.5. Let $p(\lambda) = \lambda^2 - \lambda - 1$ and let $f(\lambda) = \lambda^3 + 3\lambda^2 + 21\lambda - 11$. Then $f(\lambda)$ is irreducible over $\mathbb{Q}$. Take

$$X = \begin{bmatrix} 1 & 0 & 2 \\ -1 & 1 & 0 \\ 0 & 3 & 1 \end{bmatrix}.$$

Then $f(\lambda)$ is the characteristic polynomial of $A := p(X)$:

$$A = X^2 - X - I = \begin{bmatrix} -1 & 6 & 2 \\ -1 & -1 & -2 \\ -3 & 3 & -1 \end{bmatrix}.$$

The roots of $f(\lambda)$, calculated using the cubic formula, are

$$\begin{aligned} \mu_1 &= (\sqrt[3]{6})^2 - \sqrt[3]{6} - 1, \\ \mu_2 &= -1 - \frac{1}{2}\left[(\sqrt[3]{6})^2 - \sqrt[3]{6}\right] + \frac{1}{2}i\sqrt{3}\left[(\sqrt[3]{6})^2 + \sqrt[3]{6}\right], \\ \mu_3 &= -1 - \frac{1}{2}\left[(\sqrt[3]{6})^2 - \sqrt[3]{6}\right] - \frac{1}{2}i\sqrt{3}\left[(\sqrt[3]{6})^2 + \sqrt[3]{6}\right]. \end{aligned}$$

The composition is $f(p(\lambda)) = \lambda^6 - 3\lambda^5 + 3\lambda^4 - \lambda^3 + 18\lambda^2 - 18\lambda - 30$ and can be factored as $f(p(\lambda)) = (\lambda^3 - 6)(\lambda^3 - 3\lambda^2 + 3\lambda + 5)$. According to Theorem 3.1, there is a rational solution X to $p(X) = A$. This is the case by construction.

The roots of $f(p(\lambda))$ are as follows:

$$\begin{aligned} \gamma_1 = \sqrt[3]{6}; \gamma_2 = -\frac{1}{2}\sqrt[3]{6} - i\frac{\sqrt{3}}{2}\sqrt[3]{6}; \qquad \gamma_3 = -\frac{1}{2}\sqrt[3]{6} + i\frac{\sqrt{3}}{2}\sqrt[3]{6}; \\ \gamma_4 = -\sqrt[3]{6} + 1; \gamma_5 = 1 + \frac{1}{2}\sqrt[3]{6} + i\frac{1}{2}\sqrt{3}\sqrt[3]{6}; \quad \gamma_6 = 1 + \frac{1}{2}\sqrt[3]{6} - i\frac{1}{2}\sqrt{3}\sqrt[3]{6}. \end{aligned}$$

When these roots are inserted into $p(\lambda)$, we obtain the following: $p(\gamma_1) = \mu_1 = p(\gamma_4)$, $p(\gamma_2) = \mu_2 = p(\gamma_5)$, $p(\gamma_3) = \mu_3 = p(\gamma_6)$. The first three roots are the roots of $\lambda^3 - 6$, while the last three roots are the roots of $\lambda^3 - 3\lambda^2 + 3\lambda + 5$, which are also the eigenvalues of X.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 40, pp. 564-573, August 2024.

It is easy to see that $\mathbb{Q}(\mu_1, \mu_2, \mu_3) = \mathbb{Q}(\sqrt[3]{6}, i\sqrt{3})$. Since $\gamma_i \in \mathbb{Q}(\sqrt[3]{6}, i\sqrt{3})$ for $i = 1, \ldots, 6$, the splitting fields of $f(p(\lambda))$ and $f(\lambda)$ coincide, and hence also the Galois groups $Gal(\mathbb{Q}(\gamma_1, \ldots, \gamma_6) : \mathbb{Q})$ and $Gal(\mathbb{Q}(\mu_1, \mu_2, \mu_3) : \mathbb{Q})$ coincide. □

Connection with a constructive approach of Drazin's paper. Next, we consider how the main result of this paper connects with a more constructive approach which originates in the paper by Drazin, [3]. The setting is the same as above: A is an $n \times n$ matrix with entries in $\mathbb{Q}$, with an irreducible characteristic polynomial $f(\lambda)$, and $p(\lambda)$ is a polynomial with coefficients in $\mathbb{Q}$. We summarize the results of [3]. If $p(X) = A$ has a solution X with entries in $\mathbb{Q}$, for every eigenvalue μ_j of A ($j = 1, 2, \ldots, n$), there is at least one $\gamma_j \in \mathbb{Q}(\mu_j)$ such that $p(\gamma_j) = \mu_j$ (see [3, Proposition 2.3]).

Conversely, let $p(\gamma) = \mu$ for some eigenvalue μ of A and some $\gamma \in \mathbb{Q}(\mu)$. Since $f(\lambda)$ is irreducible, $f(\lambda)$ is the minimum polynomial of μ over $\mathbb{Q}$ and hence $[\mathbb{Q}(\mu) : \mathbb{Q}] = n$. Let $Aw = \mu w$, so w is the eigenvector of A corresponding to eigenvalue μ. Then the entries of w are in $\mathbb{Q}(\mu)$, so there is an $n \times n$ matrix W with entries in $\mathbb{Q}$ such that $Wv_n(\mu) = w$, where $v_n(\mu)$ is the vector $v_n(\mu) = \begin{bmatrix} 1 & \mu & \mu^2 & \cdots & \mu^{n-1} \end{bmatrix}^T$.

Following [3], we can now construct X as follows. From the fact that we want $p(X) = A$, and the fact that A is nonderogatory, it follows from [4] that X is a polynomial in A. Hence, we have $Xw = \gamma w$. The latter equation can also be expressed as $XWv_n(\mu) = \gamma W v_n(\mu)$. Since $\gamma \in \mathbb{Q}(\mu)$, there is an $n \times n$ matrix C with rational entries such that $\gamma W v_n(\mu) = C v_n(\mu)$. It can be shown (see [3]) that W is invertible, so solving X from $XW = C$ produces an $n \times n$ matrix X with rational entries such that $p(X) = A$.

Number of solutions. Introduce the following terminology: a solution γ of $p(\lambda) = \mu_i$ will be called *admissible for* μ_i if $\gamma \in \mathbb{Q}(\mu_i)$. Two admissible solutions γ_r and γ_s for μ_i are called *G-connected* if there exists an element g of G such that $g(\gamma_r) = \gamma_s$, that is, the set of all elements which are G-connected to γ_r is the G-orbit of γ_r.

PROPOSITION 3.6. *Let A be an $n \times n$ matrix over $\mathbb{Q}$ with irreducible characteristic polynomial $f(\lambda)$, and let $p(\lambda)$ be any polynomial over $\mathbb{Q}$. Then for every eigenvalue μ of A the number of admissible γ is the same, and this number equals the number of rational solutions to $p(X) = A$.*

Proof. Let γ_i be an admissible solution of $p(\lambda) = \mu_i$ for $1 \leq i \leq n$. We use a set of n G-connected admissible elements to construct the factor $h(\lambda)$. It cannot occur that two distinct admissible elements γ_i^1 and γ_i^2 associated with μ_i are both G-connected to the same admissible element γ_j associated with μ_j. Indeed, if $g \in G$ fixes some eigenvalue μ_i, then the restriction $g|_{\mathbb{Q}(\mu_i)}$ must be the identity, and hence g must fix each admissible element associated with μ_i since by definition they all lie inside $\mathbb{Q}(\mu_i)$. From this, it also follows that the number of admissible γ associated with an eigenvalue μ_i is the same for every eigenvalue μ_i.

Now the number of rational solutions to $p(X) = A$ is equal to the number of admissible solutions of $p(\lambda) = \mu_i$ (see [3]). Therefore, the number of solutions of $p(X) = A$ is equal to the number of admissible elements associated with any eigenvalue. □

4. The simple case. By using the idea of working separately with the irreducible parts of the characteristic polynomial of a simple matrix and using the companion-Jordan form of the matrix, we can prove the following result. Recall that a matrix is called *simple* if the algebraic multiplicity of each eigenvalue is 1. In particular, when the characteristic polynomial of the matrix is irreducible, then the matrix is simple, but the converse is not always true.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 40, pp. 564-573, August 2024.

PROPOSITION 4.1. *Let $A \in M_n(\mathbb{Q})$ be a simple matrix. Let the characteristic polynomial of A be $f(\lambda) = f_1(\lambda)f_2(\lambda)\cdots f_r(\lambda)$, for some r, where $f_i(\lambda)$ are distinct, irreducible, and of degree n_i. Let $p(\lambda) \in \mathbb{Q}[\lambda]$. Then $p(X) = A$ has a solution $B \in M_n(\mathbb{Q})$ if and only if $f_i(p(\lambda))$ has a factor of degree n_i in $\mathbb{Q}[\lambda]$, for each $1 \le i \le r$.*

Proof. Let $A \in M_n(\mathbb{Q})$ be simple with characteristic polynomial $f(\lambda) = f_1(\lambda)f_2(\lambda)\cdots f_r(\lambda)$, where $f_i(\lambda)$ is irreducible and of degree n_i. By the companion-Jordan form (see [9, Theorem 2.1], also [16]), there exists an invertible matrix $T \in M_n(\mathbb{Q})$ such that

$$A = T^{-1}(C_1 \oplus \cdots \oplus C_r)T, \tag{4.1}$$

where C_i is the $n_i \times n_i$ companion matrix of the polynomial $f_i(\lambda)$. Here, $C_1 \oplus \cdots \oplus C_r$ denotes the block diagonal matrix with block (matrix) entries $C_1, \cdots, C_r$.

Suppose $p(X) = A$ has a solution $B \in M_n(\mathbb{Q})$. Then by [9, Proposition 5.1], B is of the form:

$$B = T^{-1}(B_1 \oplus \cdots \oplus B_r)T,$$

where the sizes of B_i correspond to those of C_i. Now, from $p(B) = A$ we obtain r different equations $p(B_i) = C_i$, $1 \le i \le r$, since we can write

$$\begin{aligned} T^{-1}p\left(B_1 \oplus \cdots \oplus B_r\right)T &= p\left(T^{-1}(B_1 \oplus \cdots \oplus B_r)T\right) = p(B) = A \\ &= T^{-1}\left(C_1 \oplus \cdots \oplus C_r\right)T. \end{aligned}$$

It is easy to check the facts that $p(X_1 \oplus \cdots \oplus X_r) = p(X_1) \oplus \cdots \oplus p(X_r)$ and $T^{-1}p(X)T = p(T^{-1}XT)$ for any polynomial $p(\lambda)$ and square matrices X_i and X. Remember that C_i has irreducible characteristic polynomial $f_i(\lambda)$. Therefore, by [15, Proposition 1], or our main Theorem 3.1 above, $f_i(p(\lambda))$ has a factor of degree n_i in $\mathbb{Q}[\lambda]$ for all $1 \le i \le r$.

Conversely, let $f_i(p(\lambda))$ have a factor of degree n_i in $\mathbb{Q}[\lambda]$ for all $1 \le i \le r$. Then, again by [15, Proposition 1] or the main Theorem 3.1 above, $p(X) = C_i$ has a solution with C_i as in (4.1), that is, there is a $B_i \in M_{n_i}(\mathbb{Q})$ such that $p(B_i) = C_i$. Hence, we have $p(B_1 \oplus \cdots \oplus B_r) = C_1 \oplus \cdots \oplus C_r$ and then

$$\begin{aligned} p(B) &= p\left(T^{-1}(B_1 \oplus \cdots \oplus B_r)T\right) \\ &= T^{-1}p\left(B_1 \oplus \cdots \oplus B_r\right)T \\ &= T^{-1}\left(C_1 \oplus \cdots \oplus C_r\right)T = A. \end{aligned}$$

Thus, $p(X) = A$ has a solution B in $M_n(\mathbb{Q})$. ∎

5. The nonderogatory case. Let A be nonderogatory, and assume that the characteristic polynomial is given by $f(\lambda) = f_1(\lambda)^{d_1} \cdot f_2(\lambda)^{d_2} \cdots f_r(\lambda)^{d_r}$ with the $f_i(\lambda)$'s pairwise coprime and irreducible and of degree k_j. We follow the construction of Theorem 7.1 in [9]. Let $p(\lambda) = \sum_{i=0}^{l} p_i\lambda^i \in \mathbb{Q}[\lambda]$.

Assume that for each eigenvalue μ of A there is a solution $\gamma \in \mathbb{Q}(\mu)$ of $p(\gamma) = \mu$. If μ is a root of $f_j(\lambda)$, let G_j be the Galois group of $f_j(\lambda)$, and consider the factor $h_j(\lambda)$ of degree k_j of $f_j(p(\lambda))$ which we obtain using the explicit construction in the proof of Theorem 3.1, that is, we consider the action of G_j on γ to obtain the k_j roots of $h_j(\lambda)$. Denote these roots by $\gamma_{j,1}, \ldots, \gamma_{j,k_j}$. Now assume that for each j for which $d_j > 1$, there is at least one $\gamma \in \mathbb{Q}(\mu_j)$ such that for each pair of roots $\gamma_{j,s}$ and $\gamma_{j,t}$ we have that condition

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 40, pp. 564-573, August 2024.

(13) in [9] is satisfied, that is:

$$\sum_{m=1}^{l} p_m \sum_{i=0}^{m-1} \gamma_{j,s}^{i} \gamma_{j,t}^{m-1-i} \neq 0.$$

Then, according to Theorem 7.1, part (iii) in [9] there is a rational solution X to $p(X) = A$. This gives a completely algebraic sufficient condition for the existence of a rational solution. However, the condition is not necessary, as pointed out in [9].

Open problems. Theorem 3.1 and its proof may be used to consider several problems that are a variation of the one we discussed in this paper. To be precise, it is good to restate our interest: given a rational matrix A and a polynomial $p(\lambda) \in \mathbb{Q}[\lambda]$, we are interested here and in [9] in necessary and sufficient conditions for the existence of a rational matrix X such that $p(X) = A$ and in an explicit construction. One might also consider the following purely algebraic problem, inspired by Theorem 3.1. Given an irreducible polynomial $f(\lambda) \in \mathbb{Q}[\lambda]$ with degree n, what can be said about the set of polynomials $p(\lambda)$ for which the equivalent conditions (ii) and (iii) in Theorem 3.1 are satisfied? One might restrict this first to consider the set of polynomials of at most a fixed degree.

Alternatively, one might consider the following problem. Given a polynomial $p(\lambda) \in \mathbb{Q}[\lambda]$, find all pairs of rational matrices (X, A) such that $p(X) = A$ with the characteristic polynomial of A irreducible. Note that this is equivalent to the problem of finding irreducible $f(\lambda) \in \mathbb{Q}[\lambda]$ such that the equivalent conditions (ii) and (iii) in Theorem 3.1 are satisfied. The problem may also be considered for pairs (X, A) with A nonderogatory.

The latter problem may be viewed as a special case of a far more difficult problem, which may be described as follows. Given a rational polynomial $R(x, y)$ in two variables, can we find a pair of rational matrices (X, Y) such that $R(X, Y) = 0$? The problem stated earlier amounts to $R(x, y) = y - p(x)$, while the problem studied in recent literature has $R(x, y) = 0$ describing an elliptic curve $y^2 = p(x)$ with p of degree 3. The study of matrix points on elliptic curves over a finite field is an active area of current research. The focus is on counting the number of solutions, see [1, 11].

Acknowledgment. This work is based on research supported in part by the National Research Foundation of South Africa (Grant Number 145688).

REFERENCES

[1] A. Blaser, M. Bradley, D.A.N. Vargas and K. Xing, Sato-Tate type distributions for matrix points on elliptic curves and some K3 surfaces. *J. Number Theory* 260:173–190, 2024.
[2] J. Bewersdorff. Galois Theory for Beginners, A Historical Perspective. *American Mathematical Society*, Providence, RI, 2006.
[3] M.P. Drazin. Exact rational solutions of the matrix equation $A = p(X)$ by linearization. *Linear Algebra Appl.*, 426:502–515, 2007.
[4] J-C. Evard and F. Uhlig. On the matrix equation $f(X) = A$. *Linear Algebra Appl.*, 162–164:447–519, 1992.
[5] F. Fasi and B. Iannazzo. Computing primary solutions of equations involving primary matrix functions. *Linear Algebra Appl.*, 560:17–42, 2019.
[6] F. Fasi and B. Iannazzo. Substitution algorithms for rational matrix equations. *Electron. Trans. Numer. Anal.*, 53:500–521, 2020.
[7] F. Fasi and B. Iannazzo. The dual inverse scaling and squaring algorithm for the matrix logarithm. *IMA J. Numer. Anal.*, 42:2829–2851, 2022.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 40, pp. 564-573, August 2024.

[8] G.J. Groenewald, D.B. Janse van Rensburg, A.C.M. Ran, F. Theron, and M. van Straaten. mth roots of H-selfadjoint matrices. *Linear Algebra Appl.,* 610:804–826, 2021.
[9] G.J. Groenewald, D.B. Janse van Rensburg, A.C.M. Ran, F. Theron, and M. van Straaten. The solutions of the matrix equation $p(X) = A$, with polynomial function $p(\lambda)$ over field extensions of $\mathbb{Q}$. *Linear Algebra Appl,.* 665:107–138, 2023.
[10] N.J. Higham. Functions of Matrices: Theory and Computation. *Society for Industrial and Applied Mathematics*, Philadelphia, PA, USA, 2008.
[11] Y. Huang, K. Ono, and H. Saad. Counting matrix points on certain varieties over finite fields. arXiv:2302.04830.
[12] D.E. Otero. Extraction of mth roots in matrix rings over fields. *Linear Algebra Appl.* 128:1–26, 1990.
[13] C.C. Pinter. *A Book of Abstract Algebra*, Second edition. Dover Publications, New York, 1990.
[14] P.J. Psarrakos. On the mth roots of a complex matrix. *Electron. J. Linear Algebra*, 9:32–41, 2002.
[15] R. Reams. A Galois approach to mth roots of matrices with rational entries. *Linear Algebra Appl.*, 258:187–194, 1997.
[16] D.W. Robinson. The generalized Jordan canonical form. *Amer. Math. Monthly*, 77:392–395, 1970.
[17] S. Roman. *Advanced Linear Algebra*, Third edition. Springer, New York, 2008.
[18] W.E. Roth. A solution of the matric equation $P(X) = A$. *Transact. Amer. Math. Soc.,* 30:579–596, 1928.
[19] E. Spiegel. On the matrix roots of $f(X) = A$. *Indian J. Pure Appl. Math.,* 19:854–864, 1988.
[20] I. Stewart. *Galois Theory*, Second edition. Chapman and Hall, London, 1989.
[21] G. ten Have. Structure of the nth roots of a matrix. *Linear Algebra Appl.*, 187:59–66, 1993.
[22] J.H.M. Wedderburn. *Lectures on Matrices*. American Mathematical Society, New York, 1934.