Affine Transformations of Algebraic Numbers

D.J. Jeffrey
Ontario Research Centre for Computer Algebra
The University of Western Ontario
London, Ontario, Canada
djeffrey@uwo.ca

Pratibha*
Department of Applied Mathematics
The University of Western Ontario
London, Ontario, Canada
pratibhag@rediffmail.com

K.B. Roach
The Symbolic Computation Group
The University of Waterloo
Waterloo, Ontario, Canada
themission@att.net

ABSTRACT

We consider algebraic numbers defined by univariate polynomials over the rationals. In the syntax of MAPLE, such numbers are expressed using the `RootOf` function. This paper defines a canonical form for `RootOf` with respect to affine transformations. The affine shifts of monic irreducible polynomials form a group, and the orbits of the polynomials can be used to define a canonical form. The canonical form of the polynomials then defines a canonical form for the corresponding algebraic numbers. Reducing any `RootOf` to its canonical form has the advantage that affine relations between algebraic numbers are readily identified. More generally, the reduction minimizes the number of algebraic numbers appearing in a computation, and also allows the Maple indexed `RootOf` to be used more easily.

Categories and Subject Descriptors

G.1.5 [**Numerical Analysis**]: Roots of Nonlinear Equations—*Polynomials, methods for*

General Terms

Algorithms

Keywords

Algebraic numbers, RootOf, Affine Transformation

1. INTRODUCTION

We consider univariate polynomials over the field $\mathbb{Q}$ of rational numbers. When MAPLE computes the roots of a polynomial $p(x) \in \mathbb{Q}[x]$, it uses the `RootOf` function to represent any algebraic numbers required. MATHEMATICA uses an equivalent construction. `RootOf` has two forms: indexed and non-indexed. The non-indexed $\texttt{RootOf}(p(x), x)$ represents either one root or all of the roots of $p(x)$, depending on the context, while the newer indexed $\texttt{RootOf}(p(x), x, \text{index} = n)$ represents the nth root of $p(x)$, as determined by the MAPLE indexing scheme. MAPLE's `RootOf` will accept any polynomial $p(x) \in \mathbb{Q}[x]$ as an argument, whether $p(x)$ is irreducible or composite over $\mathbb{Q}$, and indeed it will accept non-polynomial functions. Here, however, the argument $p(x)$ is restricted to being irreducible, because then the indexing scheme is cleaner, there being no repeated roots [4].

The original motivation for the current investigation came from the calculation of series solutions for an ordinary differential equation about a regular singular point, using the method of Frobenius. Della Dora and Tournier [1] point out that computing such solutions requires a test to determine when two roots of the characteristic equation differ by an integer, and comment that this is non-trivial. It is natural to generalize this question to a general affine relationship.

The general question is the following. Given a root r of a polynomial $p(x) \in \mathbb{Q}[x]$, and a root s of a polynomial $q(x) \in \mathbb{Q}[x]$, which may or may not equal $p(x)$, determine whether an affine transformation exists between them, i.e., determine $\alpha, \beta \in \mathbb{Q}$ such that $s = \alpha r + \beta$. This is answered here by expressing each algebraic number $r \notin \mathbb{Q}$ in terms of a uniquely defined algebraic number t such that if $r = \alpha_1 t + \beta_1$ for $\alpha_1, \beta_1 \in \mathbb{Q}$, then there is an affine transformation from $s \notin \mathbb{Q}$ to r only if s is also expressed in terms of t, i.e., $s = \alpha_2 t + \beta_2$ for $\alpha_2, \beta_2 \in \mathbb{Q}$.

Example 1. The aim of this paper is not simply to repair some shortcomings in Maple's current implementation, but it is relevant to see what the situation is at present. We illustrate the shortcomings of MAPLE's current simplification of the `RootOf` function with respect to the questions just posed. Let

$$p(x) = x^{10} - 10x^9 + 40x^8 - 78x^7 + 66x^6 + 14x^5 - 89x^4 + 116x^3 - 106x^2 + 76x - 27\ . \quad (1)$$

The Maple 9.5 command `solve(p(x))` returns the results presented in table 1. (We have abbreviated `RootOf` to $\mathcal{R}$ and omitted the underscore.) It is not obvious from these results that the roots differ by an integer. As a consequence of the properties of the indexing scheme, the indexes must be selected differently for the two families of roots. Once this is done, the following statement is true.

```
R(_Z^5-10*_Z^4+40*_Z^3-79*_Z^2+76*_Z-27,index=1)
 - R(_Z^5+_Z^2+1,index=3)   =  2
```

*Present address: Information Technology Development Agency (ITDA), Government of Uttaranchal,272-B, Phase II, Vasant Vihar, DEHRADUN, INDIA 248 006

ISSAC'05, July 24–27, 2005, Beijing, China.

```
R(Z^5+Z^2+1,index = 1),
R(Z^5+Z^2+1,index = 2),
R(Z^5+Z^2+1,index = 3),
R(Z^5+Z^2+1,index = 4),
R(Z^5+Z^2+1,index = 5),
R(Z^5-10*Z^4+40*Z^3-79*Z^2+76*Z-27,index = 1),
R(Z^5-10*Z^4+40*Z^3-79*Z^2+76*Z-27,index = 2),
R(Z^5-10*Z^4+40*Z^3-79*Z^2+76*Z-27,index = 3),
R(Z^5-10*Z^4+40*Z^3-79*Z^2+76*Z-27,index = 4),
R(Z^5-10*Z^4+40*Z^3-79*Z^2+76*Z-27,index = 5)
```

Table 1: The roots of the polynomial $p(x)$ defined in equation 1 as expressed in `RootOf` notation by Maple 9.5. The function name `RootOf` has been abbreviated to $\mathcal{R}$ to save space.

MAPLE's `simplify`, `evala`, `evalb` commands cannot verify this. However, finding the limitations of particular MAPLE commands is not the point (a sufficiently expert user will be able to guide MAPLE to this simplification). The point is that working with algebraic numbers is more convenient if they are expressed in a canonical form. Notice that two things must be recognized in the above statement: the relation between the polynomials and the indexing of the roots.

2. AFFINE TRANSFORMATIONS OF ALGEBRAIC NUMBERS

Let $\mathbb{P} \subset \mathbb{Q}[x]$ be the set of monic irreducible polynomials over $\mathbb{Q}$. Further, let $\mathbb{P}_n$ be the set of monic irreducible polynomials of degree n. We consider the algebraic numbers defined by the roots of the elements of $\mathbb{P}$. From the point of view of MAPLE, this corresponds to using the output of the `factors` command, rather than the `solve` command.

Definition 1. For $x \in \mathbb{C}$ and $\alpha, \beta \in \mathbb{Q}$, an affine transformation $T(\alpha, \beta)$ of x is defined by

$$T(\alpha, \beta)x = \alpha x + \beta \ .$$

Definition 2. For $p(x) \in \mathbb{P}$, $\alpha \in \mathbb{Q}\backslash\{0\}$, $\beta \in \mathbb{Q}$, a monic affine shift $M(\alpha, \beta)$ of $p(x)$ is defined by

$$M(\alpha, \beta)p(x) = \alpha^{-\deg p} p(T(\alpha, \beta)x) = \alpha^{-\deg p} p(\alpha x + \beta) \ .$$

The inverse transformations are

$$T(\alpha, \beta)^{-1}x = \alpha^{-1}x - \alpha^{-1}\beta \ , \quad (2)$$

$$M(\alpha, \beta)^{-1}p(x) = \alpha^{\deg p} p \ \ T(\alpha, \beta)^{-1}x \quad . \quad (3)$$

We must show that the transformations $M(\alpha, \beta)$ form a group. This is obvious if M is applied to $\mathbb{Q}[x]$, but we apply it only to $\mathbb{P}$. Therefore we first prove that M is closed on $\mathbb{P}$.

THEOREM 1. *Given $p(x) \in \mathbb{P}$, $\alpha \in \mathbb{Q}\backslash\{0\}$, $\beta \in \mathbb{Q}$, then $M(\alpha, \beta)p(x) \in \mathbb{P}$.*

PROOF. First, observe that

$$M(\alpha, \beta)p(x) = \alpha^{-\deg(p)} p(\alpha x + \beta)$$

is monic. Second, suppose $M(\alpha, \beta)p(x) = a(x)b(x)$, where $a(x) \in \mathbb{Q}[x]$ and $b(x) \in \mathbb{Q}[x]$. Then

$$p(x) = \alpha^{\deg(p)} a(\alpha^{-1}x - \alpha^{-1}\beta)b(\alpha^{-1}x - \alpha^{-1}\beta) \ ,$$

where we must have that $a(\alpha^{-1}x - \alpha^{-1}\beta) \in \mathbb{Q}[x]$ and also that $b(\alpha^{-1}x - \alpha^{-1}\beta) \in \mathbb{Q}[x]$. If neither $a(x)$ nor $b(x)$ is a unit of $\mathbb{Q}[x]$ (an element of $\mathbb{Q}$), then neither $a \ \ \alpha^{-1}x - \alpha^{-1}\beta$ nor $b \ \ \alpha^{-1}x - \alpha^{-1}\beta$ is a unit of $\mathbb{Q}[x]$. □

We can in fact make the stronger statement that if $p \in \mathbb{P}_n$ then $M(\alpha, \beta)p \in \mathbb{P}_n$.

THEOREM 2. *The set of monic affine shifts $M(\alpha, \beta)$ with $\alpha \in \mathbb{Q}\backslash\{0\}$, $\beta \in \mathbb{Q}$ is a group action on $\mathbb{P}$.*

PROOF. The action is well defined because of theorem 1, and the verification of the group laws is straightforward. □

Now we establish the correspondence between affine shifts of the polynomials and affine transformations of algebraic numbers.

Definition 3. Let $p(x) \in \mathbb{P}$, then the set $\mathcal{R}(p(x), x) \subset \mathbb{C}$ is defined by

$$x \in \mathcal{R}(p(x), x) \iff p(x) = 0 \ . \quad (4)$$

This corresponds to the non-indexed MAPLE `RootOf`. The number of elements in $\mathcal{R}$ is $\deg p$.

An affine shift does not change the degree of the polynomial, so we must establish that an affine relation between algebraic numbers can only exist between numbers of the same degree. We note that an affine shift of the defining polynomial corresponds to an affine transformation of every root in the set of roots, and the set remains the same size.

LEMMA 1. *Let $p \in \mathbb{P}$ have roots $r_i \notin \mathbb{Q}$. If $q \in \mathbb{P}$ has a root $s \notin \mathbb{Q}$ such that $s = \alpha r_j + \beta$ for some j and $\alpha \in \mathbb{Q}\backslash\{0\}$, $\beta \in \mathbb{Q}$, then the roots of q are precisely $\alpha r_i + \beta$ for all i.*

PROOF. Assume $r \notin \mathbb{Q}$ and $s \notin \mathbb{Q}$ are roots of irreducible polynomials p and q such that $s = \alpha r + \beta$, $\alpha \in \mathbb{Q}$, and $\beta \in \mathbb{Q}$. Then $s \notin \mathbb{Q}$ implies $\alpha \neq 0$ and r is a root of $\alpha^{-\deg(q)} q(\alpha x + \beta) \in \mathbb{Q}[x]$. Since $q(x)$ is irreducible, $\alpha^{-\deg(q)} q(\alpha x + \beta)$ must also be irreducible by theorem 1. But r must be a root of a unique irreducible polynomial. Hence, $p(x) = \alpha^{-\deg(q)} q(\alpha x + \beta)$ which shows that if the roots of p are r_i, then the roots of q are $\alpha r_i + \beta$. □

Our aim is to relate algebraic numbers through their defining polynomials.

Definition 4. If r and s be algebraic numbers, then r and s are affinely related over $\mathbb{Q}$, written $r \equiv_{\mathbb{Q}} s$, if there exist $\alpha \in \mathbb{Q}\backslash\{0\}$, and $\beta \in \mathbb{Q}$ such that $s = T(\alpha, \beta)r$.

LEMMA 2. *Affinely related over $\mathbb{Q}$ is an equivalence relation.*

PROOF. Since the affine transformations define a group, they define a set of orbits of the algebraic numbers. Two algebraic numbers are affinely related if they are members of the same orbit. It is standard theory that orbits define equivalence classes. □

THEOREM 3. *For $p(x) \in \mathbb{P}$, $\alpha \in \mathbb{Q}\backslash\{0\}$ and $\beta \in \mathbb{Q}$,*

$$M(\alpha, \beta)p(x) = 0 \iff T(\alpha, \beta)^{-1}\mathcal{R}(p(x)) \ .$$

PROOF. By definition 2 and lemma 1. □

Theorem 3 shows that algebraic numbers are affinely related if their defining polynomials are related by corresponding affine shifts. Consequently it seems that all affine relations can be deduced by considering the orbits of the defining polynomials. There is a difficulty, however. It is possible for affine relations to exist within a `RootOf` set. This corresponds to an affine polynomial shift mapping a polynomial nontrivially onto itself. It is therefore important to decide when this can occur.

THEOREM 4. *Two different roots $r \notin \mathbb{Q}$ and $s \notin \mathbb{Q}$ of an irreducible polynomial cannot be linearly related over $\mathbb{Q}$ unless $r = T(-1, \beta)s$ for some $\beta \in \mathbb{Q}$.*

PROOF. Assume r and s are different roots of an irreducible polynomial $p(x)$ such that $s = T(\alpha, \beta)r$, $\alpha \in \mathbb{Q}$, and $\beta \in \mathbb{Q}$. Let

$$r_n = \alpha^n + \beta \sum_{i=0}^{n-1} \alpha^i$$

for $n \in \mathbb{N}$. Then $r_0 = r$ is a root of p. By lemma 1, if r_n is a root of p, then $r_{n+1} = \alpha\, r_n + \beta$ is a root of p. So, by induction, r_n is a root of p for all $n \in \mathbb{N}$. Consider the cases $\alpha = 1$ and $\alpha \neq 1$ separately.

If $\alpha = 1$, then $r_n = r + n\beta$ is a root of p for all $n \in \mathbb{N}$. The Fundamental Theorem of Algebra requires $\{r_n \,|\, n \in N\}$ be a finite set. Since $\alpha \in \mathbb{Q}$, $\alpha \neq 0$, and $\alpha \neq 1$, this implies $\beta = 0$ and $s = r$, a contradiction.

If $\alpha \neq 1$, then

$$r_n = \alpha^n \quad r + \frac{\beta}{\alpha - 1} \quad - \frac{\beta}{\alpha - 1}$$

is a root of p for all $n \in \mathbb{N}$. The Fundamental Theorem of Algebra requires $\{r_n | n \in N\}$ be a finite set. This implies $\alpha = -1$ and $s = -r + \beta$. □

Example. The irreducible polynomial $x^2 - 2x - 1$ has roots $r = 1 + \sqrt{2}$ and $s = 1 - \sqrt{2} = -r + 2$.

To summarize this section, we have shown that affine relations between algebraic numbers can be studied by considering the orbits of the defining polynomials under the group action of $T(\alpha, \beta)$. We now must define an invariant for each orbit, and this will be the number t referred to in the introduction. The set of numbers t will define a cross section in the sense of [2, 3], or in equivalent terminology a canonical form. The selection of the cross-section element is the subject of the next section.

3. DEFINITION OF A CANONICAL FORM

Definition 5. Let $p(x) = x^n + p_{n-1}\, x^{n-1} + \ldots + p_0$, where $p(x) \in \mathbb{P}$, be an irreducible polynomial. Define $p_+(x) \in \mathbb{P}$

$$p_+(x) = M(1, -p_{n-1}/n)p(x) \ .$$

If $r = \mathcal{R}(p(x), x)$, then $r_+ \equiv_{\mathbb{Q}} r$ is defined by

$$r_+ = \mathcal{R}(p_+(x), x) = T(1, p_{n-1}/n)r \ .$$

The properties of the + operator are summarized in the following theorem.

THEOREM 5. *If $\alpha \in \mathbb{Q}\backslash\{0\}$, $\beta \in \mathbb{Q}$, $p(x) \in \mathbb{P}$ is an irreducible polynomial, r an algebraic number, and $n = \deg(p)$, then*

(a) $(M(\alpha, \beta)p(x))_+ = M(\alpha, 0)p_+(x)$.

(b) $(T(\alpha, \beta)r)_+ = T(\alpha, 0)r_+$.

(c) $p_{++} = p_+$.

(d) $r_{++} = r_+$.

PROOF. For (a), since

$$\begin{aligned} M(\alpha, \beta)p(x) &= \alpha^{-n}(\alpha x + \beta)^n \\ &\quad + \alpha^{-n} p_{n-1}(\alpha x + \beta)^{n-1} + \ldots \\ &= x^n + \alpha^{-1}(p_{n-1} + n\beta)\, x^{n-1} + \ldots \ , \end{aligned}$$

we have

$$\begin{aligned} (M(\alpha, \beta)p(x))_+ &= \alpha^{-n}\, p \quad \alpha x - \frac{p_{n-1}}{n} - \beta \\ &= \alpha^{-n} \mathrm{p}_+(\alpha x) \ . \end{aligned}$$

For (b), since $\alpha r + \beta$ has irreducible polynomial

$$\begin{aligned} \alpha^n p(\alpha^{-1}x - \alpha^{-1}\beta) &= \alpha^n \ \ \alpha^{-1}x - \alpha^{-1}\beta \ ^n \\ &\quad + \alpha^n p_{n-1} \ \ \alpha^{-1}x - \alpha^{-1}\beta \ ^{n-1} \\ &\quad + \ldots \\ &= x^n + \alpha(p_{n-1} - n\beta)x^{n-1} + \ldots \end{aligned}$$

we get

$$\begin{aligned} (\alpha r + \beta)_+ &= (\alpha r + \beta) + \frac{\alpha(p_{n-1} - n\,\beta)}{n} \\ &= \alpha \ \ r + \frac{p_{n-1}}{n} \ = \alpha r_+ \ . \end{aligned}$$

Parts (c) and (d) follow from definition 5 and parts (a) and (b). □

Definition 6. Suppose $\rho \in \mathbb{Q}\backslash\{0\}$, and n is a positive integer. Let $\rho = a/b$ where $a \in \mathbb{Z}$ and $b \in \mathbb{N}$. Let $a\, b^{n-1} = k^n\, c$ where k is a positive integer, c is a nonzero integer that is not divisible by any nontrivial nth power of an integer (i.e. other than ± 1). Then $[\rho]_n$ is defined by $[\rho]_n = k/b$.

Remark. The a and b do not need to be relatively prime. To check that $[\rho]_n$ is well defined, observe that if $\rho = a_1/b_1 = a_2/b_2$, and

$$a_1 b_1^{n-1} = k_1^n c_1 \qquad \text{and} \qquad a_2 b_2^{n-1} = k_2^n c_2$$

then

$$b_1 k_2^n c_2 = b_1^n a_2 b_2^{n-1} = a_1 b_1^{n-1} b_2^n = b_2^n k_1^n c_1 \ ,$$

but b_1, b_2, k_1, k_2 are positive integers; c_1 and c_2 are nonzero integers; and neither c_1 nor c_2 is divisible by a nontrivial nth power of an integer. Therefore, by the Fundamental Theorem of Arithmetic, we must have $c_1 = c_2$ and $k_1/b_1 = k_2/b_2$.

LEMMA 3. *Suppose $\rho_1 \in \mathbb{Q}$, $\rho_1 \neq 0$, $\rho_2 \in \mathbb{Q}$, $\rho_2 \neq 0$, and n is a positive integer. Then $\mathrm{sgn}([\rho_1]_n) = 1$ and $[\rho_1^n\, \rho_2]_n = |\rho_1|\, [\rho_2]_n$.*

PROOF. First, $\mathrm{sgn}([\rho_1]_n) = 1$ is easily seen from definition 6. Second, let $\rho_i = a_i/b_i$ where $a_i \in \mathbb{Z}$ and $b_i \in \mathbb{N}$. Let $a_2\, b_2^{n-1} = k_2^n\, c_2$ where k_2 is a positive integer; c_2 is a nonzero integer; and c_2 is not divisible by any nontrivial nth power of an integer. Then

$$\rho_1^n \rho_2 = \frac{a_1^n a_2}{b_1^n b_2}$$

and

$$\begin{aligned} a_1^n a_2 (b_1^n b_2)^{n-1} &= a_1^n b_1^{n(n-1)} k_2^n c_2 \\ &= a_1 b_1^{n-1} k_2 \ {}^n \ c_2 \end{aligned}$$

implying

$$[\rho_1^n \ \rho_2]_n = \frac{|\, a_1 \,|\, b_1^{n-1} \, k_2}{b_1^n \, b_2} = \frac{|\, a_1 \,|\, k_2}{b_1 \, b_2} = |\, \rho_1 \,|\, [\rho_2]_n \ .$$

□

Definition 7. If $p(x) \in \mathbb{P}_n$, $n > 1$ is a nonlinear irreducible polynomial and $p_+(x) = x^n + a_{n-2}x^{n-2} + \ldots + a_0$, define irreducible polynomial $p_\times(x) \in \mathbb{P}$ by

$$p_\times(x) = [a_0]_n^{-n} \, p_+ \ \ [a_0]_n \, x$$

Suppose $r \notin \mathbb{Q}$ is a root of $p(x)$. Define root $r_\times \equiv_{\mathbb{Q}} r$ of $p_\times(x)$ by $r_\times = [a_0]_n^{-1} \, r_+$.

THEOREM 6. *If $\alpha \in \mathbb{Q}\backslash\{0\}$, $\beta \in \mathbb{Q}$, $p(x) \in \mathbb{P}_n$, $n > 1$ is a nonlinear irreducible polynomial, $r \notin \mathbb{Q}$ is an algebraic number, and $n = \deg(p)$, then, with $s_\alpha = \operatorname{sgn}\alpha$,*

(a) $(M(\alpha,\beta)p(x))_\times = (s_\alpha)^n p_\times(s_\alpha x)$.

(b) $(T(\alpha,\beta)r)_\times = s_\alpha r_\times$

(c) $p_{\times+} = p_{\times\times} = p_\times$

(d) $r_{\times+} = r_{\times\times} = r_\times$

PROOF. (a) Let $p_+(x) = x^n + a_{n-2}x^{n-2} + \ldots + a_0$. By theorem 5,

$$\alpha^{-n} p(\alpha x + \beta) \ {}_+ = \alpha^{-n} p_+(\alpha x) = x^n + \ldots + \alpha^{-n} a_0 \ .$$

Therefore,

$$\begin{aligned} \frac{p(\alpha x + \beta)}{\alpha^n} \ {}_\times &= \left[\frac{a_0}{\alpha^n}\right]_n^{-n} \alpha^{-n} p_+ \ \ \alpha \ \ \alpha^{-n} a_0 \ {}_n \, x \\ &= |\alpha|^n [a_0]_n^{-n} \alpha^{-n} p_+ \ \ \alpha|\alpha|^{-1} [a_0]_n x \\ &= s_\alpha^n \, [a_0]_n^{-n} \, p_+ \ \ [a_0]_n \, s_\alpha x \\ &= s_\alpha^n p_\times(s_\alpha x) \end{aligned}$$

(b) Let r have irreducible polynomial $p(x)$ and let $p_+(x) = x^n + a_{n-2}x^{n-2} + \ldots + a_0$. Then $\alpha r + \beta$ has irreducible polynomial $\alpha^n p(\alpha^{-1}x - \alpha^{-1}\beta)$. By theorem 5,

$$\alpha^n p(\alpha^{-1}x - \alpha^{-1}\beta) \ {}_+ = \alpha^n p_+(\alpha^{-1}x)$$

and $(\alpha r + \beta)_+ = \alpha r_+$. Therefore by definition 7,

$$\begin{aligned} (\alpha r + \beta)_\times &= [\alpha^n a_0]_n^{-1} \alpha r_+ = |\alpha|^{-1} [a_0]_n^{-1} \alpha r_+ \\ &= s_\alpha r_\times \ . \end{aligned}$$

(c) By construction, $p_\times = [a_0]_n^{-n} p\left([a_0]_n x - p_{n-1}/n\right)$. By theorem 5 and definition 7,

$$\begin{aligned} p_{\times+} &= [a_0]_n^{-n} p \ \ [a_0]_n x - \frac{p_{n-1}}{n} \ {}_+ \\ &= [a_0]_n^{-n} p_+ \left([a_0]_n x\right) = \mathrm{p}_\times(x) \end{aligned}$$

Also, by (a) and lemma 3,

$$p_{\times\times} = \ [a_0]_n^{-n} p \ \ [a_0]_n x - \frac{p_{n-1}}{n} \ {}_\times = p_\times(x)$$

(d) By construction, $r_\times = [a_0]^{-1} r + p_{n-1}/n$. By theorem 5 and definition 7,

$$r_{\times+} = \ [a_0]^{-1} r + p_{n-1}/n \ {}_+ = [a_0]^{-1} r_+ = r_\times$$

Also, by (b) and lemma 3

$$r_{\times\times} = \ [a_0]^{-1} r + p_{n-1}/n \ {}_\times = r_\times$$

□

Definition 8. If $p(x) \in \mathbb{Q}[x]$ and $p(-x) = p(x)$, then $p(x)$ is an **even polynomial**.

Definition 9. The complex signum of a complex number z is

$$\operatorname{csgn}(z) = \begin{cases} 0 \ , & z = 0, \\ 1 \ , & -\pi/2 < \arg z \le \pi/2, \\ -1 & \textit{otherwise.} \end{cases}$$

Definition 10. Let $p(x) \in \mathbb{P}$ be a nonlinear irreducible polynomial, $\mathrm{p}_\times(x) = x^n + a_{n-2}x^{n-2} + \ldots + a_0$. If $p_\times$ is even, define $\sigma(p)$ to be 1. If $p_\times$ is not even, define $\sigma(p)$ to be the sign of the first nonzero coefficient in the sequence a_{n-3}, a_{n-5}, ... Suppose $r \notin \mathbb{Q}$ is a root of $p(x)$. If $p_\times$ is even, define $\sigma(r)$ to be $\operatorname{csgn}(r_\times)$. If $p_\times$ is not even, define $\sigma(r)$ to be $\sigma(p)$.

LEMMA 4. *If $\alpha \in \mathbb{Q}$, $\alpha \neq 0$, $\beta \in \mathbb{Q}$, and $r \notin \mathbb{Q}$ is an algebraic number, then $\sigma(\alpha r + \beta) = s_\alpha \sigma(r)$, and $s_\alpha = \operatorname{sgn}\alpha$ as before.*

PROOF. Let r have irreducible polynomial $p(x) \in \mathbb{P}$ and $p_\times = x^n + a_{n-2}x^{n-2} + \ldots + a_0$. Then $\alpha r + \beta$ has irreducible polynomial $\alpha^n p(\alpha^{-1}x - \alpha^{-1}\beta)$. By theorem 6, $\alpha^n p(\alpha^{-1}x - \alpha^{-1}\beta) \ {}_\times = s_\alpha^n p_\times(s_\alpha x)$ and this equals $x^n + s_\alpha a_{n-1} x^{n-1} + \ldots + s_\alpha^n a_0$.
Therefore, $\alpha^n p(\alpha^{-1}x - \alpha^{-1}\beta) \ {}_\times$ is an even polynomial if and only if $\mathrm{p}_\times(x)$ is an even polynomial. By theorem 6, we know $(\alpha r + \beta)_\times = s_\alpha r_\times$. Hence, if $\alpha^n p(\alpha^{-1}x - \alpha^{-1}\beta) \ {}_\times$ is even, we get

$$\begin{aligned} \sigma(\alpha r + \beta) &= \operatorname{csgn}\left((\alpha r + \beta)_\times\right) = \operatorname{csgn}\left(s_\alpha r_\times\right) \\ &= s_\alpha \operatorname{csgn}(r_\times) = s_\alpha \sigma(r) \end{aligned}$$

If $\alpha^n p(\alpha^{-1}x - \alpha^{-1}\beta) \ {}_\times$ is not an even polynomial, we get

$$\begin{aligned} \sigma(\alpha r + \beta) &= \operatorname{sgn} \ \operatorname{sgn}(\alpha)^{2l+1} a_{n-2l-1} \\ &= \operatorname{sgn}(\alpha) \operatorname{sgn}(a_{n-2l-1}) = s_\alpha \sigma(r) \end{aligned}$$

where a_{n-2l-1} is the first nonzero coefficient in the sequence a_{n-1}, a_{n-3}, a_{n-5}, ... □

Definition 11. Suppose $p(x) \in \mathbb{P}_n$, $n > 1$ is a nonlinear irreducible polynomial. Define irreducible polynomial $p_\sigma(x) \in \mathbb{P}$ by

$$p_\sigma(x) = M(\sigma(p), 0) p_\times(x)$$

Suppose $r \notin \mathbb{Q}$ is a root of $p(x)$. Define root $r_\sigma \equiv_{\mathbb{Q}} r$ of $p_\sigma(x)$ by $r_\sigma = T(\sigma(r), 0) r_\times$.

THEOREM 7. *If $\alpha \in \mathbb{Q}$, $\alpha \neq 0$, $\beta \in \mathbb{Q}$, $p(x) \in \mathbb{P}$ is a nonlinear irreducible polynomial, and $r \notin \mathbb{Q}$ is an algebraic number, then*

(a) $(M(\alpha,\beta)p(x))_\sigma = p_\sigma(x)$

(b) $(T(\alpha,\beta)r)_\sigma = r_\sigma$

(c) $p_{\sigma+} = p_{\sigma\times} = p_{\sigma\sigma} = p_\sigma$

(d) $r_{\sigma+} = r_{\sigma\times} = r_{\sigma\sigma} = r_\sigma$

(e) $\sigma(p_\sigma) = 1$

(f) $\sigma(r_\sigma) = 1$

PROOF. (a) Let $p_\times(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_0$. By theorem 6, $\alpha^{-n}p(\alpha x+\beta)_\times = \mathrm{sgn}(\alpha)^n p_\times(\mathrm{sgn}(\alpha)x)$ and this equals $x^n + \mathrm{sgn}(\alpha)a_{n-1}x^{n-1} + \ldots + \mathrm{sgn}(\alpha)^n a_0$. Therefore, $\alpha^{-n}p(\alpha x+\beta)_\times$ is an even polynomial if and only if $p_\times(x)$ is an even polynomial, and if $\alpha^{-n}p(\alpha x+\beta)_\times$ is an even polynomial, then

$$\sigma\ \alpha^{-n}p(\alpha x+\beta)\ = \sigma(p) = 1$$

and n is even, which makes

$$\begin{aligned}\alpha^{-n}p(\alpha x+\beta)_\sigma &= \alpha^{-n}p(\alpha x+\beta)_\times \\ &= p_\times(x) = p_\sigma(x)\end{aligned}$$

If $\alpha^{-n}p(\alpha x+\beta)_\times$ is not even, then by lemma 4, it must be that $\sigma\ \alpha^{-n}p(\alpha x+\beta) = \mathrm{sgn}(\alpha)\sigma(p)$. For convenient notation, let $\sigma_1 = \sigma\ \alpha^{-n}p(\alpha x+\beta) = \mathrm{sgn}(\alpha)\sigma(p)$. Then

$$\begin{aligned}\alpha^{-n}p(\alpha x+\beta)_\sigma &= \sigma_1^n\ \alpha^{-n}p(\alpha\sigma_1 x+\beta)_\times \\ &= \sigma_1^n \mathrm{sgn}(\alpha)^n p_\times(\mathrm{sgn}(\alpha)\sigma_1 x) \\ &= \sigma(p)^n p_\times(\sigma(p)x) = p_\sigma(x)\end{aligned}$$

(b) By lemma 4, $\sigma(\alpha r+\beta) = \mathrm{sgn}(\alpha)\sigma(r)$ and by theorem 6, $(\alpha r+\beta)_\times = \mathrm{sgn}(\alpha)r_\times$. Therefore,

$$\begin{aligned}(\alpha r+\beta)_\sigma &= \sigma(\alpha r+\beta)(\alpha r+\beta)_\times \\ &= \mathrm{sgn}(\alpha)\sigma(r)\,\mathrm{sgn}(\alpha)r_\times = \sigma(r)r_\times = r_\sigma\end{aligned}$$

(c) By construction,

$$p_\sigma = \sigma(p)^n[a_0]_n^{-n}p\left(\sigma(p)[a_0]_n x - p_{n-1}/n\right)$$

By theorem 5, definition 7, and definition 11,

$$\begin{aligned}p_{\sigma+} &= \sigma(p)^n[a_0]_n^{-n}p\left(\sigma(p)[a_0]_n x - p_{n-1}/n\right)_+ \\ &= \sigma(p)^n[a_0]_n^{-n}p_+\left(\sigma(p)[a_0]_n x\right) \\ &= \sigma(p)^n p_\times(\sigma(p)x) = p_\sigma(x)\end{aligned}$$

By theorem 6, lemma 3, and definition 11,

$$\begin{aligned}p_{\sigma\times} &= \sigma(p)^n[a_0]_n^{-n}p\left(\sigma(p)[a_0]_n x - p_{n-1}/n\right)_\times \\ &= \sigma(p)^n p_\times(\sigma(p)x) = p_\sigma(x)\end{aligned}$$

By (a), $p_{\sigma\sigma} = \sigma(p)^n[a_0]_n^{-n}p\left(\sigma(p)[a_0]_n x - p_{n-1}/n\right)_\sigma$, and this equals p_σ.

(d) By construction,

$$r_\sigma = \sigma(r)[a_0]_n^{-1}r + \sigma(r)[a_0]_n^{-1}p_{n-1}/n$$

By theorem 5, definition 7, and definition 11,

$$\begin{aligned}r_{\sigma+} &= \sigma(r)[a_0]_n^{-1}r + \sigma(r)[a_0]_n^{-1}p_{n-1}/n\ _+ \\ &= \sigma(r)[a_0]_n^{-1}r_+ = \sigma(r)r_\times = r_\sigma\end{aligned}$$

By theorem 6, lemma 3, and 11,

$$\begin{aligned}r_{\sigma\times} &= \sigma(r)[a_0]_n^{-1}r + \sigma(r)[a_0]_n^{-1}p_{n-1}/n\ _\times \\ &= \sigma(r)r_\times = r_\sigma\end{aligned}$$

By (b), $r_{\sigma\sigma} = \sigma(r)[a_0]_n^{-1}r + \sigma(r)[a_0]_n^{-1}p_{n-1}/n\ _\sigma = r_\sigma$ (e) Suppose $p_\times = x^n + a_{n-1}x^{n-1} + \ldots + a_0$. Then

$$\begin{aligned}p_\sigma &= \sigma(p)^n p_\times\left(\sigma(p)x\right) \\ &= x^n + \sigma(p)a_{n-1}x^{n-1} + \ldots + \sigma(p)^n a_0\end{aligned}$$

We see that p_σ is even if and only if $p_\times$ is even. If p_σ is even, then $\sigma(p_\sigma) = 1$ by definition 10. If p_σ is not even, then $p_{\sigma\times} = p_\sigma$ by (c), so

$$\begin{aligned}\sigma(p_\sigma) &= \mathrm{sgn}\ \ \sigma(p)^{2l+1}a_{n-2l-1} \\ &= \sigma(p)\,\mathrm{sgn}(a_{n-2l-1}) = \sigma(p)^2 = 1\end{aligned}$$

where a_{n-2l-1} is the first nonzero coefficient in the sequence a_{n-1}, a_{n-3}, a_{n-5}, ...

(f) Suppose r is a root of p. Then r_σ is a root of p_σ. If p_σ is even, then by (d) and 11,

$$\begin{aligned}\sigma(r_\sigma) &= \mathrm{csgn}(r_{\sigma\times}) = \mathrm{csgn}(r_\sigma) = \mathrm{csgn}(\sigma(r)r_\times) \\ &= \sigma(r)\,\mathrm{csgn}(r_\times) = \sigma(r)^2 = 1\end{aligned}$$

If p_σ is not even, then $\sigma(r_\sigma) = \sigma(p_\sigma) = 1$ by (e). □

THEOREM 8. *Let r and s be algebraic numbers. Then $r \equiv_\mathbb{Q} s$ if and only if $r_\sigma = s_\sigma$.*

PROOF. If $r \equiv_\mathbb{Q} s$, then there exist $\alpha \in \mathbb{Q}\backslash\{0\}$, $\beta \in \mathbb{Q}$ such that $r = \alpha s + \beta$. So $r \in \mathbb{Q}$ if and only if $s \in \mathbb{Q}$. If $r \in \mathbb{Q}$ and $s \in \mathbb{Q}$, then $r_\sigma = s_\sigma = 0$. If $r \notin \mathbb{Q}$ and $s \notin \mathbb{Q}$, then $r_\sigma = (\alpha s+\beta)_\sigma = s_\sigma$ by theorem 7.

If $r_\sigma = s_\sigma$, then $r \equiv_\mathbb{Q} r_\sigma = s_\sigma \equiv_\mathbb{Q} s$ implies $r \equiv_\mathbb{Q} s$. □

Definition 12. If $p(x) \in \mathbb{P}$ is an irreducible polynomial, then $p_\sigma(x)$ is called a **canonical polynomial**. If r is an algebraic number, r_σ is a **canonical root**.

LEMMA 5. *Suppose $p(x) \in \mathbb{Q}[x]$ is an irreducible polynomial and r is a root of $p_\sigma(x)$. If p_σ is even, then $r_\sigma = \mathrm{csgn}(r)r$. If p_σ is not even, then $r_\sigma = r$.*

PROOF. By construction

$$p_\sigma(x) = \sigma(p)^n[a_0]_n^{-n}p\left(\sigma(p)[a_0]_n x - p_{n-1}/n\right)$$

and this is $x^n + 0x^{n-1} + \ldots$, which makes $r_+ = r$. Next, $p_\sigma(x) = \sigma(p)^n[a_0]_n^{-n}p_+\left(\sigma(p)[a_0]_n x\right)$, which equals

$$x^n + \sigma(p)[a_0]_n^{-1}a_{n-1}x^{n-1} + \ldots + \sigma(p)^n[a_0]_n^{-n}a_0$$

Therefore, by lemma 3,

$$r_\times = \ \sigma(p)^n[a_0]_n^{-n}a_0\ _n^{-1}\, r_+ = [a_0]_n[a_0]_n^{-1}r_+ = r$$

If p_σ is even, then $p_{\sigma\times} = p_\sigma$ is even and the relation $\sigma(r) = \mathrm{csgn}(r_\times) = \mathrm{csgn}(r)$ gives us

$$r_\sigma = \sigma(r)r_\times = \mathrm{csgn}(r)r\ .$$

If p_σ is not even, then $p_{\sigma\times} = p_\sigma$ and $\sigma(r) = \sigma(p_\sigma) = 1$ by theorem 7, implying $r_\sigma = \sigma(r)r_\times = 1 \cdot r = r$. □

THEOREM 9. *Suppose $p(x) \in \mathbb{P}$ is an irreducible polynomial. If p_σ is even, then exactly half of the roots of p_σ are canonical, and the other half are the negatives of the first half. The roots of p_σ expressed in canonical form are*

$$\pm\sqrt{s_1}, \ldots, \pm\sqrt{s_{n/2}}$$

where the $s_1, \ldots, s_{n/2}$ are the roots of $p_\sigma(\sqrt{x})$ and the roots of $p(x)$ corresponding to $\sqrt{s_1}, \ldots, \sqrt{s_{n/2}}$ are canonical and distinct. If p_σ is not even, then all of the roots of p_σ are canonical and distinct.

PROOF. If p_σ is even, then $p_\sigma(\sqrt{x})$ is a polynomial. If $s_1, \ldots, s_{n/2}$ are the roots of $p_\sigma(\sqrt{x})$ then the roots of p_σ must be

$$\pm\sqrt{s_1}, \ldots, \pm\sqrt{s_{n/2}}$$

Since $\mathrm{csgn}(\sqrt{s_i}) = 1$ for the principal branch of the square root, we see $\sqrt{s_1}, \ldots, \sqrt{s_{n/2}}$ are canonical by lemma 5. They are distinct because of irreducibility.

If p_σ is not even, then all of the roots of p_σ are canonical by lemma 5, and again they are distinct. □

4. AN ALGORITHM

The input for the algorithm is $p(x) \in \mathbb{Q}[x]$, a univariate polynomial over the rationals. The output is a list of the roots of $p(x)$ expressed in the canonical form just described.

1. Factor $p(x)$ over $\mathbb{Q}$. (The Maple `factors` command can do this.)
$$p(x) = c\prod_i p_i(x)^{e_i}$$

2. Solve the linear $p_i(x)$ to obtain the rational roots of $p(x)$.

3. For each nonlinear $p_i(x)$, obtain the canonical polynomial. If $n = \deg(p_i)$, the steps are
$$\begin{aligned} p_{i+}(x) &= M(1, p_{n-1}/n)p_i(x) \\ p_{i\times} &= M([a_0]_n, 0)p_{i+}(x) \\ p_{i\sigma} &= M(\sigma(p), 0)p_{i\times}(x) \end{aligned}$$
Each $p_i(x)$ can therefore be expressed canonically as the polynomial $M(\alpha_i, \beta_i)p_{i\sigma}(x)$ for computed α_i, β_i.

   Let B be the set of distinct $p_{i\sigma}$ that appear.

4. Each $p_{i\sigma} \in B$ that is not an even polynomial has $\deg(p_i)$ distinct roots, which in Maple `RootOf` notation are written $\mathcal{R}(p_i(x))$.

5. Each $p_{i\sigma} \in B$ that is an even polynomial has $\deg(p_i)$ distinct roots, which can be written in `RootOf` notation as $\pm\sqrt{\mathcal{R}(p_i(\sqrt{x}))}$.

6. Collect the $\mathcal{R}$ canonical forms from steps (2), (4) and (5) according to multiplicities e_i to form the final answer.

5. EXAMPLES

Example 1 can now be solved as follows. The roots become
$$\begin{aligned} x_1 &= R(z^5+z^2+1, z, \mathrm{index}=1), \\ x_2 &= R(z^5+z^2+1, z, \mathrm{index}=2), \\ x_3 &= R(z^5+z^2+1, z, \mathrm{index}=3), \\ x_4 &= R(z^5+z^2+1, z, \mathrm{index}=4), \\ x_5 &= R(z^5+z^2+1, z, \mathrm{index}=5), \\ x_6 &= 2+R(z^5+z^2+1, z, \mathrm{index}=1), \\ x_7 &= 2+R(z^5+z^2+1, z, \mathrm{index}=2), \\ x_8 &= 2+R(z^5+z^2+1, z, \mathrm{index}=3), \\ x_9 &= 2+R(z^5+z^2+1, z, \mathrm{index}=4), \\ x_{10} &= 2+R(z^5+z^2+1, z, \mathrm{index}=5). \end{aligned}$$
The statement $x_8 - x_3 = 2$, now becomes obvious, as does the statement $x_1 - x_j \notin \mathbb{Z}, j > 6$.

Example 2. Let $p(x)$ be the polynomial
$$\begin{aligned} p(x) = {} & 20736x^{18} + 179712x^{17} + 457920x^{16} - 94656x^{15} \\ & -2769344x^{14} - 3990464x^{13} + 4663468x^{12} + 13336348x^{11} \\ & +1419041x^{10} - 11999454x^9 - 12064413x^8 - 11691713x^7 \\ & +2228047x^6 + 20917493x^5 + 14885343x^4 - 1856968x^3 \\ & -8277728x^2 - 4558616x - 804752. \end{aligned}$$

The factored form of $p(x)$ is (having been obtained, for example, from Maple `factors`)
$$20736(x-1)^2 a(x)b(x)c(x)d(x)$$
where
$$\begin{aligned} a(x) &= x^4 - \tfrac{4}{3}x^3 + \tfrac{2}{3}x^2 - \tfrac{31}{27}x - \tfrac{53}{81} \\ b(x) &= x^4 + 8x^3 + 24x^2 + 40x + 16 \\ c(x) &= x^4 + 6x^3 + \tfrac{53}{4}x^2 + \tfrac{51}{4}x + \tfrac{73}{16} \\ d(x) &= x^4 - 2x^3 + \tfrac{1}{2}x^2 + \tfrac{1}{2}x + \tfrac{13}{16} \end{aligned}$$
Solving the linear factor of $p(x)$ produces the rational root 1 with multiplicity 2. Computing a_+, b_+, c_+, and d_+ gives
$$\begin{aligned} a_+(x) &= a\ \ x + \tfrac{1}{3} \ = x^4 - x - 1 \\ b_+(x) &= b(x-2) = x^4 + 8x - 16 \\ c_+(x) &= c\ \ x - \tfrac{3}{2} \ = x^4 - \tfrac{1}{4}x^2 + \tfrac{1}{16} \\ d_+(x) &= d\ \ x + \tfrac{1}{2} \ = x^4 - x^2 + 1 \end{aligned}$$

Computing $a_\times$, $b_\times$, $c_\times$, and $d_\times$ gives
$$\begin{aligned} a_\times(x) &= a_+(x) = x^4 - x - 1 \\ b_\times(x) &= \tfrac{1}{16}b_+(2x) = x^4 + x - 1 \\ c_\times(x) &= 16c_+\ \ \tfrac{1}{2}x \ = x^4 - x^2 + 1 \\ d_\times(x) &= d_+(x) = x^4 - x^2 + 1 \end{aligned}$$

Computing a_σ, b_σ, c_σ, and d_σ gives
$$\begin{aligned} a_\sigma(x) &= a_\times(-x) = x^4 + x - 1 \\ b_\sigma(x) &= b_\times(x) = x^4 + x - 1 \\ c_\sigma(x) &= c_\times(x) = x^4 - x^2 + 1 \\ d_\sigma(x) &= d_\times(x) = x^4 - x^2 + 1 \end{aligned}$$

The canonical forms for a, b, c, and d are given by
$$\begin{aligned} a(x) &= a_\sigma\ \ -x + \tfrac{1}{3} \\ b(x) &= 16a_\sigma\ \ \tfrac{1}{2}x + 1 \\ c(x) &= \tfrac{1}{16}c_\sigma(2x+3) \\ d(x) &= d_\sigma\ \ x - \tfrac{1}{2} \end{aligned}$$
and the basis B is
$$B = \{a_\sigma(x), c_\sigma(x)\} = \ x^4 + x - 1, x^4 - x^2 + 1$$

The polynomial $a_\sigma(x)$ is not even and therefore has 4 distinct algebraic roots, represented in Maple by $\mathcal{R}(a_\sigma(x))$. They will contribute 8 of the roots of the $p(x)$, specifically, $\frac{1}{3} - \mathcal{R}(x^4 + x + 1)$ from the reduction of a(x) and $-2+2\mathcal{R}(x^4+x+1)$ from the reduction of b(x). The polynomial c_σ is even and contributes 4 roots in two pairs, namely $-\frac{3}{2} \pm \frac{1}{2}\sqrt{\mathcal{R}(x^2 - x + 1)}$ and $\frac{1}{2} \pm \sqrt{\mathcal{R}(x^2 - x + 1)}$.

All possible linear relations over $\mathbb{Q}$ between roots of $p(x)$ are now explicitly obvious. This concludes the example.

6. CONCLUDING REMARKS

There remain a number of implementation questions. It should be recalled that MAPLE allows any polynomial to be an argument of `RootOf`. For an indexed `RootOf`, the index must be re-computed for the canonical polynomial. For a non-indexed `RootOf`, there is a question of what to do with a set that separates into several subsets. Even for the `RootOf` an irreducible polynomial, there are other properties that might be considered. If p_σ is a canonical polynomial with $\deg p_\sigma = n$, then we have

$$\sum_{k=1}^{n} \mathcal{R}(p_\sigma(x), x, \text{index} = k) = 0 \ . \tag{5}$$

One way to ensure that this simplification is known to the system would be to express the nth `RootOf` as the negative of the sum of the first $(n-1)$ `RootOf`s. For large n, however, this would be cumbersome, and would not apply to a non-indexed `RootOf`.

7. REFERENCES

[1] J. Della Dora and E. Tournier. Formal solutions of differential equations in the neighborhood of singular points (regular and irregular). In *Proceedings of the 1981 ACM Symposium on Symbolic and Algebraic Computation*, pages 25–29. ACM Press, 1981.

[2] M. Fels and P. Olver. Moving coframes. II. Regularization and theoretical foundations. *Acta Appl. Math.*, 55:127–208, 1999.

[3] P. J. Olver. *Classical Invariant Theory*. Cambridge University Press, 1999.

[4] J.-P. Tignol. *Galois theory of algebraic equations*. Longman, 1988.