BULL. AUSTRAL. MATH. SOC. 30C10, 26C05, 12D05, 10M05, 13F20
VOL. 47 (1993) [163–168]

POWER ROOTS OF POLYNOMIALS

JOHN BORIS MILLER

Sufficient conditions are given for the existence of an mth power root of one polynomial modulo another, over the complexes or the reals. Examples show the non-necessity of the conditions. In particular cases there can exist infinitely many square roots.

Let $K[\lambda]$ denote as usual the algebra of all polynomials in an indeterminate λ over a field K. If $p(\lambda)$ and $f(\lambda)$ belong to $K[\lambda]$, $m \in \mathbb{N}$ and

$$(p(\lambda))^m = f(\lambda),$$

we say that $p(\lambda)$ is a *power root* of $f(\lambda)$, more precisely an mth *root*.

It is not difficult to show that if this equation, with $f(\lambda)$ given in $\mathbb{C}[\lambda]$, has any solutions $p(\lambda)$ then it has precisely m solutions in $\mathbb{C}[\lambda]$. For suppose $f(\lambda)$ is monic (that is, has leading coefficient 1); then one verifies by solving for the coefficients of $p(\lambda)$ that there exists at most one solution $p(\lambda)$ which is monic; the general statement can be deduced from this.

Of more interest and abundance than power roots of polynomials are power roots of residue classes of polynomials. For any $w(\lambda) \in K[\lambda]$ let

$$\mathfrak{U}_{w,K} := K[\lambda] \mod w(\lambda)$$

denote the residue-class algebra over K of $K[\lambda]$ modulo the principal ideal generated by $w(\lambda)$; its elements will be written $[f]$, $[p]$, ... An mth root of $[f]$ is any coset $[p]$ such that $[p]^m = [f]$. Our principal result is:

THEOREM. *In the complex residue-class algebra $\mathfrak{U}_{w,\mathbb{C}}$ where degree$(w) \geqslant 1$, a sufficient condition for a class $[f]$ to possess mth roots of all orders $m \in \mathbb{N}$ is: that $f(\lambda)$ does not vanish at any multiple zero of $w(\lambda)$. In the real residue-class algebra $\mathfrak{U}_{w,\mathbb{R}}$, sufficient conditions are: that $w(\lambda)$ is real with real roots, $f(\lambda)$ is a real polynomial, $f(\lambda)$ does not vanish at any multiple zero of $w(\lambda)$, and $f(\lambda) \geqslant 0$ at every zero of $w(\lambda)$.*

The proof of the theorem will be separated into the following two lemmas, whose proofs use the existence of mth roots in algebras of matrices.

Received 17 February 1992

Lemma 1. *With $K = \mathbb{C}$, let $m, n \in \mathbb{N}$, $m \geqslant 2$, and suppose that in $\mathbb{C}[\lambda]$:*

(i) *$w(\lambda)$ is a polynomial of degree n;*

(ii) *$f(\lambda)$ is a polynomial such that $f(b) \neq 0$ for each multiple zero b of $w(\lambda)$.*

Then there exist polynomials $p(\lambda)$ and $q(\lambda)$ in $\mathbb{C}[\lambda]$ such that

$$(p(\lambda))^m = f(\lambda) + w(\lambda)q(\lambda). \tag{1}$$

Proof: The case $n = 1$, $w(\lambda) = \lambda - b$ say, is disposed of by dividing $f(\lambda)$ by $\lambda - b$ to get $f(\lambda) = -(\lambda - b)q(\lambda) + a$ say, $a \in \mathbb{C}$, and taking $\deg(p) = 0$, $p(\lambda) = a^{1/m}$.

Henceforth suppose that $n \geqslant 2$ and that $\deg(f) \geqslant 1$. Let $b_1, b_2, \ldots, b_r$ be an enumeration of the distinct zeros of $w(\lambda)$, and let $k_1, k_2, \ldots, k_r$ be their multiplicities, so that $\sum_{j=1}^{r} k_j = n$ and

$$w(\lambda) = (\lambda - b_1)^{k_1}(\lambda - b_2)^{k_2}\ldots(\lambda - b_r)^{k_r} \tag{2}$$

(without loss of generality we assume that $w(\lambda)$ has leading coefficient 1). Let

$$J_k(b) = \begin{pmatrix} b & 1 & & & \\ & b & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & b & 1 \\ & & & & b \end{pmatrix} \tag{3}$$

denote the usual $k \times k$ Jordan block matrix, and let W be the $n \times n$ block diagonal matrix

$$W = \operatorname{diag}(J_{k_1}(b_1), J_{k_2}(b_2), \ldots, J_{k_r}(b_r)). \tag{4}$$

The elementary divisors of W over $\mathbb{C}$ are

$$(\lambda - b_1)^{k_1}, (\lambda - b_2)^{k_2}, \ldots, (\lambda - b_r)^{k_r}, \tag{5}$$

and the minimal (annihilating) polynomial of W is our given polynomial $w(\lambda)$ in (2).

Using the other given polynomial $f(\lambda)$ we have

$$f(W) = \operatorname{diag}(f(J_{k_1}(b_1)), f(J_{k_2}(b_2)), \ldots, f(J_{k_r}(b_r))). \tag{6}$$

Here well-known calculations give

$$f(J_k(b)) = [f(b), \frac{f'(b)}{1!}, \ldots, \frac{f^{(k-1)}(b)}{(k-1)!}], \tag{7}$$

that is, $f(J_k(b))$ is the $k \times k$ upper triangular matrix with constant diagonals having for its top row the tuple shown in (7).

The matrix $f(W)$ in (6) has at least one matrix mth root, that is, a solution Y in $\mathfrak{M}_n(\mathbb{C})$ (the algebra of all $n \times n$ matrices over $\mathbb{C}$) of the equation

$$Y^m = f(W), \tag{8}$$

namely

$$Y = \operatorname{diag}(X_1, X_2, \ldots, X_r), \tag{9}$$

where for each j,

$$X_j^m = f\left(J_{k_j}(b_j)\right). \tag{10j}$$

Clearly, if an X_j exists for each j, then (8) follows easily from (6), (9) and (10j).

The existence of X_j is shown by its construction. See Gantmacher [**1**, pp.231–234] for the construction of square roots of matrices. For completeness we give a construction here, but prefer to rely upon the general functional calculus in the complex Banach algebra $\mathfrak{M}_n(\mathbb{C})$ with any algebra matrix norm (see [**2**, Theorem 5.2.5, pp.168–169]), since this method makes clear a commutativity property needed presently.

Assume first that $f(b_j) \neq 0$. Let ρ be any ray from the origin in the complex plane, not passing through any nonzero values among the numbers b_j, $f(b_j)$ for $j = 1, 2, \ldots, r$. Let $\Gamma(\mu)$ denote a small positively oriented circle about μ not intersecting ρ. Let h denote any holomorphic branch function of the root relation $\lambda^{1/m}$ on the plane cut along ρ from 0 to ∞. Fix j; we shall suppress the suffix j temporarily. The matrix

$$X := \frac{1}{2\pi i} \int_{\Gamma(f(b))} h(\lambda)(\lambda I - f(J_k(b)))^{-1} d\lambda \tag{11}$$

is well defined, and if $l(\lambda) := \lambda^m$ then [**2**, Theorem 5.3.2, p.171]

$$X^m = l(X) = l \circ h(f(J_k(b))) = f(J_k(b)). \tag{12}$$

Thus for each j the matrix $X_j = X$ in (11) gives a solution of (10j).

Suppose instead that $f(b_j) = 0$; by (ii) $k_j = 1$, so (10j) in this case becomes $X_j^m = O$ in $\mathfrak{M}_1(\mathbb{C})$ and we therefore take $X_j = O$. (If $k > 1$ then $J_k(0)$ has no mth root, so we exclude this possibility.)

Definition (11) shows that X belongs to the second commutant of $f(J_k(b))$, that is, it commutes with every matrix which commutes with $f(J_k(b))$. Therefore

$$X_j \smile J_{k_j}(b_j) \qquad \text{for each } j \tag{13}$$

and hence $Y \smile W$ (the symbol $\smile$ means 'commutes with').

Now the elementary divisors (5) of W are pairwise coprime, since the b_j's are distinct. This implies (see Gantmacher [1, p.222]) that the first commutant of W coincides with the set of all matrices which are expressible as a polynomial in W over K. Therefore Y is expressible as a polynomial in W over $\mathbb{C}$.

If for two matrices A and $B(\neq O)$ in $\mathfrak{M}_n(K)$, B is expressible as a polynomial in A over K, say $B = p(A)$, then there exists a unique such representing polynomial $p(\lambda)$ of least degree, call it $p_{B,A}(\lambda)$. For by using the Euclidean algorithm in $K[\lambda]$ we can show, first, that for any representing polynomial $p(\lambda)$ of least degree, its degree is less than the degree of the minimal (annihilating) polynomial of A; and secondly, that if $p(\lambda) = \alpha_0\lambda^s + \ldots$ and $q(\lambda) = \beta_0\lambda^s + \ldots$ are two distinct representing polynomials of least degree then, if $\alpha_0 \neq \beta_0$, the polynomial $(\beta_0 - \alpha_0)^{-1}(\beta_0 p(\lambda) - \alpha_0 p(\lambda))$ is a representing polynomial of lower degree, which is impossible, while if $\alpha_0 = \beta_0$, then $p(\lambda) - q(\lambda)$ is an annihilating polynomial for A, which is also impossible.

Thus in particular there exists a minimal representing polynomial $p_{Y,W}(\lambda)$,

$$Y = p_{Y,W}(W). \tag{14}$$

But then $(p_{Y,W}(W))^m - f(W) = Y^m - f(W) = O$, so $(p_{Y,W}(\lambda))^m - f(\lambda)$ is an annihilating polynomial for W and hence is a multiple of $w(\lambda)$: there exists $q(\lambda) \in K[\lambda]$ such that

$$(p_{Y,W}(\lambda))^m - f(\lambda) = w(\lambda)q(\lambda). \tag{15}$$

This completes the proof of the lemma when $\deg(f) \geqslant 1$. When $f(\lambda)$ is a constant c say, we solve (8) by taking $Y = c^{1/m}I$ and then argue as before. ◻

Lemma 2. *With $K = \mathbf{R}$, let m, n, w, f be as in Lemma 1 and suppose in addition to (i) and (ii) that*

(iii) *all zeros of $w(\lambda)$ are real, so that $w(\lambda) \in \mathbf{R}[\lambda]$; and*

(iv) *$f(\lambda) \in \mathbf{R}[\lambda]$, and $f(b) \geqslant 0$ at every zero b of $w(\lambda)$.*

Then there exist polynomials $p(\lambda)$ and $q(\lambda)$ in $\mathbb{R}[\lambda]$ such that (1) holds.

Proof: Under these conditions W is a real matrix and its elementary divisors over $\mathbf{R}$ are again (5). The matrices $f(J_k(b))$ in (7) are real, so $f(W)$ in (6) is real. Let ρ be a ray other than the positive real axis, and let h be the branch function which is real and positive on the positive real axis. Then using (iv) it can be verified that the righthand side of (11) is selfconjugate, so each X_j is real and Y in (9) is real. The rest of the argument in the proof of Lemma 1 then applies, with $K = \mathbf{R}$. ◻

The theorem follows immediately from the lemmas.

Corollary 1. *If $w(\lambda)$ has only simple zeros then every $[f]$ in $\mathfrak{U}_{w,\mathbf{C}}$ has mth roots of all orders in that algebra.*

Corollary 2. *The identity coset* $[1]$ *has mth roots of all orders m, in $\mathfrak{U}_{w,\mathbf{C}}$ and in $\mathfrak{U}_{w,\mathbf{R}}$, for every choice of polynomial $w(\lambda)$.*

We remark that in Lemma 1 the polynomial $p(\lambda)$ satisfies the same conditions as $f(\lambda)$; in Lemma 2, one of $\pm p(\lambda)$ does so.

In each lemma the proof obtains the sought power root $p(\lambda)$ as the representing polynomial of least degree of a particular matrix root of $f(W)$, for a particular matrix W constructed from $w(\lambda)$. The power root is far from being unique; see Examples 3 and 4 below.

The conditions in the theorem are sufficient, not necessary; this is shown in Example 3. But the conditions may not be omitted from the theorem; see Examples 1 and 2.

For any case of $[p]^m = [f]$ there are unique polynomials $p_0(\lambda)$ and $f_0(\lambda)$ in these cosets respectively with degrees less than n; necessarily $p_0(\lambda) = p_{Y,W}(\lambda)$. Writing $s := \deg(p_0)$ we have

$$\frac{n}{m} \leqslant s < n \text{ if } q(\lambda) \neq O, \qquad s < \frac{n}{m} \text{ if } q(\lambda) = O$$

In (15) if $f(\lambda) = f_0(\lambda)$ we have

$$\deg(q) = sm - n \text{ or } O.$$

Examples: For low values of m and n and given polynomials $w(\lambda)$ and $f(\lambda)$, one may look for power roots by substituting unknown polynomials $p(\lambda)$ and $q(\lambda)$ in equation (1), assuming minimal degrees, and attempting to solve the resulting nonlinear equations in the coefficients.

1. Take $w(\lambda) = \lambda^2(\lambda - 1)$, $f(\lambda) = \lambda(\lambda - 2)$, $m = 2$. Here $f(\lambda)$ has a zero at a multiple zero of $w(\lambda)$. We find that no square root of $[f]$ exists. Thus the condition (ii) in Lemma 1 and the theorem cannot be omitted.

2. Take $w(\lambda) = \lambda(\lambda - 2)$, $f(\lambda) = \lambda - 1$, $m = 2$. Here all conditions of Lemma 2 are satisfied except that $f(0) < 0$ at a zero 0 of $w(\lambda)$. We find that there exists no square root $[p]$ in $\mathfrak{U}_{w,\mathbf{R}}$.

3. Take $w(\lambda) = \lambda(\lambda - 1)^2$, $f(\lambda) = (\lambda - 1)^2$, $m = 2$. Again (ii) fails, but in this case equation (1) has infinitely many solutions with $\deg(p) < n = 3$, namely those

https://doi.org/10.1017/S0004972700012363 Published online by Cambridge University Press

indicated in the table

$p(\lambda)$	$q(\lambda)$
$a\lambda^2-(a+1)\lambda+1$	$a^2\lambda-2a,$
$a\lambda^2-(a-1)\lambda-1$	$a^2\lambda+2a,$

where a is arbitrary in $\mathbb{C}$ (or $\mathbb{R}$). The table includes all solutions. Distinct polynomials $p(\lambda)$ determine distinct square roots $[p]$ in $\mathfrak{U}_w$, so this $[f]$ has infinitely many square roots in $\mathfrak{U}_{w,\mathbb{C}}$, indeed in $\mathfrak{U}_{w,\mathbb{R}}$. The example also shows the non-necessity of condition (ii).

4. Take $w(\lambda)=(\lambda-a)(\lambda-b)(\lambda-c)$, $f(\lambda)=\lambda$, $m=2$.

Let τ_1, τ_2, τ_3 denote the three elementary symmetric functions on the numbers $a^{1/2}$, $b^{1/2}$, $c^{1/2}$, where $a^{1/2}$ is chosen to be either one of the two complex square roots of a, and likewise for $b^{1/2}$ and $c^{1/2}$. There are 8 square roots of $f(\lambda)$ mod $w(\lambda)$, namely all possible polynomials of the form

with
$$p(\lambda)=\left(\lambda^2+\left(\tau_2-\tau_1^2\right)\lambda-\tau_1\tau_3\right)\left(\tau_3-\tau_1\tau_2\right)^{-1},$$
$$q(\lambda)=\left(\lambda-\tau_1^2\right)\left(\tau_3-\tau_1\tau_2\right)^{-1},$$

provided $\Delta:=\tau_3-\tau_1\tau_2$ does not vanish. Now $\Delta=0$ if and only if one of $b^{1/2}+c^{1/2}$, $c^{1/2}+a^{1/2}$, $a^{1/2}+b^{1/2}$ vanishes; hence $\Delta\neq 0$ if a, b, c are distinct.

Suppose $a\neq b=c\neq 0$. We can still ensure that $\Delta\neq 0$ by choosing $b^{1/2}=c^{1/2}$, and so obtain a square root $[p]$ of $[f]$. However, there are now only 4 distinct square roots.

Suppose $a\neq b=c=0$, so that condition (ii) is violated. In this case $\Delta=0$ and indeed there exists no square root of $[f]$.

References

[1] F.R. Gantmacher, *Theory of matrices*, Vol 1 (Chelsea, New York, 1960).

[2] E. Hille and R.S. Phillips, *Functional analysis and semi-groups* (Amer. Math. Soc. Coll. Publ. 31, Providence, 1957).

Department of Mathematics
Monash University
Clayton Vic 3168
Australia