Interpolation Polynomials and Linear Algebra

Askold Khovanskii, FRSC, Sushil Singla, and Aaron Tronsgard

Presented by Pierre Milman, FRSC

Abstract. We reconsider the theory of Lagrange interpolation polynomials with multiple interpolation points and apply it to linear algebra. In particular, we show that one can evaluate a meromorphic function at a matrix, using only an interpolation polynomial.

Résumé. On reconsidère la thèorie des polynômes d'interpolation de Lagrange et l'applique à l'algèbre linéaire. En particulier, on peut évaluer une fonction méromorphe à une matrice seulement avec un polynôme d'interpolation.

1. **Introduction** Interpolation polynomials with multiple interpolation points are widely used in applied mathematics under the name Hermite–Lagrange interpolation polynomials for approximating functions or other data sets by polynomials, see for example Chapter 3 in [1]. But such polynomials are useful in pure mathematics as well. For example, one can construct Galois theory including the problem of solvability by radicals using Lagrange interpolation polynomials with simple roots as the main tool (see [3]).

One may guess that Lagrange interpolation polynomials with multiple interpolation points also have applications in pure mathematics. In the paper, we present our reconstruction of the theory of interpolation polynomials with multiple interpolation points and its applications to linear algebra.

In Section 3, we show applications of Lagrange interpolation polynomials in computing functions of matrices. From Theorem 3.6, it follows that if A is a $n \times n$ matrix, then for a rational function without poles at the eigenvalues of A, or for an entire function f, we have $f(A) = Q(A)$ where Q is the Lagrange interpolation polynomial of f with interpolation points given by the eigenvalues of A.

Explicit examples of using Lagrange interpolation polynomials to compute the inverse of a matrix, and to give the general solution to homogeneous linear differential equations are shown.

In Section 4, principal Lagrange resolvents are defined and applications are given. Corollary 4.6 shows that for computation of functions of matrices, it is enough to compute the principal Lagrange resolvents.

We end our discussions with application to linear algebra, specifically we provide a proof that all matrices can be put into Jordan normal form, as well as generalize Proposition 2.6 of [3].

This paper was written by Sushil Singla and Aaron Tronsgard who attended Askold Khovanskii's course on Topological Galois theory at the Fields Institute during the Fall, 2021. They brought to life a sketch of the theory presented in the course.

2. Preliminary Definitions and First Results

Let $\mathbb{K}$ be a field of characteristic zero. Let $\mathbb{K}[x]$ be the polynomial algebra over $\mathbb{K}$. Let $\{\lambda_1, \ldots, \lambda_k\} = \Lambda \subseteq \mathbb{K}$ be a set of k distinct elements. For all $1 \leq j \leq k$, let $m_j \in \mathbb{N}$ be a natural number associated with λ_j such that $\sum_{j=1}^{k} m_j = n$.

DEFINITION   2.1.   A polynomial $L \in \mathbb{K}[x]$ of degree less than n is called the *Lagrange interpolation polynomial* with interpolation points $\lambda_1 \ldots, \lambda_k$ with multiplicities $m_1, \ldots, m_k$ and the interpolation data

$$c_1^{(0)}, \ldots, c_1^{(m_1-1)}, \ldots, c_k^{(0)}, \ldots, c_k^{(m_k-1)}$$

if for every $\lambda_j \in \Lambda$ and $0 \leq m < m_j$, we have

$$L^{(m)}(\lambda_j) = c_j^{(m)},$$

where $L^{(0)}(x) = L(x)$ and for $m > 0$, $L^{(m)}(x)$ denotes the mth derivative of L.

We are justified in our language, defining the Lagrange interpolation polynomial due to the following result.

THEOREM 2.2.   *The Lagranage interpolation polynomial with given interpolation points and interpolation data exists and is unique.*

PROOF.   Let $L(x) = \sum_{i=0}^{n-1} a_i x^i$ be polynomials with undetermined cofficients $a_0, \ldots, a_{n-1}$. Now determining an interpolation polynomial is equivalent to solving the n equations

$$L^{(m)}(\lambda_j) = c_j^{(m)}.$$

This system of equations will have a unique solution provided the corresponding homogenous system

$$L^{(m)}(\lambda_j) = 0$$

has only the trivial solution.

A solution of the corresponding homogeneous equation is a polynomial of degree less than n such that for every $\lambda_j \in \Lambda$ and $0 \leq m < m_j$, we have $L^{(m)}(\lambda_j) = 0$ i.e. a polynomial L which has roots $\lambda_1, \ldots, \lambda_k$ with multiplities $m_1, \ldots, m_k$. By definition, $\sum_{j=1}^{k} m_j = n$. Clearly the zero polynomial satisfies these conditions. And this is only solution because if a polynomial of degree less than n has n roots counted with multiplicity, then it is identically zero. $\square$

Note that if all of the multiplicities are equal to 1, we have an explicit formula for the Lagrange interpolation polynomial:

$$(1) \qquad L(x) = \sum_{j=1}^{n} c_j \prod_{i \neq j} \frac{x - \lambda_i}{\lambda_j - \lambda_i}.$$

It will be useful for us to specify the interpolation data by giving function values, rather than as a list of points. We may also want to specify the interpolation points via eigenvalues of a linear transformation. To this end, we make the following definitions.

DEFINITION 2.3. Let $Q : \mathbb{K} \to \mathbb{K}$ be a polynomial. The Lagrange interpolation polynomial of Q with the interpolation points $\lambda_1, \ldots, \lambda_k$ with multiplicities $m_1, \ldots, m_k$ is defined as the unique polynomial L of degree less than or equal to $n := m_1 + \cdots + m_k$ such that for all $1 \leq j \leq k$ and $0 \leq m < m_j$, we have

$$L^{(m)}(\lambda_j) = Q^{(m)}(\lambda_j).$$

DEFINITION 2.4. Let $f : U \to \mathbb{C}$ be a function defined on an open subset U of $\mathbb{C}$. The Lagrange interpolation polynomial of f with the interpolation points $\lambda_1, \ldots, \lambda_k$ with multiplicities $m_1, \ldots, m_k$ is defined as the unique polynomial L of degree less than $n := m_1 + \cdots + m_k$ such that for all $1 \leq j \leq k$ and $0 \leq m < m_j$, we have

$$L^{(m)}(\lambda_j) = f^{(m)}(\lambda_j),$$

where $f^{(0)}(x) = f(x)$ and for $m > 0$, $f^{(m)}(x)$ denotes the mth derivative of f.

EXAMPLE 2.5. If $f : U \to \mathbb{C}$ is an $(m-1)$-times differentiable function. Then for any $x_0 \in U$, the Lagrange interpolation polynomial L of f with one interpolation point x_0 of multiplicity m coincides with the degree $m-1$ Taylor polynomial of f at the point x_0; that is, we have,

$$L(x) = f(x_0) + f^{(1)}(x_0)(x - x_0) + \cdots + \frac{1}{(m-1)!} f^{(m-1)}(x_0)(x - x_0)^{m-1}.$$

One very classical result on Taylor polynomials is the Lagrange remainder theorem. It states that if f is an n-times differentiable function and T_{n-1} denotes

the degree $n-1$ Taylor polynomial of f at point x_0, then for any $x \in \mathbb{R}$ there exists ξ between x and x_0 such that

$$f(x) - T_{n-1}(x) = \frac{f^{(n)}(\xi)}{n!}(x - x_0)^n.$$

It is a natural question whether we have an analogous result for the general interpolation polynomials. The answer is a resounding yes.

THEOREM 2.6.   *Let $f(x)$ be an n-times differentiable function. Let L be its interpolation polynomial with interpolation points $\lambda_1, \ldots, \lambda_k$ and multiplicities $m_1, \ldots, m_k$ such that $\sum_{j=1}^{n} m_j = n$. Then for $x_0 \in \mathbb{C}$, there exists ξ inside the convex hull U formed by the interpolation points such that*

$$f(x_0) - L(x_0) = \frac{f^{(n)}(\xi)}{n!}(x_0 - \lambda_1)^{m_1} \ldots (x_0 - \lambda_k)^{m_k}.$$

PROOF.   Consider the interpolation polynomial L_0 for f consisting of the same data as L, with the additional interpolation point $\lambda_{k+1} = x_0$ with multiplicity one. We write

$$L_0(x) = L(x) + C(x - \lambda_1)^{m_1} \ldots (x - \lambda_k)^{m_k}$$

with

$$C = \frac{f(x_0) - L(x_0)}{(x_0 - \lambda_1)^{m_1} \ldots (x_0 - \lambda_k)^{m_k}}.$$

Now, consider the function $f(x) - L_0(x)$. This function has at least $n+1$ roots in U, counting with multiplicity. By Rolle's theorem, this implies that the n-th derivative $(f - L_0)^{(n)}$ has at least one root in U, call it ξ. Moreover, we compute

$$(f - L_0)^{(n)}(x) = f^{(n)}(x) - Cn!.$$

And therefore, we have $C = \frac{f^{(n)}(\xi)}{n!}$. And since

$$f(x_0) - L(x_0) = C(x_0 - \lambda_1)^{m_1} \ldots (x_0 - \lambda_k)^{m_k},$$

this implies that

$$f(x_0) - L(x_0) = \frac{f^{(n)}(\xi)}{n!}(x_0 - \lambda_1)^{m_1} \ldots (x_0 - \lambda_k)^{m_k}.$$

$\square$

3. Computation of Functions of Matrices If interpolation points are roots of a polynomial, we have the folllowing proposition.

PROPOSITION 3.1. *Let $T \in \mathbb{K}[x]$ be a polynomial of degree n with k distinct roots $\lambda_1, \ldots, \lambda_k$ with multiplitcities $m_1, \ldots, m_k$ and $\sum_{j=1}^{k} m_j = n$.*

Then for any polynomial $Q \in \mathbb{K}[x]$ of degree at least n, a polynomial L is the Lagrange interpolation polynomial of Q with interpolation points equal to the roots of T with corresponding multiplicities if and only if

$$Q - L \equiv 0 \ (mod\ T).$$

In other words, the Lagrange polynomial L of Q with the above interpolation data is the remainder of Q by T.

PROOF. Let $L \in \mathbb{K}[x]$ be the remainder of Q by T, i.e., there exists $S \in \mathbb{K}[x]$ such that $Q = TS + L$ and the degree of L is less than or equal to n. For $1 \leq j \leq k$ and $1 \leq m < m_j$, we have $Q^{(m)}(\lambda_j) = \sum_{i=1}^{m} T^{(i)}(\lambda_j) S^{(m-i)}(\lambda_j) + L^{(m)}(\lambda_j)$. Since $T^{(i)}(\lambda_i) = 0$ for all $1 \leq i \leq m$, we get $Q^{(m)}(\lambda_j) = L^{(m)}(\lambda_j)$. And by the definition and uniqueness of the Lagrange interpolation polynomial, we get the result. $\square$

As an immediate application of the above proposition, we get the following theorem related to the computation of polynomials of operators.

THEOREM 3.2. *Let A be a linear operator on a vector space V over $\mathbb{K}$ such that A satisfies a polynomial T that splits over $\mathbb{K}$. Let $\lambda_1, \ldots, \lambda_k$ be the distinct roots of T with multiplities $m_1, \ldots, m_k$. Then for any $Q \in \mathbb{K}[x]$, we have*

$$Q(A) = R(A),$$

where R is the Lagrange interpolation polynomial of Q with interpolation points $\lambda_1, \ldots, \lambda_k$ and multiplicities $m_1, \ldots, m_k$. (Note that R is also the remainder of Q by T.)

PROOF. By the previous proposition, we know that the Lagrange interpolation polynomial R for Q with interpolation points $\lambda_1, \ldots, \lambda_k$ and multiplicities $m_1, \ldots, m_k$ is the remainder of the division of Q by T. In particular

$$Q - R \equiv 0 \mod (T).$$

We have $T(A) = 0$, and therefore $(Q - R)(A) = 0$. Or, in other words

$$Q(A) = R(A).$$

$\square$

COROLLARY 3.3.   *Let A be an $n \times n$ matrix with the eigenvalues $\lambda_1, \ldots, \lambda_k$ with multiplicities $m_1, \ldots, m_k$, such that $\sum\limits_{j=1}^{n} m_j = n$.*

Then for any polynomial Q with degree at least n, we have

$$Q(A) = R(A)$$

where R is the interpolation polynomial of Q with respect to the interpolation points $\lambda_1, \ldots, \lambda_k$ and multiplicities $m_1, \ldots, m_k$.

EXAMPLE 3.4.   Let A be $n \times n$ matrix with n distinct eigenvalues $\lambda_1, \ldots, \lambda_n$. Let $Q \in \mathbb{K}[x]$. Since the unique Lagrange interpolation polynomial L of degree less than equal to n with interpolation points $\lambda_1, \ldots, \lambda_n$ and interpolation data $Q(\lambda_1), \ldots, Q(\lambda_n)$ is

$$L(x) = \sum_{j=1}^{n} Q(\lambda_j) \prod_{i \neq j} \frac{(x - \lambda_i)}{(\lambda_j - \lambda_i)},$$

we have, $Q(A) = L(A) = \sum\limits_{j=1}^{n} Q(\lambda_j) \prod\limits_{i \neq j} \dfrac{(A - \lambda_i)}{(\lambda_j - \lambda_i)}$.

We now restrict ourselves to the case $\mathbb{K} = \mathbb{C}$. And we get the following results for functions of operators.

Let A be an operator on a vector space V over $\mathbb{C}$ such that A satisfies a polynomial $T \in \mathbb{C}[x]$. Let $\lambda_1, \ldots, \lambda_k$ be distinct roots of T with multiplicities $m_1, \ldots, m_k$.

LEMMA 3.5.   *Assume $0 \neq \lambda_j$ for all $1 \leq j \leq k$. Then A is invertible and we have*

$$A^{-1} = L(A),$$

where L is the Lagrange interpolation polynomial of the function

$$1/x$$

with interpolation points $\lambda_1, \ldots, \lambda_k$ and multiplicities $m_1, \ldots, m_k$.

PROOF.   We know that $\lambda_1, \ldots, \lambda_k$ are roots of the rational function

$$\frac{1}{x} - L(x)$$

with multiplicities $m_1, \ldots, m_k$. Multiplying by x, we retain those roots and so we have

$$1 - xL(x) \equiv 0 \mod (T).$$

And since $T(A) = 0$ this implies that

$$AL(A) = I.$$

$\square$

THEOREM 3.6.   *We have the following statements:*

(i) *Consider a rational function $f(x) = P(x)/Q(x)$ such that Q does not vanish at λ_j for all $1 \le j \le k$. Let L be an interpolation polynomial of f with the interpolation points $\lambda_1, \ldots, \lambda_k$ and multiplicities $m_1, \ldots, m_k$. Then the operator $P(A)[Q(A)]^{-1}$ is defined and it is equal to $L(A)$.*

(ii) *For an entire function $F(x)$ of complex variable x, we have*

$$F(A) = L(A),$$

where L is the Lagrange interpolation polynomial of the function F with interpolation points $\lambda_1, \ldots, \lambda_k$ and multiplicities $m_1, \ldots, m_k$.

PROOF.   (i) Since Q does not vanish on Λ we know that $(Q, T) = 1$, and thus we can find $V, U \in \mathbb{C}(x)$ such that

$$Q(x)V(x) + U(x)T(x) = 1.$$

Evaluating at A we get

$$Q(A)V(A) = I.$$

So $Q(A)$ is invertible, with $Q^{-1}(A) := V(A)$. Then, proceeding similarly to the case above, we have that

$$P(x) - L(x)Q(x) \equiv 0 \mod (T).$$

And therefore

$$P(A) = L(A)Q(A).$$

Finally, we multiply on the right by $Q(A)^{-1}$ to get

$$L(A) = P(A)Q(A)^{-1}.$$

(ii) We know that for $1 \le j \le k$ and $0 \le m \le m_j - 1$ we have

$$F^{(m)}(\lambda_j) = L^{(m)}(\lambda_j).$$

Therefore, we can write

$$F(x) - L(x) = T(x)G(x)$$

for some entire function $G(x)$. Evaluating at A we get

$$F(A) - L(A) = 0.$$

$\square$

EXAMPLE 3.7.   . Let A be $n \times n$ matrix with n distinct eigenvalues $\lambda_1, \ldots, \lambda_n$. From Example 3.4 and Theorem 3.6, we have

$$(2) \qquad \exp(A) = \sum_{j=1}^{n} \exp(\lambda_j) \prod_{i \neq j} \frac{(A - \lambda_i)}{(\lambda_j - \lambda_i)}.$$

In fact, for any entire complex valued function f, we have

$$f(A) = \sum_{j=1}^{n} f(\lambda_j) \prod_{i \neq j} \frac{(A - \lambda_i)}{(\lambda_j - \lambda_i)}.$$

Below we show application of our methods to compute inverse of a 3×3 matrix and in solving an order 3 homogeneous linear differential equation.

Computing the Inverse of a matrix As seen in Theorem 3.6, we are able to compute the inverse of a non-singular matrix A using a Lagrange interpolation polynomial. Thus, we are able to solve consistent systems of linear equations. Let

$$A = \begin{pmatrix} 9 & -15 & -25 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

We compute the characteristic polynomial for A: $T(x) = -(x+1)(x-5)^2$. Thus, we want to find the Lagrange interpolation polynomial $L(x)$ for $f(x) = 1/x$, interpolation points $\lambda_1 = 5$, $\lambda_2 = -1$ with multiplicities $m_1 = 2$ and $m_2 = 1$.

There is a useful trick for computing the interpolation polynomial with *one more* multiplicity than one we already know (which we do for the case $m_1 = m_2 = 1$). Let

$$L_0(x) = \frac{x-5}{6} + \frac{x+1}{30}$$

denote the interpolation polynomial with simple multiplicities. Then, we look for $L(x) = L_0(x) + c(x-5)(x+1)$ where $c \in \mathbb{C}$ is a constant.

In our case, we evaluate the derivative at 5 and require it to be equal to $f'(5) = \frac{-1}{25}$:

$$L'(5) = \frac{1}{6} + \frac{1}{30} + 6c$$

from which we find $c = \frac{-1}{25}$. And so $L(x) = \frac{x-5}{6} + \frac{x+1}{30} - \frac{1}{25}(x-5)(x+1)$. The relevant matrices are

$$A + I = \begin{pmatrix} 10 & -15 & -25 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \quad A - 5I = \begin{pmatrix} 4 & -15 & -25 \\ 1 & -5 & 0 \\ 0 & 1 & -5 \end{pmatrix}$$

$$(A + I)(A - 5I) = \begin{pmatrix} 25 & -100 & -125 \\ 5 & -20 & -25 \\ 1 & -4 & -5 \end{pmatrix}.$$

And we find that

$$A^{-1} = L(A) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -\frac{1}{25} & \frac{9}{25} & -\frac{3}{5} \end{pmatrix}.$$

An order 3 homogeneous linear differential equation One standard application of the matrix exponential is to ordinary linear differential equations. In particular, a general solution to the matrix differential equation

$$\mathbf{y}' = A\mathbf{y}$$

is given by $\mathbf{y}(t) = \exp(tA)\mathbf{c}$, where $\mathbf{c}$ is an arbitrary column of constants.

The standard solution to this problem is to fix a basis of the vector space so that the matrix A is in Jordan normal form, and to learn how to take the exponentials of Jordan blocks. See for example Chapters 5 and 6 of [2]. Similarly to the above computation of the inverse matrix, Theorem 3.6 allows us to compute the matrix exponential using Lagrange interpolation polynomials, without considering any special bases, and without needing any normal forms.

Consider for example the equation

$$y''' - 9y'' + 15y' + 25y = 0.$$

Written as a system of differential equations, this is

$$\mathbf{y}' = \begin{pmatrix} 9 & -15 & -25 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \mathbf{y}$$

where $\mathbf{y} = \begin{pmatrix} y'' \\ y' \\ y \end{pmatrix}$, and we see the same matrix A as in the previous example.

We are interested in computing $\exp(tA)$. Similar to before, the characteristic polynomial is $T(x) = -(x+t)(x-5t)^2$, so we want to compute the interpolation polynomial for the function $f(x) = \exp(x)$ with interpolation points $\lambda_1 = 5t$ and $\lambda_2 = -t$ with multiplicities 2 and 1 respectively. Again, we write down the interpolation polynomial with simple multiplicities

$$L_0(x) = -\frac{e^{-t}(x - 5t)}{6t} + \frac{e^{5t}(x + t)}{6t}$$

and we look for $L(x) = L_0(x) + c(x - 5t)(x + t)$, noting that c may depend on t (but not x).

We require

$$e^{5t} = f'(5t) = L'(5t) = -\frac{e^{-t}}{6t} + \frac{e^{5t}}{6t} + 6ct,$$

from which we find $ct^2 = \frac{e^{5t}(6t-1)+e^{-t}}{36}$. And therefore

$$L(x) = -\frac{e^{-t}(x-5t)}{6t} + \frac{e^{5t}(x+t)}{6t} + \frac{e^{5t}(6t-1)+e^{-t}}{36t^2}(x+t)(x-5t).$$

The relevant matrices are

$$tA - 5tI = \begin{pmatrix} 4t & -15t & -25t \\ t & -5t & 0 \\ 0 & t & -5t \end{pmatrix} \quad tA + tI = \begin{pmatrix} 10t & -15t & -25t \\ t & t & 0 \\ 0 & t & t \end{pmatrix}$$

$$(tA + tI)(tA - 5tI) = \begin{pmatrix} 25t^2 & -100t^2 & -125t^2 \\ 5t^2 & -20t^2 & -25t^2 \\ 1t^2 & -4t^2 & -5t^2 \end{pmatrix}.$$

Evaluating $L(tA) = \exp(tA)$, we find that the general solution is given by

$$y(t) = C_1 \left(-\frac{2e^{-t}}{3} + 2e^{5t} + \frac{25}{36}\left(e^{5t}(6t-1) + e^{-t}\right) \right)$$

$$+ C_2 \left(-\frac{5e^{-t}}{2} - 3e^{5t} - \frac{25}{9}\left(e^{5t}(6t-1) + e^{-t}\right) \right)$$

$$+ C_3 \left(-\frac{25e^{-t}}{6} - 5e^{5t} - \frac{125}{36}\left(e^{5t}(6t-1) + e^{-t}\right) \right),$$

where $C_1, C_2, C_3 \in \mathbb{C}$ are arbitrary constants.

The above examples show that the general procedure to compute the Lagrange interpolation polynomials is as follows. If we know the interpolation polynomial L_0 for k points $\lambda_1, \ldots, \lambda_k$ with multiplicities $m_1, \ldots, m_k$, then to find the interpolation polynomial L for the same data of interpolation polynomial with one extra multiplicity or one extra point with multiplicity 1 can be computed by taking

$$L(x) = L_0(x) + c(x - \lambda_1)^{m_1} \cdots (x - \lambda_k)^{m_k}$$

and using the interpolation data to find c.

The following result shows the method of computing the interpolation polynomial on the union of two sets of interpolation points:

Let $T_1(x) = (x - \lambda_1)^{m_1} \ldots (x - \lambda_p)^{m_p}$ and $T_2(x) = (x - \beta_1)^{\ell_1} \ldots (x - \beta_q)^{\ell_q}$ with $\lambda_i \neq \beta_j$ for all $1 \leq i \leq p$ and $1 \leq \beta_j \leq q$.

And let T_1^{-1}, T_2^{-1} denote the inverses of T_1 and T_2 modulo the ideal generated by T_2 and T_1 respectively (which exists because T_1 and T_2 has no common roots).

THEOREM 3.8.   *The Lagrange interpolation polynomial of $Q \in \mathbb{K}[x]$ with respect to $\Lambda = \Lambda_1 \cup \Lambda_2$ is equal to*

$$L = [QT_2^{-1}]_1 T_2 + [QT_1^{-1}]_2 T_1,$$

where $[QT_2^{-1}]_1$ and $[QT_1^{-1}]_2$ denote the interpolation polynomials of QT_2^{-1} and QT_1^{-1}, as rational functions, with respect to Λ_1 and Λ_2 respectively.

PROOF.   Let $n_1 = \sum_{j=1}^{p} m_j$ and $n_2 = \sum_{j=1}^{q} l_j$. First observe that as defined, L is a polynomial of degree smaller than $n_1 + n_2$. This is easily seen since $[\cdot]_1$ and $[\cdot]_2$ have degrees smaller than n_1 and n_2 respectively.

Let us examine what happens at Λ_1. We know that

$$[QT_2^{-1}]_1 - \frac{Q}{T_2}$$

has roots of multiplicities $m_1, \ldots, m_q$ at $\lambda_1, \ldots, \lambda_q \in \Lambda_1$. In particular, write

$$[QT_2^{-1}]_1(x) = \frac{Q(x)}{T_2(x)} + (x - \lambda_1)^{m_1} \ldots (x - \lambda_q)^{m_q} G(x)$$

for some rational function G which has no poles on Λ_1. Then we have

$$L(x) = Q(x) + (x - \lambda_1)^{m_1} \ldots (x - \lambda_q)^{m_q} G(x) T_2(x) + [QT_1^{-1}]_2(x) T_1(x).$$

Since T_1 has roots on Λ of the corresponding multiplicities, we see right away that for $1 \leq j \leq q$:

$$L^{(m)}(\lambda_j) = Q^{(m)}(\lambda_j).$$

for all $0 \leq m \leq m_j - 1$. Symmetric arguments will give that for $1 \leq j \leq q$:

$$L^{(\ell)}(\beta_j) = Q^{(\ell)}(\beta_j)$$

for all $0 \leq \ell \leq \ell_j - 1$.

And so by uniqueness of the Lagrange interpolation polynomial we are done.

$\square$

4.   Principal Lagrange Resolvent and Its Applications

4.1.   Definitions and basic properties In this section, we define principal Lagrange resolvents and we show that it is enough to compute the principal Lagrange interpolation polynomials to compute functions of matrices over $\mathbb{C}$.

As mentioned before that the standard method to solve the homogenous linear ordinary differential equations is in finding the Jordan normal form of matrices and learn how to take the exponentials of Jordan blocks. We have already seen the application of Lagrange interpolation in solving homogeneous linear ordinary differential equations without computing the Jordan canonical form. As an application of principal Lagrange interpolation polynomials, we provide a proof that all matrices can be put into Jordan normal form.

DEFINITION 4.1.   In the special case where $c_i^{(0)} = 1$ and all other interpolation data are 0, we define the corresponding Lagrange interpolation polynomial to be the *principal Lagrange resolvent* with respect to λ_i.

We will denote the principal Lagrange resolvent with respect to λ_i, by $\hat{T}_i$.

PROPOSITION 4.2.   *Let*

$$T(x) = \prod_{i=1}^{k}(x - \lambda_i)^{m_i}.$$

Then for $\{\hat{T}_i\}_{i=1}^{k}$ principal Lagrange resolvents on k interpolation points with multiplicities $m_1, \ldots, m_k$ summing to n, we have

(i) $\hat{T}_1 + \ldots + \hat{T}_k - 1 = 0,$
(ii) $\hat{T}_i \hat{T}_j \equiv 0 \mod (T)$ *if* $i \neq j,$
(iii) $(\hat{T}_i)^2 \equiv \hat{T}_i \mod (T),$
(iv) $(t - \lambda_i)^{m_i} \hat{T}_i \equiv 0 \mod (T)$ *for all* $1 \leq i \leq k.$

PROOF.   (i)   Note that $\lambda_1, \ldots, \lambda_k$ are distinct roots with multiplicities $m_1, \ldots m_k$ of the degree at most $(n-1)$ polynomial

$$\hat{T}_1 + \ldots + \hat{T}_k - 1.$$

Since $\sum_{i=1}^{k} m_i = n$, this implies that the polynomial is identically 0.

(ii)   Let $1 \leq i < j \leq k$. Then, for every λ_ℓ, $1 \leq \ell \leq k$ we have either $\hat{T}_i(\lambda_\ell) = 0$ or $\hat{T}_j(\lambda_\ell) = 0$ with multiplicity m_ℓ. And so every root of T is a root of $\hat{T}_i \hat{T}_j$ with at least equal multiplicity. And therefore $\hat{T}_i \hat{T}_j \equiv 0 \mod (T)$.

(iii) Consider the polynomial $(\hat{T}_i)^2 - \hat{T}_i = \hat{T}_i(\hat{T}_i - 1)$. For $\ell \neq i$, λ_ℓ is a root of $\hat{T}_i$ with multiplicity atleast m_ℓ, and therefore a root of multiplicity at least m_ℓ of $\hat{T}_i^2 - \hat{T}_i$.

For λ_i, we have $\hat{T}_i(\lambda_i) = 1$ and $\hat{T}_i^{(j)}(\lambda_i) = 0$ for all $j = 1, 2, \ldots, m_i$. Now,

$$(\hat{T}_i^2 - \hat{T}_i)^{(m)}(\lambda_i) = \sum_{p=0}^{m} \binom{m}{p} \hat{T}_i^{(p)}(\lambda_i)(\hat{T}_i - 1)^{(m-p)}(\lambda_i) = 0$$

for all $m = 0, 1, \ldots, m_i$. So, λ_i is also a root of $\hat{T}_i^2 - \hat{T}_i$ with multiplicity at least m_i. Therefore, we have $\hat{T}_i^2 - \hat{T}_i \equiv 0 \mod (T)$.

(iv)   Similarly, note that λ_j, $j \neq i$ is a root of multiplicity at least m_j for $\hat{T}_i$. And clearly, λ_i is a root of $(t - \lambda_i)^{m_i}$ with multiplicity m_i. Therefore $(t - \lambda_i)^{m_i} \hat{T}_i \equiv 0 \mod (T)$.

□

Consider now a linear operator A over a vector space V (possibly infinite dimensional) over $\mathbb{K}$. Suppose A satisfies a polynomial $T \in \mathbb{K}[x]$ of degree n. Assume T splits over $\mathbb{K}$, and has k different roots $\lambda_1, \ldots, \lambda_k$ with multiplicities $m_1, \ldots, m_k$.

Definition 4.3. The operator $L_i(A) = \widehat{T}_i(A)$ is called principal Lagrange resolvent of the operator A corresponding to the polynomial T and root λ_i, where $\widehat{T}_i(x)$ is the principal Lagrange resolvent corresponding to λ_i.

Theorem 4.4. *The principal Lagrange resolvents L_i of the operator A corresponding to an annihilating polynomial T satisfy the following.*

(i) $L_1(A) + \cdots + L_k(A) = I$, *where I is the identity matrix,*
(ii) $L_i(A)L_j(A) = 0$ *for $i \neq j$,*
(iii) $L_i^2(A) = L_i(A)$,
(iv) $(A - \lambda_i I)^{m_i} L_i(A) = 0$.

Proof.   Each of these expressions is obtained directly from evaluating the corresponding polynomials on the left hand sides of Proposition 4.2 at A. Each of these polynomials is divisible by T, and T annihilates A, so we are able to replace the   mod (T) congruence with equality. $\square$

4.2.   Formulas involving the principal resolvents As a direct corollary of Theorem 3.6 and Example 3.7, we get that for a matrix A over $\mathbb{C}$ having only one eigenvalue, functions of matrices can be computed by principal Lagrange resolvents.

Corollary 4.5. *Let A be an operator on a vector space V over $\mathbb{C}$ such that A satisfies the polynomial $(x - \lambda_1)^n$. Let $F(x)$ be an entire function or a rational function such that λ_1 is not a pole of $F(x)$. Then*

$$F(A) = T_{\lambda_1}^{n-1}(A),$$

where $T_{\lambda_1}^{n-1}(x)$ is the degree $n-1$ Taylor polynomial of F at the point λ_1. Thus,

$$F(A) = T_{\lambda_1}^{n-1}(A) = \sum_{k=0}^{n-1} \frac{1}{k!} F^{(k)}(\lambda_1)(A - \lambda_1 I)^k.$$

Using Theorem 4.4, we get the following more general result.

Corollary 4.6. *Let A be an operator on a vector space V over $\mathbb{C}$ such that A satisfies a polynomial $T(x) \in \mathbb{C}[x]$. Let $\lambda_1, \ldots, \lambda_k$ be distinct roots of T with multipicities $m_1, \ldots, m_k$. Let $F(x)$ be an entire function or a rational function such that $\lambda_1, \ldots, \lambda_k$ are not poles of $F(x)$. Then*

$$F(A) = \sum_{i=1}^{k} T_{\lambda_i}^{m_i-1}(A) L_i(A),$$

where $T_{\lambda_i}^{(m_i-1)}(x)$ is the degree $m_i - 1$ Taylor polynomial of F at point λ_i and $L_i(A)$ is the principal Lagrange resolvent of the operator A corresponding to the polynomial T and root λ_j.

PROOF.   Using Theorem 4.4, we have $F(A) = F(A)L_1(A) + \cdots + F(A)L_k(A)$. For every $1 \leq i \leq k$, $F(A)L_i(A) = F(A_i)L_i(A)$, where A_i is the restriction of operator A on the range V_i of $L_i(A)$. By Theorem 4.4(iv), we have $(A - \lambda_i I)^{m_i}(v) = 0$ for all $v \in V_i$, and so the characteristic polynomial of A_i is $(x - \lambda_i)^{m_i}$. Using Corollary 4.5, we get $F(A_i) = T_{\lambda_i}^{m_i-1}(A)$. $\square$

Remark 1: Corollary 4.6 also holds for any meromorphic function F such that $\lambda_1, \ldots, \lambda_k$ are not poles of $F(x)$.

Now, we present a formula for the principal Lagrange interpolating polynomial with the given interpolation data.

PROPOSITION 4.7.   *Let* $\Lambda = \{\lambda_1, \ldots, \lambda_k\} \subseteq \mathbb{K}$ *be a set of k distinct elements. For all $1 \leq i \leq k$, let $m_i \in \mathbb{N}$ be a natural number associated with λ_i such that $\sum_{i=1}^{k} m_i = n$. Then the principal Lagrange resolvent with respect to λ_i is given by*

$$\widehat{T_i} = T_{\lambda_i}^{m_i-1}(x) \prod_{j \neq i} (x - \lambda_j)^{m_j},$$

where $T_{\lambda_i}^{m_i-1}$ is the Taylor polynomial at λ_i of degree $m_i - 1$ of the function

$$F(x) = \frac{1}{\prod_{j \neq i} (x - \lambda_j)^{m_j}}.$$

PROOF.   By definition we know that the principal Lagrange resolvent is the Lagrange interpolation polynomial with interpolation data $c_1^{(0)} = 1$ and everything else 0. We proceed by verifying that $\hat{T}_i$ as defined is this interpolation polynomial.

First, note that $\deg \hat{T}_i = n-1$ as required. Then, for $j \neq i$ and $0 \leq m \leq m_j-1$ we easily see that

$$\hat{T}_i^{(m)}(\lambda_j) = 0.$$

And now we want to see what happens at λ_i. By definition of the Taylor polynomials we know that

$$T_{\lambda_i}^{m_1-1}(x) - F(x)$$

has a root at λ_i of multiplicity m_i. In particular, we write

$$T_{\lambda_i}^{m_1-1}(x) - F(x) = (x - \lambda_i)^{m_i} G(x)$$

for some $G(x)$ with no pole at λ_i. And then we have

$$\hat{T}_i(x) = F(x) \prod_{j \neq i} (x - \lambda_j)^{m_j} + G(x) \prod_{j=1}^{k} (x - \lambda_j)^{m_j} = 1 + G(x) \prod_{j=1}^{k} (x - \lambda_j)^{m_j}.$$

Since G does not have a pole at λ_i, we see right away that $\hat{T}_i(\lambda_i) = 1$ and

$$\hat{T}_i^{(m)}(\lambda_i) = 0$$

for all $1 \leq m \leq m_i - 1$.

And therefore, $\hat{T}_i$ is indeed the Lagrange interpolation polynomial with data $c_i^{(0)} = 1$ and all others 0.

$\square$

4.3. Application to Jordan normal form

DEFINITION 4.8. For every vector $v \in V$, the vector $v_i = L_i(A)v$ will be called the principal Lagrange resolvent of v corresponding to the root λ_i and the operator A.

Then we get the following corollary to Theorem 4.4:

COROLLARY 4.9. *Every vector $v \in V$ is representable as sum of its principal Lagrange resolvents i.e. $v = v_1 + \cdots + v_k$. Moreover, all nonzero Lagrange resolvents of v are linearly independent and satisfy $(A - \lambda_i I)^{m_i} v_i = 0$.*

Now we restrict to a special case. Let V be a finite dimensional space and T be the characteristic polynomial of A that splits over $\mathbb{K}$. By Cayley Hamilton theorem, $T(A) = 0$. We consider principal Lagrange resolvents $L_i(A)$ of A with resepct to T. Let E_i denote the range of $L_i(A)$, that is, $E_i = \{L_i(A)v : v \in V\}$. Then we have the following result.

THEOREM 4.10. *The following statements are true.*

(i) $V = E_1 \oplus \cdots \oplus E_k$.
(ii) $E_i = \{v \in V : (A - \lambda_i I)^{m_i} v = 0\} = \{v \in V : \text{there exists } p \in \mathbb{N} \text{ such that } (A - \lambda_i I)^p v = 0\}$.
(iii) *The dimension of E_i is m_i, where m_i is the multiplicity of the eigenvalue of A.*

PROOF. Corollary 4.9 is equivalent to $V = E_1 \oplus \cdots \oplus E_k$ and $E_i \subseteq \{v \in V : (A - \lambda_i I)^{m_i} v = 0\}$. Clearly, $V_i = \{v \in V : (A - \lambda_i I)^{m_i} v = 0\}$ is a subspace of V and also invariant under the action of A. So the characteristic polynomial of A restricted to V_i divides the characteristic polynomial T of A. For every $v \in V_i$, the annihilating polynomial of v is of the form $(t - \lambda_i)^p$, so the characteristic polynomial of A restricted to V_i is also of the form $(t - \lambda_i)^p$. Since $(t - \lambda_i)^p$ divides T, we have $p \leq m_i$. So the dimensions of E_i and V_i are less than equal to m_i. Since $V = E_1 \oplus \cdots \oplus E_k$ and $\sum_{i=1}^{k} m_i = n$, we must have the dimensions of E_i and V_i both equal to m_i. This proves (i) and (iii). It also proves $E_i = V_i$. Now let $v \in V$ such that there exists $p \in \mathbb{N}$ such that $(A - \lambda_i I)^p v = 0$. Let $(t - \lambda_i)^p$ itself be the annihilating polynomial of v. By definition, $(t - \lambda_i)^p$ divides characteristic polynomial T of A. Hence $p \leq m_i$. So $(A - \lambda_i I)^{m_i} v = 0$. Hence $v \in E_i$. This completes the proof of the theorem. $\square$

Now we are ready to prove the Jordan decomposition theorem for matrices whose characteristic polynomial splits. We need the following definition.

DEFINITION 4.11.   Let A be a linear operator on a vector space V. Let $v \in V$ be such that there exists $m \in \mathbb{N}$ such that $(A - \lambda I)^m(v) = 0$. Then v is said to be a generalized eigenvector of A corresponding to the eigenvalue λ. Suppose that p is the smallest positive integer for which $(A - \lambda I)^p(v) = 0$. Then the set $\{v, (A - \lambda I)(v), (A - \lambda I)^2(v), \dots (A - \lambda I)^{p-1}(v)\}$ is called the cycle of generalized eigenvectors of A corresponding to λ generated by v.

The proof of the following lemma is straightforward.

LEMMA 4.12.   *Let $v_1, \dots, v_k$ be linearly independent generalized eigenvector of A corresponding to the eigenvalue λ. Then for every i, the cycle of generalized eigenvectors generated by v_i is linearly independent.*

THEOREM 4.13.   *Let A be a linear operator on a finite-dimensional vector space V such that characteristic polynomial of A splits. Then for every eigenvalue λ of A, there exists a basis of E_λ which is the union of cycles of generalized eigenvectors of A corresponding to λ.*

PROOF.   Consider the maximal set $\mathcal{B}$ of linearly independent generalized eigenvectors of A corresponding to the eigenvalue λ. And the claim is that it is a basis for E_λ, else there will be a generalized eigenvector not belonging to $E_\lambda \setminus \text{span}(\mathcal{B})$, which will contradict maximality of $\mathcal{B}$.   □

And finally using Theorem 4.10 and the basis of E_λ consisting of cycle generalized eigenvectors A corresponding to λ, the matrix represntation of A is the Jordan decomposition of matrix A.

4.4.   Resolvents associated to families of commutative operators Finally, we remark that the construction of interpolation polynomials with multiple roots allows a direct generalization of Section 2.1.3 in [3].

Let A and B be commuting operators satisfying polynomial relations $T(x)$ and $Q(x)$ respectively. And set

$$L_{i,j} = \hat{T}_i(A)\hat{Q}_j(B)$$

where $\hat{T}_i(A)$ and $\hat{Q}_j(B)$ are the principal Lagrange resolvents of A and B corresponding to the roots λ_i and μ_j. Then we directly generalize Proposition 2.6 of [3] as follows (noting that we now allow T and Q to have repeated roots):

PROPOSITION 4.14.   *The following statements hold:*

(i)   *The generalized Lagrange resolvents $L_{i,j}$ of commuting operators A and B satisfy the following relations:*

$$\sum L_{i,j} = E, \quad L_{i_1,j_1} L_{i_2,j_2} = 0 \text{ for } (i_1, j_1) \neq (i_2, j_2),$$

$$L_{i,j}^2 = L_{i,j}, \quad AL_{i,j} = \lambda_i L_{i,j}, \quad BL_{i,j} = \mu_j L_{i,j}.$$

(ii) *Every vector $x \in V$ is representable as the sum of its generalized Lagrange resolvents:*

$$x = \sum L_{i,j} x.$$

Moreover, the nonzero resolvents $L_{i,j}x$ are linearly independent and simultaneously eigenvectors of A and B with eigenvalues λ_i and λ_j respectively.

REFERENCES

1. R.L. Burden, J.D. Faires. *Numerical Analysis, 9th Edition*. Cengage Learning. 2010.
2. M.W. Hirsch, S. Smale. *Differential Equations, Dynamical Systems, and Linear Algebra*. Academic Press. 1974.
3. A. Khovanskii. *Topological Galois Theory*. Springer Monographs in Mathematics. 2014.

University of Toronto, Toronto, Canada
e-mail: askold@math.toronto.edu

Department of Mathematics, Shiv Nadar University, Greater Noida, India 201314
e-mail: ss774@snu.edu.in

University of Toronto, Toronto, Canada
e-mail: tronsgar@math.toronto.ca