The Generalized Spectral Decomposition of a Linear Operator

Garret Sobczyk*
Universidad de las Americas
Departamento de Fisico-Matematicas
Apartado Postal #100, Santa Catarina Mártir
72820 Puebla, Pue., México
sobczyk@udlapvms.pue.udlap.mx

September 25, 1996

Abstract

The *generalized spectral decomposition* expressing a linear operator as a sum of idempotent and nilpotent operators is derived from the partial fraction decomposition of the reciprocal of its minimal polynomial, and makes possible a functional calculus of analytic functions of the operator. The classical spectral and polar decompositions for matrices are shown to be easy consequences. The idempotents and associated nilpotents provide a basis for the commutative algebra generated by an operator that is far more convenient for computations than the usual basis of powers of the operator.

(*The College Mathematics Journal*, January 1997)

*I gratefully acknowledge the support given by INIP of the Universidad de las Americas, and CONACYT (The Mexican National Council for Science and Technology) grant 3803-E.

The ring of polynomials in a given linear transformation with coefficients from a field may have quite a complicated algebraic structure; it may contain nilpotent elements and it may have zero divisors.

N. Jacobson *Lectures in Abstract Algebra,* vol 2, p. 75.

A linear operator t on a finite dimensional vector space V over a field F that contains all the eigenvalues of t generates a commutative algebra $F[t]$ with the standard basis $[1, t, t^2, \ldots, t^{m-1}]$, where 1 denotes the identity operator. But there is another basis for this algebra, consisting of idempotents and associated nilpotent operators, that is far more convenient for computations than the standard basis. The *generalized spectral decomposition* of t, the expression for t in terms of its *sprectral* basis, is directly computable from the minimal polynomial of t. Many important invariants of an operator follow from this decomposition. For example, the inverse t^{-1}, the logarithm $\log t$, and other analytic functions of t can be constructed at once, as explicit polynomials in t. The entire discussion is independent of the dimension of V or details of the action of t on this vector space— everything follows by simple algebra from the decomposition of the reciprocal of the minimal polynomial into partial fractions.

To emphasize the advantages of this approach, I show how it provides very simple proofs of standard theorems: the classical spectral theorem for self-adjoint operators and the polar decomposition theorem of a linear operator. Although the results are presented here using standard methods of abstract algebra to make the discussion logically airtight, I include many examples and indicate along the way how the ideas can be presented intuitively to students. Indeed, students enjoy calculating with the spectral basis, since everything is commutative and the basis elements are either idempotent or nilpotent. Experts as well as students should find our approach appealing because it makes avaiable powerful techniques that were here-to-fore considered beyond the reach of elementary mathematics.

The spectral basis for the algebra $F[t]$

The *minimal polynomial* of an operator t is the polynomial $\psi(x)$ of lowest degree, with leading coefficient 1, such that $\psi(t) = 0$. That is, if $\psi(x) = x^m + a_{m-1}x^{m-1} + \ldots + a_1 x + a_0$ then $\psi(t) \equiv t^m + a_{m-1}t^{m-1} + \ldots + a_1 t + a_0 I = 0$, where

I denotes the identity operator. Any other polynomial in $F[x]$ that similarly yields the zero operator when x is replaced by t is a polynomial multiple of $\psi(x)$. The equation $\psi(t) = 0$ implies that t^m is a linear combination of $\{I, t, \ldots, t^{m-1}\}$, namely $t^m = -a_{m-1}t^{m-1} - \ldots - a_1 t - a_0 I$. By induction all higher powers of t are also linear combinations of $\{I, t, \ldots, t^{m-1}\}$. The mapping that sends 1 to I and $x \to t$ extends to an homormorphism from the commutative algebra $F[x]$ of polynomials in the indeterminate x onto the algebra $F[t]$ of polynomials in the operator t. The kernel consists of all polynomial multiples of the minimal polynomial $\psi(x)$, so the algebra $F[t]$ is isomorphic to the quotient $F[x]/(\psi(x))$.

Via the homomorphism from $F[x]$ to $F[t]$, each element of F is identified with multiplication by that element. Henceforth I shall write 0 for the zero operator, 1 for the identity operator, and in general a for multiplication by any field element a. The symbol I will be reserved for the identity matrix. This abuse of notation generalizes the identification of real numbers with the complex numbers whose imaginary part is zero. A special case of our discussion is Cauchy's observation that $\mathbb{C}$ is isomorphic to the quotient $\mathbb{R}[x]/(\psi(x))$ for $\psi(x) = x^2 + 1$. The hyperbolic numbers [4] are another example of this quotient with $\psi(x) = x^2 - 1$. Rather than thinking of its elements as equivalence classes of polynomials, we shall view $F[t]$ as an extension of F, obtained by adjoining new elements (the spectral basis elements) that satisfy specified multiplication rules.

We now assume that the prime factors of $\psi(x)$ in the ring $F[x]$ are linear; that is, all the roots $\lambda_1, \lambda_2, \ldots, \lambda_r$ of $\psi(x)$ lie in F. Of course if $F = \mathbb{C}$, the field of complex numbers this will always be the case. We will later show that these roots are precisely the *eigenvalues* of t. In the meantime we will anticipate this result by referring to $\lambda_1, \ldots, \lambda_r$ as the eigenvalues of t. Suppose the distinct eigenvalues λ_i of the minimal polynomial

$$\psi(x) \equiv \prod_{i=1}^{r} (x - \lambda_i)^{m_i}$$

have the respective *algebraic multiplicities* m_i, and the degree of ψ is $m = m_1 + m_2 + \ldots + m_r$. We order the eigenvalues so that their multiplicities satisfy $1 \le m_1 \le m_2 \le \ldots \le m_r$, and let $h \ge 0$ be the number of m_i that equal 1.

Several ways are available for finding the spectral basis of t, perhaps the simplest being to use the partial fraction decomposition of the reciprocal of the minimal polynomial.

$$\frac{1}{\psi(x)} = \sum_{i=1}^{r} \left[\frac{a_0}{(x-\lambda_i)^{m_i}} + \frac{a_1}{(x-\lambda_i)^{m_i-1}} + \cdots + \frac{a_{m_i-1}}{(x-\lambda_i)} \right]$$

$$= \frac{h_1(x)}{(x-\lambda_1)^{m_1}} + \cdots + \frac{h_r(x)}{(x-\lambda_r)^{m_r}}$$

$$= \frac{h_1(x)\psi_1(x) + \cdots + h_r(x)\psi_r(x)}{\psi(x)},$$

where $h_i(x) = \sum_{\ell=0}^{m_i-1} a_\ell (x-\lambda_i)^\ell$ and $\psi_i(x) = \prod_{j\neq i}(x-\lambda_j)^{m_j}$.

Since $\sum_{i=1}^{r} h_i(x)\psi_i(x) = 1$, if we set $p_i = h_i(t)\psi_i(t)$ then in $F[t]$ $\sum_{i=1}^{r} p_i = 1$. Because $\psi_i(x)\psi_j(x)$ is divisible by $\psi(x)$ whenever $i \neq j$, we have $p_i p_j = 0$. Also $p_i^2 = (p_1 + \cdots + p_r)p_i = 1p_i = p_i$, so the operators $p_1 \ldots p_r$ are *mutually annihilating idempotents* that form a *partition of unity.* An analyst might say the p_i are a *supplementary family of orthogonal projections*, but here the algebraist's terminology is preferred, to emphasize the purely *algebraic* properties of these elements of $F[t]$.

For $1 \le i \le r$, set $q_i = (t-\lambda_i)p_i$. Then $q_i = 0$ when $m_i = 1$ (that is, for $i = 1, \ldots, h$), and when $m_i > 1$ (so $i > h$), $q_i^{m_i-1} \neq 0$ but $q_i^{m_i} = 0$. Thus $q_1 \ldots q_r$ are nilpotents in $F[t]$ with indices of nilpotency $m_1 \ldots m_r$. Also, $p_j q_i = 0$ if $j \neq i$, so $p_i q_i = (p_1 + \cdots + p_r)q_i = 1q_i = q_i$; thus the nilpotents $q_1 \ldots q_r$ are said to be *projectively related* to the idempotents $p_1 \ldots p_r$. To summarize, the commutative algebra $F[t]$ has a basis, the *spectral basis*

$$\{p_1, \ldots, p_h, \overbrace{p_{h+1}, q_{h+1}, q_{h+1}^2, \ldots q_{h+1}^{m_{h+1}-1}}^{m_{h+1}}, \ldots, \overbrace{p_r, q_r, q_r^2, \ldots, q_r^{m_r-1}}^{m_r}\} \qquad (1)$$

satisfying the properties

$$p_1 + \ldots + p_r = 1,\ p_i p_j = \delta_{ij} p_i,\ q_k^{m_k-1} \neq 0 \text{ but } q_k^{m_k} = 0,\ p_k q_k = q_k \qquad (2)$$

Note that the number of members of the basis (2) is $m_1 + \cdots + m_r = m$.

Example 1. Consider a linear operator t with the minimal polynomial $\psi(x) = (x-\lambda_1)(x-\lambda_2)^2$. Thus the degree of the minimal polynomial is $m = 3$, there are $r = 2$ distinct eigenvalues, with algebraic multiplicities

$m_1 = 1$ and $m_2 = 2$. Since $h = 1$, we have $q_1 = 0$. The spectral basis for $F[t]$ is therefore $\{p_1, p_2, q_2\}$. To express these basis vectors in terms of the standard basis $\{1, t, t^2\}$, we use the partial fraction decomposition

$$\frac{1}{(x-\lambda_1)(x-\lambda_2)^2} = \frac{\frac{1}{(\lambda_2-\lambda_1)^2}}{x-\lambda_1} + \frac{\frac{-1}{(\lambda_2-\lambda_1)^2}}{x-\lambda_2} + \frac{\frac{1}{(\lambda_2-\lambda_1)}}{(x-\lambda_2)^2}.$$

Putting the right side over the common denominator $\psi(x)$ and comparing numerators gives

$$1 = \frac{1}{(\lambda_2-\lambda_1)^2}(x-\lambda_2)^2 + \frac{1}{(\lambda_2-\lambda_1)^2}\left[-(x-\lambda_1)(x-\lambda_2) + (\lambda_2-\lambda_1)(x-\lambda_1)\right]$$

$$= \frac{(x-\lambda_2)^2}{(\lambda_2-\lambda_1)^2} + \frac{(x-\lambda_1)(2\lambda_2-\lambda_1-x)}{(\lambda_2-\lambda_1)^2}.$$

Thus

$$p_1 = \frac{(t-\lambda_2)^2}{(\lambda_2-\lambda_1)^2}, \quad q_1 \equiv 0,$$

$$p_2 = \frac{(t-\lambda_1)(2\lambda_2-\lambda_1-t)}{(\lambda_2-\lambda_1)^2}.$$

Then

$$q_2 = (t-\lambda_2)p_2 = \frac{(t-\lambda_2)(t-\lambda_1)}{(\lambda_2-\lambda_1)},$$

where the calculation requires reducing $(x-\lambda_2)\dfrac{(x-\lambda_1)(2\lambda_2-\lambda_1-x)}{(\lambda_2-\lambda_1)^2}$ modulo $\psi(x)$ and then replacing x by t in the remainder. (The partial fraction decomposition and the polynomial algebra are easy to carry out using a computer algebra system such as *Mathematica*.) It is readily verified that these polynomial expressions in t satisfy the required conditions

$$p_1 + p_2 = 1, \quad p_i p_j = \delta_{ij} p_i, \quad q_2 \neq 0 \quad \text{but} \quad q_2^2 = 0, \quad p_2 q_2 = q_2,$$

because t satisfies its minimal polynomial: $(t-\lambda_1)(t-\lambda_2)^2 = 0$.

Since the exponents $m_1, \ldots, m_r$ of the prime factors of $\psi(x)$ completely determine the multiplication table for the spectral basis, we can forget all about linear operators when computing in the algebra $F[t]$. Given any finite non-decreasing sequence of positive intergers $m_1, m_2, \ldots, m_r$, by using the properties (2) we can write down an unambiguous multiplication table for basis elements (1) to define an algebra $F\{m_1, \ldots, m_r\} \equiv F[t]$. Thus the algebra $F[t]$ in Example 1 might be denoted $F1, 2$. Students need not struggle with equivalence classes of polynomials to compute using the spectral basis in $F\{m_1, \ldots, m_r\}$ any more than they do to work in $\mathbb{C}$.

The *generalized spectral decomposition* of the linear operator t is the equation

$$t = \sum_{i=1}^{r} (\lambda_i + q_i) p_i, \tag{3}$$

expressing the operator in terms of the spectral basis (1). This is a special case of what is often called the *primary decomposition* of a linear operator—the case when the prime factors of $\psi(\lambda)$ are all powers of linear factors [1, p. 197].

To indicate how easy calculations are when performed using the spectral basis, let's find the expressions for the standard basis $\{1, t, t^2, \ldots, t^{m-1}\}$ in terms of the spectral basis (1). Using (2) and (3),

$$\begin{aligned}
1 &= p_1 + p_2 + \ldots + p_r, \\
t &= (\lambda_1 + q_1)p_1 + (\lambda_2 + q_2)p_2 + \ldots + (\lambda_r + q_r)p_r. \\
t^2 &= (\lambda_1 + q_1)^2 p_1 + (\lambda_2 + q_2)^2 p_2 + \ldots + (\lambda_r + q_r)^2 p_r \\
&\ \ \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\
t^{m-1} &= (\lambda_1 + q_1)^{m-1} p_1 + (\lambda_2 + q_2)^{m-1} p_2 + \ldots + (\lambda_r + q_r)^{m-1} p_r.
\end{aligned}$$

The powers $(\lambda_j + q_j)^k$ are easily computed by the binomial theorem. Since q_j is nilpotent with index m_j, the result may be expressed as

$$(\lambda_j + q_j)^k = \lambda_j^k + \binom{k}{1}\lambda_j^{k-1}q_j + \ldots + \binom{k}{m_j - 1}\lambda_j^{k-m_j+1}q_j^{m_j-1}, \qquad (4)$$

where $\binom{k}{j}$ are the usual binomial coefficients for $j \leq k$, and $\binom{k}{j} \equiv 0$ for $j > k$. Substituting (4) in the equations above then expresses the powers of t in terms of the spectral basis (1). This will prove useful in the next section.

Analytic functions of an operator

It is often necessary to consider analytic functions of a linear operator $t = \sum_{i=1}^{r}(\lambda_i + q_i)p_i$. The algebra $F[t] \equiv F\{m_1, \ldots, m_r\}$ is the perfect setting for such considerations. We will consider here only the case when $F = \mathbb{R}$ or $\mathbb{C}$.

Let $g(z)$ be any function which has a convergent Taylor series around each of the eigenvalues $z = \lambda_i \in \mathbb{C}$ of the operator t. Then $g(t)$ is defined to be the sum of the series obtained by replacing z by t in a Taylor series for $g(z)$, just as we evaluate polynomials in t. In particular, if $g(z) = \sum_{k=0}^{\infty} a_k z^k$ is the Taylor series around the origin, then $g(t) = a_0 + a_1 t + a_2 t^2 + \cdots$. Using formulas (2) and (4) we find

$$g(t) = a_0[p_1 + \cdots + p_r] + a_1[(\lambda_1 + q_1)p_1 + \cdots + (\lambda_r + q_r)p_r] +$$

$$a_2[(\lambda_1 + q_1)^2 p_1 + \cdots + (\lambda_r + q_r)^2 p_r] + a_3[(\lambda_1 + q_1)^3 p_1 + \cdots + (\lambda_r + q_r)^3 p_r] + \cdots$$

$$= [a_0 + a_1\lambda_1 + a_2\lambda_1^2 + a_3\lambda_1^3 + \cdots]p_1 + [a_1 + \binom{2}{1}a_2\lambda_1 + \binom{3}{1}a_3\lambda_1^2 + \cdots]q_1 +$$

$$[a_2 + \binom{3}{2}a_3\lambda_1 + \binom{4}{2}a_4\lambda_1^2 + \cdots]q_1^2 + \cdots$$

$$= g(\lambda_1)p_1 + g'(\lambda_1)q_1 + \frac{1}{2!}g''(\lambda_1)q_1^2 + \cdots + \frac{1}{(m_1 - 1)!}g^{(m_1-1)}(\lambda_1)q_1^{(m_1-1)} \quad + \cdots +$$

$$g(\lambda_r)p_r + g'(\lambda_r)q_r + \frac{1}{2!}g''(\lambda_r)q_r^2 + \cdots + \frac{1}{(m_r - 1)!}g^{(m_r-1)}(\lambda_r)q_r^{(m_r-1)}.$$

If we define

$$g(\lambda_i + q_i) \equiv g(\lambda_i)p_i + g'(\lambda_i)q_i + \ldots + \frac{1}{(m_i-1)!}g^{(m_i-1)}(\lambda_i)q_i^{m_i-1} \quad (5)$$

the result above can be written compactly as

$$g(t) = \sum_{i=1}^{r} g(\lambda_i + q_i)p_i, \quad (6)$$

For instance, if, as in Example 1, the minimal polynomial of t is $\psi(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)^2$, then by (6) the exponential e^t is given in terms of the spectral basis by

$$e^t = e^{\lambda_1}p_1 + e^{\lambda_2 + q_2}p_2$$

$$= e^{\lambda_1}p_1 + e^{\lambda_2}p_2 + \lambda_2 e^{\lambda_2}q_2.$$

If t is invertible, so $\lambda_1, \lambda_2 \neq 0$, the inverse is given by the expression

$$t^{-1} = \frac{1}{\lambda_1}p_1 + \frac{1}{\lambda_2 + q_2}p_2 = \frac{1}{\lambda_1}p_1 + \frac{\lambda_2 - q_2}{\lambda_2^2}p_2,$$

and just as easily we find

$$\log(t) = \log(\lambda_1)p_1 + \log(\lambda_2 + q_2)p_2 = \log(\lambda_1)p_1 + [\log(\lambda_2) + \frac{q_2}{\lambda_2}]p_2,$$

and

$$t^{1/2} = \lambda_1^{1/2}p_1 + (\lambda_2 + q_2)^{1/2}p_2 = \lambda_1^{1/2}p_1 + \lambda_2^{1/2}p_2 + \frac{1}{2\lambda_2^{1/2}}q_2.$$

The generalized spectral decomposition (3) therefore determines the *analytic structure* of the operator t, because (5) and (6) define $g(t)$ for any function g which is analytic on the spectrum of t. A good exercise for students with a computer algebra system is to take a given square matrix T, find its minimal polynomial and express the associated spectral basis for $F[T]$ as polynomials in T, as in Exercise 1. Then, reasoning as above, they can compute various analytic functions of the matrix as polynomials in T, and ultimately as single matrices.

The Jordan Canonical Form

With the structure equation in hand, it is easy to describe the action of t on a vector space V. The range V_i of the projection p_i is called the *generalized eigenspace* corresponding to the *eigenvalue* λ_i. Relations (2) imply that V is the direct sum $V_1 \oplus V_2 \oplus \ldots \oplus V_r$. On V_i the difference $t - \lambda_i p_i$ coincides with the restriction to V_i of the nilpotent transformstion q_i, because $tp_i = (\lambda_i + q_i)p_i = \lambda_i p_i + q_i$. Since $p_i q_i = q_i$, the range of q_i is contained in V_i, and in the chain of subspaces

$$V_i = p_i(V) \supset q_i(V) \supset q_i^2(V) \supset \ldots \supset q_i^{m_i-1}(V) \supset q_i^{m_i}(V) = \{0\}$$

the m_i inclusions are all proper, for if any were an equality then all later inclusions would be equalities as well and q_i would not be nilpotent of index m_i. Thus $dim(q_i^{m_i-1}(V)) \geq 1,\ dim(q_i^{m_i-2}(V)) \geq 2, \ldots\ dim(q_i(V)) \geq m_i - 1$, and $dim(V_i) \geq m_i$. Hence if we set $n_i \equiv\ dim(V_i)$, so $n \equiv n_1 + \ldots + n_r = \dim V$, the *heights* of the eigenvalues (indices of nilpotency) m_i satisfy $m_i \leq n_i$.

Now consider the matrix T of t restricted to V_i relative to a basis consisting of a basis for $ker(q_i|v_i)$ extended to a basis for $ker(q_i^2|V_i)$ then extended to a basis for $ker(q_i^3|V_i)$, and so on until we have a basis for $ker(q_i^{m_i}|V_i = V_i$. Since the restriction of t to V_i coincides with the restriction of $\lambda_i p_i + q_i$ to this generalized eigenspace, the matrix T is upper triangular with λ_i down the diagonal. Thus the characteristic polynomial of the restriction of t to the generalized eigenspace V_i is $\det(\lambda I - T) = \prod_{i=1}^r (\lambda - \lambda_i)^{n_i}$, where I is the $n \times n$ identity matrix. The minimal polynomial $\psi(\lambda) = \prod_{i=1}^r (\lambda - \lambda_i)^{m_i}$ therefore has the same roots as the characteristic polynomial, and is a divisor of the characteristic polynomial. The Cayley-Hamilton theorem, that a linear operator satisfies its characteristic polynomial, is an immediate consequence.

A theorem of Frobenius [3, p. 99] guarantees that if $\gamma(x)$ is the greatest common divisor of all the subdeterminants of the characteristic matrix $xI - T$ that one obtains by deleting a row and a column, then

$$\psi(x) = \frac{\det(xI - T)}{\gamma(x)}.$$

Using a computer algebra system, this formula can be used to quickly compute the minimum polynomial of t from its matrix with respect to any basis for V, provided the dimension of V is not too large.

For questions that involve the action of t on V, the structure equation can be helpful in finding a basis for V relative to which the matrix of t is in Jordan canonical form, a form that displays the geometric effect of the linear transformation in more detail than the triangular form discussed above. But because a general description of the procedure is rather tedious, I will just indicate it by a simple example.

Example 2. Find a basis for $\mathbb{R}^4$ in which

$$T = \begin{pmatrix} 15 & -3 & 9 & 3 \\ -3 & 23 & 3 & 1 \\ 3 & 1 & 5 & 15 \\ 9 & 3 & 7 & 5 \end{pmatrix}$$

assumes Jordan canonical form.

Solution. The characteristic polynomial of T is $\det(xI - T) = (x - 24)^2 x^2$, and the minimal polynomial is $\psi(x) = (x - 24)x^2$, as can easily be verified. Thus the structure equation is $T = 24P_1 + 0P_2 + Q_2$. Using our earlier formulas from Example 1, with $\lambda_1 = 24$ and $\lambda_2 = 0$, we find

$$P_1 = \frac{1}{24^2}T^2 = \frac{1}{18}\begin{pmatrix} 9 & -3 & 6 & 6 \\ -3 & 17 & 2 & 2 \\ 6 & 2 & 5 & 5 \\ 6 & 2 & 5 & 5 \end{pmatrix} \quad \text{and} \quad Q_1 = 0$$

$$P_2 = \frac{(24I - T)(T + 24I)}{24^4} = \frac{1}{18}\begin{pmatrix} 9 & 3 & -6 & -6 \\ 3 & 1 & -2 & -2 \\ -6 & -2 & 13 & -5 \\ -6 & -2 & -5 & 13 \end{pmatrix} \text{ (or use } P_2 = I - P_1\text{)},$$

$$Q_2 = \frac{1}{3}\begin{pmatrix} 9 & 3 & 3 & -15 \\ 3 & 1 & 1 & -5 \\ -15 & -5 & -5 & 25 \\ 3 & 1 & 1 & -5 \end{pmatrix} \text{ (or use } Q_2 = T - 24P_1\text{)}.$$

Since $m_1 = 1$, the Jordan chains that form a basis of V_1 are all of length 1; that is, V_1 has a basis of eigenvectors, found by choosing a basis for the column space of P_1, say $v_1 = [9, -3, 6, 6]^T$, and $v_2 = [6, 2, 5, 5]^T$. Since $m_2 = 2 = n_2$, a single Jordan chain of length m_2 forms a basis for the generalized eigenspace V_2, and it consists of a column vector that forms a basis for the column space of Q_2, together with the corresponding column vector of P_2, say $v_3 = \frac{1}{3}[3, 1, -5, 1]^T$ and $v_4 = \frac{1}{18}[3, 1, -2, -2]^T$, the *second* columns of Q_2 and P_2, respectively. Hence the matrix of T with respect to the basis v_1, v_2, v_3, v_4 has the Jordan canonical form

$$T_J = \begin{pmatrix} 24 & 0 & 0 & 0 \\ 0 & 24 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Some Basic Theorems

If $\langle\ ,\ \rangle$ is an inner product on V over the field $\mathbb{C}$ of complex numbers, the adjoint t^* of an operator t is uniquely defined by

$$\langle tx, y\rangle = \langle x, t^*y\rangle$$

for all vectors $x, y \in V$. Note also that $\langle x, ty\rangle = \langle t^*x, y\rangle$, since $\langle x, ty\rangle = \overline{\langle ty, x\rangle} = \overline{\langle y, t^*x\rangle} = \langle t^*x, y\rangle$. The operator t is called *normal* if $tt^* = t^*t$, *self-adjoint* if $t^* = t$, and *unitary* if $t^* = t^{-1}$; it follows that self-adjoint operators and unitary operators are normal.

We will need three simple lemmas for the proof of the spectral theorem for normal operators.

Lemma 1 *If t is a normal operator then $ker(t) = ker(t^*)$.*

Proof: For any $x \in V$ we have,

$$\langle tx, tx\rangle = \langle x, t^*tx\rangle = \langle x, tt^*x\rangle = \langle t^*x, t^*x\rangle,$$

from which it follows that $tx = 0$ if and only if $t^*x = 0$.

Q.E.D.

Lemma 2 *Every normal nilpotent operator q is identically* 0.

Proof: Let q be a normal nilpotent operator with index of nilpotency $k > 1$; that is, $q^{k-1} \neq 0$ but $q^k = 0$. Then for all x, $qq^{k-1}x = 0$, so by Lemma 1 $q^*q^{k-1}x = 0$. But then

$$\langle q^{k-1}x, q^{k-1}x \rangle = \langle q^*q^{k-1}x, q^{k-2}x \rangle = 0$$

for all x, which means that $q^{k-1} = 0$, a contradiction. Thus a normal nilpotent operator q must have index of nilpotency 1, i.e., $q = 0$.

Q.E.D.

Lemma 3 *A normal projection p is self-adjoint.*

Proof:

Note that the adjoint of a projection p is also a projection, for $(p^*)^2 = (p^2)^* = p^*$. The range of $1 - p$ is $ker(p)$. Thus if p is normal then by Lemma 1

$$ker(p) = ker(p^*) \text{ so } 1 - p = 1 - p^*,$$

from which it follows that $p = p^*$.

Q.E.D.

Theorem 1 (The spectral decomposition theorem for normal operators) *If t is a normal operator, then $t = \sum_{i=1}^{r} \lambda_i p_i$, where the p_i's are self-adjoint orthogonal idempotents with $\sum p_i = 1$.*

Proof: By the structure theorem

$$t = \sum_{i=1}^{r} (\lambda_i + q_i) p_i,$$

where the operators p_i and q_i are all *polynomials* in t. But a polynomial in a normal operator is normal, so the projections p_i and nilpotents q_i are themselves normal operators. Thus by Lemmas 2 and 3, each q_i is zero and each p_i is self- adjoint.

Q.E.D.

The following corollary now follows trivially

Corollary 1 *The eigenvalues of a self-adjoint operator are real.*

Proof:

Just compare the two expressions of $t^* = t$,

$$t = \sum_{i=1}^{r} \lambda_i p_i$$

where $p_i^* = p_i$, and

$$t^* = \sum_{i=1}^{r} \overline{\lambda_i} p_i$$

to conclude that $\overline{\lambda_i} = \lambda_i$ for $i = 1, \ldots, r$.

Q.E.D.

The polar factorization

As a final demonstration of the pedagogical uses of the generalized spectral decomposition (3), we will use it to give a constructive derivation of the *polar factorization* of the operator t.

Since tt^* is self-adjoint,

$$tt^* = \sum_{i=1}^{r} \lambda_{r_i}^2 p_{r_i},$$

and since

$$0 \leq \langle t^* x, t^* x \rangle = \langle tt^* x, x \rangle = \langle \lambda^2 x, x \rangle = \lambda^2 \langle x, x \rangle$$

for any eigenvector x with the eigenvalue λ^2, it follows that the eigenvalues of tt^* are all non negative. The square root of tt^* is then explicitly given by

$$h_r \equiv \sqrt{tt^*} = \sum_i \lambda_{r_i} p_{r_i},$$

for orthogonal projections p_{r_i}. Similarly,

$$h_l \equiv \sqrt{t^* t} = \sum_i \lambda_{l_i} p_{l_i},$$

is well-defined for the orthogonal projections p_{l_i}. We will shortly see that the orthogonal projections p_{r_i} and p_{l_i}, can be ordered to have the same rank and eigenvalues, *i.e.*, $\lambda_i \equiv \lambda_{r_i} = \lambda_{l_i}$.

Theorem 2 *[5] Every operator t has the polar form $t = h_r u = u h_l$, where h_r and h_l are the right and left self-adjoint operators of t defined above, and u is unitary.*

Proof: We begin by observing that if $x = \sum_{i=1}^{r} p_i x = \sum_{i=1}^{r} x_i$ is the decomposition of a typical vector into orthogonal components in the generalized eigenspaces for $t^* t$, then

$$\langle t x_i, t x_k \rangle = \langle t^* t x_i, x_k \rangle = \lambda_i^2 \langle x_i, x_k \rangle = \lambda_i^2 \delta_{ik} \langle x_i, x_k \rangle.$$

It follows that for $\lambda_i \neq 0$

$$u_i x_i \equiv \frac{1}{\lambda_i} t x_i,$$

defines an orthogonal unitary transformation u_i on the generalized eigenspace $V_i = p_i V$. If zero is an eigenvalue, say $\lambda_j = 0$, we have $\langle t x_j, t x_j \rangle = \lambda_j^2 \langle x_j, x_j \rangle = 0$, so $t x_j = 0$. In this singular case we define u_j to be any unitary map from V_j to the orthogonal complement of the sum of the (orthogonal) range sapces $u_i V_i = t V_i$, $i \neq j$.

Our polar decomposition is completed by defining $u = \sum_i u_i$ and $h_l' = \sum_i \lambda_i u_i u_i^{-1}$. Thus h_l' acts as scalar multiplication by λ_i on each of the subspaces $u_i V_i$, so it is self-adjoint. It is easily verified that $t = u h_l = h_l' u$. Then

$$h_r^2 = t t^* = h_l' u u^* h_l' = (h_l')^2$$

implies that $h_l' = h_r$. Hence h_r and h_l have the same eigenvalues and the associated projections can be reordered so that p_i' and p_i have the same rank. Q.E.D.

Corollary 2 *Every operator t has the polar decomposition $t = h_r \exp(i h_\theta) = \exp(i h_\theta) h_l$, where h_r, h_l and h_θ are self-adjoint operators.*

Proof: We need only show that a unitary operator u can always be expressed in the form $u = \exp i h_\theta$. Using our operator calculus, we define $h_\theta = -i \log u$. Since u is unitary, and therefore normal, u has the eigenprojector form $u = \sum_{i=1}^{r} \omega_i p_i$ where the p_i's are self-adjoint idempotents, and $\omega_i \overline{\omega_i} = 1$. It easily follows that

$$h_\theta^* = i \log u^* = i \log u^{-1} = -i \log u = h_\theta,$$

showing that h_θ is self-adjoint.

Q.E.D.

Aside from being conceptually straightforward, our proof constructs an orthogonal transformation u in such a way that $t = h_r u = u h_l$; it is not necessary to have *different* orthogonal transformations in the respective *left* and *right* polar decompositions, [2, p. 276].

Example 3. Find the *left* polar decomposition $T = UH_l$ of the singular matrix T of Example 2.

Solution. The first step is to calculate $H_l = \sqrt{T^*T}$. The eigenvalues of T^*T are $\tau_1^2 = 24^2, \tau_2^2 = 0, \tau_3^2 = 12^2$. Since the T^*T is self-adjoint, it follows that its minimal polynomial is $(\lambda - \tau_1^2)(\lambda - \tau_2^2)(\lambda - \tau_3^2)$. Solving the structure equation $T^*T = \tau_1^2 P_1 + \tau_2^2 P_2 + \tau_3^2 P_3$, we find

$$P_1 = \frac{(\tau_2^2 - t^*t)(\tau_3^2 - t^*t)}{(\tau_2^2 - \tau_1^2)(\tau_3^2 - \tau_1^2)} = \frac{1}{18} \begin{pmatrix} 9 & -3 & 6 & 6 \\ -3 & 17 & 2 & 2 \\ 6 & 2 & 5 & 5 \\ 6 & 2 & 5 & 5 \end{pmatrix}$$

$$P_2 = \frac{(t^*t - \tau_1^2)(\tau_3^2 - t^*t)}{(\tau_2^2 - \tau_1^2)(\tau_3^2 - \tau_2^2)} = \frac{1}{36} \begin{pmatrix} 9 & 3 & -15 & 3 \\ 3 & 1 & -5 & 1 \\ -15 & -5 & 25 & -5 \\ 3 & 1 & -5 & 1 \end{pmatrix}$$

$$P_3 = \frac{(t^*t - \tau_1^2)(t^*t - \tau_2^2)}{(\tau_3^2 - \tau_1^2)(\tau_3^2 - \tau_2^2)} = \frac{1}{36} \begin{pmatrix} 9 & 3 & 3 & -15 \\ 3 & 1 & 1 & -5 \\ 3 & 1 & 1 & -5 \\ -15 & -5 & -5 & 25 \end{pmatrix}.$$

Thus

$$H_l = \sqrt{T^*T} = 24P_1 + 12P_3 = \begin{pmatrix} 15 & -3 & 9 & 3 \\ -3 & 23 & 3 & 1 \\ 9 & 3 & 7 & 5 \\ 3 & 1 & 5 & 15 \end{pmatrix}.$$

Next, we calculate the orthogonal components U_i of the decomposition $T = UH_l$ where $U = U_1 + U_2 + U_3$. The action of U_1 and U_3 is determined by the action of T on the independent column vectors of P_1 and P_3, respectively; say $v_1 = [9, -3, 6, 6]^T$ and $v_2 = [6, 2, 5, 5]^T$ of P_1 and $v_4 = [3, 1, 1, -5]^T$ of P_3. Thus we define

$$v_1' = U_1\ v_1 \equiv \frac{1}{24} T\ v_1 = [9, -3, 6, 6]^T,$$

$$v_2' = U_1\ v_2 \equiv \frac{1}{24} T\ v_2 = [6, 2, 5, 5]^T$$

and

$$v_4' = U_3\ v_4 \equiv \frac{1}{12} T\ v_4 = [3, 1, -5, 1]^T.$$

We cannot use the action of T on the column vector $v_3 = [9, 3, -15, 3]^T$ of P_2 to define U_2 because $Tv_3 = 0$. We proceed by finding any vector v_3' which is orthogonal to v_1', v_2', v_4' and has the same magnitude as v_3, and define U_2 such that $v_3' = U_2\ v_3 = [9, 3, 3, -15]^T$. (We could equally well redefine v_3' to be $v_3'' = \omega v_3'$, where $\omega \in \mathbb{C}$ and $|\omega| = 1$, which shows that the polar decompostion is not unique in the singular case.) For this choice of the orthogonal transformation, we find

$$U = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

If we substitute $-v_3'$ for v_3' (by taking $\omega = -1$ as mentioned above), we find that $T = UH_l = U'H_l$ for

$$U' = \frac{1}{18} \begin{pmatrix} 9 & -3 & 15 & -3 \\ -3 & 17 & 5 & -1 \\ -3 & -1 & 5 & 17 \\ 15 & 5 & -7 & 5 \end{pmatrix}.$$

Appendix

The following Mathematica Package “Scientific Calculator” performs calculations in the algebras $\mathbb{C}\{m_1, m_2, \ldots, m_r\}$ and $\mathbb{C}\{m\}$.

```
(* SCIENTIFIC CALCULATOR FOR THE ALGEBRA  alg={m1,m2,...,mr} *)
(* SCIENTIFIC CALCULATOR FOR THE ALGEBRA ALG={m} where q^m=0 *)
(* Defines multiplication by truncation of Taylor Series *)
    m[a_, b_] :=If[ListQ[a],
       Normal[Table[Series[
       Table[Series[a[[i]], {q[i], 0, alg[[i]] - 1}],
              {i,Length[alg]}]*b,{q[k],0,alg[[k]]-1}][[k]],
              {k,Length[alg]}]],
       Normal[Series[Series[a,{q,0,qm-1}]*b,
                                         {q,0,qm-1}]]  ]

(* Defines basic functions of elements in alg={m1,...,mr}   *)
  (* and in the algebra  a = a0 + a1 q + ... + am-1 q^m-1  *)

power[a_,k_]:=m[a^k,a^0]
rcp[a_]:=power[a,-1]     (* Multiplicative Inverse *)
exp[a_]:=m[Exp[a],a^0]   (* Finds the exponential of a *)
log[a_]:=m[Log[a],a^0]   (* Finds the log of a *)
sin[a_]:=m[Sin[a],a^0]
asin[a_]:=m[ArcSin[a],a^0]
cos[a_]:=m[Cos[a],a^0]
acos[a_]:=m[ArcCos[a],a^0]

         (*  Example: The algebras alg={2,3} and q^3=0  *)
alg={2,3};qm=3

a = {1 + 2*q[1], 5 + 3*q[2] + q[2]^2}
b = {2 + 3*q[1], 1 + q[2] + 5*q[2]^2}
c = 3-q+2q^2
d=5+2q+7q^2

m[a,b]  (* Finds the product of a and b *)
m[c,d]  (* Finds the product of c and d *)
log[a]  (* Finds the log of a *)
        (* END SCIENTIFIC CALCULATOR *)
```

The following *Mathematica* package for calculating the spectral basis for the algebra $\mathbb{C}\{m_1, m_2, \ldots, m_r\}$ is based on elementary results from complex analysis. If $\psi(x) = \sum_{i=1}^{r}(x - x_i)^{m_i}$ and $\psi_i(x) = \prod_{j \neq i}(x - x_j)^{m_j}$, then the complex rational function $h_i(z) = 1/\psi_i(z)$ is analytic at $z = x_i$, so it has a Taylor series $a_0 + a_1(z - x_i) + \ldots + a_{m_i - 1}(z - x_i)^{m_i - 1} + O(z - x_i)^{m_i}$. Hence

$$\frac{1}{\psi(z)} - \frac{h_i(z)}{(z - x_i)^{m_i}} = \frac{1}{(z - x_i)^{m_i}\psi_i(z)} - \left(\frac{a_0}{(z - x_i)^{m_i}} + \frac{a_1}{(z - x_i)^{m_i - 1}} + \ldots + \frac{a_{m_i - 1}}{z - x_i}\right)$$

is analytic at $z = x_i$. Thus $h_i(z)/(z - x_i)^{m_i}$ is the *principal part* of $1/\psi(z)$ at its pole $z = x_i$. Since a rational function is the sum of its principal parts at all its poles, it follows that

$$\frac{1}{\psi(z)} = \sum_{i=1}^{r} \frac{h_i(z)}{(z - x_i)^{m_i}}.$$

Thus the key step in computing the partial fraction decomposition of $\frac{1}{\psi(z)}$ is to find the Taylor series expansion $h_i(z) = a_0 + a_1(z - x_i) + \ldots + a_{m_i - 1}(z - x_i)^{m_i - 1} + O(z - x_i)^{m_i}$.

If t is an operator with minimal polynomial $\psi(x)$ then $(t - x_i)^{m_i}\psi_i(t) = 0$. Since the spectral basis elements are the $p_i = h_i(t)\psi_i(t)$ and $q_i^k = (t - x_i)^k p_i$ for $1 \leq k \leq m_i - 1$, all the terms in the power series for $h_i(t)$ that have exponents greater than $m_i - 1 - k$ will vanish in the product $q_i^k = (t - x_i)^k h_i(t)\psi_i(t)$.

```
                (*  A Mathematica Package  *)
         (* POWER SERIES METHOD FOR EVALUATING POLYNOMIALS  *)
   (* Calculates p[i]=pqs[m,i,0] and q[i]^k=pqs[m,i,k]
     using power series for i=1,2, ... , Length[m],
     where m={m1,m2,...mr}                            *)
psi[m_List]:=Product[(t - x[j])^(m[[j]]),{j,Length[m]}]
psi[m_List,i_]:=psi[m]/((t-x[i])^(m[[i]]))
pqs[m_,i_,k_]:=
   Together[Normal[Series[1/psi[m,i],{t,x[i],m[[i]]-
                      (1+k)}]]]*(t-x[i])^k *psi[m,i]
   (* Calculates p[i] and q[i] for i=1,2, ... , Length[m]  *)
pqs[m_List]:=Do[p[i]=pqs[m,i,0];q[i]=pqs[m,i,1];Print[i],
                                 {i,Length[m]}]
```

Acknowledgement

I thank the referees for suggesting many improvements of the original manuscript.

References

[1] C. W. Curtis, *Linear Algebra: An Introductory Approach*, 3rd ed., Allyn and Bacon, Boston, 1974.

[2] F.R. Gantmacher, *Theory of Matrices*, translated by K. A. Hirsch, Chelsea Publishing Co., New York (1959).

[3] N. Jacobson, *Lectures in Abstract Algebra*, vol 2, Van Nostrand, Princeton, 1953.

[4] G. Sobczyk, Hyperbolic Number Plane, *The College Mathematics Journal*, September 1995.

[5] G. Sobczyk, "Structure of Factor Algebras and Clifford Algebra", *Linear Algebra and its Applications,*, Vol. 241-3, (to appear).