Cubo A Mathematical Journal
Vol.05/N$^{\underline{o}}$03 - OCTOBER 2003

Functions of Matrices

H. Bueno
Dep. de Matemática - ICEx, UFMG, Caixa Postal 702.
Belo Horizonte, MG 30.123-970, Brazil.
e-mail address: hamilton@mat.ufmg.br

Functions $f : U \subset \mathbb{C} \to \mathbb{C}$ might engender functions of square matrices, known as **functions of matrices**. Examples are the matrices A^k, A^{-1}, $\sin A$ and the flow e^{At}. These functions, despite being an essential topic in Linear Algebra, are not usually covered by the traditional presentation of the subject.

Usually, most of the attention is restricted to polynomial functions of a matrix or, when $A = P^{-1}DP$ with D diagonal, $f(A)$ is defined by $P^{-1}f(D)P$, where $f(D)$ is obtained by evaluating f in each of the diagonal entries of D.

Due to its overwhelming role in differential equations, e^{At} is usually defined by means of the power series expansion, which demands the concept of uniform convergence and, in consequence, makes the exposition accessible only to more advanced students. Moreover, e^{At} is obtained just in some simple cases (mostly when A is in Jordan canonical form) and the student gets the impression that e^{At} is a "theoretical" solution of the system $x' = Ax$. The fact that e^{At} is a polynomial (with coefficients depending on t) in the matrix A is not emphasized.

Not much attention is given to the function $f(A) = A^k$. Of course, one would probably find A^k for large k by iterating powers A^s already obtained. However, A^k has particular importance in problems where symmetry plays an essential role, thus making possible an easy derivation of the eigenvalues of A and a much simpler obtention of A^k.

The definition of the function $f(A) = A^{-1}$ is rare, although the algorithm to obtain the inverse by applying Gaussian elimination simultaneously to A and I is given in every textbook. In some problems, however, A is symmetric and has few eigenvalues. Also in this case it is worthwhile to use another method to obtain the inverse.

Our exposition of functions of matrices could be summarized as a generalization of the finite dimensional version of Dunford-Schwartz's functional calculus [8] and goes back (at least) to Gantmacher [9]. It is simple and has amazing consequences: every function $f(A)$ is a polynomial in the matrix, which can be easily obtained if one knows its eigenvalues with multiplicities. (Of course, the coefficients of the polynomial depend on the function f). This method, a standard tool in Numerical Linear Algebra, has surprisingly sunk into oblivion in the introductory texts of Linear Algebra. Well-set textbooks (see [2], [10], [11], [12], [18]) or even more advanced monographs (see [3] or [13]) do not even mention it. With this article we expect to contribute to the reappraisal of the functional calculus method in (basic) Linear Algebra.

We now describe briefly what is in the text. Section 1 is basically a theoretical remark concerning the division of a function by a polynomial. We show that under natural hypotheses this division is euclidean. (The ultimate versions of this result are the Weierstrass and Malgrange Preparation Theorems). The functional calculus, which is a consequence of this fact, is analyzed in section 2. The next section displays various examples of the use of the functional calculus and section 4 is devoted to the (elementary) proofs of the Spectral Mapping Theorem and the Spectral Theorem by the functional calculus method. Section 5 is a little more advanced and exhibits a situation where the symmetry of a matrix ensures the existence of few eigenvalues. An appendix shows that the functional calculus gives rise to a continuous homomorphism between the algebra $\mathcal{F}^k$ (of functions of class C^k) and the algebra of polynomials of matrices. This approach is basically the one used to establish the functional calculus in its full generality (see [8], [14]) and makes it possible to find error estimates in Numerical Linear Algebra.

1 The Interpolation Polynomial

Definition 1.1 *A function* $f : U \subset \mathbb{C} \to \mathbb{C}$ *(or* $f : I \subset \mathbb{R} \to \mathbb{R}$*) is* **euclidean** *with respect to the polynomial* p *if*

(i) all roots of p *are in* U *(resp., in* I*);*

(ii) if z_0 *is a zero of* p *with multiplicity* k*, then* f *has derivatives up to order* k *at* z_0.

(For reasons that will be clear later, we do distinguish between (complex) analytical and holomorphic functions, analytical meaning the existence of the (first) derivative, holomorphic meaning that it is representable by a power series).

The terminology we have used in the above definition is suggested by the following result, which is valid for functions defined either in $I \subset \mathbb{R}$ or in $U \subset \mathbb{C}$. We define the degree of the zero polynomial as $-\infty$.

Proposition 1.2 *Let f be euclidean with respect to the polynomial p. Then, there exist a function q, continuous at each root of p, and a polynomial r such that $f = qp + r$, where* $\deg r < \deg p$.

Proof: Let r be an arbitrary polynomial. We consider the function q defined (in the points of the domain of f which are not roots of p) by

$$q(z) = \frac{f(z) - r(z)}{p(z)}.$$

We will show that we can choose r with degree smaller than that of p, so that q has a continuous extension at each root of p. We note that q is as smooth as f at each point z that is not a root of p.

Let z_0 be a root of p with multiplicity k, that is,

$$p(z) = (z - z_0)^k s(z),$$

where s is a polynomial such that $s(z_0) \neq 0$. We want to find r such that the quotient

$$\frac{f(z) - r(z)}{(z - z_0)^k}$$

has a continuous extension at z_0. According to the L'Hospital rule, this will happen when

$$f(z_0) = r(z_0), \ \ f'(z_0) = r'(z_0), \ \ \ldots \ \ , \ \ f^{(k-1)}(z_0) = r^{(k-1)}(z_0).$$

It then suffices to show that there exists a polynomial r with degree smaller than that of p, that satisfies relations as the above at each root z_0 of p. The existence of such a polynomial will be shown in the lemma below. □

We denote $f^{(0)} = f$.

Lemma 1.3 *Given a function f and the values*

$$\begin{array}{ccccc} f(z_1) & f'(z_1) & \cdots & \cdots & f^{(d_1-1)}(z_1) \\ \vdots & \cdots & \cdots & \vdots & \\ f(z_\ell) & f'(z_\ell) & \cdots & f^{(d_\ell-1)}(z_\ell) & \end{array}$$

where $z_1, \ldots, z_\ell$ *are distinct, let us denote* $n = d_1 + d_2 + \ldots + d_\ell$. *Then there exists a polynomial* r, *with degree not greater than* $n-1$, *satisfying*

$$r^{(j)}(z_i) = f^{(j)}(z_i)$$

for all $i = 1, \ldots, \ell$ *and* $j = 0, \ldots, d_i - 1$.

Proof: We may assume that one of the given values is not zero. The polynomial r we look for satisfies a non-homogeneous linear system, which can be written as

$$Bz = b,$$

where z is the vector which has as coordinates the coefficients we seek for r, b is a vector whose n coordinates are the given values of f and B is the $n \times n$ matrix of the linear system thus obtained.

If B is not invertible, the associated homogeneous system has a non-trivial solution

$$z_0 = (a_0, \ldots, a_{n-1}) \in \mathbb{C}^n.$$

We consider the polynomial

$$t(z) = a_0 + a_1 z + \ldots + a_{n-1} z^{n-1},$$

which has degree not greater than $n-1$. Since z_0 satisfies the associated homogeneous system, t must be a multiple of

$$(z - z_1)^{d_1} \cdots (z - z_\ell)^{d_\ell},$$

which is absurd, since this last polynomial has degree n. So, B is invertible and the system $Bz = b$ has the unique solution z. □

The polynomial r is known as the **interpolation polynomial**.

The rest of this section will show how the results given here can be included in a basic course of one complex variable.

If we consider an analytical function $f : U \subset \mathbb{C} \to \mathbb{C}$, $z_0 \in U$ and $p(z) = z - z_0$, Proposition 1.2 shows that $f(z) = h(z)(z - z_0) + c$, with h continuous in U and analytical in $U \setminus \{z_0\}$. We clearly have $c = f(z_0)$. Supposing additionally that U is convex, it is easy to show that h has a primitive in U (see [15], 10.14), thus implying that

$$\int_\gamma h(z)dz = 0$$

for every closed path $\gamma \in U$. From this follows immediately the Cauchy Integral Formula in convex sets:

$$\int_\gamma \frac{f(z)dz}{z-z_0} = \int_\gamma h(z)dz + f(z_0)\int_\gamma \frac{dz}{z-z_0} = f(z_0)W(z_0,\gamma),$$

where $W(z_0,\gamma)$ denotes the winding number. As a consequence, one shows that **every analytical function is holomorphic** (see [6], Theorem 2.8 or [15], 10.16).

We now give a consequence of Proposition 1.2 that appears to be missing in our introductory courses of one complex variable: the algebra $\mathcal{H}$ of all analytic functions $f:\mathbb{C}\to\mathbb{C}$ is euclidean with respect to each polynomial p. More generally, we have

Proposition 1.4 *In the euclidean division*

$$f = qp + r, \quad \deg r < \deg p$$

of an analytic function $f: U\subset\mathbb{C}\to\mathbb{C}$ by a polynomial p whose roots are in U, the quotient q is analytical.

Proof: According to the proof of Proposition 1.2, the function

$$q = \frac{f-r}{p}$$

is analytic, since both the numerator and the denominator have roots at the same points and the zeroes of the numerator have multiplicity greater or equal than that of the denominator. Therefore, q has a power series expansion at each point of U. □

This result can be extended to functions $f: I\subset\mathbb{R}\to\mathbb{R}$ of class C^∞ and polynomials whose roots are all in I: the L'Hospital rule will then imply that $q\in C^\infty$.

Now, let $f: U\subset\mathbb{C}\to\mathbb{C}$ be analytical on U and γ be a closed path (chain) homologous to 0 in U. The (global) Cauchy Integral Formula follows from the (global) Cauchy Theorem[1] by the argument given above. Another application comes if we consider the quotient $f(z)/(z-z_0)^n$. For instance, if $n=2$ we have $f(z) = h(z)(z-z_0)^2 + a(z-z_0) + b$ for an analytical function h on U and constants a and b. Clearly $b = f(z_0)$ and (taking derivatives) $a = f'(z_0)$. So,

$$\begin{aligned}\int_\gamma \frac{f(z)dz}{(z-z_0)^2} &= \int_\gamma h(z)dz + f'(z_0)\int_\gamma \frac{dz}{z-z_0} + f(z_0)\int_\gamma \frac{dz}{(z-z_0)^2}\\ &= f'(z_0)W(z_0,\gamma).\end{aligned}$$

[1]Of course, we suppose that Dixon's proof of the (global) Cauchy Theorem has not been used to demonstrate this result. In this proof, the Cauchy Theorem is a consequence of the Cauchy Integral Formula. See [6], [15].

2 The Finite Dimensional Functional Calculus

We sometimes write $f(z)$ to distinguish the complex function from $f(A)$.

We use some well known facts about the minimal polynomial m of a square matrix A. It is defined as the polynomial of smallest degree and leading coefficient 1 such that $m(A) = 0$. Any polynomial that vanishes at A is a multiple of m. The roots of m are the same as those of the characteristic polynomial p of A (a result proved in section 4) and form the spectrum $\sigma(A)$ of A (i.e., the set of eigenvalues of A). If a matrix A is diagonalizable, m is a product of distinct linear factors.

Functions of a matrix A are usually defined in two situations: the function $f(z)$ is smooth in the eigenvalues of the diagonalizable matrix A and $f(A)$ is then given by $P^{-1}f(D)P$ or the function $f(z)$ is analytical and $f(A)$ is defined by a power series expansion of f. (See section 3 for examples). In both cases the function f is euclidean with respect to m.

So, in the cases above, by considering the euclidean division

$$f = qm + r, \tag{1}$$

either q is defined at each eigenvalue of the diagonal matrix D (by Proposition 1.2) and $q(A)$ is then given by $P^{-1}q(D)P$ or q is analytical in the spectrum $\sigma(A)$ (by Proposition 1.4) and $q(A)$ may also be defined (see below). It turns out in both cases that[2]

$$f(A) = r(A). \tag{2}$$

This is one of the deepest results of spectral theory: $f(A)$ is a polynomial in A, which has coefficients determined only by the values of f (and its derivatives, according to the case) in the spectrum $\sigma(A)$.

We now consider the inverse problem. Let be given a square matrix A.

Definition 2.1 *Let $m(z) = (z-\lambda_1)^{d_1}\cdots(z-\lambda_\ell)^{d_\ell}$ be the minimal polynomial of A. If the values*

$$\begin{matrix} f(\lambda_1) & f'(\lambda_1) & \cdots & \cdots & f^{(d_1-1)}(\lambda_1) \\ \vdots & \cdots & \cdots & \vdots & \\ f(\lambda_\ell) & f'(\lambda_\ell) & \cdots & f^{(d_\ell-1)}(\lambda_\ell) & \end{matrix}$$

exist, we say that f is **euclidean with respect to** *A and define*

$$f(A) = r(A),$$

where r is the interpolation polynomial given by Lemma 1.3.

[2] To be precise, one must show that $(qm+r)(A) = q(A)m(A) + r(A)$. See the appendix.

If we compare the definition above with the definition of an euclidean function $f(z)$ w.r.t. $m(z)$, we see that the requirements on f are less restrictive. Why this difference?

The answer is simple: the consideration of the abstract division

$$f(z) = q(z)m(z) + r(z),$$

imposes some smoothness conditions on $f(z)$, in order to define a function $q(z)$ that makes sense. If these requirements are fulfilled, we conclude that $r(z)$ is the interpolation polynomial, which happens to be defined under less severe conditions. But, since $m(z)$ is the minimal polynomial of A, we have $f(A) = r(A)$ no matter how we define $q(A)$.

Remark 2.2 Let $A = P^{-1}DP$ for a diagonal matrix D. The usual definition for $f(A)$ is $P^{-1}f(D)P$. Writing $\mathbb{C}^n = W_1 \oplus \cdots \oplus W_n$, where W_i is the eigenspace associated to the eigenvector v_i, we see that $f(D) = r(D)$. It follows that

$$r(A) = r(P^{-1}DP) = P^{-1}r(D)P = P^{-1}f(D)P = f(A),$$

thus showing that both definitions coincide in the case of a diagonalizable matrix A. The exponential of a generic matrix J in Jordan canonical form is explicitly calculated in [5], [17]. The same procedure may be used to define $f(J)$ changing the exponential function for any function $f(z)$ smooth enough. The expression for $f(J)$ thus obtained makes clear that $f(z)$ must be euclidean with respect to A, if one wants to define $f(J)$. As before, it can be proved that the definition $f(A) = P^{-1}f(J)P$ is equivalent to the definition 2.1. ◄

However, Definition 2.1 lacks applicability: in most cases, once the characteristic polynomial $p(z)$ of A is known, the minimal polynomial $m(z)$ still has to be computed. It would be nice if we could use $p(z)$ instead of $m(z)$ to define the function $f(A)$. We can do that: we can deal with multiples of $m(z)$ as long as the smoothness of $f(z)$ permits. In fact, if $f(z) = q(z)m(z) + r(z)$, we have defined $f(A) = r(A)$. If $s(A) = 0$ and $f(z) = q_1(z)s(z) + r_1(z)$ with $\deg s \geq \deg m$, the proof of Proposition 1.2 guarantees that

$$r_1(z) = q_2(z)m(z) + r(z). \tag{3}$$

(If λ is a root of multiplicity d of $m(z)$, we note that $r_1^{(i)}(\lambda) = f^{(i)}(\lambda) = r^{(i)}(\lambda)$, for $i = 0, \ldots, d-1$).

Equation (3) then implies that $r_1(A) = r(A)$, thus authorizing the use of any multiple $s(z)$ of the minimal polynomial $m(z)$ of A instead of $m(z)$ in the Definition 2.1.

We note that definition 2.1 has an important consequence: each function $f(A)$ commutes with the matrix A!

3 Examples

3.1 The exponential

We start with the standard definition of the function e^{At}, $t \in \mathbb{R}$. For this, we consider the exponential function $\exp : \mathbb{C} \to \mathbb{C}$, whose power series representation

$$\exp(z\tau) = e^{z\tau} = 1 + \sum_{n=1}^{\infty} \frac{z^n \tau^n}{n!},$$

converges uniformly in compact sets. If $\|A\|$ denotes the usual norm in the space $\mathcal{L}(\mathbb{C}^n, \mathbb{C}^n)$ of the linear operator $A : \mathbb{C}^n \to \mathbb{C}^n$, we claim that

$$I + \sum_{n=1}^{\infty} \frac{A^n \tau^n}{n!}$$

defines a linear operator. Indeed, the norm in $\mathcal{L}(\mathbb{C}^n, \mathbb{C}^n)$ has the property

$$\|AB\| \leq \|A\| \; \|B\|,$$

which implies $\|A^i\| \leq \|A\|^i$. It follows that, for $k = 1, 2, \ldots$,

$$\left\| I + \sum_{n=1}^{k} \frac{A^n \tau^n}{n!} \right\| \leq 1 + \sum_{n=1}^{k} \frac{\|A\|^n |\tau|^n}{n!}. \tag{4}$$

For each fixed τ, the series on the right-hand side converges. Since the space $\mathcal{L}(\mathbb{C}^n, \mathbb{C}^n)$ is complete, we have proved that

$$\exp(A\tau) = e^{A\tau} := I + \sum_{n=1}^{\infty} \frac{A^n \tau^n}{n!}$$

is a linear operator. Choosing $\tau = t \in \mathbb{R}$, we have defined e^{At}. We also note that (4) shows that the convergence is uniform if τ belongs to a compact set. Hence, term by term differentiation produces its derivatives and

$$\frac{d}{dt} e^{At} = e^{At} A.$$

Furthermore, when $t = 0$, we have

$$e^{At}\big|_{t=0} = e^0 = I.$$

These are the main properties of the exponential matrix e^{At}. In particular we see that e^{At} is a fundamental solution of the matrix system $X' = AX$, $X(0) = I$.

This definition of the exponential e^{At} is not very suitable for computation: usually one has to obtain the Jordan canonical form $J = P^{-1}AP$ of the matrix A, then e^{Jt} and (at last!) $e^{At} = Pe^{Jt}P^{-1}$. The functional calculus makes it possible to compute e^{At} easily.

Furthermore, the properties of the flow e^{At} follow immediately from the functional calculus. For example,

$$\frac{\partial}{\partial t}f(zt) = f'(zt)z \quad \Rightarrow \quad \frac{d}{dt}e^{At} = e^{At}A.$$

Remark 3.1 Although the function $f(z) = e^z$ satisfies the equation

$$e^{z+w} = e^z e^w,$$

we can not deduce that $e^{A+B} = e^A e^B$, since the simultaneously substitution of the variables z by A and w by B is not allowed by the functional calculus. However, if A and B commute, the simple knowledge that e^A is a polynomial in the variable A allows us to conclude that $e^A B = Be^A$. The proof that $e^{A+B} = e^A e^B$ iff $AB = BA$ then follows as usual (see [1]). ◄

Example 3.2 Let

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & -5 \\ 0 & 1 & -2 \end{pmatrix}.$$

The characteristic polynomial of A (which coincides with the minimal polynomial) is

$$p(z) = (z-1)(z+i)(z-i).$$

To compute e^{At}, we define the function $f(zt) = e^{zt}$. It is then enough to get a polynomial r, with degree no greater than 2, such that $r(1) = f(1t) = e^t$, $r(i) = f(it) = \cos t + i \sin t$ and $r(-i) = f(-it) = \cos t - i \sin t$. Substituting these relations in the polynomial $r(z) = az^2 + bz + c$, we find $a = (e^t)/2 - (\cos t + \sin t)/2$, $b = \sin t$ and $c = (e^t)/2 + (\cos t - \sin t)/2$. Thus,

$$e^{At} = [(e^t)/2 - (\cos t + \sin t)/2]A^2 + (\sin t)A + [(e^t)/2 + (\cos t - \sin t)/2]I,$$

which is a real matrix (as expected), although A has complex roots. ◇

Example 3.3 Let

$$A = \begin{pmatrix} 3 & -4 & -1 \\ -3 & 5 & 1 \\ 21 & -32 & -7 \end{pmatrix}.$$

The characteristic polynomial of A is

$$p(z) = (z-1)z^2.$$

To compute e^{At}, we obtain the coefficients of $r(z) = az^2 + bz + c$ so that $r(1) = e^{1t} = e^t$, $r(0) = e^{0t} = 1$ and $r'(0) = te^{0t} = t$. Thus, $c = 1$, $b = t$ and $a = e^t - t - 1$. We conclude that

$$e^{At} = (e^t - t - 1)A^2 + tA + 1I. \qquad \circ$$

The examples above show the practical advantages of using the functional calculus to compute the exponential e^{At}. As a consequence, we can deduce that the predominant role given to the Jordan canonical form in the study of the linear system $x' = Ax$ is not intrinsic: all the analysis of hyperbolic systems can be done without using it (see [4]).

3.2 Trigonometric functions

The study we have just completed is also valid for the exponential e^{iAt} (that is, the case $\tau = it$, $t \in \mathbb{R}$, in the subsection before), which generates the trigonometric functions $\sin At$ and $\cos At$. These functions are usually defined by means of the power series expansion of $\sin z$ and $\cos z$, but are also easily obtainable through the functional calculus.

The same remarks also applies to other trigonometric functions.

3.3 Logarithm

A logarithm of a matrix A is usually defined by means of the Jordan canonical form. (Of course, the hypothesis $\det A \neq 0$ is necessary). Since all the eigenvalues of A are nonzero, one usually takes logarithms of the Jordan blocks, which is accomplished through the power series expansion of $\log(1 + z)$ (see [3]). However, according Remark 2.2, a logarithm of a Jordan block can be directly defined. As before, the main shortcoming of this method is that the Jordan form of a matrix is needed to compute its logarithm.

The functional calculus allow us to get a matrix $B = \log A$, if $\det A \neq 0$. We only have to choose a branch of the function $f(z) = \log z$ that contains the spectrum $\sigma(A)$ and then evaluate $B = \log A$ through the interpolation polynomial. Of course, the matrix B depends on the branch we have chosen, but the relation $e^B = A$ follows in any case from $e^{\log z} = z$.

If all the eigenvalues of the real matrix A are positive, we can consider the real function $f(x) = \ln x$ and apply the same technique. The matrix $B = \ln A$ thus obtained is then a real solution of the equation $e^B = A$.

3.4 Square root

Let us suppose that all the eigenvalues of the real matrix A are real and non-negative. Furthermore, if 0 is an eigenvalue of A, we suppose that it is a simple root of the minimal polynomial m of A. In this case, we may use the function $f:(0,\infty)\to\mathbb{R}$, $f(x)=\sqrt{x}$ to define $\sqrt{A}$. We employ in this case the functional calculus to a function that is only continuous at the simple eigenvalue $\lambda = 0$.

However, we may define $\sqrt{A}$ if the eigenvalues of A are complex and not equal 0. We just choose a branch of the logarithm function $f(z)=\log z$ for which the square roots of all the eigenvalues of A are defined. We apply then the functional calculus to the complex function $f(z)=\sqrt{z}$.

Remark 3.4 The definition of the function $\sqrt{A}$ does not provide all solutions of the equation $B^2 = A$. If A is the 2×2 identity matrix,

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}$$

are also solutions of $B^2 = I$, besides $B = I$, the unique solution obtained by means of the real square root function. Furthermore, if $A = -I$, the equation $B^2 = A$ has the real solution

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

which does not arise from the function $\sqrt{A}$. ◀

3.5 The inverse

The classical way to obtain the inverse of a (invertible) matrix A through its characteristic (or minimal) polynomial $p(z)$ is the following: if $p(z) = z^m + a_{m-1}z^{m-1} + \ldots + a_1 z + a_0$, then

$$0 = A^m + a_{m-1}A^{m-1} + \ldots + a_1 A + a_0 I.$$

Multiplying by A^{-1}, we obtain

$$a_0 A^{-1} = -[A^m + \ldots + a_1 A].$$

Since $a_0 \neq 0$ (of course, $a_0 = \det A$ whenever $p(z)$ is the characteristic polynomial), A^{-1} is thus obtained.

For a general invertible square matrix A this method usually brings no benefit, when compared with the calculation of the inverse by Gaussian elimination. Also the functional calculus is not advantageous.

But, for instance, if the matrix A is symmetric and has few eigenvalues (as in the example of the tetrahedron in $\mathbb{R}^n$, given in the section 5), the use of the functional calculus is helpful: for instance, in that example A^{-1} is computed by a polynomial of the first degree!

4 The Spectral Theorem And The Decomposition T = D + N

In this section we show how the functional calculus is useful in the proof of abstract results. We begin with

Theorem 4.1 (Spectral Mapping Theorem) *Let f be euclidean with respect to the $n \times n$ complex matrix A. If λ is an eigenvalue of A, then $f(\lambda)$ is an eigenvalue of $f(A)$. Every eigenvalue of $f(A)$ is of the form $f(\lambda)$, where λ is an eigenvalue of A.*

Proof: Since f is euclidean with respect to A, $f(A) = r(A) = a_k A^k + \ldots + a_1 A + a_0 I$. If v is an eigenvector related to λ,

$$f(A)v = r(A)v = (a_k \lambda^k + \ldots + a_1 \lambda + a_0)v = r(\lambda)v = f(\lambda)v.$$

Let us suppose that μ is an eigenvalue of $f(A) = r(A)$. We consider the polynomial $r(z) - \mu$, which factors in $\mathbb{C}$ as

$$r(z) - \mu = a_k \prod_{i=1}^{k} (z - \lambda_i).$$

Consequently,

$$r(A) - \mu I = a_k \prod_{i=1}^{k} (A - \lambda_i I).$$

Since the left-hand side of this equation is not invertible, at least one of the factors $A - \lambda_i I$ is not invertible. Thus, λ_i is both an eigenvalue of A and a root of $r(z) - \mu$. Therefore,

$$f(\lambda_i) = r(\lambda_i) = \mu. \qquad \square$$

We now prove the seminal Spectral Theorem.

Theorem 4.2 *Let $T : \mathbb{C}^n \to \mathbb{C}^n$ be a linear operator, with characteristic polynomial*

$$p(z) = (z - \lambda_1)^{s_1} \cdots (z - \lambda_\ell)^{s_\ell},$$

where the eigenvalues λ_i, $i = 1, \ldots, \ell$ are distinct.

Then, there exist subspaces $W_1, \ldots, W_\ell$ such that

$$\mathbb{C}^n = W_1 \oplus \cdots \oplus W_\ell$$

and $T(W_i) \subset W_i$. The operator $T|_{W_i}$ has only the eigenvalue λ_i, $\dim W_i = s_i$ and $m(z) = (z - \lambda_1)^{d_1} \cdots (z - \lambda_\ell)^{d_\ell}$ for $1 \leq d_i \leq s_i$.

Furthermore, $T = D + N$, with D diagonal(izable), N nilpotent and $DN = ND$.

Proof: For each λ_i, we consider an open set $U_i \ni \lambda_i$, such that $U_i \cap U_k = \emptyset$, if $i \neq k$. We define $f_i(z) = 1$, if $z \in U_i$, $f_i(z) = 0$, if $z \in U_j$, for $j \neq i$. The functions $f_1, \ldots, f_\ell$ are euclidean with respect to p and the relations

$$f_i^2 = f_i, \quad f_i f_j = 0, \quad \text{if} \quad i \neq j \quad \text{and} \quad \sum_{i=1}^{\ell} f_i = 1$$

are valid for each $z \in \bigcup_{i=1}^{\ell} U_i$. Therefore, denoting $f_i(T)$ by π_i, we have

$$\pi_i^2 = \pi_i, \quad \pi_i \pi_j = 0, \text{ if } \quad i \neq j \quad \text{and} \quad \sum_{i=1}^{\ell} \pi_i = I, \tag{5}$$

hence showing that each π_i is a projection.

If W_i stands for the image $\pi_i(\mathbb{C}^n)$, we obtain

$$\mathbb{C}^n = W_1 \oplus \cdots \oplus W_\ell.$$

Because π_i commutes with T, it is true that $T(W_i) \subset W_i$.

Regardless of the choice of bases $\mathcal{B}_1, \ldots, \mathcal{B}_\ell$ for the spaces $W_1, \ldots, W_\ell$, respectively, T can be represented as a block matrix A with respect to the basis $\mathcal{B} = \{\mathcal{B}_1, \ldots, \mathcal{B}_\ell\}$ of $\mathbb{C}^n = W_1 \oplus \cdots \oplus W_\ell$:

$$A = \begin{pmatrix} T_1 & 0 & \cdots & 0 \\ 0 & T_2 & \cdots & 0 \\ \vdots & \ldots & \vdots & \vdots \\ 0 & 0 & \cdots & T_\ell \end{pmatrix}.$$

We now claim that the characteristic polynomial of T_i is

$$p_i(z) := (z - \lambda_i)^{s_i}$$

for $i = 1, \ldots, \ell$, thus implying that $\dim W_i = s_i$ and (according to the Cayley-Hamilton theorem) that the minimal polynomial of T_i is $(z - \lambda_i)^{d_i}$, for $1 \leq d_i \leq s_i$. We can also deduce the given expression for m.

Since $p(z) = \det(zI - A) = \det(zI - T_1) \cdots \det(zI - T_\ell)$, to prove the claim it suffices to show that the only eigenvalue of T_i is λ_i. We consider only $i = 1$, the remaining cases being similar. For a fixed $\lambda \neq \lambda_1$, we define the functions

$$g(z) = \begin{cases} g_1(z) = z - \lambda & \text{if } z \in U_1 \\ g_j(z) = 1 & \text{if } z \in U_j, \end{cases} \quad \text{and} \quad h(z) = \begin{cases} 1/(z - \lambda) & \text{if } z \in U_1 \\ 1 & \text{if } z \in U_j. \end{cases}$$

for $j = 2, \ldots, \ell$. In the construction of the projections π_i we can take the open neighborhood $U_1 \ni \lambda_1$ as small as we wish. So, we may assume that $\lambda \notin U_1$. This implies that

$$g(z)h(z) = 1.$$

thus guaranteeing that $g(A)$ has an inverse.

Now we calculate $g(A)$. For this, we note that $g(z) = \sum_{i=1}^{\ell} q_i(z) f_i(z)$. Therefore, $g(T) = (T - \lambda I)\pi_1 + \ldots + I\pi_\ell$ and, in the base $\mathcal{B}$,

$$g(A) = \begin{pmatrix} T_1 - \lambda I & 0 & \cdots & 0 \\ 0 & I & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & I \end{pmatrix}.$$

Since $g(A)$ has as inverse, λ is not an eigenvalue of T_1. This proves that the only eigenvalue of T_1 is λ_1.

We now consider the "diagonal" operator $D = \sum_{i=1}^{\ell} \lambda_i \pi_i$. (In each W_i we have $D_i := \lambda_i \pi_i = \lambda_i I$, where I is the identity operator in W_i. Because of this we call D diagonal. See, however, example 4.3).

We define $N = T - D$. Clearly, for $i = 1, \ldots, \ell$, it holds $N = h(T)$, where

$$h(z) = z - \lambda_i, \quad z \in U_i.$$

According to the Cayley-Hamilton Theorem, $(T_i - \lambda_i I)^{s_i} = 0$, thus proving that $N^k = 0$, where $k = \max\{s_1, \ldots, s_\ell\}$. So $T = D + N$, where D is diagonal and N nilpotent. (Actually, $(T_i - \lambda_i I)^{d_i} = 0$ for $d_i \in \{1, \ldots, s_i\}$. The integer d_i is the **index** of the eigenvalue λ_i).

Since $D = \sum_{i=1}^{\ell} \lambda_i \pi_i$ is a sum of polynomials in T, D commutes with T. Thus $ND = (T - D)D = D(T - D) = DN$. □

Example 4.3 Let $T : \mathbb{C}^4 \to \mathbb{C}^4$ be defined by

$$T(x_1, x_2, x_3, x_4) = (2x_1 - x_2 + x_4, 3x_2 - x_3, x_2 + x_3, -x_2 + 3x_4).$$

The characteristic polynomial of T is $p(z) = (z-3)(z-2)^3$ and one easily checks that $m(z) = (z-3)(z-2)^2$ is the minimal polynomial of T.

We first exemplify the theorem 4.2 with respect to the canonical basis of $\mathbb{C}^4$. We call A the matrix that represents T in this basis.

The projection π_1 (attached to the eigenvalue 3) is obtained by solving the system[3]

$$r(z) = az^2 + bz + c, \quad r(3) = 1, \ r(2) = 0, \ r'(2) = 0.$$

[3] To simplify computations, we have used the minimal instead of the characteristic polynomial of T.

Thus, $a = 1$, $b = -4$, $c = 4$ and

$$\pi_1 = A^2 - 4A + 4I = \begin{pmatrix} 0 & -2 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -2 & 1 & 1 \end{pmatrix}.$$

In the same manner,

$$\pi_2 = \begin{pmatrix} 1 & 2 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 2 & -1 & 0 \end{pmatrix}.$$

The relations (5) follow immediately. So,

$$\left\{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}\right\} = \mathbb{R}^4 = \left\{\begin{pmatrix} -2x_2 + x_3 + x_4 \\ 0 \\ 0 \\ -2x_2 + x_3 + x_4 \end{pmatrix}\right\} + \left\{\begin{pmatrix} x_1 + 2x_2 - x_3 - x_4 \\ x_2 \\ x_3 \\ 2x_2 - x_3 \end{pmatrix}\right\}$$
$$= W_1 \oplus W_2.$$

The matrix D is defined as

$$D = 3\pi_1 + 2\pi_2 = \begin{pmatrix} 2 & -2 & 1 & 1 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & -2 & 1 & 3 \end{pmatrix}$$

and the nilpotent matrix N as

$$N = A - D = \begin{pmatrix} 0 & 1 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & -1 & 0 \end{pmatrix}.$$

One easily checks that $N^2 = 0$ and $ND = DN$.

If we choose, for example, basis $\mathcal{B}_1 = \{w_1 = (1,0,0,1)\}$ and $\mathcal{B}_2 = \{w_2 = (1,0,0,0),\ w_3 = (0,1,0,2),\ w_4 = (0,0,1,-1)\}$ for the spaces W_1 and W_2, respectively, then T is represented by the matrix in block form

$$B = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 1 & -1 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

in the basis $\{w_1, w_2, w_3, w_4\}$. Now D stands for the diagonal matrix

$$\begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \quad \text{and} \quad N = B - D = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

also satisfies $N^2 = 0$. ◇

5 Symmetry and Eigenvalues

In this section we consider a situation where we make use of the symmetry of a matrix to obtain its eigenvalues. In two of the three situations we will examine - namely, in the case of the n-dimensional versions of the tetrahedron and octahedron - the matrix has very few eigenvalues, creating the optimum background for the application of the functional calculus.

We start considering the problem of evaluating a high power of a matrix. One of the best ways to motivate this problem is the introduction of the adjacency matrix in graph theory. Although the subject now appears in some books of Linear Algebra with emphasis on applications (see [2], [18]), we provide a brief exposition for the convenience of the reader.

A(n) (undirected) **graph** is defined as a set of points (the **vertices**) joined together by some arcs (the **edges**). Two vertices are **adjacent** if they are joined by an edge. (See Figure 1, below).

A **path of length n** joining vertices u and w is a finite sequence $v_0 = u, v_1, \ldots, v_n = w$, with v_i and v_{i+1} adjacent. The path is **closed** if $u = w$. We consider the following problem: how many paths of a given length joining v to ν do exist?

It is somehow amazing that the use of matrices can help in solving such a problem. To see how this is accomplished, we define the **adjacency matrix** A of the graph. We denote the n vertices of a graph by $v_1, \ldots, v_n$ and define the $n \times n$ matrix $A = (a_{ij})$ by

$$a_{ij} = \begin{cases} 1, & \text{if the vertices } v_i \text{ and } v_j \text{ are adjacent;} \\ 0, & \text{if not.} \end{cases}$$

We note that, by definition, the matrix A is symmetric.

Example 5.1 The adjacency matrix A_2 of a square with vertices v_1, v_2, v_3 and v_4, which are labelled counterclockwise, is

$$A_2 = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}.$$ ◇

Proposition 5.2 *Let $A = (a_{ij})$ denote the adjacency matrix of a graph and a_{ij}^L the (i,j)-entry of the matrix A^L, $L \in \{1, 2, \ldots\}$. Then, the number of paths of length L joining the vertex v_i to the vertex v_j is a_{ij}^L.*

Proof: If v_k is an arbitrary vertex, we note that a path of length $L+1$ joining v_i to v_j is nothing more than a path of length L joining v_i to v_k, followed by a path of length 1 joining v_k to v_j. This suggests a proof by induction on the length L, the case $L=1$ being true by the very definition of the adjacency matrix. We suppose that the result is true for paths of length L, i.e., a_{ik}^L is the number of paths of length L joining v_i to v_k. Now there are only two possibilities: if there exists an edge joining v_k to v_j, then $a_{ik}^L a_{kj} = a_{ik}^L$ is the number of paths of length $L+1$ joining v_i to v_j which have the form

$$v_i, \ldots, v_k, v_j.$$

On the other hand, if $a_{kj}=0$, then it is not possible to join v_i to v_j passing through v_k. It follows immediately that

$$a_{i1}^L a_{1j} + a_{i2}^L a_{2j} + \ldots + a_{in}^L a_{nj} \tag{6}$$

is the number of paths joining v_i to v_j with length $L+1$. But (6) is exactly the definition of the element (i,j) of A^{L+1}. □

So, the question we have posed leads naturally to the calculus of high powers of a matrix A. But how should we obtain A^k for a large integer k? Iteration is, of course, a solution. If this method is good for a general matrix A, in some cases the application of the functional calculus is better. We now turn our attention to one of these cases: the adjacency matrices of regular polyhedra in $\mathbb{R}^n$.

If the ambient space has dimension $n \geq 5$, there exist only three regular polyhedra: the n-dimensional versions of the cube, tetrahedron and octahedron (see [7]). Taking for granted this result, we will evaluate the eigenvalues with multiplicity of their adjacency matrices[4] following Saldanha-Tomei [16].

We start with the n-dimensional cube. If $n=2$, we have the square, whose vertices we label counterclockwise, as in example 5.1. Therefore, it is easy to calculate the eigenvalues of the adjacency matrix A_2 of the square: 2, 0 (with multiplicity 2) and -2. With the same pattern for the labelling of the vertices of two opposite faces in a cube, we see that its adjacency matrix can be written as

$$A_3 = \begin{pmatrix} A_2 & I_2 \\ I_2 & A_2 \end{pmatrix},$$

where I_2 stands for the $2^2 \times 2^2$ identity matrix. Keeping the notation, if I_{n-1} denotes the $2^{n-1} \times 2^{n-1}$ identity matrix, the adjacency matrix of the n-dimensional cube is

$$A_n = \begin{pmatrix} A_{n-1} & I_{n-1} \\ I_{n-1} & A_{n-1} \end{pmatrix}.$$

[4] This is done for all regular polyhedra in [16].

Thus[5],

$$\begin{aligned}\det(A_n - \lambda I_n) &= \det[(A_{n-1} - \lambda I_{n-1})^2 - I_{n-1}^2] \\ &= \det[(A_{n-1} - (\lambda-1)I_{n-1})(A_{n-1} - (\lambda+1)I_{n-1})] \\ &= \det[A_{n-1} - (\lambda-1)I_{n-1}].\det[A_{n-1} - (\lambda+1)I_{n-1}].\end{aligned}$$

So,

$$\lambda \in \sigma(A_n) \Leftrightarrow (\lambda - 1) \in \sigma(A_{n-1}) \text{ or } (\lambda+1) \in \sigma(A_{n-1}).$$

By induction, the spectrum of A_n is

$$\{n, n-2, \ldots, -n+2, -n\}$$

with multiplicities

$$\binom{n}{0}, \binom{n}{1}, \ldots, \binom{n}{n-1}, \binom{n}{n},$$

respectively. (The multiplicities can also be calculated by the method given below).

To evaluate the eigenvalues with multiplicities of the n-dimensional versions of the tetrahedron and octahedron, we will make full use of their symmetry.

Before considering the general case, let us deal with the octahedron in $\mathbb{R}^3$. The problem we consider will be solved by an abstract formulation following the same pattern.

Example 5.3 If we hang up the octahedron by a vertex v, the other vertices line up at different heights: v forms the top level, then the 4 vertices at the central level and finally the vertex opposite to v. (See Figure 1). We enumerate the vertices following a definite pattern: top level, central level counterclockwise and bottom level. We denote by A the adjacency matrix of the octahedron thus obtained.

Let V be the set of vertices of the octahedron. In order to attach its adjacency matrix A to the set V, we consider A as an operator on U, the space of functions from V to $\mathbb{R}$ (a space clearly isomorphic to $\mathbb{R}^6$). The function u_i assuming the value 1 at the vertex v_i and 0 at the other vertices will be represented by the canonical vector $e_i \in \mathbb{R}^6$.

If R stands for an isometry of the octahedron, it is geometrically clear that $AR = RA$ (isometries preserve adjacency). Let u_0 be an eigenfunction of A attached to the eigenvalue λ. Other eigenfunctions are attached to the same eigenvalue:

$$ARu_0 = RAu_0 = R\lambda u_0 = \lambda Ru_0. \tag{7}$$

Since $u_0(\nu) \neq 0$ for a vertex ν, we can choose an isometry R that takes the vertex ν to the vertex v_1. As a consequence, $u := Ru_0$ satisfies $u(v_1) \neq 0$.

[5] The first equality is a well known property of the determinant of block matrices.

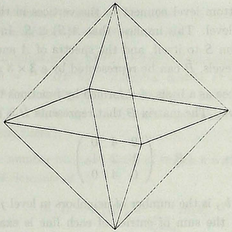

Figure 1: The octahedron is an example of a graph: vertices joined by edges. When suspended by one vertex, the vertices line up at 3 different heights.

Now, let R be an isometry maintaining fixed the first level (so, it is a permutation of the vertices of the second level). Considering the space generated by the eigenfunctions thus obtained, we can replace u by its average $\bar{u}$ in the second level and still get a genuine eigenfunction, since it satisfies $\bar{u}(v_1) = u(v_1) \neq 0$. If $u(v_i) = b_i$ for $i = 2, \ldots, 5$, computing $Ru(v_i)$ for these isometries (there are 8 isometries fixing the first level of the octahedron: 4 rotations and 4 reflections, as in the case of a square), one checks that u is constant on levels[6]. Note that, if $Ru = u$ for all R fixing v, then u is already constant on levels.

Let us denote by $S \subset U$ the subspace of functions that are constant on each level. We have showed that each eigenvalue of A is attached to an eigenvector in S.

The regularity of the polyhedron implies that the form of the matrix

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 & 0 \end{pmatrix}$$

is also "regular": the top and bottom levels are represented by the first and last line of this matrix, the other lines representing the central level. The top level connects to all vertices in the central level, but not with the bottom level or itself. Each vertex in the central level connects to the top and bottom level and is connected two times by other vertices

[6] This property follows from the fact that the subgroup of such isometries acts transitively on V.

in the same level. The bottom level connect to the vertices in the central level, but not with itself or with the top level. This implies[7] that $A(S) \subset S$. In particular, A induces a linear transformation $\tilde{B}$ from S to itself, and the spectra of A and $\tilde{B}$ coincide. Since the octahedron has 3 different levels, $\tilde{B}$ can be represented by a 3×3 matrix.

To represent $\tilde{B}$ we choose as a basis of S formed by functions taking the value 1 in one level and 0 in each other level. The matrix B that represents $\tilde{B}$ is then

$$B = \begin{pmatrix} 0 & 4 & 0 \\ 1 & 2 & 1 \\ 0 & 4 & 0 \end{pmatrix}.$$

We note that, in this basis, b_{ij} is the number of neighbors in level j of an arbitrary element in level i. In consequence, the sum of entries of each line is exactly the number of the neighbors of an arbitrary vertex.

The spectrum of B (and, therefore, of A) is $\{4, -2, 0\}$. ◇

Now we consider the general case. Let V be the set of vertices of one of these polyhedra and, as before, we consider its adjacency matrix A acting on the space U of functions from V to $\mathbb{R}$.

We fix a vertex v and suppose that the eigenfunction u of A attached to the eigenvalue λ satisfy $u(v) \neq 0$.

Let $\mathcal{S}_v$ be the subgroup of all isometries of the polyhedron that fix v. For each $\nu \in V$, we define the **levels** of V as the orbits

$$\{R\nu : R \in \mathcal{S}_v\}$$

(The reader might want to check the distinct levels are classes of an equivalence relation).

Since A commutes with each element $R \in \mathcal{S}_v$, considering the average of the values of Ru on each level, we can change the eigenfunction u and assume that u is constant on each level. Since the value of u in v remains fix, we have $u \neq 0$.

Let $S \subset U$ be the set of functions taking constant values on levels. Then S is an invariant subspace under A, hence inducing a linear transformation $\tilde{B}$ from S to itself. The spectra of A and $\tilde{B}$ are clearly the same.

In order to represent $\tilde{B}$ as an matrix, we choose as a basis the set of vectors taking the value 1 in one level and 0 in any other level. So, $\tilde{B}$ will be represented by a $r \times r$ matrix B, r standing for the number of distinct levels. The element $b_{ij} \in B$ is the number of neighbors

[7] We can also deduce that $(1, 1, \ldots, 1)$ is an eigenvector of A, attached to the eigenvalue 4, which is maximal.

in level j of an arbitrary element in level i. Therefore, the sum of entries in each line of B is exactly the number of neighbors of an arbitrary vertex.

To compute the multiplicities m_i of the distinct eigenvalues λ_i of A, we observe that, since A is symmetric, $A = P^{-1}DP$ for a diagonal matrix D and an invertible matrix P. Therefore, $A^k = P^{-1}D^kP$, thus showing that the same basis diagonalizes simultaneously the powers A^k. Thus, if $\lambda_1, \ldots, \lambda_j$ are the distinct eigenvalues of A, the diagonal of D^k is λ_1^k (m_1 times), $\ldots, \lambda_j^k$ (m_j times).

So, if ℓ_k denotes the number of closed paths of length k with base point v, we have ($|V|$ stands for the number of elements of V)

$$|V|.\ell_k = tr(A^k) = \sum_{\lambda_i \neq \lambda_j} m_i \lambda_i^k.$$

Once the λ_i are known, the m_i are solutions of a linear system obtained by inserting small values of k.

We now calculate the eigenvalues with multiplicities of the adjacency matrix of the n-dimensional versions of the tetrahedron and octahedron. In the case of the tetrahedron, we have

$$B = \begin{pmatrix} 0 & n \\ 1 & n-1 \end{pmatrix}.$$

The spectrum of the adjacency matrix of the tetrahedron is, therefore, $\{-1, n\}$. To evaluate the multiplicity of these eigenvalues, we consider the paths of length 1 joining v to v (that is, a_{11} in the matrix A). We obtain the equation[8]

$$(-1)m_1 + (n)m_2 = 0.$$

An easy computation shows that the element c_{11} of the matrix $C = A^2$ is n. We hence get a second equation:

$$(-1)^2 m_1 + (n^2)m_2 = n^2 + n = n.(\text{number of vertices}).$$

Solving the system, we obtain $m_1 = n$ and $m_2 = 1$.

In the same way, for the octahedron

$$B = \begin{pmatrix} 0 & 2n-2 & 0 \\ 1 & 2n-4 & 1 \\ 0 & 2n-2 & 0 \end{pmatrix}.$$

It follows that the spectrum is $\{2n-2, 0, -2\}$. The multiplicities, which are calculated as before, are 1, n and $n-1$, respectively.

[8] The Perron-Frobenius theorem (see [13]) implies that the multiplicity of the greater eigenvalue is 1, thus reducing the number of equations needed.

The calculus of powers of the adjacency matrix of the tetrahedron and octahedron by use of the functional calculus is very attractive: in each case, the minimal polynomial of the matrix A has (at most) degree 3, since A is diagonalizable. Therefore, each power A^k will be obtained by calculating the coefficients of the interpolation polynomial, which has degree (at most) two.

6 Appendix: a Homomorphism of Algebras

Let A be a $n \times n$ matrix, $m(z) = (z - \lambda_1)^{d_1} \cdots (z - \lambda_\ell)^{d_\ell}$ the minimal polynomial of A and $k = \max\{d_1 - 1, \dots, d_\ell - 1\}$.

There exists a natural homomorphism between $\mathcal{P}$, the algebra of polynomials (with coefficients in $\mathbb{R}$ or $\mathbb{C}$, according to the case) and $\mathcal{P}(A)$, the algebra of matrices formed by evaluating each polynomial $p \in \mathcal{P}$ at A. In other words, there exists a **linear** application

$$\begin{array}{rccc} \phi: & \mathcal{P} & \to & \mathcal{P}(A) \\ & p & \mapsto & p(A) \end{array}$$

which also satisfies

$$\phi(pq) = p(A)q(A) = \phi(p)\phi(q).$$

Its kernel is the set of multiples of the polynomial m. Euclidean division shows that the algebra $\mathcal{P}(A)$ is constituted by polynomials in the matrix A with degree smaller than that of m. By definition, the homomorphism ϕ is onto.

Let K be a compact set such that $\sigma(A) \subset K$. We endow $\mathcal{P}$ with the norm

$$\|p\|_{C^k(K)} = \max_{z \in K}\{|p(z)|, \dots, |p^{(k)}(z)|\}.$$

Clearly, convergence in this norm implies convergence in the semi-norm

$$\|p\|_{\mathcal{P}} = \max\{|p(\lambda_1)|, \dots, |p^{(d_1-1)}(\lambda_1)|, \dots, |p(\lambda_\ell)|, \dots, |p^{(d_\ell-1)}(\lambda_\ell)|\}.$$

(The reader might want to verify that there is no integer $j < k$ such that convergence in the norm $\|\cdot\|_{C^j(K)}$ implies convergence in the semi-norm[9] $\|\cdot\|_{\mathcal{P}}$).

If we consider $\mathcal{P}(A)$ with the usual euclidean topology of the isomorphic space $\mathbb{R}^{n^2}$ (or $\mathbb{C}^{n^2}$, according to the case), the homomorphism ϕ is continuous. Indeed, the polynomial (in A) $p(A) - q(A)$ has coefficients that depend only on the values that the polynomials p and q (and their derivatives, up to the order k) attain at the eigenvalues $\{\lambda_1, \dots, \lambda_\ell\}$ of the

[9] If $\mathcal{I}$ denotes the set of multiples of the minimal polynomial $m(z)$, then $\|\cdot\|_{\mathcal{P}}$ is a norm in the quotient space $\mathcal{P}/\mathcal{I}$.

matrix A, following Definition 2.1. It follows immediately that $p(A)$ will be near $q(A)$, if p and q are near in the norm $\|\cdot\|_{C^k(K)}$.

We denote by $\mathcal{F}^k$ the algebra of all functions f defined and of class C^k in all points of $\sigma(A)$. Of course, $\mathcal{P}$ is a sub-algebra of $\mathcal{F}^k$, and we consider $\mathcal{F}^k$ with the norm already defined. The functions in $\mathcal{F}^k$ are euclidean with respect to A.

If we define $\Phi : \mathcal{F}^k \to \mathcal{P}(A)$ by $\Phi(f) = f(A)$, it can be checked that Φ is an algebra-homomorphism that extends ϕ. Furthermore, the argument given above shows that Φ is also continuous.

$$\begin{array}{ccc} & \phi & \\ \mathcal{P} & \rightarrow & \mathcal{P}(A) \\ \downarrow & \nearrow & \\ \mathcal{F}^k & \Phi & \end{array}$$

The kernel of Φ is constituted by the functions $f \in \mathcal{F}^k$ that have zero remainder when divided by m.

Acknowledgments. The author wishes to thank George Svetlichny and Carlos Tomei, who teach such kind of things. Carlos is responsible for many improvements in the preliminary version of this paper. The author also thanks Helder Rodrigues and Michel Spira for all the help in the preparation of this article.

References

[1] AMANN, H., *Ordinary Differential Equations – an Introduction to Nonlinear Analysis*, Walter de Gruyter, Berlin, 1990.

[2] ANTON, H. AND RORRES, C., *Elementary Linear Algebra: Applications version*, 6th. edition, Wiley, New York, 1991.

[3] BELLMAN, R., *Introduction to Matrix Analysis*, McGraw-Hill Book Company, New York, (1960). (Republished in the series "Classics in Applied Mathematics", vol. 12, SIAM, 1995).

[4] BUENO, H., *Equações Diferenciais Ordinárias – 1a. parte*, Lecture notes, Departamento de Matemática da UFMG, 2001.

[5] CODDINGTON, E.A. AND LEVINSON, N., *Theory of Ordinary Differential Equations*, McGraw-Hill, New York, 1955.

[6] CONWAY, J.B., *Function of One Complex Variable*, Springer, New York, 1975.

[7] COXETER, H.S.M., *Regular Polytopes*, 3rd. Edition, Dover, New York, 1973.

[8] DUNFORD, N. AND SCHWARTZ, J.T., *Linear operators I*, Interscience, New York, 1968.

[9] Gantmacher, F.R., *The Theory of Matrices*, vol. 1 and 2, Chelsea Publishing Co., New York, 1959.

[10] Hirsch, M. and Smale, S., *Differential Equations, Dynamical Systems and Linear Algebra*, Academic Press, New York, 1974.

[11] Hoffman, K. and Kunze, R., *Linear Algebra*, 2nd. edition, Prentice Hall, Englewoods Cliffs, 1971.

[12] Lang, S., *Linear Algebra*, 3rd. Edition, Springer-Verlag, New York, 1989.

[13] Lax, P.D., *Linear Algebra*, Wiley-Interscience Publication, New York, 1997.

[14] Reed, M. and Simon, B., *Methods of Modern Mathematical Physics*, vol. I, Academic Press, New York, 1972.

[15] Rudin, W., *Real and Complex Analysis*, 3rd. Edition, McGraw-Hill, New York, 1987.

[16] Saldanha, N.C. and Tomei, C., *Spectra of Regular Polytopes*, Discrete & Computational Geometry **7** (1992), 403–414.

[17] Sotomayor, J., *Lições de Equações Diferenciais Ordinárias*, IMPA, Rio de Janeiro, 1979.

[18] Strang, G., *Linear Algebra and its Applications*, 3rd. edition, Harcourt, Fort Worth, 1988.