Putzer's Algorithm

Norman Lebovitz

September 8, 2016

1 Putzer's algorithm

The differential equation

$$\frac{dx}{dt} = Ax, \tag{1}$$

where A is an $n \times n$ matrix of constants, possesses the fundamental matrix solution $\exp(At)$, which reduces to the identity at $t = 0$. The spectrum of A, and its invariant subspaces, determine the time dependence of this solution. These are particularly important in the theorems regarding the decisiveness of the spectrum of the linear problem for the stability of the corresponding nonlinear problem. The traditional way of determining this time dependence consists of first transforming A to canonical form (Jordan form, say). The canonical forms simplify the problem of obtaining estimates regarding time behavior of $\exp(At)$.

There is another, less traditional way of expressing the time dependence of the fundamental matrix solution, called Putzer's algorithm, which relies less on the development of the theory of canonical forms and proceeds more directly to the problem of the time dependence of the solutions of equation (1). It *does* rely on an important theorem of linear algebra, the Cayley-Hamilton theorem:

Theorem 1.1 *(Cayley-Hamilton) Every matrix satisfies its characteristic equation* $p_A(\mu) \equiv |A - \mu I| = 0$.

In other words, if

$$p_A(\mu) = (-1)^n \left(\mu^n + c_1\mu^{n-1} + \cdots + c_{n-1}\mu + c_n\right)$$

then

$$p_A(A) \equiv (-1)^n \left(A^n + c_1 A^{n-1} + \cdots + c_{n-1} A + c_n I \right) = 0.$$

We state and prove Putzer's algorithm in this section and provide discussion, examples and applications subsequently.

Theorem 1.2 *Let A be an $n \times n$ matrix of constants, real or complex, and let its eigenvalues be $\mu_1, \mu_2, \ldots, \mu_n$ in any order and without regard to multiplicities. Define the $n-$component, time-dependent vector $r(t)$ by the succession of first-order initial-value problems*

$$\frac{dr_1}{dt} = \mu_1 r_1,\ r_1(0) = 1; \frac{dr_j}{dt} = \mu_j r_j + r_{j-1}(t),\ r_j(0) = 0,\ j = 2, 3, \ldots, n. \quad (2)$$

Define the matrices $P_0, P_1, \ldots, P_n$ recursively by the formulas

$$P_0 = I,\ P_j = \prod_{k=1}^{j} \left(A - \mu_j I \right), j = 1, 2, \ldots, n. \quad (3)$$

Then

$$e^{At} = \sum_{j=0}^{n-1} r_{j+1}(t) P_j. \quad (4)$$

Proof: Define $\Phi(t) = \sum_{j=0}^{n-1} r_{j+1}(t) P_j$. Note that $\Phi(0) = I$. Its t derivative is

$$\dot{\Phi} = \sum_{j=0}^{n-1} \dot{r}_{j+1}(t) P_j = \mu_1 r_1 P_0 + \sum_{j=1}^{n-1} \left(\mu_{j+1} r_{j+1} + r_j \right) P_j. \quad (5)$$

On the other hand

$$A\Phi = r_1(t) A P_0 + \sum_{j=1}^{n-1} r_{j+1}(t) A P_j.$$

Now $AP_j = P_{j+1} + \mu_j P_j$ for $j = 1, 2, \ldots n-1$, by definition. Therefore we have

$$A\Phi = r_1(t) A P_0 + \left(\mu_2 r_2 + r_1 \right) P_1 + \ldots \left(\mu_n r_n + r_{n-1} \right) P_{n-1} + r_n P_n.$$

This is term-by-term the same as the expression (5) except for the last term. But we recall that $P_n = 0$ by the Cayley-Hamilton theorem, so $\dot{\Phi} = A\Phi$ and $\Phi(0) = I$: by the uniqueness theorem for the initial-value problem, equation (4) holds. $\square$.

Remarks

1. If the minimal polynomial M (see §3.2 below) has lower degree than the characteristic polynomial, it can replace the characteristic polynomial in this algorithm. Then equation (4) above is replaced by

$$e^{At} = \sum_{j=0}^{m-1} r_{j+1}(t) P_j \tag{6}$$

   where the matrix factors P_j are now those appearing in the minimal polynomial.

2. Equations (2) above can be rewritten in integral form:

$$r_1(t) = e^{\mu_1 t}, \quad r_{j+1}(t) = e^{\mu_{j+1} t} \int_0^t e^{-\mu_{j+1} s} r_j(s)\, ds \tag{7}$$

   for $j = 1, \ldots, n-1$.

2 Examples

Example 2.1

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

The characteristic polynomial is $(\mu^2+1)(\beta-\mu)$ *with roots* $\pm i, \beta$. *Straightforward if tedious calculations lead to*

$$r_1(t) = e^{it}, \quad r_2(t) = \sin(t), \quad r_3(t) = \left(e^{\beta t} + \beta \sin(t) - \cos(t)\right)/(1+\beta^2).$$

The matrix $\exp(At)$ *then becomes, according to the formula above,*

$$\cos(t) I + \sin(t) A + r_3(t)\left(A^2 + I\right).$$

If β *is real, this is of course real, despite the occurrence of complex eigenvalues – and therefore of complex coefficients* β_j *– in intermediate expressions.*

Example 2.2

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

The characteristic polynomial is $p(\mu) = \mu^2 + 1$ *with eigenvalues* $\mu = \pm i$. *We easily find*

$$r_1(t) = e^{it}, \quad r_2(t) = \sin(t)$$

and

$$e^{At} = \cos(t)I + \sin(t)A.$$

In the preceding example it is easy to check that $A^2 = -I$; this makes it easy to use the power-series expression for $\exp(At)$ to arrive at this conclusion.

In the next example the minimal polynomial differs from the characteristic polynomial.

Example 2.3

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

The characteristic polynomial is $p(\mu) = -(\mu + 3)(\mu - 3)^2$ *but the minimal polynomial is easily found to be* $m(\mu) = \mu^2 - 9$. *If we order the characteristic roots as* $(\mu_1, \mu_2, \mu_3) = (3, 3, -3)$ *we find, after some simple calculations that*

$$r_1(t) = e^{3t}, \quad r_2(t) = te^{3t}, \quad r_3(t) = \frac{1}{6}te^{3}t - \frac{1}{36}e^{3t} + \frac{1}{36}e^{-3t}.$$

With

$$P_0 = I, \quad P_1 = A - 3I, \quad P_2 = (A - 3I)^2$$

we may form the expression for $\exp(At)$ *given in equation (4) above. Separating these into powers of* A *leads to the formula*

$$e^{At} = \frac{1}{2}(e^{3t} + e^{-3t})I + \frac{1}{6}(e^{3t} - e^{-3t})A = \cosh(3t)I + \frac{1}{3}\sinh(3t)A. \tag{8}$$

These formulas can of course also be obtained directly from the definition of the exponential $\exp(At)$ along with the observation that $A^2 = 9I$.

It may be of some interest to carry out the preceding example using only the characteristic roots 'that matter:' $\mu = \pm 3$. For this purpose we reorder the roots (and the operators P_j) so that $\mu_1 = -3, \mu_2 = 3, \mu_3 = 3$. This results in the formulas

$$r_1(t) = e^{-3t}, \quad r_2(t) = \frac{1}{6}(e^{3t} - e^{-3t})$$

where we have omitted the calculation for $r_3(t)$ since its operator coefficient, $A^2 - 9I$, vanishes. Forming the matrix $\exp(At)$ according to the formula (4) now results in

$$e^{At} = e^{-3t} I + \frac{1}{6}(e^{3t} - e^{-3t})(A + 3I) = \cosh(3t) I + \frac{1}{3}\sinh(3t) A,$$

agreeing with the result above and with significantly less computation.

Example 2.4

$$A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

This clearly has the eigenvalues $\mu = 3, \mu = 2(2)$, where we have indicated that the eigenvalue $\mu = 2$ has multiplicity two. The characteristic polynomial is $p(\mu) = (\mu - 3)(\mu - 2)^2$. With the ordering

$$P_0 = I, \quad P_1 = A - 3I, \quad P_2 = (A - 3I)(A - 2I)$$

we find

$$r_1(t) = e^{3t}, r_2(t) = e^{3t} - e^{2t}, r_3(t) = e^{3t} - (1 + t)e^{2t}.$$

Putzer's algorithm now provides the matrix exponential in the form

$$\begin{aligned} e^{At} &= e^{3t} I + \left(e^{3t} - e^{2t}\right) P_1 + \left(e^{3t} - (1+t)e^{2t}\right) P_2 && (9) \\ &= \begin{pmatrix} r_1 & 0 & 0 \\ 0 & r_1 - r_2 & r_2 - r_3 \\ 0 & 0 & -r_2 \end{pmatrix} = \begin{pmatrix} e^{3t} & 0 & 0 \\ 0 & e^{2t} & te^{2t} \\ 0 & 0 & e^{2t} \end{pmatrix}. && (10) \end{aligned}$$

One can also write this on collecting powers of A as

$$(4e^{3t} - 3e^{2t} - 6te^{2t})I + (-4e^{3t} + 4e^{2t} + 5te^{2t})A + (e^{3t} - (1+t)e^{2t})A^2, \quad (11)$$

but this seems unnecessarily awkward, at least in this example.

3 Estimates

Estimates like those normally drawn from transformation to Jordan canonical form can be made straightforwardly with the aid of the Putzer algorithm. We start with the following rather crude estimate.

3.1 A Simple Bound

Theorem 3.1 *Choose the order of the $\{\mu_j\}$ so that the real parts do not decrease: if $\mu_j = \rho_j + i\sigma_j$, then $\rho_j \leq \rho_{j+1}$. Then*

$$|r_j(t)| \leq \frac{t^{j-1}}{(j-1)!} e^{\rho_j t} \text{ for } t > 0, j = 1, 2, \ldots, n.$$

Proof: Since $r_1(t) = \exp(\mu_1 t)$ it holds for $j = 1$. Assume that it holds for j. Then by equation (7) above

$$|r_{j+1}(t)| \leq e^{\rho_{j+1} t} \int_0^t e^{-\rho_{j+1} s} \frac{s^{j-1}}{(j-1)!} e^{\rho_j s} \, ds.$$

In the integral we have the factor

$$e^{(\rho_j - \rho_{j+1})s} \leq 1$$

because the ordering of the eigenvalues makes the exponent negative. Therefore

$$|r_{j+1}(t)| \leq e^{\rho_{j+1} t} \frac{1}{(j-1)!} \int_0^t s^{j-1} \, ds = \frac{t^j}{j!} e^{\rho_{j+1} t},$$

proving the assertion by induction. □

Since ρ_n is the largest of the real parts, this result gives the simple estimate

$$r_j(t) \leq (t^{j-1}/(j-1)!) \exp(\rho_n t).$$

Since for any exponent j and any $\rho' > 0$ it is true that $t^j < k \exp(\rho' t)$ for $t \geq 0$, for some choice of positive constant k, we further have

$$r_j(t) \leq k \exp((\rho_n + \rho')t), \quad t \geq 0.$$

From this it is easy to derive the following estimate:

Corollary 3.1 *If the real parts of all the eigenvalues are negative, there are positive constants k, K and ρ such that*

$$|r_j(t)| \le ke^{-\rho t} \text{ and } \|e^{At}\| \le Ke^{-\rho t}.$$

This corollary will be of use in proving that linear, exponential, asymptotic stability is decisive for nonlinear stability. We shall also need to prove that linear, exponential *instability* is decisive for nonlinear *instability*, but for this we shall need some considerations of a linear-algebraic character.

3.2 The Minimal Polynomial

The matrix A satisfies its characteristic equation $P_n(A) = 0$ where

$$P_n(\lambda) = (\lambda - \lambda_1)^{n_1} \cdots (\lambda - \lambda_r)^{n_r},$$

and we have departed from the previous notation in explicitly noting the multiplicities, and have denoted by $\lambda_1, \ldots, \lambda_r$ the distinct eigenvalues ($r \le n$). The matrix A *may* satisfy a polynomial equation of lower degree. As an extreme example, the unit matrix I has the characteristic polynomial $(\lambda - 1)^n$ but satisfies the equation $A - I = 0$. We call $M(A) = A - I$, or $M(\lambda) = \lambda - 1$, its minimal polynomial. See also example (2.3) above. Any matrix A has a minimal polynomial

$$M(\lambda) = (\lambda - \lambda_1)^{m_1} \cdots (\lambda - \lambda_r)^{m_r} \tag{12}$$

in which $m_i \le n_i$ for $i = 1, 2, \ldots, r$; it is therefore of degree not greater than n (i.e. $m_1 + m_2 + \cdots + m_r = m \le n$) such that $M(A) = 0$ and there is no polynomial of lower degree for which this is so. It may be that $m = n$ and the minimal polynomial is the same as the characteristic polynomial[1], but to treat the general case we need to allow for them to be different. For the remainder of this section we refer to the minimal polnomial (12), and the relation of its distinct eigenvalues (denoted by λ) to the (possibly repeated) eigenvalues (denoted by μ) of Theorem 1.2 is

$$\lambda_1 = \mu_1 = \cdots = \mu_{m_1};\ \lambda_2 = \mu_{m_1+1} = \cdots = \mu_{m_1+m_2};\ \text{etc.} \tag{13}$$

The principal effect of the multiplicity of the eigenvalues is to modify the purely exponential growth (or decay) by algebraically growing factors. The remainder of this section is devoted to the following theorem characterizing of maximal growth.

[1]For example, this is necessarily the case if all eigenvalues are distinct.

Theorem 3.2 *Let $\lambda_1 = \rho_1 + i\sigma_1$ be an eigenvalue of A of largest real part.*

1. *There is positive constant K such that, for all $t \geq 0$*
$$\|e^{At}\| \leq K(1 + t^{m_1 - 1})e^{\rho_1 t}, \text{ and} \tag{14}$$

2. *there is a positive constant k and initial data ξ_0 with $\|\xi_0\| = 1$ such that, for infinitely many values t_n, $t_n \to +\infty$,*
$$\|e^{At_n}\xi_0\| \geq k(1 + t_n^{m_1 - 1})e^{\rho_1 t_n}. \tag{15}$$

Here m_1 is the exponent of the factor $\lambda - \lambda_1$ in the minimal polynomial (equation 12 above).

There is arbitrariness in the choice of norm. A choice appropriate to the complex vector space C^n rather than R^n is convenient since the coefficients $\{r_j(t)\}$ in Putzer's algorithm may be complex even though the matrix A is real. We'll make repeated us of the following:

Lemma 3.1 *Define, for any real or complex $\lambda \neq 0$,*
$$q_k(t, \lambda) = \int_0^t s^k e^{\lambda s}\, ds, \quad k = 0, 1, 2, \ldots. \tag{16}$$

Then
$$q_k(t, \lambda) = e^{\lambda t} p_k(t, \lambda) + c \tag{17}$$
where p_k is a polynomial of degree k and c is a constant.

The proof is elementary and is omitted. One finds for p_k and c the following:
$$p_k(t, \lambda) = (-1)^k \frac{k!}{\lambda^{k+1}} \sum_{j=0}^{k} \frac{(-\lambda t)^j}{j!} \text{ and } c = (-1)^{k+1} \frac{k!}{\lambda^{k+1}}. \tag{18}$$

We prove Theorem 3.2 in this section, beginning with equation (14). In equation (6), arrange the eigenvalues so that λ_1 (the one with largest real part ρ_1) comes first, i.e., that the first m_1 terms in the sum refer to λ_1. We find from equation (7) for the first m_1 coefficients the formulas
$$r_j(t) = \frac{t^{j-1}}{(j-1)!} e^{\lambda_1 t}, \quad j = 1, 2, \ldots, m_1.$$

The contribution of the first m_1 terms to the sum in equation (4), is clearly dominated as follows:

$$\| \sum_{j=1}^{m_1} r_j(t) P_{j-1} \| \leq K_1(1 + t + t^2 + \cdots + t^{m_1 - 1}) e^{\rho_1 t} \tag{19}$$

for some $K_1 > 0$. If $\lambda_1 = \rho_1$ is real, we pause after this estimate and go on to the next eigenvalue. Supposing instead that $\lambda_1 = \rho_1 + i\sigma_1$ is complex ($\sigma_1 \neq 0$), we choose $\lambda_2 = \overline{\lambda_1} = \rho_1 - i\sigma_1$, and, assuming the next m_1 terms in equation (4) refer to $\overline{\lambda_1}$, we use equation (7). For the next coefficient $r_{m_1+1}(t)$ we find

$$r_{m_1+1}(t) = e^{\overline{\lambda_1} t} \int_0^t e^{-\overline{\lambda_1} s} r_{m_1}(s)\, ds = e^{\overline{\lambda_1} t} \int_0^t e^{(\lambda_1 - \overline{\lambda_1})s} \frac{s^{m_1 - 1}}{(m_1 - 1)!}\, ds.$$

Applying Lemma 3.1 to the integral then gives

$$r_{m_1+1}(t) = e^{\lambda_1 t} p^{(0)}_{m_1 - 1}(t) + q_0 e^{\overline{\lambda_1} t}$$

where $p^{(0)}_{m_1-1}$ is a polynomial of degree $m_1 - 1$ and q_0 is a constant. Subsequent applications of Lemma 3.1 to obtain $r_{m_1+2}, \ldots, r_{2m_1}$ give expressions like that above,

$$r_{m_1+j} = e^{\lambda_1 t} p^{(j)}_{m_1-1}(t) + q_{j-1}(t) e^{\overline{\lambda} t} \tag{20}$$

with a succession of polynomials $p^{(j)}_{m_1-1}$ all of degree $m_1 - 1$, and a succession of polynomials q_j with degrees increasing from zero (for r_{m_1+1}) to $m_1 - 1$ (for r_{2m_1}). Then estimating the first $2m_1$ terms in the expression (4) gives us an estimate exactly like that of equation (19) above, although with a different choice of the constant K_1.

We pass on to the next eigenvalue $\rho_2 + i\sigma_2$ where $\rho_2 < \rho_1$ and find the next coefficients

$$r_{2m_1+1}, \ldots, r_{2m_1+m_2}$$

with the aid of equations (7) and (17). The results are sums of exponentials times polynomials. The sums involving the preceding eigenvalues $\lambda_1, \overline{\lambda_1}$ persist (the coefficients of their polynomials are altered but not the degrees).The new eigenvalue λ_2 generates, according to Lemma 3.1, a polynomial factor of degree $m_2 - 1$. It is possible that the most rapidly growing terms $t^{m_2-1} \exp(\rho_2 t)$ include higher powers of t than in preceding terms (this will be so if $m_2 > m_1$), but there is a constant C such that

$t^{m_2-1}\exp(\rho_2 t) < Ct^{m_1-1}\exp(\rho_1 t)$ for all $t > 0$. Thus all the terms in the expression corresponding to λ_2 are dominated by expressions like that on the right-hand side of equation (19) above. The block of terms in the expression (4) belonging to λ_2 and beyond are *all* estimated in the manner of equation (19), but with K_1 in general replaced by a larger constant. This gives an estimate of the form

$$\|\sum_{j=1}^{m} r_j(t)P_{j-1}\| \leq K(1+t+t^2+\cdots+t^{m_1-1})e^{\rho_1 t} \tag{21}$$

valid for all $t \geq 0$.

The formula above is essentially the same as that of the estimate (14) above. To make them agree exactly it is sufficient to note that, for any non-negative integer n, the expression

$$\frac{1+t+t^2+\cdots+t^n}{1+t^n} \tag{22}$$

has a finite maximum (K, say) on $[0,\infty)$.

> We are using the principle that a continuous function on a compact set has a maximum (and a minimum) there, but the interval $[0,\infty)$ is not compact, so the principle appears to fail. It's easy to rectify this. Here are a couple of ways:
>
> 1. Let $\tau = t^n/(1+t^n)$, or $t = (\tau/(1-\tau))^{1/n}$, so that τ runs from 0 to 1 as t runs from 0 to ∞. The expression (22) becomes
>
> $$1+\tau^{1/n}(1-\tau)^{(n-1)/n}+\cdots+\tau^{(n-1)/n}(1-\tau)^{1/n}.$$
>
> This can be evaluated in the limit $\tau = 1$ (where it has the value 1) and this value adjoined to the function. It is now a continuous function on a compact interval, to which the principle in question may be applied.
>
> 2. Consider a continuous function f defined and continuous on $[0,\infty)$ and having a limit b (say) as $t \to \infty$. Then, given $\epsilon > 0$ there exists $T > 0$ such that $|f(t) - b| < \epsilon$ if $t > T$. On $[0,T]$ f is continuous and therefore bounded by (say) M. It follows that for all $t \geq 0$
>
> $$f(t) \leq \max(M, b+\epsilon)$$
>
> providing an upper bound on $[0,\infty)$. A lower bound can be found in a similar way.

We now verify the second estimate of Theorem 3.2, that of equation (15).

Choose an eigenvalue, say λ_1 with multiplicity m_1 in the minimal polynomial M (equation 12). It must have at least one eigenspace of dimension m_1 associated with it. We see this as follows. There is some vector $v \neq 0$ such that

$$(A-\lambda_1 I)^{m_1-1}\cdots(A-\lambda_r I)^{m_r}v \neq 0$$

for otherwise there would be a polynomial of lower degree m_1-1 that vanishes and M would not be minimal. Put

$$\xi_1 = (A-\lambda_2 I)^{m_2}\cdots(A-\lambda_r I)^{m_r}v.$$

Then, if $m_1 > 1$, define the m_1 vectors

$$\xi_1, \quad \xi_2 = (A-\lambda_1 I)\xi_1, \quad \ldots, \quad \xi_{m_1} = (A-\lambda_1 I)\xi_{m_1-1}. \tag{23}$$

That these vectors are linearly independent is seen as follows. If not, there would exist $c_1, c_2, \ldots, c_{m_1}$ (not all zero) such that

$$c_1\xi_1 + c_2\xi_2 + \cdots + c_{m_1}\xi_{m_1} = 0.$$

Operating on this with the operator $(A-\lambda_1 I)^{m_1-1}$, we see that only the first term remains and that $c_1 = 0$. Returning to the equation now and applying instead $(A-\lambda_1 I)^{m_1-2}$ we next infer that $c_2 = 0$; and so on. In this way one sees that all the coefficients must vanish so the vectors are linearly independent. It is easy to see that the subspace spanned by these vecors is invariant under A.

Because of this invariance, we can find solutions of the basic linear equation $\dot{x} = Ax$ in this subspace. Seeking it in the form

$$x(t) = c_1(t)\xi_1 + c_2(t)\xi_2 + \cdots + c_{m_1}(t)\xi_{m_1} \tag{24}$$

with initial data $c_j(0) = \gamma_j$ for $j = 1, 2, \ldots, m_1$ one easily finds

$$\begin{aligned} c_{m_1}(t) = \gamma_{m_1}e^{\lambda_1 t}, \quad c_{m_1-1}(t) &= (\gamma_{m_1-1} + \gamma_{m_1}t)\, e^{\lambda_1 t}, \quad \ldots, \\ c_1(t) &= \left(\gamma_1 + \gamma_2 t + \cdots + \gamma_{m_1}\tfrac{t^{m_1-1}}{(m_1-1)!}\right) e^{\lambda_1 t}. \end{aligned} \tag{25}$$

These considerations apply to any eigenvalue and an associated invariant subspace. We now suppose that λ_1 is the eigenvalue with largest real part and the invariant subspace is the largest of the invariant subspaces belonging to it, of dimension m_1.

Suppose first that $\lambda_1 = \rho_1$ is real. Then the vectors $\{\xi_j\}_1^{m_1}$ may also be chosen real, and we achieve a real solution of maximal growth on taking (for example) $\gamma_i = 0$ if $i = 1, 2, \ldots, m_1 - 1$ and $\gamma_{m_1} = 1$. The resulting formula for the solution $\xi(t)$ becomes

$$\xi(t) = \left(\frac{t^{m_1-1}}{(m_1-1)!}\xi_1 + \cdots + t\xi_{m_1-1} + \xi_{m_1} \right) e^{\rho_1 t}$$

and its norm is

$$\| \frac{t^{m_1-1}}{(m_1-1)!}\xi_1 + \cdots + t\xi_{m_1-1} + \xi m_1 \| e^{\rho_1 t}.$$

Consider the non-negative function

$$\phi(t) = \frac{\| \frac{t^{m_1-1}}{(m_1-1)!}\xi_1 + \cdots + t\xi_{m_1-1} + \xi m_1 \|}{1 + t^{m_1-1}}.$$

It is continuous on $[0, \infty)$ and therefore has a maximum and minimum there (see the argument attending equation (22) above). Its minimum, ϕ_* say, cannot vanish. For it is achieved at some point t_* and if ϕ_* vanished it would follow that

$$\frac{t_*^{m_1-1}}{(m_1-1)!}\xi_1 + \cdots + t_*\xi_{m_1-1} + \xi_{m_1} = 0.$$

This is a relation of linear dependence and is not possible since the set $\{\xi_j\}_1^{m_1}$ is linearly independent. Thus, for all $t \geq 0$,

$$\| \frac{t^{m_1-1}}{(m_1-1)!}\xi_1 + \cdots + t\xi_{m_1-1} + \xi m_1 \| \geq \phi_*(1 + t^{m_1-1}) \tag{26}$$

where $\phi_* > 0$. This provides a vector solution $\xi(t)$ of the kind proposed in Theorem 2, equation (15), except for the condition that $\xi(0) = 1$; a different choice of the constant γ_{m_1} can be made to satisfy this condition. This result is stronger than that of equation (15) in that it holds for all $t > 0$, not just for a sequence $\{t_j\}$.

Suppose now that λ_1, the eigenvalue of A with (algebraically) greatest real part ρ_1, is complex ($\sigma_1 \neq 0$). There may be more than one string of vectors like that of equation (23); if so we choose a string of maximal length m_1. For real matrices A, if λ_1 is complex, its complex conjugate $\overline{\lambda}_1$ also occurs, and possesses the corresponding eigenspace spanned by the vectors

$$\overline{\xi}_1, \quad \overline{\xi}_2 = (A - \overline{\lambda}_1 I)\overline{\xi}_1, \quad \cdots, \quad \overline{\xi}_{m_1} = (A - \overline{\lambda}_1 I)\overline{\xi}_{m_1-1}.$$

In the solution (24) above again choose $\gamma_j = 0$ if $j = 1, \ldots, m_1 - 1$, and choose $\gamma_{m_1} = 1$. This provides the solution

$$x(t) = \left(\frac{t^{m_1-1}}{(m_1-1)!}\xi_1 + \cdots + t\xi_{m_1-1} + \xi_{m_1} \right) e^{\lambda_1 t}.$$

Since $\lambda_1 = \rho_1 + i\sigma_1$ is complex, so also is $\xi_k = \eta_k + i\zeta_k$. The $2m_1$ complex vectors $\{\xi_j, \overline{\xi}_j\}_1^{m_1}$ are linearly independent over C and therefore the $2m_1$ real vectors $\{\eta_j, \zeta_j\}_1^{m_1}$, consisting of their real and imaginary parts, are linearly independent over R. From the equation for $x(t)$ above we can form the real solution $y(t) = (x(t) + \overline{x}(t))/2$ or

$$y(t) \quad \equiv \quad = e^{\rho_1 t} \left[\cos(\sigma_1 t) \left(\frac{t^{m_1-1}}{(m_1-1)!}\eta_1 + \cdots + t\eta_{m_1-1} + \eta_{m_1} \right) - \right. \tag{27}$$

$$\left. \sin(\sigma_1 t) \left(\frac{t^{m_1-1}}{(m_1-1)!}\zeta_1 + \cdots + t\zeta_{m_1-1} + \zeta_{m_1} \right) \right] \tag{28}$$

At times $t_j = (\pi/\sigma_1)j$ this real solution has norm

$$\|y(t_j)\| = e^{\rho t_j} \| \frac{t_j^{m_1-1}}{(m_1-1)!}\eta_1 + \cdots + t_j\eta_{m_1-1} + \eta_{m_1} \|.$$

The factor of $\exp(\rho t_j)$, is bounded below as in the inequality (26) above. This provides the estimate (15) above except for the initial condition, and this can be satisfied with a different choice of the constant γ_{m_1}.

3.3 Rationale

We address here how Putzer was led to this result. He first considered a more transparent result. It's clear from the use of the C-H theorem that you should be able to express $\exp(At)$ in finitely many powers of A:

$$e^{At} = \sum_{k=0}^{n-1} p_k(t) A^k. \tag{29}$$

Differentiating this with respect to t on the one hand, and using the formula

$$\frac{d}{dt}e^{At} = Ae^{At}$$

on the other, together with an application of the C-H theorem, provides a system of first-order differential equations for the coefficients $\{p_k\}$; evaluation at *zero* (where $\exp(At)$ reduces to the identity), provides initial data for this system. This works nicely and provides an expression for $\exp(At)$ via equation (29).

Putzer presents, without motivation, the alternative expression (4) given above. It can be motivated in the following way.

The space of linear operators on a vector space of dimension n is itself a vector space of dimension n^2. For a given operator A with minimal polynomial of degree m, the powers $\{A^k\}$ form a subspace of dimension m: the operators $\{A^k\}_0^{m-1}$ form a basis for this subspace. This underlies the formula (29). What underlies the formula (4) is that the operators $\{P_k\}_0^{m-1}$ likewise form a basis for this subspace. To see this, it suffices to show that they are linearly independent. To this end consider the expression of linear dependence

$$c_0P_0 + c_1P_1 + \cdots + c_{m-1}P_{m-1} = 0. \tag{30}$$

Operate on this sum with the operator

$$(A - \mu_2 I)\cdots(A - \mu_m I).$$

It is easy to see that every term but the first disappears from this, whereas the coefficient of c_0 cannot vanish, or there would be a vanishing polynomial of degree less than m. This can't happen since m is the degree of the minimal polynomial, so $c_0 = 0$. We now return to equation (30) with $c_0 = 0$ and apply the operator

$$(A - \mu_3 I)\cdots(A - \mu_m I)$$

and infer, by the same reasoning as above, that $c_1 = 0$. We continue successively in this way to find that all the coefficients in equation (30) must vanish, verifying the linear independence of the operators $\{P_k\}_0^{n-1}$.