STABILITY OF PERIODIC LINEAR DELAY–DIFFERENTIAL EQUATIONS AND THE CHEBYSHEV APPROXIMATION OF FUNDAMENTAL SOLUTIONS

ED BUELER AND ERIC BUTCHER

ABSTRACT. These notes are a general introduction and exposition. They contain more theory and more appendices than are likely to appear in any paper but components will be published.

We start with relatively well-known theory including an important variation–of–parameters formula whose (period)=(delay) form is equation (7). Next we review the Sinha–Wu [SW] method of approximation of fundamental solutions to ODEs. We give additional results on the accuracy of Picard iteration and Chebyshev approximation.

We extend the method to the approximation of the "infinite–dimensional Floquet transition matrix" U in (7). The stability of the DDE is approximately determined by the eigenvalues of the approximating matrix to U.

Examples are given in [A] which show the effectiveness of the resulting approximation. V. Averina's thesis [A] should be regarded as part II of these notes. See [BM] for numerical results using the ideas of these notes.

CONTENTS

1. Theory for periodic linear delay–differential equations

These notes address the solution of, and the stability of, the system of delay–differential equations (DDEs), with fixed *delay* $\tau > 0$,

$$(1) \qquad \dot{x}(t) = A(t)x(t) + B(t)x(t - \tau).$$

Here $x(t) \in \mathbb{R}^n$ and A, B are continuous "$n \times n$ matrix" functions of time $t \geq 0$. That is, A and B are linear operators on $\mathbb{R}^n$ which depend continuously on t. Because we will be interested in the periodic case, we assume A and B are defined for all $t \in \mathbb{R}$ and are bounded.

We start by deriving a variation–of–parameters formula for the solution of (1). First recall the existence of a fundamental solution $\Phi(t)$ to the ODE system $\dot{x} = A(t)x$. That is, $\Phi(t)$ is an $n \times n$ matrix solution to the initial value problem

$$(2) \qquad \dot{\Phi} = A(t)\Phi, \quad \Phi(0) = I,$$

where I is the identity on $\mathbb{R}^n$.

Lemma 1. *For all $t \in \mathbb{R}$, $\Phi(t)$ is invertible. If $\Psi^\top = \Phi^{-1}$, where $(\cdot)^\top$ is the matrix transpose, then*

$$(3) \qquad \dot{\Psi} = -A(t)^\top \Psi, \quad \Psi(0) = I.$$

See appendix A for the proof.

We can use Φ and Ψ as integrating factors to write an integral equation "solution" to (1). In fact, (4) below is a solution to (1) in the sense that only delayed values $x(s - \tau)$ for $s \leq t$ determine $x(t)$. It will be convenient to have $t_0 \neq 0$ in later sections.

Lemma 2. *If x solves* (1), *Φ solves* (2), *and Ψ solves* (3) *then*

$$(4) \qquad \boxed{x(t) = \Phi(t)\Psi(t_0)^\top x(t_0) + \int_{t_0}^t \Phi(t)\Psi(s)^\top B(s)x(s - \tau)\, ds.}$$

Proof. Define $y(t) \in \mathbb{R}^n$ by $x(t) = \Phi(t)y(t)$, or equivalently $y(t) = \Psi(t)^\top x(t)$. Then

$$\dot{x}(t) = \dot{\Phi}(t)y(t) + \Phi(t)\dot{y}(t) = A(t)\Phi(t)y(t) + \Phi(t)\dot{y}(t).$$

On the other hand, by (1)

$$\dot{x}(t) = A(t)x(t) + B(t)x(t - \tau) = A(t)\Phi(t)y(t) + B(t)\Phi(t - \tau)y(t - \tau).$$

Thus by cancelling the term $A(t)\Phi(t)y(t)$,

$$\dot{y}(t) = \Psi(t)^\top B(t)\Phi(t - \tau)y(t - \tau).$$

This can be solved by integration:

$$y(t) - y(t_0) = \int_{t_0}^t \Psi(s)^\top B(s)\Phi(s - \tau)y(s - \tau)\, ds$$

or

$$\Phi(t)^{-1}x(t) = \Psi(t_0)^\top x(t_0) + \int_0^t \Psi(s)^\top B(s)x(s-\tau)\,ds$$

as claimed. □

Equation (4) should rightly be called a "variation–of–parameters" (or "convolution" or "Duhamel's principle") solution to the delay differential equation (1). In fact, (4) has precisely the same form as the variation–of–parameters solution to $\dot{x}(t) = A(t)x(t) + u(t)$ where $u(t)$ is a given nonhomogeneity.

Equation (4) is the foremost method for actually solving delay differential equations. In particular, (4) determines $x(t)$ for $t \in (k\tau, (k+1)\tau]$ if $x(t)$ is known for $t \in ((k-1)\tau, k\tau]$.

Before considering the periodicity of $A(t)$ and $B(t)$ we note $\Phi(t)$ in (2) and $\Psi(t)^T$ in (3) can be regarded as special cases of a two–parameter linear map $\Omega_s^t : \mathbb{R}^n \to \mathbb{R}^n$ with $s, t \in \mathbb{R}$.

Definition. Let $v \in \mathbb{R}^n$. Define $\Omega_s^t v$ as the value $\xi(t)$ of the solution to $\frac{d}{d\theta}\xi(\theta) = A(\theta)\xi(\theta)$ with initial condition $\xi(s) = v$.

The uniqueness of solutions of ODEs makes Ω_s^t well–defined. It is easy to prove that Ω_s^t is linear. Conceptually, Ω_s^t advances the solution forward from time s to time t. However, we do not assume $s \le t$. Note Ω_s^t is the general form of a "state transition matrix." In physics terms, Ω_s^t is a "propagator." See appendix A for a proof of the following "group" property of Ω_s^t.

Lemma 3. $\Omega_s^t \Omega_r^s = \Omega_r^t$.

Note that $\Phi(t) = \Omega_0^t$. By the lemma $\Omega_t^0 \Omega_0^t = \Omega_0^0 = I$ so $\Phi(t)^{-1} = \Omega_t^0$. Thus we have $\Psi(s)^\top = \Phi(s)^{-1} = \Omega_s^0$ and $\Phi(t)\Psi(s)^\top = \Omega_0^t \Omega_s^0 = \Omega_s^t$. It follows that (4) is equivalent to

$$(5) \qquad x(t) = \Omega_{t_0}^t x(t_0) + \int_{t_0}^t \Omega_s^t B(s)x(s-\tau)\,ds.$$

Assumption 1. From now on we suppose $A(t)$ and $B(t)$ are periodic with period $T > 0$.

Lemma 4. *Under the above assumption for A,*

$$\Omega_{s-T}^{t-T} = \Omega_s^t.$$

The idea is that advancing the solution from s to t uses $A(\theta)$ for $\theta \in [s, t]$, and advancing the solution from $s - T$ to $t - T$ uses exactly the same values of A, since A is periodic. The proof is again in appendix A.

By this lemma, if A and B are periodic then the fundamental solutions Φ and Ψ satisfy

$$\boxed{\Phi(t)\Psi(s)^\top = \Phi(t-T)\Psi(s-T)^\top.}$$

The following two corollaries explain the significance of the *Floquet transition matrix* $\Phi(T)$. It contains all the information necessary to determine (for instance) the long–term stability of (2).

Corollary 1. *If $\Phi(t)$ solves* (2) *then*

$$\Phi(nT + \eta) = \Phi(\eta)\Phi(T)^n.$$

In particular, (2) *is (asymptotically) stable if and only if all the eigenvalues of $\Phi(T)$ are inside the unit circle.*

Proof. From the boxed equation above,

$$\Phi(nT + \eta)\Psi(T)^\top = \Phi((n-1)T + \eta)\Psi(0)^\top = \Phi((n-1)T + \eta).$$

Thus $\Phi(nT + \eta) = \Phi((n-1)T + \eta)\Phi(T)$ since $\Psi(t)^\top = \Phi(t)^{-1}$. Induct. $\qquad\square$

Corollary 2 (Floquet theorem)**.** *Consider*

$$(6) \qquad\qquad\qquad\qquad \dot{x} = A(t)x.$$

A nonzero solution to (6) *of the form $x(t) = e^{\lambda t}p(t)$ with $p(t)$ periodic of period T exists if and only if $e^{\lambda T}$ is an eigenvalue of $\Phi(T)$ and $p(0)$ is an eigenvector of $\Phi(T)$.*

Proof. By definition of the fundamental solution Φ, $x(t) = \Phi(t)x(0)$ solves (6) with initial condition $x(0)$. By corollary 1, if $t = nT + \eta$ for some integer n then $x(t) = \Phi(\eta)\Phi(T)^n x(0)$.

If $x(0)$ is an eigenvector of $\Phi(T)$ with eigenvalue $e^{\lambda T}$ then

$$x(t) = \Phi(\eta)(e^{\lambda T})^n x(0) = e^{\lambda(nT+\eta)}e^{-\lambda\eta}\Phi(\eta)x(0) = e^{\lambda t}p(\eta)$$

where $p(t) = e^{-\lambda t}\Phi(t)x(0)$. Note $p(kT+\eta) = e^{-\lambda\eta}\Phi(\eta)(e^{-\lambda T})^k\Phi(T)^k x(0) = e^{-\lambda\eta}\Phi(\eta)x(0) = p(\eta)$ so p is periodic with period T.

Conversely, suppose $x(t) = e^{\lambda t}p(t)$ is a nonzero solution of (6) with $p(t)$ periodic. Since also $x(T) = \Phi(T)p(0)$, it follows that $\Phi(T)p(0) = e^{\lambda T}p(T) = e^{\lambda T}p(0)$. $\qquad\square$

Corollary 3 (Liapunov–Floquet theorem)**.** *Suppose $\Phi(T) = e^{CT}$ for fixed matrix C. Then the matrix function $L(t) = \Phi(t)e^{-Ct}$ is invertible and T–periodic. Furthermore, if $x = L(t)y$ then $\dot{x} = A(t)x$ if and only if $\dot{y} = Cy$.*

Proof. Note L is invertible because Φ is invertible. Also,

$$L(t + T) = \Phi(t + T)e^{-C(t+T)} = \Phi(t)\Phi(T)e^{-CT}e^{-Ct} = L(t).$$

Note

$$\dot{L}(t) = \dot{\Phi}(t)e^{-Ct} - \Phi(t)e^{-Ct}C = A(t)L(t) - L(t)C.$$

Now suppose $x = L(t)y$. If $\dot{x} = A(t)x$ then $\dot{x} = \dot{L}y + L\dot{y} = ALy - LCy + L\dot{y}$ on the one hand and $\dot{x} = ALy$ on the other, so $-LCy + L\dot{y} = 0$ which is equivalent to $\dot{y} = Cy$ because L is invertible.

If $\dot{y} = Cy$ then $\dot{x} = \dot{L}y + L\dot{y} = (AL - LC)y + LCy = ALy = Ax.$ $\qquad\square$

Our goal is the stability of (1) not (2), so we return to the delayed case. We specialize here to the simplest situation in which $\tau = T$. It is relatively easy to work in the case $\tau = MT$ (for M an integer). See [HL]. The case $T = N\tau$ (for N an integer) is addressed in appendix C. We are interested in the latter case because it is characteristic of milling.

Assumption 2. $\tau = T$, that is, the delay and the period of the coefficients are equal.

By this assumption and since B is periodic the integral in (5) can be rewritten

$$\int_{t_0}^{t} \Omega_s^t B(s)x(s-\tau)\,ds = \int_{t_0}^{t} \Omega_{s-\tau}^{t-\tau} B(s-\tau)x(s-\tau)\,ds$$

$$= \int_{t_0-\tau}^{t-\tau} \Omega_{s'}^{t-\tau} B(s')x(s')\,ds'.$$

For stability analysis we need to think of the solution to (1) as defining a linear map from a space of functions back to the same space of functions. In particular, we can then talk sensibly about the eigenvalues of that map.

Definition.
$$C[a,b] = \left\{ x : [a,b] \to \mathbb{R}^n \,\middle|\, x \text{ is continuous} \right\}.$$

Note that elements of $C[a,b]$ are $\mathbb{R}^n$–valued functions. That is, if one chooses a basis for $\mathbb{R}^n$ then each element of $C[a,b]$ is really a list of n continuous scalar functions on $[a,b]$.

Let $\theta = t - \tau$ for $t \in [0,\tau]$. Note $\Omega_0^t = \Omega_{-\tau}^{t-\tau} = \Omega_{-\tau}^{\theta}$. Let $t_0 = 0$. Then (5) can be written

$$(5') \qquad (\tilde{U}x)(\theta) = \Omega_{-\tau}^{\theta} x(0) + \int_{-\tau}^{\theta} \Omega_s^{\theta} B(s)x(s)\,ds$$

which defines a linear map $\tilde{U} : C[-\tau,0] \to C[-\tau,0]$. Equation (5') is equation (3.5) in [HL] section 8.3.

It is convenient to shift $\tilde{U}$ to the interval $[0,\tau]$. It will also make sense for what follows to use Φ and Ψ instead of Ω. Thus for $t \in [0,\tau]$, define

$$(7) \qquad \boxed{(Ux)(t) = \Phi(t)\left\{ x(\tau) + \int_0^t \Psi(s)^\top B(s)x(s)\,ds \right\},}$$

a linear map

$$U : C[0,\tau] \to C[0,\tau].$$

Note how the end value $x(\tau)$ is used. Equation (7) makes sense in the case $B = 0$, that is, no delay. In fact, note that $(U\cdot)(T)$ "is" the Floquet transition matrix $\Phi(T)$ if $B = 0$.

The initial value problem

$$(8) \qquad \dot{x}(t) = A(t)x(t) + B(t)x(t-\tau), \qquad x(s-\tau) = \varphi(s) \text{ for } s \in [-\tau, 0],$$

for the differential–delay equation (1) is solved by iterating U as follows. Let $f(t) = \varphi(t - \tau)$ for $t \in [0, \tau]$. Then:

$$x(t) = (Uf)(t) \qquad\qquad\qquad \text{for } t \in [0, \tau],$$
$$x(t) = (U^2 f)(t - \tau) \qquad\qquad\qquad \text{for } t \in [\tau, 2\tau],$$
$$x(t) = (U^3 f)(t - 2\tau) \qquad\qquad\qquad \text{for } t \in [2\tau, 3\tau],$$

and so on. This is the "method of steps."

Note $A(t)$ in equation (1) will depend on parameters ("system parameters"). This means Φ, Ψ, and Ω depend on those parameters. Also, $B(t)$ depends on parameters, so U in fact depends on all the parameters which appear in (1). We point out these dependencies because our goal is a parameter–dependent symbolic stability analysis.

We assert that U is the "delay Floquet transition matrix". It is an operator on the infinite dimensional space $C[0, \tau]$ so "matrix" is a little silly. Its existence depends essentially on the periodicity of $A(t)$ and $B(t)$, so "Floquet" is appropriate.

The stability question: Are all the eigenvalues of U inside the unit circle?

To actually possess eigenvalues, a matrix M must be square. That is, M must be the matrix of a linear operator from a vector space back to itself. The operator U is "square" because is maps $V = C[0, \tau]$ to itself.

For nonconstant coefficient problems one does not expect U to have an analytical expression. One approximates U instead. Suppose, in fact, that $V_l \subset V$ is a finite–dimensional subspace of dimension nl. (Note that V contains $\mathbb{R}^n$–valued functions so that a subspace V_1 containing only constant functions is of dimension n.) For this paper V_l will be the span of the shifted Chebyshev polynomials $\{T_0^*(t), \ldots, T_{l-1}^*(t)\}$ defined in section 3. For the sake of argument, V_l could also be a Fourier subspace, for instance the span of $\{1, \cos(\pi t), \cos(2\pi t), \ldots, \cos((l-1)\pi t)\}$.

In any case, U maps V_l into a subspace of V of dimension at most nl. A square matrix approximation to U would be

$$U \approx P_{V_l} U\big|_{V_l},$$

where $U\big|_{V_l}$ is the restriction of U to V_l (i.e. so V_l is the input space to U), and P_{V_l} is a projection of V onto V_l (so V_l is the output space). This would replace U by an nl by nl matrix. The stability question for U is to be approximately answered in terms of the (parameter dependent!) eigenvalues of this $nl \times nl$ matrix approximation to U.

If we are to believe that U is really an infinite matrix, then the reader should rightly be skeptical that a mere $nl \times nl$ approximation (for any finite l) will carry enough of the information in U so as to give adequate stability information about the delay differential equation (1). In this regard, at least to the mathematician, the following is the fundamental theorem of linear delay differential equations:

Theorem 1. $U : V \to V$ *is a compact (or completely–continuous) operator.*

A very general proof is in [HL] (lemma 2.3). However, a specific proof based on the variation–of–parameters formula is given in appendix D. By definition, a compact operator takes bounded subsets to sets which have compact closure.

The definition of "compact operator" may not suggest the important point which is that U is well approximated by finite matrices. Any compact operator on V is the operator norm limit of finite rank operators. At this level of abstraction we want to find a finite dimensional subspace V_l so that $U_l = P_{V_l} U|_{V_l}$ is a good approximation to U in operator norm, and the fact that U is compact is what makes this possible. In particular, we claim that the eigenvalues of U are approximated by those of U_l for sufficiently large l.

2. APPROXIMATION BY PICARD ITERATION

Picard iteration is commonly used for the purpose of proving existence and unique-ness of solutions of ODE initial value problems. That is not its only use. Here we develop a form suited to approximation of matrix–valued solutions of linear ODEs. In particular we wish to have a polynomial expression for the fundamental solution because the goal is a *symbolic* stability analysis.

The fundamental solution $\Phi(t)$ solves (2) or equivalently

$$(9) \qquad \Phi(t) = I + \int_0^t A(s)\Phi(s)\,ds.$$

Start Picard iteration with $\Phi^{(0)}(t) = I$, which is a poor approximation but is correct at $t = 0$. Let

$$\Phi^{(1)}(t) = I + \int_0^t A(s_0)\Phi^{(0)}(s_0)\,ds_0 = I + \int_0^t A(s_0)\,ds_0.$$

Continue on by induction, so that

$$(10) \qquad \Phi^{(p)}(t) = I + \int_0^t A(s_{p-1})\Phi^{(p-1)}(s_{p-1})\,ds_{p-1}$$

$$= I + \int_0^t A(s_{p-1})\,ds_{p-1} + \int_0^t A(s_{p-1})\int_0^{s_{p-1}} A(s_{p-2})ds_{p-2}\,ds_{p-1}$$

$$+ \cdots + \int_0^t A(s_{p-1})\int_0^{s_{p-1}} A(s_{p-2})\ldots\int_0^{s_1} A(s_0)\,ds_0\,ds_1\ldots ds_{p-1}.$$

Better notation uses:

Definition. If $F = F(t)$ is a continuous $n \times n$ matrix–valued function then let

$$(G_A F)(t) = \int_0^t A(s)F(s)\,ds.$$

We can write

$$\Phi^{(p)}(t) = I + (G_A\,I)(t) + (G_A(G_A\,I))(t) + \cdots + (\underbrace{G_A\ldots(G_A}_{p \text{ times}}\,I)\ldots))(t),$$

or even

$$(11) \qquad \boxed{\Phi^{(p)}(t) = \left[(1 + G_A + (G_A)^2 + \cdots + (G_A)^p)I\right](t).}$$

Abstractly, the limit of (11) is $\Phi = (1 - G_A)^{-1}I$ where 1 is the identity mapping on the function space of $n \times n$ matrix–valued functions and I is the constant $n \times n$ matrix function which is always the identity. Equation (9) is $\Phi = I + G_A\Phi$ in these terms. The inverse $(1 - G_A)^{-1}$ in the space of operators can be calculated by the usual power series $1 + G_A + G_A^2 + \ldots$. At the level of the matrix–valued functions themselves this is Picard iteration. In particular, the series has good convergence properties. For instance, it produces the series of the exponential e^{At} in the constant–coefficient

$(A(t) = A)$ case. The good convergence properties follow from the fact that G_A is a compact operator (though it has no eigenvalues!).

Similarly,

(12)
$$\Psi^{(p)}(t) = \left[(1 - G_{A^\top} + (G_{A^\top})^2 - \cdots + (-1)^p (G_{A^\top})^p)I\right](t)$$

is the Picard approximation for the adjoint equation (3).

The following lemma shows that actual convergence to the correct $\Phi(t)$ (respectively, $\Psi(t)$) happens as $p \to \infty$. We simultaneously derive an estimate for p sufficient to achieve a certain accuracy. (And we now use, for the first time in this paper, vector and matrix *norms*. It may be useful to review appendix B.)

Lemma 5. *Let $\|\cdot\|_r$ be the matrix norm for $1 \le r \le \infty$. For t in any finite interval $[0, b]$, $\Phi^{(p)}(t) \to \Phi(t)$ as $p \to \infty$. In fact,*

(13)
$$\|\Phi(t) - \Phi^{(p)}(t)\|_r \le \sum_{q=p+1}^\infty \overline{a}_r^q \frac{t^q}{q!}$$

where

$$\overline{a}_r = \max_{0 \le s \le b} \|A(s)\|_r.$$

A similar result holds for convergence of $\Psi^{(p)}(t)$.

Proof. The content of the second sentence is this: One can show that for fixed t the matrices $\{\Phi^{(p)}(t)\}_{p=0}^\infty$ form a Cauchy sequence in space of matrices. One then shows that the limit of the sequence as $p \to \infty$ satisfies the ODE IVP (2) and thus by uniqueness the limit must be $\Phi(t)$.

In fact, one can show the sequence is Cauchy by showing (13), as follows. Let $n_q(t) = \|\Phi^{(q)}(t) - \Phi^{(q-1)}(t)\|_r$ for $q \ge 1$ and let $n_0(t) = 1$. Then Picard iteration (10) shows

$$n_q(t) \le \int_0^t \|A(s)\|_r n_{q-1}(s)\, ds$$

for $q \ge 1$. Thus $n_q(t) \le \overline{a}_r \int_0^t n_{q-1}(s)\, ds$. Easy induction gives $n_q(t) \le \overline{a}_r^q t^q / q!$.

The sequence $\{\Phi^{(p)}(t)\}$ is Cauchy because $\sum_{q=j}^k x^q / q! \to 0$ as $j, k \to \infty$ for any x. Equation (13) follows from $\|\Phi(t) - \Phi^{(p)}(t)\|_r \le \sum_{q=p+1}^\infty \|\Phi^{(q)}(t) - \Phi^{(q-1)}(t)\|_r$. $\square$

This lemma shows that the following *conservative* algorithm will estimate p given a specified $TOL > 0$. Here TOL represents how closely we wish to approximate $\Phi(t)$ by $\Phi^{(p)}(t)$ in the interval $t \in [0, b]$. Use of the resulting p will guarantee that all entries of $\Phi^{(p)}(t)$ will be within TOL of the desired values. Because $\Psi(t)$ is computed from $A(s)^\top$ and because $\|A(s)\|_1 = \|A(s)^\top\|_\infty$, the algorithm assures equal accuracy for $\Psi^{(p)}(t)$.

Algorithm 1. Inputs: $A(s)$ *for $s \in [0, 1]$ and $TOL > 0$.*

   (1) *Calculate (or find upper bounds for)* $\|A(s)\|_1 = \max_j(\sum_i |a_{ij}(s)|)$ *for each*
       $s \in [0,1]$ *and* $\|A(s)\|_\infty = \max_i(\sum_j |a_{ij}(s)|)$ *for each* $s \in [0,1]$.
   (2) *Calculate (or find upper bounds for)* $\overline{a}_1 = \max_{0 \le s \le 1} \|A(s)\|_1$ *and* $\overline{a}_\infty = \max_{0 \le s \le 1} \|A(s)\|_\infty$.
   (3) *Define* $f_1(p) = \exp(\overline{a}_1) - \sum_{q=0}^{p}(\overline{a}_1)^q/q!$ *and* $f_\infty(p) = \exp(\overline{a}_\infty) - \sum_{q=0}^{p}(\overline{a}_\infty)^q/q!$.
       Calculate $f(p) = \max\{f_1(p), f_\infty(p)\}$ *for successive values of* p *until* $f(p) <$
       TOL.

   Output: p *such that* $\Phi^{(p)}(t)$ *is entrywise within TOL of* $\Phi(t)$ *and* $\Psi^{(p)}(t)$ *is entrywise
within TOL of* $\Psi(t)$ *for all* $t \in [0,1]$.

   Note that the function $f(p)$ is easily calculated exactly for values of p large enough
for any practical approximation purpose. However, if $2|a| < p + 2$ then[1]

$$\sum_{q=p+1}^{\infty} \frac{a^q}{q!} \le 2\frac{a^{p+1}}{(p+1)!},$$

so f can be easily estimated from the first term of the sum, at least for large p which
is the case of interest.

   As an approximation scheme we intend to use (11) and (12). We regard p as a
parameter of the scheme, which we can choose with confidence based on a combination
of the above worst–case analysis and numerical experiment.

Example. Consider the Mathieu equation $y'' + (a + \epsilon \cos t)y = 0$. As a $T = 1$ periodic
first–order system (i.e. put in state–space form and the normalized) we have

$$(14) \qquad \dot{x}(t) = A(t)x \quad \text{where} \quad A(t) = \begin{pmatrix} 0 & 2\pi \\ -2\pi(a + \epsilon\cos(2\pi t)) & 0 \end{pmatrix}.$$

Clearly $\|A(s)\|_1 = \|A(s)\|_\infty$ for all s. Let $\overline{a} = \max_{0 \le s \le 1}\|A(s)\|_1$. Note

$$\overline{a} = \max_{s \in [0,1]}\{2\pi|a + \epsilon\cos(2\pi s)|, 2\pi\} = 2\pi\max\{|a| + |\epsilon|, 1\}.$$

Approximate Floquet multipliers for this system are shown in Table 2 of [SB] for
various p. In the worst–case analysis given here, we want to know what p will ap-
proximate $\Phi(1)$ within *TOL* in the part of the parameter plane where $|a| < 1.5$ and
$|\epsilon| < 1.5$. (See also figure 5 in [SB].) Thus $\overline{a} = 6\pi$. Let $TOL = .001$. We hope
that the eigenvalues of the Floquet matrix $\Phi(1)$ will be roughly that accurate. The
algorithm says we find p so that $f(p) = e^{\overline{a}} - \sum_{q=0}^{p}(\overline{a})^q/q!$ is less than TOL. In fact
we find:

p	:	10	20	30	40	50	54	55	60
$f(p)$	:	1.5(8)	5.2(7)	9.6(5)	1.0(3)	0.11	1.6(−3)	5.5(−4)	1.7(−6),

where "5.5(−4)" denotes $.00055 = 5.5 \times 10^{-4}$. These values for $f(p)$ show $p = 55$
iterations will work.

   Now, Picard iteration has been used in [SW] and in [SB] to approximate funda-
mental solutions using (11). The number 55 above compares to $p = 24$ in [SB] which

---

[1]Approximate the series by an obvious geometric series.

was chosen based on comparing the Picard approximation visually to the "exact" Φ computed with Runge–Kutta. We can come much closer to this lower value of p by replacing the norm of $A(t)$ in the above by the maximum magnitude of the time–dependent eigenvalues of $A(t)$. This replacement is "heuristic" because the noncommutative nature of most systems like (2) means that the eigenvalues of $A(t)$ are not the exact exponential rates present in the solution of (2). See [A].

In any case, $\lambda(t) = \pm i 2\pi \sqrt{a + \epsilon \cos(2\pi t)}$ are these eigenvalues and with the above bounds on a, b they attain a maximum magnitude of $\bar{a} = 2\pi\sqrt{3} \approx 10.9$ and we find

p :	10	20	25	30	32	33	35
$f(p)$:	2.8(4)	2.2(2)	3.7	2.5(−2)	2.7(−3)	8.7(−4)	8.0(−5)

so $p = 33$ steps of Picard iteration is likely to suffice for an approximation of $\Phi(t)$ within $TOL = .001$.

The Picard techniques of this section should be compared to two alternatives. A direct integral equation technique is used in [BM] and it is more efficient for numerical purposes, but does not satisfy the symbolic stability need for a polynomial expression for U in the system parameters. On the other hand, see [A] for thoughts on the use of the Magnus expansion which can be regarded as "Picard iteration in the exponent." Use of such a technique still (presumably) requires a polynomial approximation of the exponential.

3. Chebyshev approximation of differentiable functions

Next we consider approximating each entry of the coefficient matrices $A(t)$ and $B(t)$ and of the solution $\Phi(t)$ and $\Psi(t)$ by linear combinations of shifted Chebyshev polynomials. With these approximations, all Picard iteration computations of the previous section will be replaced by algebra with parameter dependent matrices of Chebyshev coefficients.

Assumption 3. For simplicity suppose $\tau = T = 1$.

Definition. Let T_j^* be the jth Chebyshev polynomial T_j shifted to the interval $[0, 1]$. That is, $T_j^*(t) = T_j(2t - 1)$.

See [R] for facts about Chebyshev polynomials.

Because $T_j(x) = \cos(j \arccos x)$, clearly $|T_j^*(t)| \leq 1$ for any $t \in [0, 1]$. Also $T_j^*(1) = 1$ for all j. Note $T_0^*(t) = 1$ and $T_1^*(t) = 2t - 1$. The polynomials can be computed by the recursion

$$T_{n+1}^*(t) = 2(2t - 1)T_n^*(t) - T_{n-1}^*(t),$$

so $T_2^*(t) = 8t^2 - 8t + 1$, $T_3^*(t) = 32t^3 - 48t^2 + 18t - 1$, etc.

The set $\{T_j^*(t)\}_{j \geq 0}$ is orthogonal with respect to $(t - t^2)^{-1/2}$ on the interval $[0, 1]$:

$$(15) \qquad \int_0^1 \frac{T_i^*(t)T_j^*(t)}{\sqrt{t - t^2}} \, dt = \begin{cases} 0, & i \neq j \\ \pi, & i = j = 0, \\ \pi/2, & i = j = 1, 2, \ldots \end{cases}$$

The polynomials $T_j^*(t)$ could thus be built by Gram–Schmidt applied to the set of monomials $\{t^n\}$, but the recursion is faster as a practical method.

If $f(t) = \sum_{j=0}^{\infty} b_j T_j^*(t)$ is a scalar function expanded in shifted Chebyshev polynomials, then

$$(16) \qquad b_j = \frac{2}{\pi} \int_0^1 f(t)T_j^*(t)(t - t^2)^{-1/2} \, dt, \quad j \neq 0,$$

and $b_0 = \pi^{-1} \int_0^1 f(t)(t - t^2)^{-1/2} \, dt$. (Compare this expression to (17) below.)

Such expansions are possible because the set $\{T_j^*(t)\}_{j=0}^{\infty}$ is a "basis"[2] for continuous scalar functions on $[0, 1]$. We mean, in particular, that every continuous function can be written as a *limit* of a sequence of finite linear combinations of Chebyshev polynomials. One can be specific about the nature of the limit. It is clearest that the limit works in $L^2([0, 1], (t - t^2)^{-1/2} \, dt)$ [R, RS]. For instance, if f is continuous then a sequence of finite linear combinations of the $T_j^*(t)$ can be found so that the

---

[2]Recall that the phrase "$\mathcal{S}$ is a *basis* for the vector space V" means that any element $v \in V$ can be written as a finite linear combination of elements of $\mathcal{S}$. On the other hand, it is clearly *not* the case that every continuous function is a finite linear combination of Chebyshev polynomials. The word "finite" appears because, from the definition of a vector space, one adds together elements two at a time and any finite number by associativity.

weighted mean–square error goes to zero. The coefficients b_j above are exactly so that the weighted mean square error goes to zero as fast as possible.

However, it is also a general fact that polynomials can *uniformly* approximate continuous functions.[3] An outstanding feature of Chebyshev polynomials is that they give good uniform approximations. We need these good uniform approximations because we need good entry–by–entry approximations of Φ *at each time* in constructing U. Thus it is not really good enough to know that the average error over an interval is small. We will show that the Chebyshev approximations do have small uniform error.

What is the most practical way to find finite expansions (approximations) in Chebyshev polynomials? It is possible to use the integrals (16) to find the constants for a best approximation in the smallest mean square error sense, but there is a different approximation procedure for which the uniform error is known to be small for differentiable functions and is easier to quantify. This faster procedure is based on *interpolation* rather than orthogonality with respect to the weight $(t - t^2)^{-1/2}$. The uniform error of both the smallest mean–square and interpolating polynomials can be analyzed and is comparable. However, it is generally easier to find the expansion constants for the interpolating polynomials.

Let m be a positive integer. Let $t_1, \ldots, t_m$ be the zeros of $T_m^*(t)$—explicitly, $2t_k - 1 = \cos(\pi(k - 1/2)/m)$. Suppose f is a continuous function on $[0, 1]$. In [NR][4] we find "discrete orthogonality relations" which, when translated to the shifted polynomials case, become the formulas:

$$\sum_{k=1}^{m} T_i^*(t_k)T_j^*(t_k) = \begin{cases} 0, & i \neq j \\ m, & i = j = 0, \\ m/2, & i = j = 1, 2, \ldots, m - 1 \end{cases}$$

(17) $$c_j = \frac{2}{m} \sum_{k=1}^{m} f(t_k)T_j^*(t_k), \quad j = 0, \ldots, m - 1$$

(18) $$f(t) \approx P_f^{m-1}(t) \equiv \frac{c_0}{2} + \sum_{j=1}^{m-1} c_j T_j^*(t).$$

It turns out that the approximation (18) is exact at the zeros t_k, that is, $f(t_k) = P_f^{m-1}(t_k)$ for $k = 1, \ldots, m$ if the c_j are determined by (17).[5] From now on, $P_f^{m-1}(t)$ is the *interpolating polynomial* described in (18). It is an $m - 1$ degree polynomial which agrees with f at the m points t_k which are the roots of $T_m^*(t)$.

How good an approximation to f is the polynomial P_f^{m-1} *away from* the "nodes" t_k? A result is:

---

[3]The Weierstrass approximation theorem.

[4]Specifically, [NR] pp. 190–191. See also exercises 1.5.26 and 1.5.27 in [R].

[5]See appendix A for a proof of this fact and of the discrete orthogonality relation.

Proposition 1 (corollary 8.11 of [BF] translated to the shifted case)**.** *If f has m continuous derivatives on $t \in [0,1]$ then*

$$\max_{t\in[0,1]} |f(t) - P_f^{m-1}(t)| \le \frac{1}{2^{2m-1}m!} \max_{t\in[0,1]} |f^{(m)}(t)|.$$

All polynomial interpolation schemes have a factor $1/m!$ as above, but an indication of the benefits of using Chebyshev polynomials is the factor $1/2^m$.[6] This factor is as small as it can be for polynomial interpolation.

The above proposition *is* useful in estimating m necessary to approximate the entries of $A(t)$ (and $B(t)$) accurately. It is also useful for estimating in advance the m necessary for the method of the next section to achieve a certain accuracy.

Assumption 4. The entries of $A(t)$ and $B(t)$ are functions which have as many continuous derivatives as needed in the following analysis.

Example. Again consider the Mathieu equation in [SB]: $y'' + (a + \epsilon \cos t)y = 0$. Recall $A(t)$ in (14), with nonconstant entry

$$f(t) = a_{21}(t) = -2\pi(a + \epsilon \cos(2\pi t)).$$

Note that even derivatives of f follow the rule:

$$f^{(2k)}(t) = \pm(2\pi)^{(2k+1)}\epsilon \cos(2\pi t).$$

Suppose one wishes to approximate f within $TOL > 0$. We want to choose k so that

$$\left| \frac{1}{2^{4k-1}(2k)!}(2\pi)^{(2k+1)}\epsilon \cos(2\pi t) \right| \le \frac{|\epsilon|\pi^{2k+1}}{2^{2k-2}(2k)!} < TOL.$$

If $TOL = 10^{-3}$ and $\epsilon = 1.5$ then I find that $k = 5$ is the first number for which the right–hand inequality is true. Thus if one uses $m = 2k = 10$ then the Chebyshev polynomial $P_f^{m-1}(t) = P_f^9(t)$ found from (18) will be within 10^{-3} of $a_{21}(t)$ at every $t \in [0,1]$. Since the other entries of $A(t)$ are constant, this clearly suffices for approximating $A(t)$.

There is a suggestive analysis in [SW] which addresses the more difficult problem of predicting the error, in advance, of an approximation to the fundamental solution $\Phi(t)$ by Chebyshev polynomials. That analysis suggests that the error $\epsilon(t) = \|\Phi(t) - \Phi^{m-1}(t)\|$, where $\Phi^{m-1}(t)$ is the Chebyshev approximation using the first m polynomials, can be estimated by calculating the $m+1$ polynomial approximation $\Phi^m(t)$ and using

$$\epsilon(t) \le \|\Phi^m(t) - \Phi^{m-1}(t)\|.$$

As it stands this can't be true because the right side will be zero at up to m points in the interval, which would mean that one knows the exact solution to any linear

---

[6]A factor of $1/2^{m-1}$ appears simply because the interval here is $[0,1]$.

differential equation. However, this heuristic does get the right order of magnitude in some examples.[7]

The following example analysis is based on an obvious necessary condition for accuracy of any scheme for approximating solutions of differential equations by Chebyshev polynomials. Namely, if $P^{m-1}(t)$ is a Chebyshev approximation to the solution of a constant coefficient ODE then m must be sufficiently large to approximate the exponential which solves the ODE. This idea can then be applied to nonconstant coefficient ODE, as in the second example.

Example. Consider the scalar initial value problem $\dot{y} = ay$, $y(0) = 1$ with solution $f(t) = e^{at}$ on the interval $t \in [0,1]$. Suppose we wish to approximate the solution by a sum of m shifted Chebyshev polynomials. By proposition 1 we choose m to satisfy

$$\frac{1}{2^{2m-1}m!}|a|^m e^{\operatorname{Re} a} \le TOL.$$

For example, if $a = 5$ and $TOL = 10^{-8}$ then $m = 15$ is the smallest value. Note that if $a = -5$ and $TOL = 10^{-8}$ then $m = 11$—clearly the sign of a only effects the e^a term in this analysis.

Recalling the Picard iteration of the previous section, it is interesting to compare the size of p in these cases. In particular the condition is $f(p) \le TOL$ (where $\bar{a}_1 = \bar{a}_\infty = |a|$) but by the comment following 1 it suffices to find p so that[8]

$$2\frac{|a|^{p+1}}{(p+1)!} \le TOL.$$

For our case $p = 25$ is the minimum for $a = \pm 5$ and $TOL = 10^{-8}$.

In general, comparison of the condition for p and for m to achieve a certain accuracy shows that $p > m$ in most cases—a factor of 1.5 to 3 seems typical in practice.

Example. Let us return to the Mathieu equation ([SB])

$$\dot{x} = A(t)x \quad \text{where} \quad A(t) = \begin{pmatrix} 0 & 2\pi \\ -2\pi(a + \epsilon \cos(2\pi t)) & 0 \end{pmatrix}.$$

We noted in an example in the previous section that $\bar{a} = \max_{0 \le s \le 1} \|A(s)\|_1 = \max_{0 \le s \le 1} \|A(s)\|_\infty = 2\pi \max\{|a| + |\epsilon|, 1\}$. We can regard $\bar{a}$ as a representative exponential rate because (by Gronwall's inequality)

$$\|x(t)\| \le \|x(0)\| \exp\left(\int_0^t \|A(s)\|_r \, ds\right) \le \|x(0)\|e^{\bar{a}t}$$

---

[7]It is possible to make a careful but extremely conservative analysis that suggests that the error is at most $\ln(m)/m$. That is not what is seen in practice.

[8]This requires $2|a| < p + 2$, i.e. $p > 8$, which is no restriction in practice.

$(r = 1, \infty)$. In keeping with the example of the previous section, suppose $|a|, |b| < 1.5$ and $TOL = .001$. Then $\bar{a} = 6\pi \approx 18.8$ and the criterion

$$\frac{1}{2^{2m-1}m!}|\bar{a}|^m e^{\bar{a}} \le TOL$$

is satisfied first by $m = 30$.

On the other hand, if the maximum magnitude of the time–dependent eigenvalues is used as the representative exponential rate then $\bar{a} = 2\pi\sqrt{3} \approx 10.9$ and $m = 19$ suffices. Compare the above values for m with $p = 55$, $p = 30$, respectively, in the previous section.

One property of Chebyshev polynomials has not yet been emphasized. Namely, recall that $|T_j^*(t)| \le 1$ if $0 \le t \le 1$. The proof of the following is therefore easy:

Lemma 6. *If*

$$f(t) = \sum_{j=0}^{m-1} c_j T_j^*(t)$$

and if $l \le m$ then

$$\left| f(t) - \sum_{j=0}^{l-1} c_j T_j^*(t) \right| \le \sum_{j=l}^{m-1} |c_j|$$

for all $t \in [0,1]$.

That is, the truncation of Chebyshev expansions is very nice because the maximum of the sum of the lost terms is at most the sum of the lost coefficients since the polynomials have maximum height 1.

Example. Approximate $f(t) = t^3$ on the interval $t \in [0,1]$ by a linear combination of $T_0^*(t)$, $T_1^*(t)$, $T_2^*(t)$. This means approximating t^3 by a quadratic, of course. One can easily show:

$$t^3 = \frac{1}{32}T_3^*(t) + \frac{6}{32}T_2^*(t) + \frac{15}{32}T_1^*(t) + \frac{10}{32}T_0^*(t).$$

Thus

$$t^3 \approx \frac{6}{32}T_2^*(t) + \frac{15}{32}T_1^*(t) + \frac{10}{32}T_0^*(t) = \frac{1}{32}\left[48t^2 - 18t + 1\right]$$

by truncating. Graphing t^3 versus this quadratic approximation is illuminating: the two graphs differ only by $\frac{1}{32}$ uniformly on $[0,1]$. But then this was the point of the lemma!

4. PRACTICAL CHEBYSHEV APPROXIMATION OF FUNDAMENTAL SOLUTIONS

We want to approximate the fundamental solution $\Phi(t)$ using Picard iteration, and then expand each entry in shifted Chebyshev polynomials, as in [SB] and [BS]. Similarly for $\Psi(t)$.

Let m be a fixed number of shifted Chebyshev polynomials. Let $\mathbf{T}(t) = (T_0^*(t)\ T_1^*(t)\ \ldots\ T_{m-1}^*(t))^\top$ be an $m \times 1$ column vector of the polynomials. Our notation for approximating functions by shifted Chebyshev polynomials looks like

$$(19) \qquad f(t) \approx \sum_{r=0}^{m-1} a_r T_r^*(t) = \mathbf{T}(t)^\top \mathbf{a}.$$

where $\mathbf{a}$ is the $m \times 1$ column vector of coefficients.

In this section we carefully develop the theory of "Chebyshev operational matrices." Also, we want to expand matrix–valued functions of $t \in [0,1]$, not mere scalar functions. In particular, the Picard approximation $\Phi^{(p)}(t)$ in equation (11), for $t \in [0,1]$, is a matrix–valued function. The following notation will therefore be needed.

Definition. If $\mathbf{b}_{ij}$ are an $m \times 1$ column vectors and $\mathbf{B}$ is the $nm \times n$ matrix

$$\mathbf{B} = \begin{pmatrix} \mathbf{b}_{11} & \mathbf{b}_{12} & \ldots & \mathbf{b}_{1n} \\ \mathbf{b}_{21} & \mathbf{b}_{22} & \ldots & \mathbf{b}_{2n} \\ \multicolumn{4}{c}{\ldots\ldots\ldots} \\ \mathbf{b}_{n1} & \mathbf{b}_{n2} & \ldots & \mathbf{b}_{nn} \end{pmatrix}$$

then $\mathbf{B}'$ and $\mathbf{B}^\circ$ are the matrices

$$\mathbf{B}' = \begin{pmatrix} \mathbf{b}_{11}^\top & \mathbf{b}_{12}^\top & \ldots & \mathbf{b}_{1n}^\top \\ \mathbf{b}_{21}^\top & \mathbf{b}_{22}^\top & \ldots & \mathbf{b}_{2n}^\top \\ \multicolumn{4}{c}{\ldots\ldots\ldots} \\ \mathbf{b}_{n1}^\top & \mathbf{b}_{n2}^\top & \ldots & \mathbf{b}_{nn}^\top \end{pmatrix}, \qquad \mathbf{B}^\circ = \begin{pmatrix} \mathbf{b}_{11} & \mathbf{b}_{21} & \ldots & \mathbf{b}_{n1} \\ \mathbf{b}_{12} & \mathbf{b}_{22} & \ldots & \mathbf{b}_{n2} \\ \multicolumn{4}{c}{\ldots\ldots\ldots} \\ \mathbf{b}_{1n} & \mathbf{b}_{2n} & \ldots & \mathbf{b}_{nn} \end{pmatrix}$$

which are $n \times nm$ and $nm \times n$, respectively.

Note that if $\mathbf{A}$ is of similar structure to $\mathbf{B}$ above then $\mathbf{A}'\mathbf{B}$ is an $n \times n$ matrix whose (i,j) entry is $\sum_{k=1}^n \mathbf{a}_{ik}^\top \mathbf{b}_{kj}$. Compare this to the product AB of two $n \times n$ matrices A, B with entries a_{ij}, b_{ij}, where the (i,j) entry of the product is $\sum_{k=1}^n a_{ik}b_{kj}$. The significance of $\mathbf{B}^\circ$ will be explained momentarily.

We also need the concept of *Kronecker product* of matrices, as follows. If A is a $p \times q$ matrix and B is an $r \times s$ matrix then $A \otimes B$ is the $pr \times qs$ matrix with general entry

$$(A \otimes B)_{r(i_1-1)+i_2,s(j_1-1)+j_2} = a_{i_1 j_1} b_{i_2 j_2}$$

where $A = (a_{ij})$ and $B = (b_{ij})$. Equivalently, $A \otimes B$ is a $pr \times qs$ matrix which has $p \times q$ blocks of size $r \times s$:

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ a_{21}B & a_{22}B & \dots & a_{2n}B \\ \multicolumn{4}{c}{\dots\dots\dots} \\ a_{n1}B & a_{n2}B & \dots & a_{nn}B \end{pmatrix}.$$

For example we will use $\hat{\mathbf{T}}(t) = I \otimes \mathbf{T}(t)$ where I is the $n \times n$ identity matrix. Then $\hat{\mathbf{T}}(t)$ is the $nm \times n$ matrix

$$\hat{\mathbf{T}}(t) = \begin{pmatrix} \mathbf{T}(t) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{T}(t) & \dots & \mathbf{0} \\ \multicolumn{4}{c}{\dots\dots\dots} \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{T}(t) \end{pmatrix},$$

where every entry is an $m \times 1$ column vector.

Note $\hat{\mathbf{T}}(t)' = I \otimes \mathbf{T}(t)^{\top}$. This is a special case of the general rule that if $\mathbf{M} = A \otimes B$ then $\mathbf{M}' = A \otimes B^{\top}$. By contrast, $(A \otimes B)^{\top} = A^{\top} \otimes B^{\top}$. If A is symmetric then $(A \otimes B)^{\top} = (A \otimes B)'$.

Also, let $\mathbf{I} = I \otimes \mathbf{e}$ where $\mathbf{e} = (1\,0\,0\,\dots\,0)^{\top}$. That is,

$$\mathbf{I} = \begin{pmatrix} \mathbf{e} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{e} & \dots & \mathbf{0} \\ \multicolumn{4}{c}{\dots\dots\dots} \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{e} \end{pmatrix}.$$

Then $I = \hat{\mathbf{T}}(t)'\mathbf{I}$ since $T_0^*(t) = 1$.

Suppose—that is to say, let us make the approximation—that each entry of $\Phi^{(p)}(t)$ is expanded in m shifted Chebyshev polynomials. That is, suppose

$$\boxed{\Phi^{(p)}(t) = \hat{\mathbf{T}}(t)'\mathbf{\Phi}}$$

for $\mathbf{\Phi}$ an $nm \times n$ constant matrix. In particular,

$$\mathbf{\Phi} = \begin{pmatrix} \mathbf{f}_{11} & \mathbf{f}_{12} & \dots & \mathbf{f}_{1n} \\ \mathbf{f}_{21} & \mathbf{f}_{22} & \dots & \mathbf{f}_{2n} \\ \multicolumn{4}{c}{\dots\dots\dots} \\ \mathbf{f}_{n1} & \mathbf{f}_{n2} & \dots & \mathbf{f}_{nn} \end{pmatrix},$$

where each $\mathbf{f}_{ij}$ is an $m \times 1$ column vector of constants. The (i, j) entry of $\Phi^{(p)}(t)$ is approximated by $\mathbf{T}(t)^{\top}\mathbf{f}_{ij}$.

Similarly, we expand $A(t)$ and $B(t)$ in the delay–differential equation (1) as

$$\boxed{A(t) = \hat{\mathbf{T}}(t)'\mathbf{A}, \qquad B(t) = \hat{\mathbf{T}}(t)'\mathbf{B}.}$$

Generally, expanding an $n \times n$ matrix function $M(t)$ looks like

$$M(t) = \hat{\mathbf{T}}(t)'\mathbf{M} = \mathbf{M}'\hat{\mathbf{T}}(t).$$

Note that $\hat{\mathbf{T}}(t)$ almost commutes with $\mathbf{M}$ in this sense because of the special form of $\hat{\mathbf{T}}(t)$. Thus if $M(t) = \hat{\mathbf{T}}(t)'\mathbf{M}$ then $M(t)^\top = \hat{\mathbf{T}}(t)'(\mathbf{M}')^\top$. It will eventually be convenient to have notation for this which is $\mathbf{M}^\circ = (\mathbf{M}')^\top = (\mathbf{M}^\top)'$. The point is that $\mathbf{M}^\circ$ are the coefficients for $M(t)^\top$.

We will integrate matrix–valued functions which have been expanded in shifted Chebyshev polynomials as follows. See [A] for the proof of the following lemma.

Lemma 7 (the integration operational matrix $\mathbf{G}$). *If $M(t)$ is a continuous $n \times n$ matrix–valued function of $t \in [0,1]$, and if $M(t) = \hat{\mathbf{T}}(t)'\mathbf{M}$ then*

$$\int_0^t M(s)\,ds = \int_0^t \hat{\mathbf{T}}(s)'\mathbf{M}\,ds = \hat{\mathbf{T}}(t)'\mathbf{G}'\mathbf{M} + \epsilon_{\mathbf{G}}(t)$$

where $\mathbf{G} = I \otimes G$ is an $nm \times nm$ matrix and G is the $m \times m$ matrix

$$G = \begin{pmatrix}
\frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & \dots & 0 \\
-\frac{1}{8} & 0 & \frac{1}{8} & 0 & 0 & \dots & 0 \\
-\frac{1}{6} & -\frac{1}{4} & 0 & \frac{1}{12} & 0 & \dots & 0 \\
\frac{1}{16} & 0 & -\frac{1}{8} & 0 & \frac{1}{16} & \dots & 0 \\
-\frac{1}{30} & 0 & 0 & -\frac{1}{12} & 0 & \dots & 0 \\
& & \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & & & & \\
& & \dots\dots\dots\dots\dots\dots\dots\dots & & & & \frac{1}{4(m-1)} \\
\frac{(-1)^m}{2m(m-2)} & 0 & 0 & 0 & \dots & \frac{-1}{4(m-2)} & 0
\end{pmatrix}.$$

See [A] *for an entrywise formula.*

In the scalar case $n = 1$, that is, $M(t) = f(t)$, the error term $\epsilon_G(t)$ is exactly $\frac{a_{m-1}}{4m}T_m^*(t)$ if $f(t) = \sum_{r=0}^{m-1} a_r T_r^*(t)$. In general, the error term is a matrix with entries

$$\epsilon_{\mathbf{G}}(t)_{ij} = \frac{(\mathbf{m}_{ij})_{m-1}}{4m}T_m^*(t)$$

if $\mathbf{M} = (\mathbf{m}_{ij})$ and $\mathbf{m}_{ij}$ is an $m \times 1$ column vector with entries $(\mathbf{m}_{ij})_k$ for $k = 0, \dots, m-1$. Thus the error is always $o(\frac{1}{m})$ as $m \to \infty$.

Using the "product operational matrix", on the other hand, involves a more significant error. First, the product of two shifted Chebyshev polynomials is

$$T_r^*(t)T_k^*(t) = \frac{1}{2}\left(T_{r+k}^*(t) + T_{|r-k|}^*(t)\right),$$

which follows from $T_r(x) = \cos(r\theta)$ and the trigonometric identity

$$\cos(r\theta)\cos(k\theta) = \frac{1}{2}\left(\cos((r+k)\theta) + \cos(|r-k|\theta)\right).$$

Suppose $f(t) = \mathbf{T}(t)^\top \mathbf{a} = \sum_{r=0}^{m-1} a_r T_r^*(t)$ and $g(t) = \mathbf{T}(t)^\top \mathbf{b} = \sum_{r=0}^{m-1} b_r T_r^*(t)$. Then

$$(20) \qquad f(t)g(t) = \frac{1}{2} \sum_{r,k=0}^{m-1} a_r b_k T_{r+k}^*(t) + \frac{1}{2} \sum_{r,k=0}^{m-1} a_r b_k T_{|r-k|}^*(t).$$

We will rearrange these sums into the form

$$f(t)g(t) = \sum_{l=0}^{2m-2} T_l^*(t) \sum_{k=0}^{m-1} (\mathbf{P}_f)_{lk} b_k,$$

and the expression for $(\mathbf{P}_f)_{lk}$ will be the exact product operational matrix. In particular, (20) can be written as four sums:

$$f(t)g(t) = \sum_{l=0}^{m-1} T_l^*(t) \sum_{k=0}^{l} \frac{1}{2} a_{l-k} b_k \qquad\qquad [1;\ l = r + k]$$

$$+ \sum_{l=0}^{m-1} T_l^*(t) \sum_{k=l}^{m-1} \frac{1}{2} a_{k-l} b_k \qquad\qquad [2;\ l = k - r]$$

$$+ \sum_{l=1}^{m-1} T_l^*(t) \sum_{k=0}^{m-l-1} \frac{1}{2} a_{k+l} b_k \qquad\qquad [3;\ l = r - k]$$

$$+ \sum_{l=m}^{2m-2} T_l^*(t) \sum_{k=l-m+1}^{m-1} \frac{1}{2} a_{l-k} b_k \qquad\qquad [4;\ l = r + k].$$

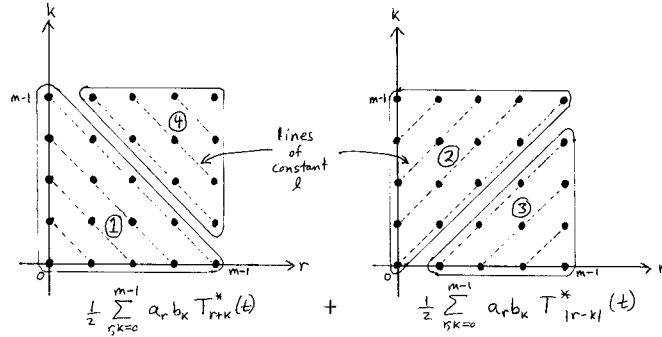This decomposition corresponds to the groupings below:

FIGURE 1. Regroup (20) this way.

Thus

$$(21) \qquad\qquad\qquad f(t)g(t) = \tilde{\mathbf{T}}(t)^\top \mathbf{P}_f \mathbf{b}$$

where

$$\mathbf{P}_f = \begin{pmatrix} a_0 & a_1/2 & a_2/2 & a_3/2 & \ldots & a_{m-1}/2 \\ a_1 & a_0 + a_2/2 & (a_1 + a_3)/2 & (a_2 + a_4)/2 & \ldots & a_{m-2}/2 \\ a_2 & (a_1 + a_3)/2 & a_0 + a_4/2 & (a_1 + a_5)/2 & \ldots & a_{m-3}/2 \\ \multicolumn{6}{c}{\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots} \\ a_{m-1} & a_{m-2}/2 & a_{m-3}/2 & a_{m-4}/2 & \ldots & a_0 \\ 0 & a_{m-1}/2 & a_{m-2}/2 & a_{m-3}/2 & \ldots & a_1/2 \\ 0 & 0 & a_{m-1}/2 & a_{m-2}/2 & \ldots & a_2/2 \\ 0 & 0 & 0 & a_{m-1}/2 & \ldots & a_3/2 \\ \multicolumn{6}{c}{\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots} \\ 0 & 0 & 0 & 0 & \ldots & a_{m-1}/2 \end{pmatrix}$$

and $\tilde{\mathbf{T}}(t)$ is a column vector with the first $2m-1$ shifted Chebyshev polynomials $T_0^*(t)$ through $T_{2m-2}^*(t)$. The matrix $\mathbf{P}_f$ has $2m - 1$ rows and m columns, which follows because the product of two $m - 1$ degree polynomials is a $2m - 2$ degree polynomial.

As with the integration operational matrix $\mathbf{G}$ above, computation will be done with square matrices, and we *write*

$$(22) \qquad \text{``} f(t)g(t) = \mathbf{T}(t)^\top \mathbf{Q}_f \mathbf{b} \text{''}$$

when using the square $m \times m$ matrix

$$(23) \qquad \mathbf{Q}_f = \begin{pmatrix} a_0 & a_1/2 & a_2/2 & a_3/2 & \ldots & a_{m-1}/2 \\ a_1 & a_0 + a_2/2 & (a_1 + a_3)/2 & (a_2 + a_4)/2 & \ldots & a_{m-2}/2 \\ a_2 & (a_1 + a_3)/2 & a_0 + a_4/2 & (a_1 + a_5)/2 & \ldots & a_{m-3}/2 \\ \multicolumn{6}{c}{\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots} \\ a_{m-1} & a_{m-2}/2 & a_{m-3}/2 & a_{m-4}/2 & \ldots & a_0 \end{pmatrix}$$

which is the first m rows of $\mathbf{P}_f$.

This approximation involves dropping the last $m - 1$ rows of $\mathbf{P}_f$ or equivalently dropping $\frac{m(m-1)}{2}$ terms from the product (21). The truncated terms have absolute values at most

(24)
$$\frac{1}{2} \left\{ |b_1||a_{m-1}| + |b_2|(|a_{m-1}| + |a_{m-2}|) + \cdots + |b_{m-1}|(|a_{m-1}| + |a_{m-2}| + \cdots + |a_1|) \right\}.$$

This is a sum of products $|a_r b_k|$ where $r + k \geq m$. Each product is small if the expansions of $f(t), g(t)$ are good in the sense of having rapidly decaying coefficients.

How fast do the (shifted) Chebyshev coefficients a_j decay as $j \to \infty$? There are likely many useful answers to this question, but one of the quickest is to first note that if $f(t) = \sum_{j=0}^{\infty} a_j T_j^*(t)$ then a_j is given by an integral which can be regarded as a Fourier coefficient:

$$a_j = \sqrt{\frac{2}{\pi}} \int_0^1 \frac{f(t) T_j^*(t)}{\sqrt{t - t^2}} \, dt \quad \overset{(x=2t-1,\, \cos\theta = x)}{=} \quad \sqrt{\frac{2}{\pi}} \int_0^\pi f((\cos\theta + 1)/2) \cos(j\theta) \, d\theta.$$

That is, if $F(\theta) = f((\cos\theta+1)/2)$ then the Fourier coefficients of F are the Chebyshev coefficients of f.

Next, integration by parts is the key to a well–known argument that says if F is d times differentiable then the Fourier coefficients c_j of F decay at the rate $O(j^{-d})$. (Note f is d times continuously differentiable if and only if f is d times continuously differentiable, so we only address the argument to F.) Integrating by parts once gives

$$\int_0^\pi F(\theta)\cos(j\theta)\,d\theta = -\frac{1}{j}\int_0^\pi F'(\theta)\sin(j\theta)\,d\theta,$$

and integrating by parts (an even number) d times gives

$$\int_0^\pi F(\theta)\cos(j\theta)\,d\theta = \frac{1}{j^d}\int_0^\pi F^{(d)}(\theta)\cos(j\theta)\,d\theta.$$

If the derivative $F^{(d)}$ is continuous then we know (by the Riemann–Lebesgue lemma, specifically) that the last integral goes to zero as $j \to \infty$ and thus $a_j \to \infty$ faster than the rate j^{-d}. That is, $a_j = o(j^{-d})$.

If both f and g have d continuous derivatives then the above allows us to estimate the lost terms (24) in using $\mathbf{Q}_f$ instead of $\mathbf{P}_f$. That is, we now can say each product $a_r b_k$ has maximum size $c_f c_g r^{-d} k^{-d}$. On the other hand $rk \geq (m-k)k$ in this context, and it is easy to see that $(m-k)k \geq m-1$. Thus we drop $\frac{m(m-1)}{2}$ terms each of which can be estimated

$$\frac{1}{2}|a_r b_k| \leq \frac{c_f c_g}{2} r^{-d} k^{-d} \leq \frac{c_f c_g}{2}(m-1)^{-d}$$

which gives a total error

$$\epsilon_Q \leq \frac{m(m-1)}{2} \cdot \frac{c_f c_g}{2}(m-1)^{-d} \leq Cm^{-d+2}.$$

We have proven:

Proposition 2. *If f, g are d times differentiable then the order of the total error in using $\mathbf{Q_a}$ instead of the exact $\mathbf{P_a}$ is at most $O(m^{-d+2})$, which goes to zero as m increases as long as $d > 2$.*

In terms of notation, we will multiply *matrix–valued* functions which have been expanded in shifted Chebyshev polynomials as follows:

Lemma 8 (the product operational matrix $\mathbf{Q}$). *If $R(t), S(t)$ are two continuous $n \times n$ matrix–valued functions of $t \in [0,1]$, and if $R(t) = \hat{\mathbf{T}}(t)'\mathbf{R}$, $S(t) = \hat{\mathbf{T}}(t)'\mathbf{S}$, and if the (i,j) entry of $R(t)$ is expanded as $r_{ij}(t) = \mathbf{T}(t)^\top \mathbf{r}_{ij}$, then*

$$R(t)S(t) = \hat{\mathbf{T}}(t)'\mathbf{R}\hat{\mathbf{T}}(t)'\mathbf{S} = \mathbf{R}'\hat{\mathbf{T}}(t)\hat{\mathbf{T}}(t)'\mathbf{S} = \hat{\mathbf{T}}(t)'\mathbf{Q}_R\mathbf{S} + \epsilon_{\mathbf{Q}}(t)$$

where $\mathbf{Q}_R$ is an $nm \times nm$ matrix whose (i,j)th $m \times m$ submatrix is $\mathbf{Q}_{r_{ij}}$ in (23). Note $R(t)I = \hat{\mathbf{T}}(t)'\mathbf{Q}_R I = \hat{\mathbf{T}}(t)'\mathbf{R}$. The error term $\epsilon_{\mathbf{Q}}(t)$ is estimated by the proposition above.

As examples, $(G_A I)(t) = \int_0^t A(s)\, ds \approx \hat{\mathbf{T}}(t)' \mathbf{G}' \mathbf{A}$ and

$$(G_A(G_A I))(t) = \int_0^t A(s_1) \int_0^{s_1} A(s_0)\, ds_0\, ds_1 \approx \hat{\mathbf{T}}(t)' \mathbf{G}' \mathbf{Q}_A \mathbf{G}' \mathbf{A}$$

if $A(t) = \hat{\mathbf{T}}(t)' \mathbf{A}$. Step–by–step, and pretending the above approximations are exact, we have:

$$\int_0^t A(s_1) \int_0^{s_1} A(s_0)\, ds_0\, ds_1 = \int_0^t A(s_1) \int_0^{s_1} \hat{\mathbf{T}}(s_0)' \mathbf{A}\, ds_0\, ds_1$$

$$= \int_0^t A(s_1) \left[\hat{\mathbf{T}}(s_1)' \mathbf{G}' \mathbf{A} \right] ds_1 = \int_0^t \hat{\mathbf{T}}(s_1)' \mathbf{A} \hat{\mathbf{T}}(s_1)' \mathbf{G}' \mathbf{A}\, ds_1$$

$$\stackrel{\text{lemma } 8}{=} \int_0^t \hat{\mathbf{T}}(s_1)' \mathbf{Q}_A \mathbf{G}' \mathbf{A}\, ds_1 = \hat{\mathbf{T}}(t)' \mathbf{G}' \mathbf{Q}_A \mathbf{G}' \mathbf{A}.$$

Thus from (11) $\Phi^{(2)}(t) = \hat{\mathbf{T}}(t)'\left(\mathbf{I} + \mathbf{G}'(\mathbf{A} + \mathbf{Q}_A \mathbf{G}' \mathbf{A})\right)$, as an example.

In general, if

$$\boxed{\mathbf{W}_A = \mathbf{Q}_A \mathbf{G}'}$$

then

$$\Phi^{(p)}(t) = \hat{\mathbf{T}}(t)'\left(\mathbf{I} + \mathbf{G}'(\mathbf{A} + \mathbf{W}_A(\mathbf{A} + \cdots + \mathbf{W}_A(\mathbf{A} + \mathbf{W}_A \mathbf{A})\ldots)))\right),$$

that is,

(25) $$\boxed{\boldsymbol{\Phi} = \mathbf{I} + \mathbf{G}'(\mathbf{A} + \mathbf{W}_A(\mathbf{A} + \cdots + \mathbf{W}_A(\mathbf{A} + \mathbf{W}_A \mathbf{A})\ldots)).}$$

This is the analogue of (32) in [SB], though the expansion there is written as a sum of powers of (essentially) $\mathbf{W}_A$. Note (25) represents only an approximation to the best Chebyshev coefficients of $\Phi^{(p)}(t)$ because each use of $\mathbf{G}$ and $\mathbf{Q}_A$ involves truncation of an expansion. Equation (25) is written out in the manner one would want to compute—think Hörner's method—with only $p + 1$ matrix multiplications.

Recall that $\Psi^{(p)}(t) = \hat{\mathbf{T}}(t)' \boldsymbol{\Psi}$ is the Picard approximation defined by (12). From the definition of $\mathbf{A}^\circ$, $A(t)^\top = \hat{\mathbf{T}}(t)' \mathbf{A}^\circ$. Thus $\mathbf{Q}_{A^\top}$ satisfies $A(t)^\top S(t) = \hat{\mathbf{T}}(t)' \mathbf{A}^\circ \hat{\mathbf{T}}(t)' \mathbf{S} = \hat{\mathbf{T}}(t)' \mathbf{Q}_{A^\top} \mathbf{S}$ if $S(t) = \hat{\mathbf{T}}(t)' \mathbf{S}$ is any $n \times n$ matrix function of t. Let $\mathbf{W}_{A^\top} = \mathbf{Q}_{A^\top} \mathbf{G}'$. Then

(26) $$\boxed{\boldsymbol{\Psi} = \mathbf{I} - \mathbf{G}'(\mathbf{A}^\circ - \mathbf{W}_{A^\top}(\mathbf{A}^\circ - \cdots - \mathbf{W}_{A^\top}(\mathbf{A}^\circ - \mathbf{W}_{A^\top} \mathbf{A}^\circ)\ldots)).}$$

Expressions (25) and (26) are the approximations to Φ and Ψ used in the computation of U in the following section. Compare to [SW, BS] where slightly different arrangements of symbols are used in studying stability of periodic linear ODEs.

The above formulas are implemented in a *Mathematica* program by V. Averina [A].

5. APPROXIMATION OF U

In the context of equation (7), if the functions x, Ux on $[0,1]$ are expanded in some finite "basis", for instance m shifted Chebyshev polynomials, then U can be approximated by a matrix, as we will do in this section. We construct this approximation in a manner allowing U to depend on the parameters which appear in equation (1).

The stability question is whether the eigenvalues of U are inside the unit circle or not. For particular parameter values one may numerically approximate the eigenvalues or pseudospectra ([TB]) of U. For a symbolic parameter–dependent computation of U, various polynomial stability criteria (Routh–Hurwitz via a Cayley transform, Schur–Cohn, etc.; see [A]) can be applied to determine stability. Because these criteria involve the calculation of determinants, which is factorial in the size of the matrix, this stability criterion stage is seen to be the most computationally difficult if one wants symbolic results [A].

In any case the result of the previous section will be matrices $\mathbf{\Phi}$, $\mathbf{\Psi}$ of coefficients for the fundamental solutions $\Phi(t)$, $\Psi(t)$ to (2), (3) respectively. These approximations are based on p Picard iterations and expansion in the first m shifted Chebyshev polynomials.

Recall that $V = C[0,1]$ is the space on which U acts. In particular the initial function for equation (1) comes from this space. We need to "approximate the space." We again choose shifted Chebyshev polynomials for the job, but there is no automatic reason for the choice. In fact, the approximation of U by choosing an l–dimensional space V_l to approximate V is a more–or–less separate issue from choosing m, the number of shifted Chebyshev polynomials used to approximate the fundamental solutions.

If we suppose $l \le m$ then we can take $\mathbf{\Phi}$, $\mathbf{\Psi}$ resulting from part II and "truncate them" to have l coefficients. As mentioned in section 3, this involves a quite small error because we are using Chebyshev polynomials. Computationally, since

$$\mathbf{\Phi} = \begin{pmatrix} \mathbf{f}_{11} & \dots & \mathbf{f}_{1n} \\ & \dots & \\ \mathbf{f}_{n1} & \dots & \mathbf{f}_{nn} \end{pmatrix}$$

where each $\mathbf{f}_{ij}$ is an $m \times 1$ column vector of constants, we can truncate $\mathbf{f}_{ij}$ to a column vector of its first l entries.

Assumption 5. Let $l \le m$. We assume from now on that the dimension m has been replaced by l. Thus $\Phi(t) = \hat{\mathbf{T}}(t)'\mathbf{\Phi}$, $\Psi(t) = \hat{\mathbf{T}}(t)'\mathbf{\Psi}$, and $B(t) = \hat{\mathbf{T}}(t)'\mathbf{B}$ where

$$\hat{\mathbf{T}}(t), \mathbf{\Phi}, \mathbf{\Psi}, \mathbf{B} \quad \text{are} \quad nl \times n.$$

Recalling (7), we expand $x(t)$ and $(Ux)(t)$ in shifted Chebyshev polynomials:

$$x(t) = \hat{\mathbf{T}}(t)'\mathbf{v},$$

$$(Ux)(t) = \hat{\mathbf{T}}(t)'\mathbf{w}.$$

Then (7) becomes

$$
\begin{aligned}
\hat{\mathbf{T}}(t)'\mathbf{w} \;&=\; \hat{\mathbf{T}}(t)'\boldsymbol{\Phi}\left\{\hat{\mathbf{T}}(1)'\mathbf{v} + \int_0^t (\hat{\mathbf{T}}(s)'\boldsymbol{\Psi})^\top \hat{\mathbf{T}}(s)'\mathbf{B}\hat{\mathbf{T}}(s)'\mathbf{v}\,ds\right\} \\
&=\; \hat{\mathbf{T}}(t)'\boldsymbol{\Phi}\left\{\hat{\mathbf{T}}(1)'\mathbf{v} + \int_0^t (\boldsymbol{\Psi}^\circ)'\hat{\mathbf{T}}(s)\hat{\mathbf{T}}(s)'\mathbf{B}\hat{\mathbf{T}}(s)'\mathbf{v}\,ds\right\} \\
&\stackrel{\text{lemma 8}}{=}\; \hat{\mathbf{T}}(t)'\boldsymbol{\Phi}\left\{\hat{\mathbf{T}}(1)'\mathbf{v} + \int_0^t \hat{\mathbf{T}}(s)'\mathbf{Q}_{\Psi^\top}\mathbf{B}\hat{\mathbf{T}}(s)'\mathbf{v}\,ds\right\}.
\end{aligned}
$$

Let $\mathbf{Z} = \mathbf{Q}_{\Psi^\top}\mathbf{B}$, and let $Z(t) = \hat{\mathbf{T}}(t)'\mathbf{Z}$. Then by lemma 8 again and lemma 7, the integral above becomes

$$
\int_0^t \hat{\mathbf{T}}(s)'\mathbf{Z}\hat{\mathbf{T}}(s)'\mathbf{v}\,ds = \int_0^t \hat{\mathbf{T}}(s)'\mathbf{Q}_Z\mathbf{v}\,ds = \hat{\mathbf{T}}(t)'\mathbf{G}'\mathbf{Q}_Z\mathbf{v}.
$$

We use lemma 8 yet again:

$$
\hat{\mathbf{T}}(t)'\mathbf{w} = \hat{\mathbf{T}}(t)'\boldsymbol{\Phi}\left\{\hat{\mathbf{T}}(1)'\mathbf{v} + \hat{\mathbf{T}}(t)'\mathbf{G}'\mathbf{Q}_Z\mathbf{v}\right\} = \hat{\mathbf{T}}(t)'\left\{\boldsymbol{\Phi}\hat{\mathbf{T}}(1)'\mathbf{v} + \mathbf{Q}_\Phi\mathbf{G}'\mathbf{Q}_Z\mathbf{v}\right\}.
$$

Thus (7) can be written as a linear operation $\mathbf{v} \to \mathbf{w}$:

$$
\mathbf{w} = \left[\boldsymbol{\Phi}\hat{\mathbf{T}}(1)' + \mathbf{Q}_\Phi\mathbf{G}'\mathbf{Q}_Z\right]\mathbf{v}.
$$

Finally, we have an approximation to U, which was a linear operator from $C[0,1]$ to $C[0,1]$:

(27)
$$
\boxed{U \approx \boldsymbol{\Phi}\hat{\mathbf{T}}(1)' + \mathbf{Q}_\Phi\mathbf{G}'\mathbf{Q}_Z,}
$$

where $Z(t) = \hat{\mathbf{T}}(t)'\mathbf{Z}$ and $\mathbf{Z} = \mathbf{Q}_{\Psi^\top}\mathbf{B}$.

The matrix on the right in (27) is an $nl \times nl$ matrix. Only the $l \to \infty$ limit is exact, of course (assuming $p, m \to \infty$ as well!). We see that the need for $l \le m$ followed from the truncation and product ("$\mathbf{Q}$") operations which went into building (27).

Again, our intent is to use a *parameter–dependent* form of the approximation in (27). So the above has to be done symbolically, as in [SB] for the non–delay case. Thus we compute the characteristic polynomial of our approximation to U and then apply stability criteria. This has been done successfully in [A], which the reader should see next.

References

[A]   V. Averina, *Symbolic stability of delay differential equations*, Master's Thesis U.A.F. 2002.

[BD]   W. Boyce & R. DiPrima, *Elementary Differential Equations*, Wiley. There exists n so that nth edition is recent as you wish.

[BF]   R. L. Burden & D. Faires, *Numerical Analysis*, 6th ed., Brooks/Cole, 1997.

[BM]   E. Butcher, H. Ma, E. Bueler, V. Averina, & Z. Szabo, Stability of time–periodic delay–differential equations via Chebyshev polynomials, *Preprint* (2002).

[BS]   E.A. Butcher & S. C. Sinha, Symbolic computation of local stability and bifurcation surfaces for nonlinear time–periodic dynamical systems, *Nonlinear Dynamics* (1998) **17**, 1–21.

[HL]   J. K. Hale & S. M. Verduyn Lunel, *Introduction to Functional Differential Equations*, Springer, 1993.

[NR]   W. Press, et.al., *Numerical Recipes in C*, 2nd ed., Cambridge, 1992.

[RS]   Michael Reed & Barry Simon *Functional Analysis* Methods of Modern Mathematical Physics I, Revised ed. , Academic Press, 1980.

[R]   T. J. Rivlin, *Chebyshev polynomials*, 2nd ed., Wiley, 1990.

[SB]   S. C. Sinha & E. A. Butcher, Symbolic computation of fundamental solution matrices for linear time–periodic dynamical systems, *J. of Sound and Vibration* (1997) **206**(1), 61–85.

[SW]   S. C. Sinha & D.-H. Wu, An efficient computational scheme for the analysis of periodic systems, *J. of Sound and Vibration* (1991) **151**(1), 91–117.

[TB]   L. N. Trefethen & D. Bau, *Numerical Linear Algebra*, SIAM, 1997.

Appendix A: Some proofs

Proof of lemma 1. In introductory ODE books like [BD], the first claim would be stated as "if the Wronskian is not zero at one time then it is never zero." In any case, we prove it by contradiction. Suppose that for some time $\bar{t} \neq 0$, $\Phi(\bar{t})$ is not invertible, so that there exists some *nonzero* $v \in \mathbb{R}^n$ with $\Phi(\bar{t})v = 0$. We are supposing that a solution $u(t) = \Phi(t)v$ exists which is nonzero at $t = 0$ but is zero at $t = \bar{t}$.

Assume for convenience that $\bar{t} > 0$—the $\bar{t} < 0$ case is similar. Let $w(s) = \Phi(\bar{t} - s)v$ for $s \geq 0$, i.e. solve backwards from $t = \bar{t}$. Let $\tilde{A}(s) = -A(\bar{t} - s)$. Then

$$\frac{d}{ds}w(s) = -A(\bar{t} - s)\Phi(\bar{t} - s)v = \tilde{A}(s)w(s)$$

and $w(0) = 0$. This is a linear ODE initial value problem for w. Note the solution is nonzero when $s = \bar{t}$, since $w(\bar{t}) = \Phi(0)v = v$, even though the initial value is zero. We have a contradiction to the uniqueness theorem for linear systems (as in [BD], for instance). Thus $\Phi(\bar{t})$ is invertible.

Next, note that

$$0 = \frac{d}{dt}(I) = \frac{d}{dt}\left(\Phi(t)\Phi(t)^{-1}\right) = A(t)\Phi(t)\Phi(t)^{-1} + \Phi(t)\frac{d}{dt}\left(\Phi^{-1}(t)\right)$$

so $\frac{d}{dt}\left(\Phi^{-1}(t)\right) = -\Phi^{-1}A(t)$. Let $\Psi^\top = \Phi^{-1}$. Then $\frac{d}{dt}\left(\Psi^\top\right) = -\Psi^\top A(t)$, i.e. $\frac{d}{dt}\Psi = -A(t)^\top\Psi$, after taking transposes. $\qquad\square$

Proof of lemma 3. Let $v \in \mathbb{R}^n$. Assume for simplicity that $r < s < t$. (The lemma holds for any r, s, t!) For this proof, let $\dot{}$ denote $\frac{d}{d\theta}$ and consider the initial value problems:

$$\dot{\nu} = A(\theta)\nu, \qquad \nu(r) = v,$$
$$\dot{\xi} = A(\theta)\xi, \qquad \xi(s) = \nu(s),$$
$$(\dagger) \qquad \dot{\zeta} = A(\theta)\zeta, \qquad \zeta(r) = v.$$

Note that $\nu(s) = \Omega_r^s v$, $\xi(t) = \Omega_s^t(\nu(s)) = \Omega_s^t\Omega_r^s v$, and $\zeta(t) = \Omega_r^t v$. The two functions

$$\zeta(\theta) \quad \text{and} \quad \tilde{\zeta}(\theta) = \begin{cases} \nu(\theta) & \text{if } r \leq \theta \leq s, \\ \xi(\theta) & \text{if } s \leq \theta \leq t \end{cases}$$

on the interval $r \leq \theta \leq t$ are both solutions to the initial value problem $(\dagger)$. Thus they are have the same value at $\theta = t$, that is, $\Omega_s^t\Omega_r^s v = \Omega_r^t v$. $\qquad\square$

Proof of lemma 4. In terms of the definition of Ω_s^t, consider the solutions to the initial value problems $\frac{d}{d\theta}\nu = A(\theta)\nu(\theta)$, $\nu(s) = v$ and $\frac{d}{d\theta'}\xi = A(\theta')\xi(\theta')$, $\xi(s - T) = v$. Let $\theta = \theta' + T$ and note $A(\theta' + T) = A(\theta')$. In θ', the equation for ν is $\frac{d}{d\theta'}\nu = A(\theta')\nu$ with $\nu(\theta' = s - T) = v$, so $\nu(\theta = t) = \nu(\theta' = t - T) = \xi(\theta' = t - T)$. Thus $\Omega_s^t v = \nu(t) = \xi(t - T) = \Omega_{s-T}^{t-T} v$. $\qquad\square$

Next we fill in some proofs having to do with discrete orthogonality relation satisfied by the shifted Chebyshev polynomials $T_j^*(t)$.

First note that since $T_j(x) = \cos(j \arccos x)$,

$$T_j^*(t_k) = \cos\left(\frac{j(2k-1)\pi}{2m}\right)$$

if $t_k = \frac{1}{2}(1 + \cos((2k-1)\pi/(2m)))$ is a zero of $T_m^*(t)$. Also recall the identity $\cos a \cos b = (1/2)\cos(a+b) + (1/2)\cos(a-b)$.

Proposition.

$$\sum_{k=1}^{m} T_l^*(t_k)T_j^*(t_k) = \frac{1}{2}\sum_{k=1}^{m} \cos((l+j)(2k-1)\pi/(2m)) + \cos((l-j)(2k-1)\pi/(2m))$$

$$= \begin{cases} 0, & l \neq j \\ m, & l = j = 0, \\ m/2, & l = j = 1, 2, \ldots, m-1. \end{cases}$$

Proof. We consider the sum $S(p) = \sum_{k=1}^{m} \cos(p(2k-1)\pi/(2m))$ for $p = l+j$ and $p = l - j$.

Note that $S(0) = m$.

We use the identity $\cos a = (1/2)(e^{ia} + e^{-ia})$, and we sum the finite geometric series

$$\sum_{k=1}^{m} e^{ip(2k-1)\pi/(2m)} = e^{ip\pi/(2m)}\frac{1 - \left(e^{ip\pi/m}\right)^m}{1 - e^{ip\pi/m}},$$

assuming $-2m < p < 2m$ and $p \neq 0$. Thus

$$S(p) = \frac{e^{ip\pi/(2m)}\left(1 - e^{ip\pi}\right)}{2(1 - e^{ip\pi/m})} + \frac{e^{-ip\pi/(2m)}\left(1 - e^{-ip\pi}\right)}{2(1 - e^{-ip\pi/m})} = \frac{\sin(p\pi)}{\sin(p\pi/(2m)))} = 0,$$

that is, if $p \neq 0$.

The orthogonality relation follows by going through the cases. $\square$

Proposition. *If c_j ($j = 0, 1, \ldots, m-1$) are given by (17) and $P(t) = \frac{c_0}{2} + \sum_{j=1}^{m-1} c_j T_j^*(t)$ then $P(t_l) = f(t_l)$ where t_l ($l = 1, \ldots, m$) are the zeros of $T_m^*(t)$.*

Proof. Recalling $T_j^*(t_k) = \cos(j(2k-1)\pi/(2m))$ and $\cos a \cos b = (1/2)\cos(a+b) + (1/2)\cos(a-b)$ we write

$$P(t_l) = \frac{1}{m}\sum_{k=1}^{m} f(t_k)\left[1 + 2\sum_{j=1}^{m-1}\cos(j(2k-1)\pi/(2m))\cos(j(2l-1)\pi/(2m))\right]$$

$$= \frac{1}{m}\sum_{k=1}^{m} f(t_k)\left[1 + \sum_{j=1}^{m-1}\cos(j(k+l-1)\pi/m) + \cos(j(k-l)\pi/m)\right].$$

We claim that the term in square brackets above simplifies to $m\delta_{kl}$. In fact, we sum[9] $\Sigma(p) \equiv \sum_{j=1}^{m-1} \cos(jp\pi/m)$ to

$$\Sigma(p) = \frac{1}{2}\frac{\sin(p\pi - p\pi/(2m))}{\sin(p\pi/(2m))} - \frac{1}{2}$$

if $p \neq 0$. Note $\Sigma(0) = m - 1$. But $\sin(p\pi - p\pi/(2m)) = \pm\sin(p\pi/(2m))$ where we choose $+$ if p is an odd integer and $-$ if p is an even integer. Thus $\Sigma(p) = 0$ if p is odd and $\Sigma(p) = -1$ if p is even.

Now note that if k, l are integers then one of $\{k + l - 1, k - l\}$ is odd and one even. The term in square brackets therefore simplifies to zero if $k \neq l$. If $k = l$ it simplifies to m as claimed. $\qquad\square$

Appendix B: Vector and matrix norms

Definition. If $y \in \mathbb{R}^n$ is a **vector** and if $1 \leq r < \infty$ let

$$\|y\|_r \equiv \left(\sum_{i=1}^{n} |y_i|^r\right)^{1/r}.$$

There is a separate case: if $r = \infty$ let $\|y\|_\infty = \max_{1 \leq i \leq n} |y_i|$.

Definition. If $M : \mathbb{R}^n \to \mathbb{R}^n$ is linear, i.e. M is an $n \times n$ **matrix**, and if $1 \leq r \leq \infty$, let

$$\|M\|_r \equiv \max_{y \text{ s.t. } \|y\|_r=1} \|My\|_r.$$

Generally, for instance in the popular $r = 2$ case, it is hard to compute the norm $\|M\|_r$ given the entries of m_{ij} of M. However, the $r = 1$ and $r = \infty$ matrix norm cases are relatively easy. In fact,

$$\|M\|_1 = \max_{1 \leq j \leq n} \|m_{\cdot j}\|_1 = \max_{1 \leq j \leq n} \left(\sum_{i=1}^{n} |m_{ij}|\right)$$

is the maximum column norm (or "maximum of the column sums" if the absolute value is understood) and

$$\|M\|_\infty = \max_{1 \leq i \leq n} \|(m_{i\cdot})^\top\|_1 = \max_{1 \leq i \leq n} \left(\sum_{j=1}^{n} |m_{ij}|\right)$$

is the "maximum of the row sums." The proof is in [TB].

Note that $\|M^\top\|_1 = \|M\|_\infty$ and $\|M^\top\|_\infty = \|M\|_1$. For the purposes of the current paper, this fact is convenient. By looking at both $\|A(t)\|_1$ and $\|A(t)\|_\infty$ we can produce more–or–less complete norm information on both Φ and Ψ.

The important properties of the matrix norm are that

- $\|I\|_r = 1$,

---

[9]Using another finite geometric series argument—see the previous proof.

- $|M_{ij}| \leq \|M\|_r$ for any matrix entry M_{ij}
- $\|MN\|_r \leq \|M\|_r \|N\|_r$, and
- $\|\alpha M + \beta N\|_r \leq |\alpha| \|M\|_r + |\beta| \|N\|_r$, for $\alpha, \beta \in \mathbb{R}$, so
- $\left\| \int_a^b M(t)\, dt \right\|_r \leq \int_a^b \|M(t)\|_r\, dt$.

APPENDIX C: THE $T = N\tau$ CASE.

In this appendix we consider (1) assuming $A(t)$, $B(t)$ are T periodic and that $T = N\tau$ for $N = 1, 2, 3, \ldots$ where τ is the delay period, of course. Certain milling problems have N equal to the number of flutes, for instance.
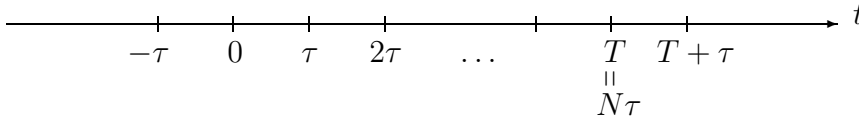
First we define functions $x_k(t)$, $k = 1, \ldots, N$, on the intervals $(k-1)\tau \leq t \leq k\tau$ by equation (4) with $t_0 = (k-1)\tau$:

$$(28) \qquad x_k(t) = \Phi(t)\Psi((k-1)\tau)^\top x_{k-1}((k-1)\tau)$$

$$+ \int_{(k-1)\tau}^t \Phi(t)\Psi(s)^\top B(s) x_{k-1}(s-\tau)\, ds,$$

where $x_0(t) = \varphi(t)$ is the initial function defined on $[-\tau, 0]$.

We want to think of (28) as defining N maps $U_k : C[0,\tau] \to C[0,\tau]$. The stability of (1) will be determined by the eigenvalues of the product

$$U = U_N U_{N-1} \ldots U_2 U_1 : C[0,\tau] \to C[0,\tau].$$

(This assertion follows from showing that for $T = N\tau \leq t \leq (N+1)\tau = T + \tau$,

$$x(t) = \Phi(t-T)\Psi(0)^\top x_N(T) + \int_T^t \Phi(t-T)\Psi(s-T)^\top B(s-T)x(s-\tau)\, ds,$$

so $x(t)$ for $T \leq t \leq T + \tau$ is computed using the same information Φ, Ψ, B as was $x(t)$ for $0 \leq t \leq \tau$. That is, $x_{N+1} = U_1 x_N$. Thus x for any time can by calculated by a product of U's with a final application of at most N of the U_k.)

For actual computation, i.e. Chebyshev approximation, it will be simplest to normalize $\tau = 1$. It will also be easier to "break" $A(t)$, $B(t)$, $\Phi(t)$, $\Psi(t)$ into N pieces each defined on the interval $[0, \tau = 1]$. That is,

$$A_k(r) = A(k - 1 + r), \qquad 0 \leq r \leq 1,$$
$$B_k(r) = B(k - 1 + r), \qquad 0 \leq r \leq 1,$$

$$\Phi_k(r) = \Phi(k - 1 + r), \qquad 0 \le r \le 1,$$
$$\Psi_k(r) = \Psi(k - 1 + r), \qquad 0 \le r \le 1,$$

for $k = 1, 2, \ldots, N$.

The natural way to construct Φ_k, Ψ_k is to recognize that

$$\Phi_1(r) \quad \text{solves} \quad \dot{\Phi}_1(r) = A_1(r)\Phi_1(r), \quad \Phi_1(0) = I, \quad 0 \le r \le 1,$$
$$\Phi_k(r) \quad \text{solves} \quad \dot{\Phi}_k(r) = A_k(r)\Phi_k(r), \quad \Phi_k(0) = \Phi_{k-1}(1), \quad 0 \le r \le 1,$$
$$\Psi_1(r) \quad \text{solves} \quad \dot{\Psi}_1(r) = -A_1(r)^\top \Psi_1(r), \quad \Psi_1(0) = I, \quad 0 \le r \le 1,$$
$$\Psi_k(r) \quad \text{solves} \quad \dot{\Psi}_k(r) = -A_k(r)^\top \Psi_k(r), \quad \Psi_k(0) = \Psi_{k-1}(1), \quad 0 \le r \le 1,$$

for $k = 2, \ldots, N$.

Also let $\tilde{x}_k(r) = x(k - 1 + r)$. In these terms,

$$\tilde{x}_k(r) = \Phi_k(r)\Psi_k(0)^\top \tilde{x}_{k-1}(1) + \int_0^r \Phi_k(r)\Psi_k(\sigma)^\top B_k(\sigma)\tilde{x}_{k-1}(\sigma)\, d\sigma$$

where $t = k - 1 + r$ and $s = k - 1 + \sigma$.

If $y \in C[0, 1]$ we define

$$(29) \qquad (U_k y)(r) = \Phi_k(r)\Psi_k(0)^\top y(1) + \int_0^r \Phi_k(r)\Psi_k(\sigma)^\top B_k(\sigma)y(\sigma)\, d\sigma.$$

Equation (29) is analogous to (7). It defines a linear map $U_k : C[0, 1] \to C[0, 1]$. Again note that we are interested for the purposes of stability in the composition (matrix product) $U = U_N U_{N-1} \ldots U_2 U_1$.

In fact, if $V_l \subset V = C[0, 1]$ is an nl dimensional approximation subspace (for instance, it is the span of the first l shifted Chebyshev polynomials) then $\tilde{U} = P_{V_l} U|_{V_l}$ is an $nl \times nl$ approximation to U. We assert that we can approximately determine the stability of (1) by finding the largest (in magnitude) eigenvalue of $\tilde{U}$.

In these terms the method of steps for (1) with initial condition $x(t) = \varphi(t)$, $-\tau = -1 \le t \le 0$, looks like:

$$
\begin{aligned}
f(t) &= \varphi(t - 1) & &\text{for } t \in [0, 1],\\
x(t) &= (U_1 f)(t) & &\text{for } t \in [0, 1],\\
x(t) &= (U_2 U_1 f)(t - 1) & &\text{for } t \in [1, 2],\\
&\quad \cdots\\
x(t) &= (U_k U_{k-1} \ldots U_2 U_1 f)(t - (k - 1)) & &\text{for } t \in [k - 1, k],\\
&\quad \cdots
\end{aligned}
$$

$$x(t) = (Uf)(t - (N - 1)) \qquad\qquad \text{for } t \in [N - 1, N],$$
$$x(t) = (U_1 U f)(t - (N)) \qquad\qquad \text{for } t \in [N, N + 1],$$
$$x(t) = (U_2 U_1 U f)(t - (N + 1)) \qquad\qquad \text{for } t \in [N + 1, N + 2],$$
$$\cdots$$

APPENDIX D: THE COMPACTNESS OF U.

> Mathematicians are like Frenchmen: whatever you say to them they translate into their own language and forthwith it is something entirely different. *Goethe*

This appendix addresses an abstract issue which nonetheless is a relevant question: Is the infinite dimensional matrix U effectively approximate–able by any finite matrix? Because of the comments in section 1, this reduces to showing U is a compact operator on $C[0, 1]$.

We need three facts about compact (completely–continuous) operators on Banach spaces X:[10]

(i) the sum of any two compact operators is compact;
(ii) any finite rank (i.e. the range is finite–dimensional) operator is compact;
(iii) if A is any bounded operator on X and K is compact on X then both AK and KA are compact.

For this appendix $X = C[0, 1]$ and we assume $T = \tau = 1$ for simplicity.

As an example of a finite rank operator, consider

$$P_1 : y \mapsto y(1)$$

for $y \in C[0, 1]$. Then P_1 has range which is always a constant function, thus it has rank one.

Now, it is well–known that integral operators with nice kernels on finite intervals tend to be compact. In particular, if $k(s)$ is a continuous function of $s \in [0, 1]$ then

$$(Ky)(t) = \int_0^t k(s) y(s) \, ds$$

is compact on $C[0, 1]$. In fact,

$$|(Ky)(t_1) - (Ky)(t_2)| \leq \int_{t_1}^{t_2} |k(s)||y(s)| \, ds \leq \left(\int_{t_1}^{t_2} |k(s)| \, ds \right) \max_{0 \leq s \leq 1} |y(s)|.$$

Since k is continuous, thus bounded, for all $\epsilon > 0$ there exists $\delta > 0$ such that $\int_{t_1}^{t_2} |k(s)| \, ds < \epsilon$ if $|t_1 - t_2| < \delta$. Thus $|(Ky)(t_1) - (Ky)(t_2)| \leq \epsilon \max_{0 \leq s \leq 1} |y(s)|$. If $B_M = \{ y \in C[0, 1] : \max_{0 \leq s \leq 1} |y(s)| \leq M \}$ then the set of all functions $K[B_M]$ are

---

[10]General theory of compact operators can be found in section IV.5 of [RS].

equicontinuous. By Ascoli's theorem (I.28 of [RS]), $K[B_M]$ has compact closure so K is compact.

Now we consider U defined by (7). We see that

$$U = M_\Phi \left(P_1 + K_{\Psi,B} \right)$$

where P_1 is as above and

$$(K_{\Psi,B}x)(t) = \int_0^t \Psi(s)^\top B(s) y(s) \, ds$$

is an integral of the type analyzed above, with $k(s) = \Psi(s)^\top B(s)$. Finally, M_Φ is a multiplication operator

$$M_\Phi : x \mapsto \Phi x$$

for $x \in C[0,1]$.

We assert that M_Φ is bounded because Φ is a bounded function. (One can find an explicit bound on Φ, Ψ as in section 2.) Similarly, $K_{\Psi,B}$ is compact by the above. By the general facts (i), (ii), (iii) above we conclude U is compact.

DEPT. OF MATHEMATICAL SCIENCES, UAF
E-mail address: ffelb@uaf.edu

DEPT. OF MECHANICAL ENGINEERING, UAF
E-mail address: ffeab@uaf.edu