THE SENSITIVITY OF THE MATRIX EXPONENTIAL*

CHARLES VAN LOAN†

Abstract. In this paper we examine how the matrix exponential e^{At} is affected by perturbations in A. Elementary techniques using log norms and the Jordan and Schur factorizations indicate that e^{At} is least sensitive when A is normal. Through the formulation of an exponential condition number, more insight is gained into the complicated connection between the condition of the eigensystem of A and the sensitivity of e^{At}.

1. Introduction. The exponential of an $n \times n$ matrix A is defined by $e^{At} = \sum_{k=0}^{\infty} (At)^k/k!$ where $t \geq 0$. The importance of this matrix function in applied mathematics is derived from the fact that it is the unique solution to the initial value problem $AX(t) = (d/dt)X(t)$, $X(0) = I$. Many methods exist for computing the matrix exponential [8]. A rigorous assessment of these algorithms demands an understanding of the sensitivity of e^{At} because one cannot fault an algorithm for rendering an inaccurate e^{At} if for that particular A, the exponential problem was inherently "ill-conditioned". In this paper we hope to contribute to the understanding of how e^{At} is affected by perturbations in A.

Our basic approach is to investigate upper bounds for $\phi(t)$ where

$$(1.1) \qquad \phi(t) = \frac{\|e^{(A+E)t} - e^{At}\|}{\|e^{At}\|}.$$

When A and E commute, the bounding of $\phi(t)$ is trivial since

$$e^{(A+E)t} - e^{At} = e^{At}(e^{Et} - I) = e^{At}(tE) \sum_{k=0}^{\infty} (Et)^k/(k+1)!$$

and thus

$$(1.2) \qquad AE = EA \Rightarrow \phi(t) \leq \|E\| t\, e^{\|E\|t}.$$

If A and E fail to commute, then $e^{(A+E)t} \neq e^{At} e^{Et}$ and the analysis of $\phi(t)$ becomes considerably harder. It proves convenient to work with the following identity which appears in Bellman [1]:

$$(1.3) \qquad e^{(A+E)t} = e^{At} + \int_0^t e^{A(t-s)} E e^{(A+E)s}\, ds.$$

Manipulation of this equation gives

$$(1.4) \qquad \phi(t) \leq \frac{\|E\|}{\|e^{At}\|} \int_0^t \|e^{A(t-s)}\|\, \|e^{(A+E)s}\|\, ds.$$

To proceed further, we must be able to bound the norm of a matrix exponential. Some of the ways this can be done are described in § 2. These results are then

applied to (1.4) in § 3 giving upper bounds for $\phi(t)$. The bounds indicate that e^{At} is least sensitive when A is normal and thus has a perfectly conditioned eigensystem. However, the results in § 4 do not answer the important question of whether poorly conditioned eigensystems imply sensitivity of the exponential.

In an attempt to more fully characterize e^{At} sensitivity, the Fréchet derivative proves useful. Recall that when it exists, the Fréchet derivative of a map $F(A)$ from $C^{n \times n}$ to $C^{n \times n}$ is another map $D(F(A)): C^{n \times n} \to C^{n \times n}$ satisfying

$$F(A + E) = F(A) + D(F(A))E + o(\|E\|).$$

If we substitute

$$e^{(A+E)s} = e^{As} + \int_0^s e^{A(s-r)} E e^{(A+E)r} \, dr$$

into (1.3) we find

$$e^{(A+E)t} = e^{At} + \int_0^t e^{A(t-s)} E e^{As} + o(\|E\|).$$

Thus, the Fréchet derivative, $D(e^{At})$, of the map $F(A) = e^{At}$ is given by

$$D(e^{At})E = \int_0^t e^{A(t-s)} E e^{As} \, ds.$$

The norm of the *operator* $D(e^{At})$ effectively quantifies the rate of change of e^{At} with respect to A. This precise measure of sensitivity can be incorporated in an "exponential condition number" and is the subject of our discussion in § 4.

We now summarize our notation. Let $\mathbb{C}^{n \times n}$ denote the set of complex $n \times n$ matrices. If $A = (a_{ij}) \in \mathbb{C}^{n \times n}$, then

$$A^* = (\bar{a}_{ji}),$$

$$\lambda(A) = \{\lambda \,|\, \det(A - \lambda I) = 0\},$$

$$\|A\| = \max\{|\lambda| \,|\, \lambda^2 \in \lambda(A^*A)\},$$

$$\kappa(A) = \|A\| \|A^{-1}\| \quad (0 \notin \lambda(A)),$$

$$\alpha(A) = \max\{\operatorname{Re}(\lambda) \,|\, \lambda \in \lambda(A)\}.$$

We have chosen to work with the 2-norm for convenience. Most of the results we present apply with little or no modification when other norms are used.

2. Bounding e^{At}. In this section we summarize and compare various ways in which $\|e^{At}\|$ can be bounded.

(a) *Power series.* By taking norms in $e^{At} = \sum_{k=0}^{\infty} (At)^k / k!$ we trivially obtain

(2.1) $$\|e^{At}\| \leq e^{\|A\|t}.$$

(b) *Log norms.* Dahlquist [2] has shown that if

$$\mu(A) = \max\{\mu \,|\, \mu \in \lambda((A^* + A)/2)\},$$

then

(2.2) $$\|e^{At}\| \leq e^{\mu(A)t}.$$

The scalar $\mu(A)$ is the log norm of A with respect to the 2-norm. A general discussion of log norms may be found in Strom [10]. Because

$$A = YBY^{-1} \Rightarrow e^{At} = Ye^{Bt}Y^{-1}$$

we have the following corollary to (2.2):

(2.3) $A = YBY^{-1} \Rightarrow \|e^{At}\| \leq \kappa(Y) e^{\mu(B)t}.$

(c) *Jordan canonical form.* Recall the Jordan decomposition theorem which states that if $A \in \mathbb{C}^{n \times n}$, then there exists an invertible $X \in \mathbb{C}^{n \times n}$ such that

(2.4) $X^{-1}AX = J_{m_1}(\lambda_1) \oplus \cdots \oplus J_{m_p}(\lambda_p) \equiv J$

where

(2.5) $J_k \equiv J_{m_k}(\lambda_k) = \begin{bmatrix} \lambda_k & 1 & & 0 \\ & \lambda_k & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_k \end{bmatrix} \in \mathbb{C}^{m_k \times m_k}, \qquad k = 1, \cdots, p.$

The matrix X is not unique but we shall always assume that it is chosen such that $\kappa(X)$ is minimized.

It is well known [7] that if the Jordan canonical form (JCF) of A is specified by (2.4) and (2.5) then

(2.6) $e^{At} = X[e^{J_1 t} \oplus \cdots \oplus e^{J_p t}]X^{-1}$

where

(2.7) $e^{J_k t} = e^{\lambda_k t} \begin{bmatrix} 1 & t & t^2/2 & \cdots & t^r/r! \\ & 1 & t & & \vdots \\ & & \ddots & & \vdots \\ & & & & t \\ 0 & & & & 1 \end{bmatrix}, \qquad r = m_k - 1.$

Using the fact that $\|B\| \leq q \max |b_{ij}|$ for $B \in \mathbb{C}^{q \times q}$ it is easy to show from (2.7) that

$$\|e^{J_k t}\| \leq m_k |e^{\lambda_k t}| \max_{0 \leq j \leq m_k - 1} t^j/j!$$

By taking norms in (2.6) and defining $m = \max\{m_1, \cdots, m_p\}$ we obtain

(2.8) $\|e^{At}\| \leq m\kappa(X) e^{\alpha(A)t} \max_{0 \leq r \leq m-1} t^r/r!$

(d) *Schur decomposition bound.* The Schur decomposition states that there exists a unitary $Q \in \mathbb{C}^{n \times n}$ such that

(2.9) $Q^*AQ = D + N$

where

$$D = \text{diag}(\lambda_i),$$

$$N = (n_{ij}), \qquad\qquad\qquad n_{ij} = 0, \quad i \geq j.$$

Notice that $\lambda(A) = \{\lambda_1, \cdots, \lambda_n\}$. If we substitute D for A and N for E in (1.3) we get

$$e^{(D+N)t} = e^{Dt} + \int_0^t e^{D(t-t_1)} N e^{(D+N)t_1} \, dt_1.$$

Using this formula to expand $e^{(D+N)t_1}$ we obtain

$$e^{(D+N)t} = e^{Dt} + \int_0^t e^{D(t-t_1)} N e^{Dt_1} \, dt_1 + \int_0^t \int_0^{t_1} e^{D(t-t_1)} N e^{D(t_1-t_2)} N e^{(D+N)t_2} \, dt_2 \, dt_1.$$

Clearly, a repetition of this process gives

$$e^{(D+N)t} = e^{Dt} + \sum_{k=1}^{n-1} A_k(t) + R_n(t)$$

where

$$A_k(t) = \int_0^t \int_0^{t_1} \cdots \int_0^{t_{k-1}} e^{D(t-t_1)} N e^{D(t_1-t_2)} N \cdots N e^{Dt_k} \, dt_k \cdots dt_1$$

and

$$R_n(t) = \int_0^t \int_0^{t_1} \cdots \int_0^{t_{n-1}} e^{D(t-t_1)} N \cdots N e^{D(t_{n-1}-t_n)} N e^{(D+N)t_n} \, dt_n \cdots dt_1.$$

Now the matrix $[e^{D(t-t_1)}N] \cdots [e^{D(t_{n-1}-t_n)}N]$ is zero because it is the product of n, $n \times n$ *strictly* upper triangular matrices and thus, $R_n(t) = 0$. Hence,

$$(2.10) \quad e^{(D+N)t} = e^{Dt} + \sum_{k=1}^{n-1} \int_0^t \cdots \int_0^{t_{k-1}} e^{D(t-t_1)} N \cdots N e^{Dt_k} \, dt_k \cdots dt_1.$$

By taking norms in this and noting that $\|e^{Ds}\| = e^{\alpha(A)s} (s \geq 0)$ the following result is obtained:

$$(2.11) \qquad\qquad \|e^{At}\| \leq e^{\alpha(A)t} \sum_{k=0}^{n-1} \frac{\|Nt\|^2}{k!}.$$

(e) *Other bounds.* For completeness we mention some other ways that bounds for $\|e^{At}\|$ can be obtained. By using the fact that

$$e^{At} = \lim_{k \to \infty} \left(I - \frac{At}{k} \right)^{-k}$$

Kato [5] has shown that if

$$\beta \geq \alpha(A)$$

and if for all sufficiently large k

$$\|(\gamma I - A)^{-k}\| \leq c|\gamma - \beta|^{-k}, \qquad \mathrm{Re}\,(\gamma) > \beta,$$

then

$$\|e^{At}\| \leq c\, e^{\beta t}.$$

In a sense, this result replaces the problem of bounding $\|e^{At}\|$ with that of bounding powers of the inverse of $(\gamma I - A)$. We will not pursue the analysis of $\|(\gamma I - A)^{-k}\|$ because the bounds one gets are similar to the ones we already have.

Gantmacher [4] has derived some interesting bounds relating to the "matric-ant" of the system $A(t)X(t) = (d/dt)X(t)$. If we specialize his results to the constant coefficient problem we obtain

$$|f_{ij}| \leq \delta_{ij} + (e^{nt\hat{a}} - 1)$$

where $A = (a_{ij}) \in \mathbb{C}^{n \times n}$, $\hat{a} = \max |a_{ij}|$, and $e^{At} = (f_{ij})$. From this it can be shown that

$$\|e^{At}\| \leq 1 + (e^{nt\hat{a}} - 1) = e^{nt\hat{a}}.$$

This result isn't much different from (2.1) because $\hat{a} \leq \|A\| \leq n\hat{a}$.

We conclude this section with an example to illustrate some of the bounds given above. If

$$A = \begin{pmatrix} -1+\delta & 4 \\ 0 & -1-\delta \end{pmatrix}, \quad \delta = 10^{-6},$$

then (2.1), (2.8), (2.11), (2.2), and (2.3) give respectively

(a) $\|e^{At}\| \leq e^{4.25t}$,

(b) $\|e^{At}\| \leq 4 \times 10^{6} \cdot e^{(-1+\delta)t}$,

(c) $\|e^{At}\| \leq (1 + 4t) e^{(-1+\delta)t}$,

(d) $\|e^{At}\| \leq e^{(-1+(4+\delta^2)^{1/2})t}$,

(e) $\|e^{At}\| \leq 4e^{(-1+(.25+\delta^2)^{1/2})t}$ $(Y = \text{diag}(1, 4))$.

Table 1 compares these bounds for selected values of t.

TABLE 1

t	$\|e^{At}\|$	Power (a)	Jordan (b)	Schur (c)	Log Norm (d)	Log Norm (e)
0	1.0E+00	1.0E+00	1.0E+00	1.0E+00	1.0E+00	1.0E+00
5	1.3E-01	1.5E+09	2.7E+04	1.4E-01	1.5E+02	3.3E-01
10	1.8E-03	2.5E+18	1.8E+02	1.9E-03	2.2E+04	2.7E-02
15	1.8E-05	5.7E+27	1.2E+00	1.9E-05	3.3E+06	2.2E-03
20	1.6E-07	9.0E+36	8.2E-03	1.7E-07	4.8E+08	1.8E-04
25	1.3E-09	1.4E+46	5.6E-05	1.4E-09	7.2E+10	1.4E-05
30	1.1E-11	2.2E+54	3.7E-07	1.1E-11	1.1E+13	1.2E-06

We see from Table 1 that some of the upper bounds may fail to decay along with e^{At}. As is well known, the asymptotic behavior of e^{At} depends upon the sign of $\alpha(A)$:

$$\lim_{t \to \infty} e^{At} = 0 \Leftrightarrow \alpha(A) < 0.$$

Hence, the Jordan and Schur bounds (2.8) and (2.11) decay precisely when e^{At} decays while the power series bound (2.1) grows regardless of the sign of $\alpha(A)$. The log norm bound (2.2) may or may not exhibit the proper limiting behavior. This is because it is possible for $\mu(A)$ to be positive even though $\alpha(A)$ is negative as the above example shows. However, it can be shown that if $\alpha(A) < 0$, then it is always possible to choose Y so that $\mu(YAY^{-1}) < 0$. The bound (e) depicts this.

The example also shows the possible advantage of the Schur bound to the Jordan bound when A has an ill-conditioned eigensystem (i.e. $\kappa(X)$ large). However, there are examples where (2.8) is sharper than (2.11). In general, the effectiveness of one bound relative to another depends upon A and t. However, when A is normal, then $\alpha(A) = \mu(A)$ in (2.2), $m = \kappa(X) = 1$ in (2.8), and $N = 0$ in (2.11) thus implying that $\|e^{At}\| \leq e^{\alpha(A)t}$. Since we have $\|e^{At}\| \geq |e^{\lambda t}|$ for all $\lambda \in \lambda(A)$, $\|e^{At}\| \geq e^{\alpha(A)t}$ implying that

$$(2.12) \qquad A^*A = AA^* \Rightarrow \|e^{At}\| = e^{\alpha(A)t}.$$

When A is not normal, it is possible for $\|e^{At}\|$ to grow initially even though $\alpha(A)$ is negative. In this case the factor $m\kappa(X) \max_{0 \leq j \leq m-1} t^j/j!$ in (2.8) and the factor $\sum_{k=0}^{n-1} \|Nt\|^k/k!$ in (2.11) are necessary to accomodate the "hump" in the graph of $\|e^{At}\|$.

Whether or not $\|e^{At}\|$ grows at all depends upon the sign of $\mu(A)$:

$$\sup_{t \geq 0} \|e^{At}\| = 1 \Leftrightarrow \mu(A) \leq 0.$$

This result follows from (2.2) and the fact that $\mu(A)$ is the derivative of $f(t) = \|e^{At}\|$ at $t = 0$.

3. Perturbation bounds. In this section we ostensibly substitute the results of the previous section into (1.4). To simplify this process it is convenient to establish the following lemma.

LEMMA 1. *If $M(t)$ is monotone increasing on $[0, \infty)$ and $\|e^{At}\| \leq M(t) e^{\beta t}$ for all $t \geq 0$, then*

$$\phi(t) = \frac{\|e^{(A+E)t} - e^{At}\|}{\|e^{At}\|} \leq \|E\| t M(t)^2 \exp[(\beta - \alpha(A) + \|E\|M(t))t].$$

Proof. A well known result from semigroup theory (see Kato [5, p. 495]) states that if $\|e^{As}\| \leq c e^{\beta s}$ for all $s \in [0, t]$, then

$$\|e^{(A+E)s}\| \leq c e^{(\beta + c\|E\|)s}, \qquad s \in [0, t].$$

By using the monotonicity of $M(t)$ we thus have

$$\int_0^t \|e^{A(t-s)}\| \|e^{(A+E)s}\| \, ds \leq \int_0^t M(t-s) e^{\beta(t-s)} M(t) e^{(\beta + M(t)\|E\|)s} \, ds$$

$$\leq t M(t)^2 e^{(\beta + M(t)\|E\|)t}.$$

Hence, from (1.4) we have

$$\phi(t) \leq \frac{\|E\|}{\|e^{At}\|} \int_0^t \|e^{A(t-s)}\| \, \|e^{(A+E)s}\| \, ds$$

$$\leq \frac{\|E\|}{e^{\alpha(A)t}} t M(t)^2 \, e^{(\beta + M(t)\|E\|)t}$$

from which the Lemma follows. $\square$

It is now a simple matter to apply the results of the previous section to obtain upper bounds for $\phi(t)$.

THEOREM 1. (See Levis [6] for a similar result.)

$$(3.1) \qquad \phi(t) = t\|E\| \, e^{(\|A\| - \alpha(A) + \|E\|)t}.$$

Proof. By virtue of (2.1) we can apply Lemma 1 with $M(t) = 1$ and $\beta = \|A\|$. $\square$

THEOREM 2.

$$(3.2) \qquad \phi(t) \leq t\|E\| \, e^{(\mu(A) - \alpha(A) + \|E\|)t}.$$

Proof. By using (2.2) we can set $M(t) = 1$ and $\beta = \mu(A)$ and invoke Lemma 1. $\square$

THEOREM 3. *If the Jordan decomposition of A is given by (2.4) and (2.5), then*

$$(3.3) \qquad \phi(t) \leq t\|E\| M_J(t)^2 \, e^{M_J(t)\|E\|t}$$

where $M_J(t) = m\kappa(X) \max_{0 \leq j \leq m-1} t^j/j!$

Proof. $M_J(t)$ is monotone increasing and from (2.8), $\|e^{At}\| \leq M_J(t) \, e^{\alpha(A)t}$. Lemma 1 is thus applicable with $M(t) = M_J(t)$ and $\beta = \alpha(A)$. $\square$

THEOREM 4. *If the Schur decomposition of A is given by (2.9) then*

$$(3.4) \qquad \phi(t) \leq t\|E\| M_S(t)^2 \, e^{M_S(t)\|E\|t}$$

where $M_S(t) = \sum_{k=0}^{n-1} \|Nt\|^k/k!$

Proof. From (2.11), $\|e^{At}\| \leq M_S(t) \, e^{\alpha(A)t}$ and thus the theorem follows by setting $M(t) = M_S(t)$ and $\beta = \alpha(A)$ in Lemma 1. $\square$

By using (2.12), the following result can be obtained as a corollary to any of Theorems 2, 3, or 4.

$$(3.5) \qquad A^*A = AA^* \Rightarrow \phi(t) \leq t\|E\| \, e^{\|E\|t}.$$

We remind the reader that the upper bounds appearing in (3.1)–(3.4) are just that—upper bounds. They are not necessarily accurate measures of the sensitivity of e^{At}. However, because these bounds are smaller when A is normal, they suggest that there is a connection between the normality of A and the sensitivity of e^{At}. We hope to shed some light on this connection in the next section.

In the mean time, we conclude this section by mentioning an entirely different approach to the problem of bounding $\phi(t)$. This approach utilizes the definition

$$e^{At} = \frac{1}{2\pi i} \oint_\Gamma e^{zt}(zI - A)^{-1} \, dz$$

which may be found in MacDuffee [7]. Here Γ is a smooth closed curve encircling the spectrum $\lambda(A)$ of A. If Γ also encloses $\lambda(A+E)$, then

$$e^{(A+E)t} - e^{At} = \frac{1}{2\pi i} \oint_\Gamma e^{zt}(zI - A - E)^{-1}E(zI - A)^{-1}\,dz.$$

One can take norms and proceed to derive an upper bound for $\phi(t)$. An analysis of this type may be found in Fair and Luke [3]. We shall not pursue the matter further, however, because the results are no better than the ones already obtained.

4. The exponential condition of a matrix. In [9] Rice gave a general theory of condition. With this theory it is possible to measure the sensitivity of a map F from metric space X to metric space Y at a point $A \in X$. In this section we investigate the sensitivity of e^{At} by applying Rice's definitions to $F(A) = e^{At}$.

Before we begin, it is instructive to look at the idea of condition in a more familiar setting. In the matrix inversion problem $(A \to A^{-1})$, $\kappa(A) = \|A\|\,\|A^{-1}\|$ is defined as the "condition of a matrix with respect to inversion". That $\kappa(A)$ measures the sensitivity of A^{-1} can be seen from the following inequality:

$$(4.1)\qquad \frac{\|(A+E)^{-1}-A^{-1}\|}{\|A^{-1}\|} \le \frac{\|E\|}{\|A\|}\frac{\kappa(A)}{1-\|E\|\,\|A^{-1}\|}\qquad (\|E\|\,\|A^{-1}\|<1).$$

It is always possible to choose E such that the above upper bound is attained. Thus, if $\kappa(A)$ is large, it is possible for a small change in A to induce a relatively large change in A^{-1}. It is in this sense that $\kappa(A)$ measures the sensitivity of the map $A \to A^{-1}$.

Returning to the e^{At} problem, we now formulate a relevant condition number consistent with Rice's theory.

DEFINITION. The *exponential condition of a matrix A at time t* is defined by

$$\nu(A,t) = \lim_{\delta \to 0} \nu_\delta(A,t)$$

where

$$\nu_\delta(A,t) = \sup_{\|E\|\le\delta\|A\|} \frac{\|e^{(A+E)t}-e^{At}\|}{\delta\|e^{At}\|}.$$

Geometrically, $r = \delta\|e^{At}\|\nu_\delta(A,t)$ is the radius of the smallest sphere in $\mathbb{C}^{n\times n}$ which is centered at e^{At} and encloses the image of the set $\{B\,|\,\|A-B\|\le\delta\|A\|\}$ under the mapping $B \to e^{Bt}$. When small relative changes in A produce relatively large alterations in e^{At}, $\nu_\delta(A,t)$ is correspondingly large.

Our investigation of the condition number $\nu(A,t)$ begins with the following theorem.

THEOREM 5. *If $D(e^{At})$ denotes the Fréchet derivative of $F(A) = e^{At}$ then*

$$\nu(A,t) = \frac{\|D(e^{At})\|\,\|A\|}{\|e^{At}\|}$$

where

$$\|D(e^{At})\| = \sup_{\|E\|\le 1}\left\|\int_0^t e^{A(t-s)}Ee^{As}\,ds\right\|.$$

Proof. As we noted in the introduction,

$$e^{(A+E)t} - e^{At} = D(e^{At})E + o(\|E\|)$$

where

$$D(e^{At})E = \int_0^t e^{A(t-s)} E e^{As}\, ds.$$

Thus,

$$\nu(A, t) = \lim_{\delta \to 0} \sup_{\|E\| \leq \delta\|A\|} \frac{\|D(e^{At})E + o(\|E\|)\|}{\delta\|e^{At}\|}$$

$$= \lim_{\delta \to 0} \sup_{\|E\| \leq \delta\|A\|} \left[\frac{\|D(e^{At})E\|}{\delta\|A\|} \frac{\|A\|}{\|e^{At}\|} + \frac{o(\delta\|A\|)}{\delta\|e^{At}\|} \right]$$

$$= \sup_{\|E\| \leq 1} \|D(e^{At})E\| \frac{\|A\|}{\|e^{At}\|} = \|D(e^{At})\| \frac{\|A\|}{\|e^{At}\|}.$$

COROLLARY 1.

$$\nu(A, t) \geq t\|A\| \quad \text{for all } t \geq 0.$$

Proof.

$$\nu(A, t) \geq \left\| \int_0^t e^{A(t-s)} I e^{As}\, ds \right\| \frac{\|A\|}{\|e^{At}\|} = t\|A\|. \quad \square$$

COROLLARY 2. *If A is normal then $\nu(A, t) = t\|A\|$.*

Proof. In view of the previous corollary, all we must show is that when A is normal, $\nu(A, t) \leq t\|A\|$. But this result follows easily by taking norms in the expression for $\nu(A, t)$ and then using (2.12):

$$\nu(A, t) \leq \int_0^t \|e^{A(t-s)}\| \|e^{As}\|\, ds \frac{\|A\|}{\|e^{At}\|}$$

$$= \int_0^t e^{\alpha(A)(t-s)} e^{\alpha(A)s}\, ds \frac{\|A\|}{e^{\alpha(A)t}} = t\|A\|. \quad \square$$

These corollaries tell us that when A is normal, the condition number $\nu(A, t)$ is as small as possible. It is rather more difficult to identify those A for which $\nu(A, t) \gg t\|A\|$. This is in contrast to the matrix inversion problem where the ill-conditioned matrices can be neatly characterized through the singular value decomposition.

Nevertheless, we can identify a class of problems for which $\nu(A, t)$ is inordinately large. If A is an $n \times n$ strictly upper triangular matrix, then

$$e^{At} = \sum_{k=0}^{n-1} \frac{(At)^k}{k!},$$

and thus, for any matrix E,

$$\int_0^t e^{A(t-s)} E e^{As}\, ds = \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} \frac{A^j E A^k}{(j+k+1)!} t^{j+k+1}.$$

Hence, if $\|E\| = 1$ and $A^{n-1}EA^{n-1} \neq 0$,

$$\nu(A, t) \geq \left\| \int_0^t e^{A(t-s)} E e^{As} ds \right\| \frac{\|A\|}{\|e^{At}\|} = \frac{O(t^{2n-1})}{O(t^{n-1})} \|A\| = O(t^n)\|A\|.$$

For such matrices, $\nu(A, t)$ is large in the sense that it grows as $t^n\|A\|$ instead of as $t\|A\|$ when A is normal. As a specific example, suppose $A = (a_{ij}) \in \mathbb{C}^{n \times n}$ is defined by

$$a_{ij} = \begin{cases} 1, & i < j, \\ 0, & \text{otherwise.} \end{cases}$$

In Table 2 we present some lower bounds for $\nu(A, t)$ for the case $n = 12$. These results were obtained by setting E in the above to be $e_n e_1^T$ where e_k is the kth column of the $n \times n$ identity.

TABLE 2

t	Lower Bound for $\nu(A, t)$
0	0.0×10^0
3	9.7×10^1
6	3.5×10^3
9	3.8×10^4
12	2.2×10^5

$(n = 12)$

Of course, in this example A has a defective eigensystem (it has only one eigenvector). However, since $\nu(A, t)$ is a continuous function of A, we know that there are nondefective matrices near to A with comparably large exponential condition numbers.

We want to conclude this section by establishing a result analogous to (4.1). To do this we need to define the following functions:

$$\hat{\nu}(A, t) = \max_{0 \leq s \leq t} \nu(A, s),$$

$$\theta(A, t) = \int_0^t \|e^{-As}\| \|e^{As}\| ds.$$

The function $\hat{\nu}(A, t)$ can be regarded as an exponential condition number of A over the *interval* $[0, t]$. We also remark that $\theta(A, t)$ is monotone increasing and that obviously $\theta(A, t) \geq t$. (Incidentally, $\theta(A, t) \equiv t$ if and only if $A = \lambda I + S$ where $S^* = -S$.)

THEOREM 6. *If* $\theta(A, t)\|E\| < 1$, *then*

$$(4.2) \qquad \frac{\|e^{(A+E)t} - e^{At}\|}{\|e^{At}\|} \leq \frac{\|E\|}{\|A\|} \frac{\hat{\nu}(A, t)}{1 - \|E\|\theta(A, t)}.$$

Proof. From (1.3)

$$e^{(A+E)u} - e^{Au} = \int_0^u e^{A(u-s)} E e^{As} ds + \int_0^u e^{A(u-s)} E(e^{(A+E)s} - e^{As}) ds$$

and thus

$$\phi(u) = \frac{\|e^{(A+E)u} - e^{Au}\|}{\|e^{Au}\|} \leq \frac{\|\int_0^u e^{A(u-s)} E e^{As} ds\|}{\|e^{Au}\|}$$

$$+ \|E\| \int_0^u \|e^{-As}\| \|e^{As}\| \frac{\|e^{(A+E)s} - e^{As}\|}{\|e^{As}\|} ds$$

$$\leq \frac{\|E\|}{\|A\|} \nu(A, u) + \|E\| \int_0^u \|e^{-As}\| \|e^{As}\| \phi(s) \, ds.$$

Let $t_1 \in [0, t]$ be such that $M = \max_{0 \leq s \leq t} \phi(s)$. By setting $t_1 = u$ in the above inequality we thus obtain

$$M \leq \frac{\|E\|}{\|A\|} \nu(A, t_1) + \|E\| M \theta(A, t_1).$$

Since $\nu(A, t_1) \leq \hat{\nu}(A, t)$ and $\theta(A, t_1) < \theta(A, t)$ we have

$$\phi(t) \leq M \leq \frac{\|E\|}{\|A\|} \frac{\hat{\nu}(A, t)}{1 - \theta(A, t)\|E\|}. \quad \Box$$

When A is normal we may replace $\hat{\nu}(A, t)$ with $\nu(A, t)$ in (4.2) and so once again we observe smaller bounds for the normal matrix case.

5. Conclusion. We have seen that the matrix exponential problem is relatively well-conditioned when A is a normal matrix and more poorly conditioned when A has a defective eigensystem. Perhaps some refinement of the analysis would yield a more precise connection between the eigenstructure of A and the sensitivity of e^{At}.

REFERENCES

[1] R. BELLMAN, *Introduction to Matrix Analysis*, McGraw-Hill, New York, 1969.
[2] G. DAHLQUIST, *Stability and error bounds in the numerical integration of ordinary differential equations*, Trans. Royal Institute of Technology, No. 130, Stockholm, Sweden, 1959.
[3] W. FAIR AND Y. LUKE, *Pade approximations to the operator exponential*, Numer. Math., 14 (1970), pp. 379–382.
[4] F. R. GANTMACHER, *Application of the Theory of Matrices*, Interscience, New York, 1959.
[5] T. KATO, *Perturbation Theory for Linear Operators*. Springer-Verlag, New York, 1966.
[6] A. H. LEVIS. *Some computational aspects of the matrix exponential*. IEEE Trans. Automatic Control, AC-14 (1969), pp. 410–411.
[7] C. C. MACDUFFEE, *The Theory of Matrices*, Chelsea, New York, 1956.
[8] C. B. MOLER AND C. F. VAN LOAN. *Nineteen dubious ways to compute the matrix exponential*. SIAM Rev., to appear.
[9] J. R. RICE. *A theory of condition*. this Journal, 3 (1966), pp. 287–310.
[10] T. STROM. *On logarithmic norms*, this Journal, 12 (1975). pp. 741–753.