MATH. SCAND. 18 (1966), 113–123

MATRIX THEOREMS FOR PARTIAL DIFFERENTIAL AND DIFFERENCE EQUATIONS

JOHN MILLER and GILBERT STRANG

We want to reexamine the Cauchy problem for systems with constant coefficients, together with the matrix questions which arise after a Fourier transformation. Our main results are in fact purely matrix-theoretic, so that after motivating those results in the following paragraphs, we hardly need to mention partial differential equations again. We do hope, however, that our ideas will prove to be useful locally in studying certain systems with variable coefficients; such an application will of course require a much fuller discussion of differential operators.

After Fourier transformation, a linear Cauchy problem for a constant-coefficient system of first order in time looks like

$$\frac{\partial \hat{u}}{\partial t} = P(\omega)\hat{u}, \qquad \hat{u}(\omega,0) = \hat{f}(\omega)\,, \tag{1}$$

where $\omega=(\omega_1,\ldots,\omega_d)$, $\hat{u}=(\hat{u}_1(\omega,t),\ldots,\hat{u}_m(\omega,t))$, and the $m\times m$ matrix P is the symbol of the given differential operator. To stay within the framework of the Fourier transform, we study (1) in the Hilbert spaces $L_2(H)$, normed by

$$\|u\|_H^2 = \int_{\mathbf{R}^d} (H(\omega)\hat{u}(\omega), \hat{u}(\omega))\, d\omega\,. \tag{2}$$

Here H is a measurable Hermitian matrix function, normalized by the requirement $H \geqq I$, that is, $H-I$ shall be non-negative definite.

Let us call (1) *well posed over* $L_2(H)$ provided that for some α,

$$\|u(t)\|_H = \|e^{P(\omega)t}f\|_H \leqq e^{\alpha t}\|f\|_H \tag{3}$$

for all $t\geqq 0$ and all initial data f. Substituting into (2), this can be made more explicit:

$$e^{P^*(\omega)t}\, H(\omega)\, e^{P(\omega)t} \leqq e^{2\alpha t} H(\omega)\,. \tag{4}$$

Received September 20, 1965.

This paper developed from the first author's M.I.T. thesis; the second author was supported by the Office of Naval Research and by the Stanford Linear Accelerator Center.

Differentiating at $t=0$, we come to a still simpler equivalent condition; for almost all ω,

$$H(\omega)P(\omega)+P^*(\omega)H(\omega) \leqq 2\alpha H(\omega)\,. \tag{5}$$

(To recover (4), post-multiply by $\exp(P(\omega)-\alpha)t$, pre-multiply by its adjoint, and integrate.)

Our definition (3) is stronger than the usual one, which permits a constant factor M on the right side. Nothing is changed, however, since if (1) is well-posed in this weaker sense with respect to $L_2(H_1)$, there is an H_2 equivalent to H_1 uniformly in ω, such that (3) holds on $L_2(H_2)$. We shall point out later how this follows from Theorem III; in fact, it is the chief result of the Kreiss theorems, which our work extends.

A simple example will illustrate the problem we solve here. Consider the system

$$\frac{\partial}{\partial t}\begin{pmatrix}u_1\\u_2\end{pmatrix}=\begin{pmatrix}0&0\\1&0\end{pmatrix}\frac{\partial}{\partial x}\begin{pmatrix}u_1\\u_2\end{pmatrix},\qquad \begin{pmatrix}u_1\\u_2\end{pmatrix}=\begin{pmatrix}f_1\\f_2\end{pmatrix}\quad \text{at } t=0\,, \tag{6}$$

which has the solution

$$u_1(t)=f_1,\qquad u_2(t)=f_2+t\frac{df_1}{dx}\,. \tag{7}$$

Because of the derivative in u_2, we may choose $f\in L_2=L_2(I)$ such that $u\notin L_2$, and the system (6) fails to be well-posed over L_2. Nevertheless, with respect to the larger norm

$$\|u\|_H^2=\int_{-\infty}^{\infty}\left(|u_1|^2+\left|\frac{\partial u_1}{\partial x}\right|^2+|u_2|^2\right)dx=\int_{-\infty}^{\infty}\left(\begin{pmatrix}1+\omega^2&0\\0&1\end{pmatrix}\hat{u}(\omega),\,\hat{u}(\omega)\right)d\omega \tag{8}$$

we no longer lose a derivative, and (6) becomes well-posed. In fact, condition (5) reduces in this case to

$$\begin{pmatrix}1+\omega^2&0\\0&1\end{pmatrix}\begin{pmatrix}0&0\\i\omega&0\end{pmatrix}+\begin{pmatrix}0&-i\omega\\0&0\end{pmatrix}\begin{pmatrix}1+\omega^2&0\\0&1\end{pmatrix}=\begin{pmatrix}0&-i\omega\\i\omega&0\end{pmatrix}\leqq\begin{pmatrix}1+\omega^2&0\\0&1\end{pmatrix}. \tag{9}$$

This gives the precise value $\alpha=\frac{1}{2}$:

$$\|u(t)\|_H \leqq e^{\frac{1}{2}t}\,\|f\|_H\,. \tag{10}$$

What we want is to associate with more general systems such a canonical *subspace, namely the largest $L_2(H)$ over which the problem is well-posed.* Without the maximality requirement, this question has been treated independently by Birkhoff and others (see [1]).

It is no trouble to bound the possible values of α from below. If some $P(\omega)$ has the eigenvalue λ with eigenvector v, we must have from (5) that

$$((HP+P^*H)v,v) \leqq 2\alpha(Hv,v) ,$$

which yields

$$\operatorname{Re}\lambda \leqq \alpha .$$

Therefore α is not less than

$$\sigma = \operatorname*{ess\,sup}_{\omega} \sup_{j} \operatorname{Re}\lambda_j(P(\omega)) , \tag{11}$$

and we must impose on the symbol P the *Petrowsky–Gårding condition* $\sigma < \infty$. Subtracting a constant multiple of the identity, we shall in fact suppose $\sigma < 0$. Now fixing $\alpha = 0$, there is no doubt that we can construct $H(\omega)$ to satisfy (5). The delicate problem is to keep H as small as possible; this we achieve, up to a constant K' depending only on the order m, in Theorem III. The corresponding space $L_2(H)$ is consequently maximal; if (1) is also well-posed over $L_2(H')$, then

$$u \in L_2(H') \Rightarrow u \in L_2(H) \text{ and } \|u\|_H \leqq K'(m)\|u\|_{H'} . \tag{12}$$

The theory of partial difference operators leads to a closely related matrix problem. In place of (1) we have

$$\hat{u}(\omega,t+k,k) = A_k(\omega)\hat{u}(\omega,t,k), \qquad \hat{u}(\omega,0,k) = \hat{f}(\omega) . \tag{13}$$

Without discussing such systems fully, we recall that the A_k are known as amplification matrices, and that the time-step k ranges over some interval $0 < k < k_0$. The analogue of (5), equivalent to the condition (3) on $L_2(H_k)$, is simply

$$A_k^*(\omega)\, H_k(\omega)\, A_k(\omega) \leqq e^{2\alpha k}\, H_k(\omega) . \tag{14}$$

The exponent α is to be independent of k, and again there is a lower bound, namely

$$\sigma' = \operatorname*{ess\,sup}_{\omega} \sup_{j,k} \frac{\log|\lambda_j(A_k(\omega))|}{k} . \tag{15}$$

Therefore we impose on (13) the *von Neumann condition* $\sigma' < \infty$. By a simple manipulation, we may achieve $\sigma' < 0$ and fix $\alpha = 0$ as before. Thus our two matrix problems can be very concisely stated: *given suitable P and A, to construct two corresponding matrices $H \geqq I$ as small as possible so that*

$$HP+P^*H \leqq 0 \quad \textit{and} \quad A^*HA \leqq H ,$$

respectively. Since the second problem is perhaps the more familiar, and its solution leads to a solution of the first, it will be treated in full detail. We need the definitions

$$|v| = (|v_1|^2 + \ldots + |v_m|^2)^{\frac{1}{2}}, \qquad |A| = \sup_{|v|=1} |Av|, \qquad \varrho(A) = \max_{1 \leq j \leq m} |\lambda_j(A)| .$$

THEOREM I. *For a suitable constant $K(m)$, depending only on the order m of the matrix A, each of the following statements implies the next:*

(i) $A^*HA \leqq H$ *for some* $H \geqq I$ *with* $(Hv,v)^{\frac{1}{2}} = C(v)$ *for* $|v| = 1$.

(ii) $|SAS^{-1}| \leqq 1$ *for some* S *with* $|S^{-1}| \leqq 1$ *and* $|Sv| = C(v)$ *for* $|v| = 1$.

(iii) $|A^n v| \leqq C(v)$ *for all* $n \geqq 0$ *and* $|v| = 1$.

(iv) $|(zI - A)^{-1} v| \leqq C(v)/(|z| - 1)$ *for all complex* $|z| > 1$ *and all* $|v| = 1$.

(v) $A^*HA \leqq \frac{1}{4}(1 + \varrho(A))^2 H \leqq H$ *for some* $H \geqq I$ *with* $(Hv,v)^{\frac{1}{2}} \leqq K(m)C(v)$ *for all* $|v| = 1$.

This theorem is very close to one originally proved by Kreiss [4], and studied subsequently by Morton [8] and Morton and Schechter [9]. We ought to emphasize that although these authors discuss families of matrices, their results are really quantitative versions (just as Theorem I is) of an easy theorem about single matrices:

$$\sup_{n \geqq 0} |A^n| < \infty \iff |SAS^{-1}| \leqq 1 \text{ for some } S .$$

The factor $\frac{1}{4}(1 + \varrho(A))^2$ is an embellishment which has proved useful in applications, but the essential point is that in the circuit from (i) to (v), only a factor $K(m)$ is lost at the last step. We should clarify those respects in which this conclusion is new:

a) The previous estimates in (v), established by induction on m, had a power of $C^{p(m)}$ in place of C, with $p(m) \to \infty$ as $m \to \infty$. Our improvement becomes important when C is not uniformly bounded with respect to ω, that is, (13) is not stable in the Lax–Richtmyer sense over L_2.

b) We estimate the action of H on each vector v, where earlier there appeared only the single constant $C = \sup C(v)$. It follows that the H in (v) is minimal in a stronger sense than just in norm: if $H' \geqq I$ and $A^*H'A \leqq H'$, then $H \leqq K^2(m)H'$: For the proof, we simply set $C(v) = (H'v,v)^{\frac{1}{2}}$, and go from (i) to (v); the parallel argument in the exponential case yields (12).

c) We shall prove (iv) $\Rightarrow$ (v) by the explicit construction (from the elements and eigenvalues of A) of a suitable matrix H, which possesses the following additional property:

For some S with $S^*S = H$, $A' = SAS^{-1}$ is upper triangular, with $A'_{ij} = 0$ unless λ_i and λ_j are in the same cluster (see below), and $|A'_{ij}| \leqq \frac{1}{2}(1 - \max(|\lambda_i|, |\lambda_j|))$ whenever $i \neq j$.

A trivial modification of S replaces this constant $\frac{1}{2}$ by any other, say $1/2m$, so that the absolute row and column sums (the l_∞ and l_1 norms of A) may also be reduced to $\frac{1}{2}(1+\varrho(A))$.

It remains to determine the behavior of the best constant $K(m)$. Our constant (which we don't compute) grows roughly like m^m, while examples of McCarthy and Schwartz [6] show that it must grow at least as fast as some power of log m; this leaves a wide gap. It is not surprising that $K(m) \to \infty$ in view of the Foguel–Halmos counterexamples [2], [3] to the Nagy conjecture.

2.

In this section we establish the first three implications in Theorem I. These are easy steps, valid also for operators on Hilbert space.

With $H = S^*S$, the equivalence of (i) and (ii) follows from that of the inequalities

$$\begin{aligned} (A^*HAv,v) &\leqq (Hv,v) \quad \text{for all } v\ , \\ |SAv|^2 &\leqq |Sv|^2 \quad \text{for all } v\ , \\ |SAS^{-1}w|^2 &\leqq |w|^2 \quad \text{for all } w\ . \end{aligned}$$

In the applications, (ii) corresponds to a change of variables and (i) to a new norm. In one respect the use of H is to be preferred; it may depend more smoothly on some relevant parameters than does an improperly chosen S. The positive square root $S = H^{\frac{1}{2}}$ is as smooth as H, but a diagonalizing S may not be, although the latter change of variables looks especially desirable. Mizohata [7] points out how this possibility can arise, when $d = 2$, from the multiple-connectedness of the unit circle in $\boldsymbol{R}^2$; there is no difficulty in his context with H.

To show that (ii) implies (iii), we compute

$$(16) \qquad |A^n v| = |S^{-1}(SAS^{-1})^n Sv| \leqq |S^{-1}|\, |SAS^{-1}|^n\, |Sv| \leqq C(v)\,.$$

Finally, given (iii), we have for $|z| > 1$

$$(17) \qquad |(zI-A)^{-1}v| = \left| \sum_0^\infty \frac{A^n v}{z^{n+1}} \right| \leqq \sum_0^\infty \frac{C(v)}{|z|^{n+1}} = \frac{C(v)}{|z|-1}\,.$$

3.

Before coming to the final step in Theorem I, we warm up with a more special result of the same kind, which shows how the geometry of the eigenvalues enters the problem.

THEOREM II. *Suppose the resolvent condition* (iv) *holds, and the eigenvalues of A satisfy*

$$\delta|\lambda_i-\lambda_j| \geqq 1-|\lambda_j| \quad \textit{for all distinct } i,j\,. \tag{18}$$

Then $A^*HA \leqq \varrho^2(A)H \leqq H$ *for some* $H \geqq I$ *with*

$$(Hv,v)^{\frac{1}{2}} \leqq m(2+4m\delta)(1+2\delta)^{2m-3}C(v)$$

for $|v|=1$. *Furthermore, there exists* S *such that* $H=S^*S$ *and* SAS^{-1} *is diagonal.*

Proof. From (iv) it is clear that no eigenvalue lies outside the unit circle, so $\varrho(A)\leqq 1$. Although (18) admits repeated eigenvalues of modulus one, suppose for the present that the eigenvalues are distinct. Then we construct the projections

$$L_i = \prod_{j\neq i} \frac{A-\lambda_j}{\lambda_i-\lambda_j}\,, \qquad 1\leqq i\leqq m\,. \tag{19}$$

Applying L_i to the eigenvectors $v_1,\ldots,v_m$, we find $L_iv_j=\delta_{ij}v_j$, so there are the standard identities

$$L_i^2 = L_i, \qquad L_iL_j = 0 \text{ for } i\neq j \tag{20}$$

$$\sum_1^m L_i = I, \qquad \sum_1^m \lambda_i L_i = A\,. \tag{21}$$

Now define the Hermitian matrix H by

$$H = m\sum_1^m L_i^*L_i\,. \tag{22}$$

From (20) and (21) we have

$$A^*HA = \sum \bar{\lambda}_j L_j^* m \sum L_i^*L_i \sum \lambda_k L_k = m\sum |\lambda_i|^2 L_i^*L_i \leqq \varrho^2(A)H\,.$$

To prove $H\geqq I$ we need only (21) and the Schwarz inequality:

$$|v|^2 = |\sum L_iv|^2 \leqq (\sum |L_iv|)^2 \leqq m\sum |L_iv|^2 = (Hv,v)\,. \tag{24}$$

From (22),

$$(Hv,v) \leqq m^2 \max|L_iv|^2 \tag{25}$$

and the crucial estimate is that of $|L_iv|$. We use the resolvent condition in the most natural way, by expanding

$$L_i = \sum_{k=1}^m b_{ik}(z_kI-A)^{-1}\,. \tag{26}$$

We shall choose $z_k=1/\bar{\lambda}_k$; if $|\lambda_k|$ is 0 or 1, then it is no longer true that $1<|z_k|<\infty$, and a simple limiting argument is required in what follows. To compute the b_{ik}, apply (26) to the eigenvectors; for each i,

$$(27)\qquad \delta_{ij} = \sum_{k=1}^{m} b_{ik}(\bar{\lambda}_k^{-1}-\lambda_j)^{-1}, \qquad 1 \leqq j \leqq m\,.$$

Solving this system, we get

$$(28)\qquad \frac{|b_{ii}|}{|z_i|-1} = (1+|\lambda_i|)\prod_{j\neq i}\frac{|1-\bar{\lambda}_j\lambda_i|^2}{|\lambda_j-\lambda_i|^2}$$

and, for $k \neq i$,

$$(29)\qquad \frac{|b_{ik}|}{|z_k|-1} = (1+|\lambda_k|)(1+|\lambda_i|)\frac{1-|\lambda_i|}{|\lambda_k-\lambda_i|}\,\frac{|1-\bar{\lambda}_k\lambda_i|}{|\lambda_k-\lambda_i|}\prod_{j\neq i,k}\frac{|1-\bar{\lambda}_j\lambda_i|\,|1-\bar{\lambda}_k\lambda_j|}{|\lambda_j-\lambda_i|\,|\lambda_j-\lambda_k|}\,.$$

For any distinct i and j,

$$(30)\qquad \left|\frac{1-\bar{\lambda}_j\lambda_i}{\lambda_j-\lambda_i}\right| = \left|\bar{\lambda}_j+(1+|\lambda_j|)\,\frac{1-|\lambda_j|}{\lambda_j-\lambda_i}\right| \leqq 1+2\delta\,.$$

Putting the pieces together,

$$(31)\qquad \begin{aligned} |L_i v| &\leqq \sum |b_{ik}|\,|(z_k I-A)^{-1}v| \\ &\leqq \sum \frac{|b_{ik}|\;C(v)}{|z_k|-1} \\ &\leqq [2(1+2\delta)^{2m-2}+(m-1)\,4\delta\,(1+2\delta)^{2m-3}]\,C(v)\,. \end{aligned}$$

Simplifying the last term and using (25),

$$(32)\qquad (Hv,v)^{\frac{1}{2}} \leqq m(2+4m\delta)(1+2\delta)^{2m-3}C(v)\,.$$

To complete the theorem, we introduce the left (*row*) eigenvectors r_k, so that

$$(33)\qquad Av_j = \lambda_j v_j, \qquad r_k A = \lambda_k r_k\,.$$

Multiplying the first by r_k and the second by v_j, there is the familiar biorthogonality condition

$$(34)\qquad r_k v_j = (0) \quad \text{for} \quad j \neq k\,.$$

Since v_j cannot be orthogonal also to r_j, we may fix the eigenvectors by the normalization

$$(35)\qquad m|v_j|^2 = 1 \quad \text{and} \quad r_j v_j = (1), \qquad 1 \leqq j \leqq m\,.$$

It follows that

$$(36)\qquad L_i = v_i r_i\,,$$

since both sides, applied to v_j, give $\delta_{ij}v_j$.

Now let the rows of S be $r_1,\ldots,r_m$, so that SAS^{-1} is diagonal. By matrix multiplication

$$(37)\qquad S^*S = (r_1^* \ldots r_m^*)\begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix} = \sum_1^m r_i^* r_i \, .$$

Using (35) and (36), this is precisely

$$(38)\qquad \sum_1^m r_i^*(mv_i^*v_i)r_i = m\sum L_i^*L_i = H \, .$$

Finally, we have to return and admit eigenvalues λ_i of modulus one and multiplicity $M > 1$. From the resolvent condition (iv), λ_i possesses M linearly independent corresponding eigenvectors; to prove this, one puts A in Jordan form to compute the resolvent $(zI-A)^{-1}$, and then lets z approach λ_i. The eigenvectors may still be chosen to satisfy (34).

Let us number the eigenvalues so that $\lambda_1,\ldots,\lambda_N$ are distinct, and the rest are duplicates of these. Then instead of (19) we want

$$(39)\qquad L_i = \prod_{\substack{j=1 \\ j\neq i}}^{N} \frac{A-\lambda_j}{\lambda_i-\lambda_j}, \qquad i=1,\ldots,N \, .$$

Simply replacing m by N in all the equations (20) to (32), the first part of the proof continues to hold. In place of (35) and (36), we have

$$N|v_j|^2 = 1 \quad \text{and} \quad r_jv_j = (1), \qquad 1 \leqq j \leqq m \, ,$$
$$L_i = v_{i_1}r_{i_1} + \ldots + v_{i_M}r_{i_M} \, ,$$

where $\lambda_{i_1},\ldots,\lambda_{i_M}$ are the appearances of the eigenvalue λ_i. Then we may once more identify

$$(40)\qquad S^*S = \sum_1^m r_i^*r_i = \sum_1^m r_i^*(Nv_i^*v_i)r_i = N\sum_1^N L_i^*L_i = H \, .$$

To bound the "condition number" $\nu = |S|\,|S^{-1}|$ in diagonalizing more general matrices B, we may set $A = B/|B|$ and $C(v) = 1$. Then the condition on the eigenvalues μ_i of B becomes

$$\delta|\mu_i - \mu_j| \geqq |B| - |\mu_j| \quad \text{for } i \neq j \, .$$

The resulting estimate $\nu \leqq m(2+4m\delta)(1+2\delta)^{2m-3}$ can be improved by looking more closely at our proof.

Theorem II suggests that the proper measure of distance between eigenvalues involves the ratios $(1-|\lambda_j|)/|\lambda_i - \lambda_j|$. When these are small, we may safely diagonalize. For poorly separated eigenvalues, however, Theorem I requires a new construction, which developed from the observation that $H=\sum_0^\infty A^{*n}A^n$ immediately yields $A^*HA \leqq H$.

4.

To complete the proof of Theorem I, it remains to show that (iv) implies (v). From (iv) we know the eigenvalues satisfy $|\lambda_j| \leqq 1$; we shall put them into clusters as Morton [8] has done. Into the cluster C_1 goes an eigenvalue, say λ_1, of largest modulus, together with all others that can be connected to λ_1 by a chain of eigenvalues, each link having length less than $(1-|\lambda_1|)/4m$. The cluster C_2 is formed in the same way from the remaining eigenvalues, and so on until every eigenvalue enters one of the clusters $C_1,\ldots,C_r$. Of course $r \leqq m$; when $r=m$, our basic constructions coincide with those in Theorem II. An eigenvalue of modulus one and multiplicity M appears alone in M clusters.

Let λ_α be the eigenvalue of largest modulus in C_α from which the cluster was formed. Then since each $\lambda_i \in C_\alpha$ is connected to λ_α by a chain with fewer than m links,

$$|\lambda_i-\lambda_\alpha| \leqq (m-1)\frac{1-|\lambda_\alpha|}{4m} \leqq \frac{1-|\lambda_\alpha|}{4}\,. \tag{41}$$

In fact, the point of constructing the clusters is to achieve

$$|\lambda_i-\lambda_j| \leqq \frac{1-|\lambda_j|}{4} \quad \text{or} \quad |\lambda_i-\lambda_j| \geqq \gamma(m)(1-|\lambda_j|) \tag{42}$$

according as λ_i and λ_j are in the same cluster or not; the computation of γ can be copied from [8].

Let us suppose that

$$\varrho(A) < 1 \quad \text{and} \quad \lambda_i \neq \lambda_j \text{ for } i \neq j\,, \tag{43}$$

and remove this hypothesis later by a continuity argument.

We want to associate with each cluster several matrices from which to construct H. Recalling the projections L_i defined in (19), let

$$I_\alpha = \sum L_i, \quad A_\alpha = \sum \lambda_i L_i, \quad B_\alpha = \sum \frac{2(\lambda_i-\lambda_\alpha)}{1-|\lambda_\alpha|} L_i\,, \tag{44}$$

summing over the indices i such that $\lambda_i \in C_\alpha$. Define

$$H_\alpha = I_\alpha^* I_\alpha + \sum_1^\infty (B_\alpha^*)^n (B_\alpha)^n\,. \tag{45}$$

From (20), I_α acts like the identity relative to C_α, and matrices associated with different clusters are orthogonal. In particular, we write down

$$(46)\qquad \begin{aligned} &I_\alpha^2 = I_\alpha, \quad I_\alpha B_\alpha = B_\alpha, \quad I_\alpha A_\alpha = A_\alpha; \\ &I_\alpha A_\beta = B_\alpha A_\beta = H_\alpha A_\beta = 0, \quad \alpha \neq \beta . \end{aligned}$$

From the definitions it follows that

$$(47)\qquad A_\alpha = \lambda_\alpha I_\alpha + \tfrac{1}{2}(1-|\lambda_\alpha|) B_\alpha$$

$$(48)\qquad I_\alpha^* H_\alpha I_\alpha = H_\alpha, \quad B_\alpha^* H_\alpha B_\alpha = H_\alpha - I_\alpha^* I_\alpha \leqq H_\alpha .$$

Then from the appropriate triangle inequality

$$(49)\qquad A_\alpha^* H_\alpha A_\alpha \leqq (|\lambda_\alpha| + \tfrac{1}{2}(1-|\lambda_\alpha|))^2 H_\alpha \leqq \tfrac{1}{4}(1+\varrho(A))^2 H_\alpha .$$

From (21) we see at once that

$$(50)\qquad \sum_1^r I_\alpha = I, \quad \sum_1^r A_\alpha = A .$$

Now the matrix we want is just

$$(51)\qquad H = m \sum_1^r H_\alpha .$$

Combining the last three equations with (46),

$$(52)\quad A^*HA = \sum A_\alpha^* m \sum H_\beta \sum A_\gamma = m \sum A_\alpha^* H_\alpha A_\alpha \leqq \tfrac{1}{4}(1+\varrho(A))^2 H .$$

To see that $H \geqq I$ we use the Schwarz inequality to compute

$$(53)\qquad |v|^2 = \left| \sum_1^r I_\alpha v \right|^2 \leqq r \sum |I_\alpha v|^2 \leqq m \sum (H_\alpha v, v) = (Hv, v) .$$

The essential problem is to bound

$$(54)\qquad (Hv,v) = m \sum_\alpha \left(|I_\alpha v|^2 + \sum_1^\infty |B_\alpha^n v|^2 \right) .$$

There are two means of carrying out this estimate. Conceptually, the simplest possible approach is to expand I_α and B_α^n as sums of resolvents, just as L_i was expanded in (26), and then apply (iv). This leads to the best *self-contained* proof of our result, and the details (which involve a good deal of algebraic manipulation) will appear elsewhere. Here we adopt a more economical alternative; with some minor refinements, the estimates we need can be lifted from those made by Morton [8]. We denote his epuations by an added asterisk.

Morton's final result is

$$(55)\qquad \text{(iv)} \Rightarrow |A^n v| \leqq K_1(m) \sup C(v) ,$$

but his proof works without requiring the supremum on the right side, by noticing the action on each v in (13*)–(16*) and (18*). Furthermore, his estimate of $A^n v$ is found precisely by bounding the contribution from each cluster; thus when $n=0$, that is, $\nu=0$ in (18*),

$$|I_\alpha v| \leqq K_2(m)\, C(v) \tag{56}$$

and also when $n>0$,

$$|A_\alpha{}^n v| \leqq K_2(m)\, C(v)\,. \tag{57}$$

Now we introduce one more matrix associated with C_α:

$$D_\alpha = A_\alpha + \lambda_\alpha(I-I_\alpha)\,. \tag{58}$$

From the identities (46), we know

$$D_\alpha{}^n = A_\alpha{}^n + \lambda_\alpha{}^n(I-I_\alpha), \qquad n>0\,. \tag{59}$$

According to (56) and (57),

$$|D_\alpha{}^n v| \leqq K_3(m)\, C(v), \qquad n \geqq 0\,. \tag{60}$$

Then the implication (iii) $\Rightarrow$ (iv) gives

$$|(zI-D_\alpha)^{-1}v| \leqq \frac{K_3(m)\, C(v)}{|z|-1}, \qquad |z|>1\,. \tag{61}$$

Manipulating with the definitions, we find

$$(zI-B_\alpha)^{-1} = \tfrac{1}{2}(1-|\lambda_\alpha|)(z_\alpha I-D_\alpha)^{-1}, \qquad z_\alpha = \lambda_\alpha + \tfrac{1}{2}(1-|\lambda_\alpha|)z\,. \tag{62}$$

Let z lie on the circle Z_α of radius 1 about the point $4\lambda_\alpha/|\lambda_\alpha|$ (or 4, if $\lambda_\alpha=0$). The minimum of $|z_\alpha|$ on this circle occurs when z is closest to the origin, and an easy computation gives

$$|z_\alpha|-1 \geqq \tfrac{1}{2}(1-|\lambda_\alpha|), \qquad z \text{ on } Z_\alpha\,. \tag{63}$$

Thus it follows from (61)–(63) that

$$|(zI-B_\alpha)^{-1}v| \leqq K_3(m)\, C(v), \qquad z \text{ on } Z_\alpha\,. \tag{64}$$

From (44), the eigenvalues μ_i of B_α are

$$\mu_i = \frac{2(\lambda_i-\lambda_\alpha)}{1-|\lambda_\alpha|}, \quad \lambda_i \in C_\alpha; \qquad \mu_i=0,\ \lambda_i \notin C_\alpha\,. \tag{65}$$

Applying (41), all these eigenvalues satisfy

$$|\mu_i| \leqq \tfrac{1}{2}\,. \tag{66}$$

Using only (64) and (66), we will obtain the required bound (70); this result may have some independent interest. Looking a second time at

Morton's argument, we put all the μ_i into one cluster, so his $X=1$. Denoting by D^p a divided difference formed at some $p+1$ of the points μ_i, (11*) becomes

$$|D^p(z^n)| \leqq n^p(\tfrac{1}{2})^{n-p} . \tag{67}$$

To bound $P(z)v=(zI-B_\alpha)^{-1}\prod(z-\mu_i)v$ on Z_α, we have only to multiply the estimate (64) by 6^m. Then because P is actually a polynomial of degree less than m, we conclude (from Cauchy's formula as in (14*)–(16*), or otherwise) that its differences are bounded in the same way:

$$|D^q(P(z)v)| \leqq K_4(m)\, C(v) . \tag{68}$$

As in (4*), $B_\alpha{}^n v$ is just the divided difference of order $m-1$ of the product $z^n P(z)v$ formed at the μ_i. Constructing a Leibniz rule, this divided difference is the sum of 2^{m-1} products, each bounded by

$$|D^p(z^n)D^{m-p-1}(P(z)v)| \leqq n^{m-1}(\tfrac{1}{2})^n K_5(m)\, C(v) . \tag{69}$$

Consequently

$$|B_\alpha{}^n v| \leqq n^{m-1}(\tfrac{1}{2})^n K_6(m)\, C(v) . \tag{70}$$

Substituting (70) and (56) into (54), the infinite series converges to give the final estimate

$$(Hv,v)^{\frac{1}{2}} \leqq K(m)\, C(v) . \tag{71}$$

We still have to eliminate the hypothesis (43). It is easy to choose M (after triangularizing A, for example) so that

$$A_\varepsilon = (1-\varepsilon)A+\varepsilon^2 M$$

satisfies (43) as $\varepsilon \to 0_+$. Then for $|v|=1$ it follows from (iv) that

$$|(zI-(1-\varepsilon)A)^{-1}v| \leqq \frac{C(v)}{|z|-(1-\varepsilon)} \leqq \min\left(\frac{C(v)}{|z|-1}, \frac{C}{\varepsilon}\right) \tag{72}$$

for $|z|>1$, where $C=\sup C(v)$. (The uniform boundedness theorem applied to (iv) assures that $C(v)$ can be reduced if necessary so that $C<\infty$.) Therefore

$$\begin{aligned} |(zI-A_\varepsilon)^{-1}v| &= \left|\sum_0^\infty [\varepsilon^2(zI-(1-\varepsilon)A)^{-1}M]^n(zI-(1-\varepsilon)A)^{-1}v\right| \\ &\leqq \frac{1}{1-\varepsilon C|M|}\,\frac{C(v)}{|z|-1} . \end{aligned} \tag{73}$$

Since (43) holds for A_ε, there is an $H_\varepsilon \geqq I$ with

$$A_\varepsilon{}^* H_\varepsilon A_\varepsilon \leqq \tfrac{1}{4}(1+\varrho(A_\varepsilon))^2 H_\varepsilon; \qquad (H_\varepsilon v,v)^{\frac{1}{2}} \leqq \frac{K(m)\,C(v)}{1-\varepsilon C|M|} . \tag{74}$$

As $\varepsilon \to 0$, some subsequence of H_ε converges by compactness to an $H \geqq I$, and taking the limit in (74) gives (v).

5.

In this section, we establish the italicized statement about S which follows Theorem I. Again we start by assuming (43), and we recall the left eigenvectors r_k defined in (33). Suppose we now number the eigenvalues in the order that they fall into clusters, and let C_1 contain $\lambda_1, \ldots, \lambda_q$. We want to prove that $H_1 = S_1{}^* S_1$, where the first q rows of S_1 are linear combinations of $r_1, \ldots, r_q$, and the other $m-q$ rows are zero. From the definition (45),

$$(75) \qquad H_1 v_k = 0 \quad \text{for } k > q, \qquad \operatorname{rank}(H_1) = q\,.$$

Writing $H_1{}^{\frac{1}{2}}$ for the non-negative definite square root,

$$(76) \qquad |H_1{}^{\frac{1}{2}} v_k|^2 = (H_1 v_k, v_k) = 0 \quad \text{for } k > q\,.$$

By (34), $r_1, \ldots, r_q$ span the orthogonal complement of the space generated by $v_{q+1}, \ldots, v_m$. Therefore each row of $H_1{}^{\frac{1}{2}}$ is a combination of $r_1, \ldots, r_q$. Let V be the space spanned by the columns of $H_1{}^{\frac{1}{2}}$. We construct orthonormal bases $u_1, \ldots, u_q$ and $u_{q+1}, \ldots, u_m$ for V and $V^\perp$. Taking the u_i as the rows of a unitary matrix U_1, we have shown that $S_1 = U_1 H_1{}^{\frac{1}{2}}$ has the required properties; of course

$$S_1{}^* S_1 = H_1{}^{\frac{1}{2}} U_1{}^* U_1 H_1{}^{\frac{1}{2}} = H_1\,.$$

For every C_α, we construct in the same way an S_α satisfying $H_\alpha = S_\alpha{}^* S_\alpha$; row j of S_α is non-zero if and only if $\lambda_j \in C_\alpha$. Then defining $\bar{S} = m^{\frac{1}{2}} \sum S_\alpha$, and recalling the multiplication rule (37), we have $\bar{S}^* \bar{S} = H$.

Let $\bar{A} = \bar{S} A \bar{S}^{-1}$. Since the first row of $\bar{S}$ is by construction a combination of $r_1, \ldots, r_q$, and $r_k A = \lambda_k r_k$, the same is true of the first row of $\bar{S}A$. This must coincide with the first row of $\bar{A}\bar{S}$, which is a combination with weights $\bar{A}_{1j}$ of the rows of $\bar{S}$. The first q rows of $\bar{S}$ contribute a combination of $r_1, \ldots, r_q$, and the last $m-q$ a combination of $r_{q+1}, \ldots, r_m$. By the linear independence of the r_k the latter contribution is zero; then because the rows of $\bar{S}$ are independent, $\bar{A}_{1j} = 0$ for $j > q$. In the same way, $\bar{A}_{ij} = 0$ whenever λ_i and λ_j are in different clusters. Therefore

$$\bar{A} = \begin{pmatrix} \bar{A}_1 & & 0 \\ & \ddots & \\ 0 & & \bar{A}_r \end{pmatrix},$$

the square block $\bar{A}_\alpha$ on the diagonal corresponding to the cluster C_α.

With a final unitary similarity $\bar{U}$ of the same block form, we triangularize each $\bar{A}_\alpha$ separately. Thus with $S=\bar{U}\bar{S}$, we have $H=S^*S$, and $A'=SAS^{-1}$ has the required (triangular, block diagonal) form.

We have still to estimate the off-diagonal entries of A'. Denoting by a prime the result of applying the similarity S, we conclude from the reasoning of the previous paragraph that A_α', I_α', B_α', and L_i', $\lambda_i \in C_\alpha$, all have zero entries outside block α. Since I_α' is the sum of the right number of mutually orthogonal projections L_i', we know that I_α' is just the identity matrix in its block. Therefore by (47) the off-diagonal entries are introduced through B_α'. According to (48), $|B_\alpha'| \leqq 1$, and the same must be true of all its entries. Then the off-diagonal entries of A_α' are bounded by

$$\tfrac{1}{2}(1-|\lambda_\alpha|) \leqq \tfrac{1}{2}(1-|\lambda_i|), \quad \lambda_i \in C_\alpha .$$

Again we must circumvent (43). Recall that the sequence $A_\varepsilon \to A$ led to a subsequence $H_\varepsilon \to H$; for each H_ε we have seen how to construct S_ε, and taking a further subsequence, we get $S_\varepsilon \to S$, where $S^*S=H$. Unless (43) is violated by a repeated eigenvalue of modulus one, the clusters for A_ε and A coincide for small ε. Therefore the limit matrix S gives an $A'=SAS^{-1}$ with the right properties. In case A has a repeated eigenvalue with $|\lambda_j|=1$, we still know A' is upper triangular and $|A'|=1$; but from this the off-diagonal entries in the rows containing λ_j must vanish, and once more A' is all right.

An alternative construction of S, leading to a more computational argument than we have given in this section, can be derived from the formula

$$(77) \qquad H_\alpha = \sum_{n=0}^{\infty} \left(\sum \bar{\mu}_i L_i^*\right)^n \left(\sum \mu_j L_j\right)^n = \sum_{n=0}^{\infty} \left(\sum \bar{\mu}_i^n L_i^*\right)\left(\sum \mu_j^n L_j\right) = \sum (1-\bar{\mu}_i\mu_j)^{-1} L_i^* L_j ,$$

summing over the projections corresponding to eigenvalues in C_α.

It is worth remarking that in (v), H and S cannot be made continuous functions of A. The family

$$A_\gamma = \begin{pmatrix} e^{i\gamma} & |\gamma| \\ 0 & 1 \end{pmatrix}, \qquad \gamma \text{ real},$$

satisfies (iv) with some $C(v)$ independent of γ. Since the eigenvalues of A_γ have modulus one, A_γ must be diagonal with respect to H_γ to satisfy $A_\gamma^* H_\gamma A_\gamma \leqq H_\gamma$. However, one of the eigenvectors of A_γ is discontinuous at $\gamma=0$, from which one easily verifies that H_γ is too.

6.

With the definitions

$$\tau(P) = \max \operatorname{Re} \lambda_j(P), \qquad \operatorname{Re} P = \tfrac{1}{2}(P+P^*)\,, \tag{78}$$

we can state the analogues of Theorems I and II for the exponential case.

THEOREM III. *For a suitable $K'(m)$ depending only on the order m of the matrix P, each of the following statements implies the next:*

(i′) $HP+P^*H \leqq 0$ *for some* $H \geqq I$ *with* $(Hv,v)^{\frac{1}{2}} = C(v)$ *for* $|v|=1$.

(ii′) $\operatorname{Re} SPS^{-1} \leqq 0$ *for some* S *with* $|S^{-1}| \leqq 1$ *and* $|Sv| = C(v)$ *for* $|v|=1$.

(iii′) $|e^{Pt}v| \leqq C(v)$ *for all* $t \geqq 0$ *and* $|v|=1$.

(iv′) $|(zI-P)^{-1}v| \leqq C(v)/\operatorname{Re} z$ *for* $\operatorname{Re} z > 0$ *and* $|v|=1$.

(v′) $HP+P^*H \leqq \tau(P)H \leqq 0$ *for some* $H \geqq I$ *with* $(Hv,v)^{\frac{1}{2}} \leqq K'(m)C(v)$ *for* $|v|=1$.

THEOREM IV. *Suppose* (iv′) *holds, and the eigenvalues of P satisfy*

$$\delta|\lambda_i - \lambda_j| \geqq -\operatorname{Re}\lambda_j \quad \textit{for all distinct } i,j\,. \tag{79}$$

Then

$$HP+P^*H \leqq 2\tau(P)H \leqq 0$$

for some $H \geqq I$ *with*

$$(Hv,v)^{\frac{1}{2}} \leqq m(2+4m\delta)(1+2\delta)^{2m-3}C(v) \qquad \textit{for } |v|=1\,.$$

Furthermore, there exists S such that $H = S^*S$ and SAS^{-1} is diagonal.

Before discussing the proofs, we redeem our earlier promise to apply Theorem III to a system which is well-posed over $L_2(H_1)$ in the weaker sense

$$\|u(t)\|_{H_1} \leqq Me^{\alpha t}\|f\|_{H_1} \qquad \text{or} \qquad e^{P^*(\omega)t}\,H_1(\omega)\,e^{P(\omega)t} \leqq M^2e^{2\alpha t}\,H_1(\omega)\,.$$

If we choose S so that $S^*S = H_1$, and set $Q = S(P-\alpha I)S^{-1}$, this becomes simply

$$|e^{Q(\omega)t}| \leqq M\,.$$

Because (iii′) $\Rightarrow$ (v′), there is an $H(\omega)$ such that

$$HQ+Q^*H \leqq 0, \qquad I \leqq H \leqq (K'(m)M)^2I\,. \tag{80}$$

Substituting back for Q, and setting $H_2 = S^*HS$, (80) is the same as

$$H_2P+P^*H_2 \leqq 2\alpha H_2, \qquad H_1 \leqq H_2 \leqq K'(m)MH_1\,.$$

Thus $H_2(\omega)$ provides an equivalent norm for $L_2(H_1)$, and on $L_2(H_2)$ the problem is well-posed in the stricter sense that (5) holds, implying

$$\|u(t)\|_{H_2} \leqq e^{\alpha t}\,\|f\|_{H_2}\,.$$

Turning now to the proofs, Theorem IV goes almost exactly as Theorem II did; one makes the choice $z_k = -\bar{\lambda}_k$ in (26), as in the original paper by Kreiss [5], and recomputes (28)–(30).

The first three implications in Theorem III are also easy. With $H = S^*S$, (i′) and (ii′) are equivalent as before. Then (i′) $\Rightarrow$ (iii′) just as (5) $\Rightarrow$ (4); we integrate the inequality

$$e^{P^*(t)}(HP+P^*H)e^{Pt} = \frac{d}{dt}(e^{P^*t}He^{Pt}) \leqq 0$$

from 0 to t, to find

$$e^{P^*t}He^{Pt} \leqq H .$$

From the properties of H given in (i′), we have (iii′):

$$|e^{Pt}v|^2 \leqq (e^{P^*t}He^{Pt}v, v) \leqq (Hv,v) = C^2(v) .$$

The step (iii′) $\Rightarrow$ (iv′) involves the Laplace transform in place of the power series in (17):

$$|(zI-P)^{-1}v| = \left|\int_0^\infty e^{-zt}e^{Pt}v\,dt\right| \leqq C(v)\int e^{-t\operatorname{Re}z}dt = C(v)/\operatorname{Re}z .$$

The cluster C_1' is now formed by starting with an eigenvalue λ_1 of largest real part (necessarily $\leqq 0$ by (iv′)) and connecting to it those eigenvalues which can be reached with links of length less than $-\operatorname{Re}\lambda_1/4m$. Then $C_2',\ldots,C_r'$ are formed consecutively in the same way. In analogy with (43) we may temporarily assume that

$$\tau(P) < 0 \quad \text{and} \quad \lambda_i \neq \lambda_j \text{ for } i \neq j , \tag{81}$$

and then remove this restriction as before. Now we can define

$$I_\alpha = \sum L_i, \qquad P_\alpha = \sum \lambda_i L_i, \qquad G_\alpha = \sum \frac{2(\lambda_\alpha - \lambda_i)}{\operatorname{Re}\lambda_\alpha} L_i , \tag{82}$$

summing over indices i such that $\lambda_i \in C_\alpha'$. Next we let

$$H_\alpha = I_\alpha^* I_\alpha + \sum_1^\infty (G_\alpha^*)^n (G_\alpha)^n, \qquad H = m\sum H_\alpha . \tag{83}$$

From the orthogonality of the L_i, it follows as usual that

$$H_\alpha I_\alpha + I_\alpha^* H_\alpha = 2H_\alpha . \tag{84}$$

Obviously for $n \geqq 0$

$$(G_\alpha^*)^n (G_\alpha - I_\alpha)^* (G_\alpha - I_\alpha)(G_\alpha)^n \geqq 0 \tag{85}$$

or in other words,

$$(86) \quad (G_\alpha{}^*)^n(G_\alpha)^{n+1}+(G_\alpha{}^*)^{n+1}(G_\alpha)^n \leqq (G_\alpha{}^*)^{n+1}(G_\alpha)^{n+1}+(G_\alpha{}^*)^n(G_\alpha)^n ,$$

where $(G_\alpha)^0$ and $(G_\alpha{}^*)^0$ are to be interpreted as I_α and $I_\alpha{}^*$. Summing (86) from 0 to ∞,

$$(87) \qquad H_\alpha G_\alpha + G_\alpha{}^* H_\alpha \leqq 2H_\alpha .$$

From (82) we have

$$(88) \qquad P_\alpha = \lambda_\alpha I_\alpha - \tfrac{1}{2}\mathrm{Re}\lambda_\alpha\, G_\alpha ,$$

so that (84) and (87) yield

$$(89) \qquad H_\alpha P_\alpha + P_\alpha{}^* H_\alpha \leqq \mathrm{Re}\lambda_\alpha H_\alpha \leqq \tau(P) H_\alpha .$$

Summing on α and using orthogonality,

$$(90) \qquad HP+P^*H \leqq \tau(P)H .$$

The inequality $H \geqq I$ is (53), and we have now to estimate (Hv,v). This time there are three possibilities. The first two—to expand I_α and $G_\alpha{}^n$ as sums of resolvents, or to repeat the argument of Theorem I with appropriate changes—would be safe but tedious. Therefore we shall try to derive the estimate from Theorem I itself, using only some essential remarks about its proof. In fact, we now give a complete proof of the last step in Theorem III *without* using the H defined explicitly in (83), and then identify the new $\bar{H}$ with that H.

For a given positive integer k, let $w=e^{z/k}$, so that $\mathrm{Re}z>0 \Leftrightarrow |w|>1$. Then as in (73)

$$(91) \qquad \begin{aligned} |(wI-e^{P/k})^{-1}v| &= k|(zI-P+F_{k,z})^{-1}v| \\ &\leqq \frac{k|(zI-P)^{-1}v|}{1-|(zI-P)^{-1}|\,|F_{k,z}|} \\ &\leqq \frac{C(v)}{\mathrm{Re}z/k}\,\frac{1}{1-C|F_{k,z}|/\mathrm{Re}z} \\ &\leqq \frac{1}{|w|-1}\,\frac{C(v)}{1-C|F_{k,z}|/\mathrm{Re}z} , \end{aligned}$$

where we used $\mathrm{Re}u \leqq |e^u|-1$. Estimating the perturbation $F_{k,z}$,

$$(92) \qquad |F_{k,z}| = k|(e^{z/k}-1-z/k)I-(e^{P/k}-I-P/k)| = O(1/k)$$

as $k \to \infty$, uniformly for z in a compact set Z. If $\mathrm{Re}z>0$ in Z, we have

$$(93) \qquad C_k(v) = \sup\nolimits_Z \frac{C(v)}{1-C|F_{k,z}|/\mathrm{Re}z} \to C(v) \quad \text{as } k \to \infty .$$

We want to deduce from (91) that Morton's result (55) holds for

$A = e^{P/k}$, in the strong form

$$|e^{Pn/k}v| \leqq K_1(m)\, C_k(v) \quad \text{for } n \geqq 0, \quad |v| = 1\,. \tag{94}$$

Then because (iii) implies (v), Theorem I will provide an explicit $H_k \geqq I$ such that

$$(H_k v, v)^{\frac{1}{2}} \leqq K(m)\, K_1(m)\, C_k(v) \tag{95}$$

$$e^{P^*/k} H_k e^{P/k} \leqq \tfrac{1}{4}(1+\varrho(e^{P/k}))^2 H_k = \tfrac{1}{4}(1+e^{\tau(P)/k})^2 H_k\,. \tag{96}$$

As $k \to \infty$, some subsequence H_{k_j} converges to a limit $\bar{H} \geqq I$, with

$$(\bar{H}v, v)^{\frac{1}{2}} \leqq K(m)\, K_1(m)\, C(v) = K'(m)\, C(v)\,. \tag{97}$$

Expanding (96) in powers of k, subtracting H_k, multiplying by k, and taking the limit as $k_j \to \infty$, we get

$$\bar{H}P + P^*\bar{H} \leqq \tau(P)\bar{H}\,. \tag{98}$$

All this is justified if, in proving (94) by applying Morton's argument to $e^{P/k}$, we actually need the estimate (91) only for z in a compact set Z (independent of k) in the right half-plane. It turns out that this is indeed the case. Morton uses the resolvent condition only in the contour integrations (14*), where $w = e^{z/k}$ lies on one of the circles with

$$\begin{aligned} \text{radius} &= \delta_\alpha = 1 - e^{\operatorname{Re}\lambda_\alpha/k} \leqq -\operatorname{Re}\lambda_\alpha/k\,, \\ \text{center} &= (1+2\delta_\alpha)\, e^{i\operatorname{Im}\lambda_\alpha/k}\,. \end{aligned} \tag{99}$$

On this circle it is easy to bound z, independently of k, in terms of $\operatorname{Re}\lambda_\alpha$ and $\operatorname{Im}\lambda_\alpha$.

To make the identification $\bar{H} = H$, we want to match the clusters C_α' derived from P with the clusters C_α derived from $e^{P/k}$, k large. Clearly λ_α of maximum real part corresponds to $e^{\lambda_\alpha/k}$ of maximum modulus, and also the ratios which arise in forming clusters satisfy

$$\frac{1-|e^{\lambda_\alpha/k}|}{4m\,|e^{\lambda_i/k} - e^{\lambda_j/k}|} \to \frac{\operatorname{Re}\lambda_\alpha}{4m\,|\lambda_i - \lambda_j|} \quad \text{as } k \to \infty\,. \tag{100}$$

Therefore $\lambda_i \in C_\alpha'$ if and only if $e^{\lambda_i/k} \in C_\alpha$, k large, if we exclude eigenvalues of equal real part (which may make the choice of λ_α ambiguous) and also exclude the possibility that the limiting ratio in (100) is one.

With these exceptions,

$$B_\alpha = \sum \frac{2|e^{\lambda_i/k} - e^{\lambda_\alpha/k}|}{1-|e^{\lambda_\alpha/k}|} L_i \to G_\alpha = \sum \frac{2(\lambda_\alpha - \lambda_i)}{\operatorname{Re}\lambda_\alpha} L_i \tag{101}$$

and $\bar{H} = \lim H_k = H$. In the excluded cases, as in the case when (81) fails, the proper estimate for (Hv, v) follows by a continuity argument.

Repeating the proof in Section 5, we can describe a further property of H:

For some S with $S^*S = H$, $P' = SPS^{-1}$ is upper triangular, with $P'_{ij} = 0$ unless λ_i and λ_j are in the same cluster C_α', and $|P'_{ij}| \leqq \frac{1}{2}\min(-\mathrm{Re}\,\lambda_i, -\mathrm{Re}\,\lambda_j)$ whenever $i \neq j$.

There is one additional consequence of our method of proof which is significant in the applications to partial differential equations: *The conclusions in* (v) *and* (v′) *may be changed to*

$$A^*HA \leqq \tfrac{1}{2}(2-\theta+\theta\varrho(A))^2 H \quad \text{and} \quad HP+P^*H \leqq \theta\tau(P)H\,,$$

where $0 \leqq \theta < 2$ *and the constants* K *and* K' *depend on* θ *as well as* m.

For the proof, one divides our B_α and G_α by $2-\theta$, and alters the clusters so that (66) will hold; then the few remaining changes are straightforward. It follows that our space $L_2(H)$, over which (1) is to be well-posed, does not depend on the multiple of the identity which was subtracted in order to make $\sigma < 0$. In other words, the minimal renorming families $H(\omega)$ used to achieve (3) are equivalent for any two choices $\alpha > \sigma$.

7.

We want finally to extend Theorem I to apply to matrices such that $\varrho(A) = 1$ but A^n is unbounded; this occurs if and only if some eigenvalue of modulus one has a non-simple elementary divisor, and consequently too few corresponding eigenvectors. The standard example is

$$A_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

It is easy to see that all the conditions (i)–(v) fail for A_1, no matter how large $C(v)$ is chosen; in particular, $|A_1{}^n|$ grows like n and the resolvent has a double pole at $z = 1$. This relationship between the growth of A^n as $n \to \infty$ and that of the resolvent as $|z| \to 1$ is made precise by

THEOREM V. *There exist constants* $\alpha(s)$ *and* $\beta(s)$ *depending on* $s > 0$, *such that with* $A_\varepsilon = \varepsilon A$ *and the constant* $K(m)$ *as in Theorem I, each of the following statements implies the next:*

(i″) *For* $\frac{1}{2} < \varepsilon < 1$, $A_\varepsilon{}^* H_\varepsilon A_\varepsilon \leqq H_\varepsilon$ *for some* $H_\varepsilon \geqq I$ *with* $(H_\varepsilon v, v)^{\frac{1}{2}} \leqq C(v)/(1-\varepsilon)^s$ *for* $|v| = 1$.

(ii″) *For* $\frac{1}{2} < \varepsilon < 1$, $|S_\varepsilon A_\varepsilon S_\varepsilon{}^{-1}| \leqq 1$ *for some* S_ε *with* $|S_\varepsilon{}^{-1}| \leqq 1$ *and* $|S_\varepsilon v| \leqq C(v)/(1-\varepsilon)^s$ *for* $|v| = 1$.

(iii″) $|A^n v| \leqq \alpha(s)(n+1)^s C(v)$ *for* $n \geqq 0$ and $|v| = 1$.

(iv″) $|(zI-A)^{-1}v| \leqq \dfrac{\alpha(s)\,\beta(s)\,|z|^s C(v)}{(|z|-1)^{s+1}}$ for $|z|>1$ and $|v|=1$.

(v″) *For* $\frac{1}{2}<\varepsilon<1$, *there exists* $H_\varepsilon \geqq I$ *such that* $A_\varepsilon^* H_\varepsilon A_\varepsilon \leqq H_\varepsilon$ *and*

$$(H_\varepsilon v, v)^{\frac{1}{2}} \leqq \frac{\alpha(s)\,\beta(s)\,K(m)\,C(v)}{(1-\varepsilon)^s} \quad \text{for } |v| = 1\,.$$

PROOF. The first two conditions are equivalent as before with $H_\varepsilon = S_\varepsilon^* S_\varepsilon$. Given (ii″), we have for $|v|=1$

$$\text{(103)} \qquad \begin{aligned} |A_\varepsilon^n v| &\leqq C(v)/(1-\varepsilon)^s, \qquad \tfrac{1}{2}<\varepsilon<1\,, \\ |A^n v| &\leqq C(v)/\varepsilon^n(1-\varepsilon)^s\,. \end{aligned}$$

Maximizing the denominator with respect to ε,

$$\text{(104)} \qquad |A^n v| \leqq \alpha(s)(n+1)^s C(v), \qquad n \geqq 0\,.$$

It follows that (iv″) holds; for $|z|>1$,

$$\begin{aligned} |(zI-A)^{-1}v| &= \left|\sum_0^\infty \frac{A^n v}{z^{n+1}}\right| \leqq \alpha(s)\,C(v)\sum_0^\infty \frac{(n+1)^s}{|z|^{n+1}} \\ &\leqq \frac{\alpha(s)\,\beta(s)\,C(v)\,|z|^s}{(|z|-1)^{s+1}}\,. \end{aligned}$$

In order to apply Theorem I, we compute

$$\text{(106)} \qquad \begin{aligned} |(zI-A_\varepsilon)^{-1}v| &= \left|\frac{1}{\varepsilon}\left(\frac{z}{\varepsilon}I-A\right)^{-1} v\right| \\ &\leqq \frac{\alpha(s)\,\beta(s)\,C(v)\,|z/\varepsilon|^s}{\varepsilon(|z/\varepsilon|-1)^{s+1}} \\ &= \frac{\alpha(s)\,\beta(s)\,C(v)\,|z|^s}{(|z|-\varepsilon)^{s+1}} \\ &\leqq \frac{\alpha(s)\,\beta(s)\,C(v)}{|z|-1}\left(\frac{|z|}{|z|-\varepsilon}\right)^s \leqq \frac{\alpha(s)\,\beta(s)\,C(v)}{(|z|-1)(1-\varepsilon)^s}\,. \end{aligned}$$

Now the last step in Theorem I yields (v″).

Without writing down the obvious exponential analogue of Theorem V, we remark that it would yield an accurate estimate of the growth in time of $e^{P(\omega)t}$. Let us refer to the example (6), in which

$$e^{P(\omega)t} = \begin{pmatrix} 1 & 0 \\ i\omega t & 1 \end{pmatrix}$$

grows linearly ($s=1$), while the estimate (10) grows exponentially. To avoid such an overestimate, one has to study the system in a sequence of norms

$$H_\varrho(\omega) = \begin{pmatrix} 1+\varrho\omega^2 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \varrho \to \infty .$$

From the inequalities

$$\|u(t)\|_{H_\varrho} \leqq e^{t/2\varrho} \|f\|_{H_\varrho}$$

one can in fact recover the linear growth; the analogue of Theorem V provides a corresponding sequence of norms for an arbitrary system satisfying the Petrowsky–Gårding condition.

REFERENCES

1. G. Birkhoff, *Well-set Cauchy problems and C_0-semigroups*, J. Math. Anal. Appl. 8 (1964), 303–324.
2. S. R. Foguel, *A counterexample to a problem of Sz.-Nagy*, Proc. Amer. Math. Soc. 15 (1964), 788–790.
3. P. R. Halmos, *On Foguel's answer to Nagy's question*, Proc. Amer. Math. Soc. 15 (1964), 791–793.
4. H.-O. Kreiss, *Über die Stabilitätsdefinition für Differenzengleichungen die partielle Differentialgleichungen approximieren*, BIT 2 (1962), 153–181.
5. H.-O. Kreiss, *Über Matrizen die beschränkte Halbgruppen erzeugen*, Math. Scand. 7 (1959), 71–80.
6. C. A. McCarthy and J. T. Schwartz, *On the norm of a finite Boolean algebra of projections, and applications to theorems of Kreiss and Morton*, Comm. Pure Appl. Math. 18 (1965), 191–202.
7. S. Mizohata, *Systèmes hyperboliques*, J. Math. Soc. Japan 11 (1959), 205–233.
8. K. W. Morton, *On a matrix theorem due to H.-O. Kreiss*, Comm. Pure Appl. Math. 17 (1965), 375–380.
9. K. W. Morton and S. Schechter, *On the stability of finite difference matrices*, SIAM J. Numer. Anal. Ser. B. 2 (1965), 119–128.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MASSACHUSETTS, U.S.A.