ON THE DIAGONALIZATION OF A CERTAIN CLASS OF OPERATORS

A. Hinkkanen

1. Introduction and results. We consider infinite dimensional matrices of bounded linear operators acting on a complex separable Hilbert space $\mathcal{H}$. The space $\mathcal{H}$ is assumed to have a suitable orthonormal basis, with respect to which the matrix elements are defined. We shall also perform formal operations on the matrix elements to form new matrices, which do not necessarily correspond to bounded operators without extra assumptions. Unless it is clear enough from the context, we will indicate if a matrix is related to an operator or if it should be understood only as a formal array of complex numbers.

First we give some definitions. Let $H = (h_{ij})$, $i, j = 1, 2, \ldots$, be an infinite-dimensional matrix of an operator. We are concerned about finding, under suitable assumptions, an invertible matrix U such that UHU^{-1} is diagonal. We look for U in the form $U = \exp A$. If H is hermitian, then A will be antihermitian and U unitary.

We write H^* for the adjoint of H, that is, $H_{ij}^* = \bar{h}_{ji}$. We define the matrix H^D by

$$H_{ij}^D = \begin{cases} h_{ij}, & i = j, \\ 0, & i \neq j, \end{cases}$$

and set $H^K = H - H^D$. If B is a matrix and $b_{ii} \neq b_{jj}$ for $i \neq j$, we define H/B as the formal matrix whose entries are given by

$$(H/B)_{ij} \equiv \left(\frac{H}{B}\right)_{ij} = \frac{h_{ij}}{b_{ii} - b_{jj}}, \quad i \neq j,$$
$$= 0, \quad i = j.$$

We note that H/B does not depend on B^K or on H^D.

The commutator of the matrices A and B is $[A, B] = AB - BA$, provided that the products AB and BA are well defined. We have

(1.1) $$[H/B, B^D] = -H^K.$$

The Hilbert–Schmidt norm $\|H\|_2$ of H is given by

$$\|H\|_2^2 = \sum_{i,j=1}^{\infty} |h_{ij}|^2,$$

and the operator norm $\|H\|$ by

$$\|H\| = \sup\{\|Hx\| \mid \|x\| = 1\}.$$

Here $\|x\|^2 = \sum_{i=1}^{\infty} |x_i|^2$ for $x \in \mathcal{H}$.

We prove the following theorems.

THEOREM 1. *Assume that for all* i, j,

(1.2) $|h_{j+1,j+1}| \le \rho |h_{jj}|$

and

(1.3) $|h_{ij}|^2 \le |h_{ii}h_{jj}|s^2(ij)^{-2}$ $(i \ne j)$

hold for some ρ $(0 < \rho < 1)$ *and some* s,

(1.4) $0 < s \le (0.03)(1-\rho)$.

Then the sequence A_n *of matrices defined by* $A_0 = 0$,

(1.5) $A_{n+1} = A_n + (e^{A_n}He^{-A_n})/H^D$, $n \ge 0$,

converges in the Hilbert-Schmidt norm to a matrix A *such that with* $U = \exp A$, *the matrix* UHU^{-1} *is diagonal.*

THEOREM 2. *Assume that* (1.2) *holds for some* ρ, $0 < \rho < 1$, *and that*

(1.6) $|h_{ij}|^2 \le |h_{ii}h_{jj}|s^2(i-j)^{-4}$ $(i \ne j)$

holds for some positive s *with*

(1.7) $s \le (1-\rho)/200$.

Then A_n *given by* (1.5) *converges in the operator norm to a matrix* A *such that* $e^A He^{-A}$ *is diagonal.*

Let us compare our results to those in Kato's book [4, pp. 74–84, 371–373]. There it is considered what happens to a single, possibly multiple, eigenvalue in a perturbation of the matrix. One gets an expansion for the new eigenvalue and eigenvector in terms of the initial matrix and the perturbation parameter, say b, and finitely many such series can be made to converge simultaneously in b under suitable circumstances. In our case, we could take H^D to be the initial matrix and $H^K = bV$ the perturbation, where V is the matrix of a suitable bounded operator acting on the Hilbert space, to obtain expansions for finitely many eigenvalues for $|b| < b_0$. What we have done in Theorems 1 and 2 is to make strong assumptions about the elements of H^K and to require the eigenvalues of H^D to be simple and well separated. As we shall see after stating Theorem 3, conditions of that kind are not unreasonable in practice. Then we have obtained an expansion for the *whole* matrix A such that with $U = \exp A$ the matrix UHU^{-1} is diagonal, and the expansion converges to A in some norm in a space of matrices (or operators). This also gives an expansion for UHU^{-1} (cf. the end of Section 2).

The results of R. Turner that generalize earlier results of K. Friedrichs in [3, Theorem 2, p. 2453] (see also [7]) are somewhat analogous to our Theorems 1 and 2. It follows from (1.2) and (1.3) that $\|H/H^D\|_2 < \infty$. In [3], a condition is considered which, under the assumption (1.2), means that $\|H/H^D\|_2$ is small enough. Then it is shown that there exists an operator U such that UHU^{-1} is diagonal in the same basis as is H^D. The purpose in [3] is to prove in this way

that H is a spectral operator of scalar type, which is also a consequence of our Theorems 1 and 2, while our main purpose is to construct an explicit diagonalization. One may further note that in Theorem 2 we have assumptions guaranteeing that $\|H/H^D\|$ is finite, while $\|H/H^D\|_2$ may be infinite.

We will also consider a problem in function theory. Let G be a bounded simply connected domain in the plane, and let $B^2(G)$ denote the Hilbert space of all $L^2(G)$-functions (with respect to the area measure) which are analytic in G (that is, the Bergman space of G). Assume that $z_i \in G$ and $\Delta_i > 0$ are such that the z_i are distinct, $z_i \to w \in G$, $|z_i - w| \le r$ for a suitable $r > 0$, and $\sum \Delta_i < \infty$. For functions ϕ, ψ in $B^2(G)$ we define

$$(1.8) \qquad [\phi, \psi] = \sum_{i=1}^{\infty} \overline{\phi(z_i)} \, \psi(z_i) \Delta_i.$$

Let (ϕ_i) be a suitable orthonormal basis of $B^2(G)$, to be specified in Section 4. Then the matrix H given by

$$(1.9) \qquad h_{mn} = [\phi_m, \phi_n]$$

is hermitian, and $\|H\|_2 < \infty$. There exists a unitary matrix (and operator) U such that the matrix UHU^* is diagonal. We define

$$(1.10) \qquad \psi_k = \sum_j \overline{u_{kj}} \, \phi_j,$$

so that (ψ_k) also is an orthonormal basis for $B^2(G)$.

THEOREM 3. *Suppose that* $f \in B^2(G)$. *Under the above assumptions, we have*

$$(1.11) \qquad f(z) = \sum_{k=1}^{\infty} \frac{[\psi_k, f]}{[\psi_k, \psi_k]} \psi_k(z), \quad z \in G.$$

We note that the right-hand side of (1.11) depends on the values of f at z_i, but not elsewhere.

In Theorem 3 we consider the problem of determining $f(z)$ for $z \in G$, if $f(z_i)$ is known for a sequence z_i converging to an inner point of G. In principle, the $f(z_i)$ determine f uniquely. We have reduced the problem to the diagonalization of a hermitian matrix H that can be diagonalized by the general theory of completely continuous operators and that satisfies (1.2) for some ρ ($0 < \rho < 1$) but neither (1.3) nor (1.6). Hence we cannot obtain an explicit solution to this problem by Theorem 1 or 2.

We may note that H satisfies $|h_{ij}|^2 \le h_{ii} h_{jj}$ (cf. (4.1)), which is a much weakened form of (1.6). Since the proofs of Theorems 1 and 2 do not take into account possible cancellations in the expansion for A, there might be a chance that by more accurate estimates one could substantially weaken the conditions for H in Theorem 2, making them applicable to the situation of Theorem 3.

In Section 2 we describe the formal series resulting from (1.5), and in Section 3 we prove Theorems 1 and 2. In Section 4 we prove Theorem 3.

2. The series of Theorem 1.
The definition (1.5) can be motivated as follows. Given H, our aim is to find a matrix A such that

(2.1) $(e^A H e^{-A})^K = 0.$

Using the expansion [2, p. 49],

(2.2) $e^A H e^{-A} = \dfrac{1}{0!} H + \dfrac{1}{1!} [A, H] + \dfrac{1}{2!} [A, [A, H]] + \cdots$

we can write (2.1) in the form

$$-[A, H^D] = (e^A H e^{-A})^K - [A, H^D]$$

(2.3)

$$= \left(H + [A, H^K] + \dfrac{1}{2!} [A, [A, H]] + \cdots \right)^K,$$

noting that $[A, H^D]^D = 0$. We form an iterative sequence by $A_0 = 0$ and

$$-[A_{n+1}, H^D] = (e^{A_n} H e^{-A_n})^K - [A_n, H^D], \quad n \ge 0.$$

But by (1.1) this has the solution

$$A_{n+1} = (e^{A_n} H e^{-A_n}) / H^D + A_n$$

(2.4)

$$= \left(H + [A_n, H^K] + \dfrac{1}{2!} [A_n, [A_n, H]] + \cdots \right) \Big/ H^D,$$

which is (1.5). The point of using (2.3) and (2.4) is that in Theorems 1 and 2, H^D is in a certain sense large if compared to H^K, and neither term H^D nor $[A_n, H^D]$ appears on the right-hand side of (2.4).

In this section we do not deal with questions of convergence, and all the series should be considered as formal. It follows by induction from (2.4) that, as a formal series, A_n is the sum of all terms of the following kind. We define terms of level $m \ge 1$ so that the only term of level 1 is H/H^D, and if $m \ge 2$ then a term T of level m is obtained as follows. Take any $k \ge 1$ and any terms $C_1, C_2, \ldots, C_k$ of level $\le m - 1$, at least one C_i being of level $m - 1$. Then set

$$T = \dfrac{1}{k!} [C_1, [C_2, \ldots, [C_k, H] \ldots]/H^D,$$

where H is replaced by H^K if $k = 1$. Then A_n is the sum of all terms of level $\le n$. In particular, if $m > n$, then A_m contains all the terms in A_n. This fact, which is perhaps not an obvious consequence of (2.4), will be used in Section 3.

Let us form all the terms of all levels $m \ge 1$ as above, except that for each term the last division by H^D is omitted. Let C be the sum of these terms, so that $A = C/H^D$. This implies that $e^A H e^{-A} - [A, H^D] = C$ as a formal series, because of the way C is obtained. Hence by (1.1),

$$(e^A H e^{-A})^K = [A, H^D] + C^K = [C/H^D, H^D] + C^K = -C^K + C^K = 0,$$

as required. In fact, $e^A H e^{-A} = C^D$.

3. Proof of Theorems 1 and 2.

Proof of Theorem 1. The assumptions (1.2) and (1.3) serve mainly two purposes: they guarantee that $\|H/H^D\|_2 < \infty$, and that the upper bound in (1.3) is convenient for matrix multiplication. First we prove a lemma.

LEMMA 1. *Assume that*

(3.1)
$$|p_{ij}|^2 \le |h_{ii}h_{jj}|s^2s_1^2(ij)^{-2}, \quad i \ne j,$$
$$|t_{ij}|^2 \le |h_{ii}h_{jj}|s^2s_2^2(ij)^{-2},$$

where s satisfies (1.4) *and* s_1 *and* s_2 *are positive. Then for* $Q = [P/H^D, T]$, *we have*

(3.2)
$$|q_{ij}|^2 \le |h_{ii}h_{jj}|s^2r^2s_1^2s_2^2(ij)^{-2},$$

where $r = 0.1$.

We may assume that $s_1 = s_2 = 1$. We estimate $((P/H^D)T)_{ik}$. We note that $|h_{ii} - h_{jj}| \ge ||h_{ii}| - |h_{jj}||$ and obtain

$$\left| \left(\left(\frac{P}{H^D} \right) T \right)_{ik} \right| \le \sum_{j \ne i} \left| \frac{p_{ij}t_{jk}}{h_{ii} - h_{jj}} \right|$$

$$\le \sum_{j < i} \frac{\sqrt{|h_{ii}h_{kk}|}}{ikj^2} \frac{s^2|h_{jj}|}{|h_{jj}|(1 - \rho^{i-j})} + \sum_{j > i} \frac{\sqrt{|h_{ii}h_{kk}|}}{ikj^2} \frac{s^2|h_{jj}|}{|h_{ii}|(1 - \rho^{j-i})}$$

$$\le \frac{\sqrt{|h_{ii}h_{kk}|}\,s}{ik(1 - \rho)} \frac{s\pi^2}{6} \le \frac{\sqrt{|h_{ii}h_{kk}|}\,sr}{2ik},$$

since $s\pi^2/(6(1 - \rho)) < r/2$ by (1.4). Similarly, we also have

$$|(T(P/H^D))_{ik}| \le \sqrt{|h_{ii}h_{kk}|}\,sr(2ik)^{-1},$$

so that (3.2) follows.

3.1. Now we turn to the proof of Theorem 1. With $r = 0.1$, we define

$$f(x) = (1 + r^{-1})(e^{rx} - 1) - x$$
$$= rx + (rx)^2/2! + \cdots + rx^2/2! + r^2x^3/3! + \cdots.$$

Further we define $f_1(r) = 0$ and

$$f_{n+1}(r) = f(1 + f_n(r)), \quad n \ge 1.$$

We can write $A_n = P_n/H^D$, where $P_n^D = 0$. Now (2.4) gives

$$P_{n+1} = \left(H + \frac{1}{1!}[A_n, H^K] + \frac{1}{2!}[A_n, [A_n, H]] + \frac{1}{3!}[A_n, [A_n, [A_n, H]]] + \cdots \right)^K.$$

Writing $H = H^K + H^D$ we write each term, starting with the second-order term, as the sum of two terms; for example,

$$[A_n, [A_n, H]] = [A_n, [A_n, H^K]] + [A_n, [A_n, H^D]].$$

By (1.1) we have

$$[A_n, H^D] = [P_n/H^D, H^D] = -P_n^K = -P_n.$$

Using this and substituting P_n/H^D for each remaining A_n, we obtain

(3.3)
$$P_{n+1} = H^K + [P_n/H^D, H^K]^K + \frac{1}{2!}[P_n/H^D, [P_n/H^D, H^K]]^K + \cdots +$$
$$- \frac{1}{2!}[P_n/H^D, P_n]^K - \frac{1}{3!}[P_n/H^D, [P_n/H^D, P_n]]^K - \cdots.$$

Alternatively, we can derive (3.3) from (1.5), (2.2) and (1.1). Then the term P_n given by (1.5) is cancelled by the term $-P_n$ arising from (2.2) and (1.1).

We make the induction assumption

$$(3.4) \qquad |(P_n)_{ik}| \le \sqrt{|h_{ii}h_{kk}|}\, s(1+f_n(r))(ik)^{-1}, \quad i \ne k,$$

which is true for $n = 1$. Applying Lemma 1 to (3.3) we obtain for $i \ne k$,

$$|(P_{n+1})_{ik}| \le \sqrt{|h_{ii}h_{kk}|}\, s(ik)^{-1}(1+r(1+f_n(r)))$$

$$+\frac{1}{2!}r^2(1+f_n(r))^2 + \cdots + \frac{1}{2!}r(1+f_n(r))^2$$

$$+\frac{1}{3!}r^2(1+f_n(r))^3 + \cdots)$$

$$= \sqrt{|h_{ii}h_{kk}|}\, s(ik)^{-1}(1+f(1+f_n(r)))$$

$$= \sqrt{|h_{ii}h_{kk}|}\, s(ik)^{-1}(1+f_{n+1}(r)).$$

This proves (3.4) for all $n \ge 1$.

Write $g(x) = f(1+x) - x$. Then $g''(x) > 0$ for $x > 0$ and $g(0) = f(1) > 0$. We have $g'(x) < 0$ for $0 \le x < x_1$ and $g'(x) > 0$ for $x > x_1$, where $x_1 = r^{-1}\log[2/(r+1)] - 1$. Since $g(x_1) < 0$, g has one zero between 0 and x_1, say x_0, and $g(x) > 0$ for $0 \le x < x_0$.

Write $h(x) = f(1+x)$. Then $h'(x) > 0$ for $x \ge 0$, so that h is strictly increasing and maps $[0, x_0)$ onto $[f(1), x_0)$, since $f(1+x_0) = x_0$. We have $f_1(r) = 0 \in [0, x_0)$, and if $f_n(r) \in [0, x_0)$ then $f_{n+1}(r) = f(1+f_n(r)) > f_n(r)$, and $f_{n+1}(r) \in [0, x_0)$. Thus the sequence $f_n(r)$ is increasing and bounded from above. Hence $f_n(r)$ has a limit, which, in fact, equals $x_0 \approx 0.203$. In any case, for every $\epsilon > 0$ there is an n_ϵ such that, for $m, n \ge n_\epsilon$,

$$(3.5) \qquad |f_m(r) - f_n(r)| < \epsilon.$$

3.2. Next we want to show that for every $\epsilon > 0$ there exists n_ϵ such that if $m, n \ge n_\epsilon$ then, for all i, j,

$$(3.6) \qquad |(A_m - A_n)_{ij}| \le \sqrt{|h_{ii}h_{jj}|}\, s(ij|h_{ii} - h_{jj}|)^{-1}\epsilon.$$

But this follows from $A_n = P_n/H^D$ and (3.5), when we take into account that the estimates obtained in the previous subsection are for the terms of level at most n separately. Thus we actually have derived the upper bound

$$|(P_m - P_n)_{ij}| \le \sqrt{|h_{ii}h_{jj}|}\, s(ij)^{-1}(1+f_m(r) - 1 - f_n(r)).$$

If, for example, $m > n$, then the series for P_m (or $f_m(r)$) contains all the terms of P_n (or $f_n(r)$), so that the remaining terms in $f_m(r)$ form an upper bound for the remaining terms in P_m. This is because for all terms C in P_m there is a term t_C in $f_m(r)$ such that, for $i \ne j$, $|C_{ij}| \le \sqrt{|h_{ii}h_{jj}|}\, s(ij)^{-1}t_C$.

Since

$$\sum_{ij} \left(\frac{\sqrt{|h_{ii}h_{jj}|}\, s}{|h_{ii} - h_{jj}|ij} \right)^2 \equiv N^2 < \infty,$$

(A_n) is a Cauchy sequence in the Hilbert–Schmidt norm, converging to a Hilbert–Schmidt operator A.

3.3 From (3.6) we deduce that

(3.7) $$|(A - A_n)_{ij}| \le \sqrt{|h_{ii} h_{jj}|}\, s(ij|h_{ii} - h_{jj}|)^{-1}\epsilon$$

for $n \ge n_\epsilon$. Since $\|A_{n+1} - A_n\|_2 \to 0$ we have, by (1.5),

$$\|(e^{A_n} H e^{-A_n})/H^D\|_2 \to 0.$$

We want to prove

(3.8) $$\|(e^{A_n} H e^{-A_n})/H^D - (e^A H e^{-A})/H^D\|_2 \to 0,$$

since this then implies $\|(e^A H e^{-A})/H^D\|_2 = 0$, or $(e^A H e^{-A})^K = 0$.

To prove (3.8), write

$$E_n = e^{A_n} H e^{-A_n} - e^A H e^{-A}$$

$$= [A_n - A, H] + \frac{1}{2!}([A_n, [A_n, H]] - [A, [A, H]]) + \cdots.$$

To estimate each difference

$$[A_n, [A_n, \dots, [A_n, H] \dots] - [A, [A, \dots, [A, H] \dots]$$

we rewrite it as a sum of differences changing each time one A_n to A. Thus, for example,

$$[A_n, [A_n, H]] - [A, [A, H]] = [A_n, [A_n, H]] - [A_n, [A, H]]$$

$$+ [A_n, [A, H]] - [A, [A, H]]$$

$$= [A_n, [A_n - A, H]] + [A_n - A, [A, H]].$$

Then we use (3.7) together with

$$|(A_n)_{ij}| \le \sqrt{|h_{ii} h_{jj}|}\, s(1 + x_0)(ij|h_{ii} - h_{jj}|)^{-1},$$

$i \ne j$, $1 + x_0 \approx 1.203$, which is a consequence of (3.4) and the discussion thereafter. The same estimate holds for A. By writing $A_n = P_n/H^D$, $A = P/H^D$, $P_n^D = P^D = 0$, and applying Lemma 1 we thus obtain, for $n \ge n_\epsilon$,

$$|(E_n)_{ij}| \le \epsilon \sqrt{|h_{ii} h_{jj}|}\, s(ij)^{-1}$$

$$\cdot \left[\sum_{k=1}^\infty \frac{k}{k!} r^k (1 + x_0)^{k-1} + 1 + \sum_{k=2}^\infty \frac{k}{k!} r^{k-1}(1 + x_0)^{k-1} \right].$$

The expression in brackets equals

$$r e^{r(1 + x_0)} + 1 + e^{r(1 + x_0)} - 1 \equiv M.$$

Hence $\|E_n/H^D\|_2 \le \epsilon MN$, which proves (3.8).

Theorem 1 is proved.   □

REMARK. Since eigenvectors are determined only up to a non-zero complex coefficient, one must use some normalization to get unique eigenvectors. In

Theorem 1, such a condition is $(A/H^D)^D = 0$ for any A. (Note that to solve $-[B, H^D] = A^K$ for B, we can choose B^D arbitrarily. For simplicity, we have taken $B^D = 0$.)

3.4 Proof of Theorem 2. We make use of the following result of Schur.

LEMMA 2 [6, p. 6]. *If we have*

$$\sum_{i=1}^{\infty} |p_{ij}| \le M_1 \quad \text{for all } j$$

and

$$\sum_{j=1}^{\infty} |p_{ij}| \le M_2 \quad \text{for all } i,$$

then

(3.9) $\|P\|^2 \le M_1 M_2.$

The following lemma can be proved in the same way as Lemma 1, by using also the fact that

$$\sum_{j=i+1}^{k-1} (i-j)^{-2}(j-k)^{-2} \le L^{-2}[\pi^2/3 + 4L^{-1}(1+\log(L-1))] \le L^{-2}(\pi^2/3 + 12/5),$$

where $L = k - i \ge 2$.

LEMMA 3. *Assume that*

(3.10)
$$|p_{ij}|^2 \le |h_{ii} h_{jj}| s^2 (i-j)^{-4}, \quad i \ne j,$$
$$|t_{ij}|^2 \le |h_{ii} h_{jj}| s^2 \min(1, (i-j)^{-4}),$$

where s satisfies (1.7). *Then for $Q = [P/H^D, T]$, we have*

(3.11) $|q_{ij}|^2 \le |h_{ii} h_{jj}| s^2 r^2 \min(1, (i-j)^{-4}),$

where $r = 0.1$.

It is this lemma that determines the condition (1.7).

Now the proof of Theorem 2 can proceed in the same way as that of Theorem 1 after Lemma 1. The only difference is that the finiteness of the Hilbert-Schmidt norm of H/H^D, or any matrix P/H^D where P satisfies (3.1), is replaced by the finiteness of the operator norm of any matrix P/H^D where P satisfies (3.10).
Namely, if (3.10) holds, then

$$\sum_{\substack{j=1 \\ j \ne i}}^{\infty} \left| \frac{p_{ij}}{h_{ii} - h_{jj}} \right| \le \sum_{j<i} \frac{s}{1-\rho} \frac{1}{(i-j)^2} \left| \frac{h_{ii}}{h_{jj}} \right|^{1/2} + \sum_{j>i} \frac{s}{1-\rho} \frac{1}{(j-i)^2} \left| \frac{h_{jj}}{h_{ii}} \right|^{1/2}$$

$$\le \frac{s\rho\pi^2}{3(1-\rho)}$$

for all i, and

$$\sum_{\substack{i=1 \\ i \neq j}}^{\infty} \left| \frac{p_{ij}}{h_{ii} - h_{jj}} \right| \leq \frac{s\rho\pi^2}{3(1-\rho)}$$

for all j. Thus Lemma 2 implies $\|P/H^D\| < \infty$. This concludes the proof of Theorem 2. $\qquad\square$

3.5. AN EXAMPLE. Since we can have $h_{ij} = 0$ for $i > n$ or $j > n$, Theorems 1 and 2 can be applied also to finite-dimensional matrices. We take $n = 2$ and set

$$H = \begin{pmatrix} \alpha & C \\ D & \beta \end{pmatrix}, \qquad A = \begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix}.$$

Since

$$\exp \begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix} = \begin{pmatrix} \cosh \sqrt{ab} & \sqrt{a/b}\ \sinh \sqrt{ab} \\ \sqrt{b/a}\ \sinh \sqrt{ab} & \cosh \sqrt{ab} \end{pmatrix},$$

a straightforward calculation shows that $(UHU^{-1})^K = 0$ $(U = \exp A)$ implies

(3.12) $a = \sqrt{C/D}\ \frac{1}{2} \arctan 2\sqrt{CD}/(\alpha - \beta)$,

(3.13) $b = \sqrt{D/C}\ \frac{1}{2} \arctan 2\sqrt{CD}/(\alpha - \beta)$.

Note that $\sqrt{a/b}\ \sinh \sqrt{ab}$ and $\sqrt{b/a}\ \sinh \sqrt{ab}$ are well-defined by the power series, and that the same applies to a and b given by (3.12–3.13) provided that $2|CD|^{1/2} < |\alpha - \beta|$.

The inspection of formula (2.4) gives the same result up to the coefficients in the series (3.12–3.13), which are difficult to calculate because of the complicated structure of (2.4).

We observe that if three of the α, β, C, D are fixed and the fourth varies, the a and b are analytic functions of the fourth. If $|C|^2 + |D|^2 > 0$, then H can be made diagonal if and only if

(3.14) $(\alpha - \beta)^2 + 4CD \neq 0.$

Thus the values of the fourth variable that do not satisfy (3.14) for the fixed three parameters are exactly the essential singularities of a and b.

4. Proof of Theorem 3. We denote by $B^2(G)$ the Hilbert space of analytic functions in G for which

$$\|f\|_2^2 = \int_G |f|^2\, dA < \infty.$$

Here dA is the area measure. The inner product is

$$(f, g) = \int_G \bar{f} g\, dA.$$

We remark that we do not associate any Hilbert space with the scalar product (1.8).

If F maps G conformally onto the unit disk D, and if $F(w) = 0$, then the functions

$$\phi_j(z) = (j/\pi)^{1/2} F(z)^{j-1} F'(z), \quad j \geq 1,$$

are a complete orthonormal system in $B^2(G)$ [1, p. 12].

Assume that $2|F'(w)| = M > 0$, and that $R < 1/\sqrt{2}$. We can find r, $0 < r < 1$, such that $\{|z - w| \leq r\} \subset G$ and that $|F'(z)| \leq M$ and $|F(z)| \leq R$ for $|z - w| \leq r$. We can assume that $|z_i - w| \leq r$ for $i \geq 1$, since this can be achieved by removing a finite number of points z_i.

Clearly the matrix H given by (1.9) is hermitian. Moreover, we have

(4.1) $$|h_{mn}|^2 \leq h_{mm} h_{nn}$$

by Schwarz's inequality, and

$$h_{n+1,n+1} = \sum_k \frac{n+1}{\pi} |F'(z_k)|^2 |F(z_k)|^{2n} \Delta_k$$

$$\leq 2R^2 \sum_k \frac{n}{\pi} |F'(z_k)|^2 |F(z_k)|^{2(n-1)} \Delta_k$$

$$= 2R^2 h_{nn}.$$

Therefore

$$\sum_{mn} |h_{mn}|^2 \leq \left(\sum_m h_{11} (2R^2)^{m-1} \right)^2 = h_{11}^2 (1 - 2R^2)^{-2} < \infty,$$

since $2R^2 < 1$.

Now it follows from [5, p. 16] that there exists a unitary matrix U such that UHU^* is diagonal. Then the functions $\psi_k = \sum_j \bar{u}_{kj} \phi_j$ form a complete orthonormal system in $B^2(G)$.

Next we calculate $[\psi_k, \psi_l]$ for $k \neq l$. By [1, p. 23] there is a constant C such that $|g(z)| \leq C \|g\|_2$, $g \in B^2(G)$, for $|z - w| \leq r$. We have

$$[\psi_k, \psi_l] = \left[\sum_{j=1}^J \bar{u}_{kj} \phi_j + \sum_{j=J+1}^\infty \bar{u}_{kj} \phi_j, \sum_{m=1}^M \bar{u}_{lm} \phi_m + \sum_{m=M+1}^\infty \bar{u}_{lm} \phi_m \right].$$

If $|g_i(z)| \leq Y_i$ for $|z - w| \leq r$, $i = 1, 2$, then $|[g_1, g_2]| \leq Y_1 Y_2 \sum \Delta_i$. Since, for $|z - w| \leq r$,

$$\left| \sum_{j=J}^\infty \bar{u}_{kj} \phi_j(z) \right|^2 \leq C^2 \sum_{j=J}^\infty |u_{kj}|^2 \to 0$$

as $J \to \infty$, we obtain

$$[\psi_k, \psi_l] = \sum_{jm} u_{kj} h_{jm} u_{ml}^* = 0, \quad k \neq l,$$

by the definition of U.

We can write

(4.2) $$f(z) = \sum_{k=1}^\infty a_k \psi_k(z), \quad z \in G,$$

with $\sum |a_k|^2 < \infty$ [1, p. 7]. It remains to determine the a_k. We have

$$[\psi_m, f] = a_m[\psi_m, \psi_m] + \left[\psi_m, \sum_{k=K}^{\infty} a_k \psi_k\right]$$

for any $K > m$. We note that since $\|\psi_m\|_2 = 1 > 0$, the function ψ_m cannot vanish at every z_i, so that $[\psi_m, \psi_m] > 0$.

We have

$$\left|\left[\psi_m, \sum_{k=K}^{\infty} a_k \psi_k\right]\right|^2 \le C^4 \left(\sum_{k=K}^{\infty} |a_k|^2\right)\left(\sum \Delta_i\right)^2 \to 0$$

as $K \to \infty$. Therefore $a_m = [\psi_m, f]/[\psi_m, \psi_m]$, so that (4.2) now implies (1.11). Theorem 3 is proved.   □

REFERENCES

1. S. Bergman, *The kernel function and conformal mapping*, Amer. Math. Soc., New York, N.Y., 1950.
2. J. D. Bjorken and S. D. Drell, *Relativistic quantum mechanics*, McGraw-Hill, New York, 1964.
3. N. Dunford and J. T. Schwartz, *Linear operators* III, Interscience, New York, 1971.
4. T. Kato, *Perturbation theory for linear operators*, 2nd edition, Springer, Berlin, 1980.
5. R. Schatten, *Norm ideals of completely continuous operators*, Springer, Berlin, 1970.
6. I. Schur, *Bemerkungen zur Theorie der beschränkten Bilinearformen mit unendlich vielen Veränderlichen*, J. Reine Angew. Math. 140 (1911), 1–28.
7. R. Turner, *Perturbation of compact spectral operators*, Comm. Pure Appl. Math. 18 (1965), 519–541.

Department of Mathematics
Imperial College
London SW7 2BZ, U.K.

Current address:

Department of Mathematics
University of Michigan
Ann Arbor, Michigan 48109–1003