VARIOUS AVERAGING OPERATIONS ONTO SUBALGEBRAS[1]

BY

CHANDLER DAVIS[2]

The operations dealt with in this paper are functions defined on a self-adjoint algebra over the complex numbers and with values in a subalgebra of it. (The subalgebra is arbitrary.) They are called "averaging" because the formal properties which define them (Definition 2.1) are possessed by the conditional expectations of I. E. Segal.[3] If the subalgebra is just the constants, the averaging operations are exactly the states; if the subalgebra is the whole algebra, the only averaging operation is the identity. In case the whole algebra is commutative, the study of averagings has been carried very far by G. Birkhoff [2] and J. L. Kelley [10].

Another class of operations (for which I claim no novelty except the name) is introduced, and elementary properties set forth, in §3. I had begun studying this class for somewhat different reasons, but it turns out to be closely related to the averaging operations—indeed, to be a subclass, and that subclass which behaves least like the classical conditional expectations which occur in the abelian case. The main result of this paper is the expression (in §4) of an arbitrary averaging operation in terms of these two special types.

The effect of averaging operations on the spectrum of a hermitian operator is the subject of §§6–7. Theorem 7.2, a simple extension of a theorem of Hardy, Littlewood, and Pólya, may have independent interest.

All the algebras in this paper are finite-dimensional. Extension of some of the results to arbitrary von Neumann algebras is the idea which leads me often to express things in terms of algebras, commutors, etc., when some proofs would be a little shorter using only matrices. [*Added March* 31, 1959. The program of characterizing noncommutative conditional expectations, in analogy to the work of Birkhoff, Moy, and others in the commutative case, was initiated by M. Nakamura and his colleagues; see especially [13]. I regret that I was in ignorance of this work when I wrote the present paper. Their results deal with the infinite-dimensional case; they do not seem to contain my main results here as specializations.]

I thank P. R. Halmos for several helpful conversations on the subject of this paper.

Received July 14, 1958.

[1] Presented to the American Mathematical Society, August 27, 1958.

[2] National Science Foundation Fellow.

[3] Further justification for the terminology appears later, e.g., Definition 3.4, Theorem 6.2, Theorem 4.1. No doubt "*positive* averaging operation" would be a better term; or "abstract expectation".

1. Some notations

Throughout the paper, $\mathfrak{A}$ denotes a finite-dimensional self-adjoint complex algebra with elements A, B, $\cdots$. There will usually be no objection to considering a particular faithful representation of $\mathfrak{A}$ by operators on a hilbert space of finite dimensionality n, and assuming the representation is of multiplicity 1.

For projection P, $\tilde{P} = 1 - P$.

If $\mathfrak{R}$ is a subalgebra (i.e., self-adjoint subalgebra) of $\mathfrak{A}$, $\mathfrak{R}'$ is the algebra of all those elements of $\mathfrak{A}$ which commute with every element of $\mathfrak{R}$ (the *commutor* of $\mathfrak{R}$ relative to $\mathfrak{A}$). Regard this as an operation on subalgebras of $\mathfrak{A}$ to subalgebras of $\mathfrak{A}$. Subalgebras of $\mathfrak{A}$, made into a partly ordered set by the relation of inclusion, form a lattice. Elementary properties follow:

PROPOSITION 1.1. *The relative commutor operation is a polarity* in the sense of Birkhoff's book [1, Chapter IV, §5]. $\mathfrak{R} = \mathfrak{R}''$ *if and only if* $\mathfrak{R} \supseteqq \mathfrak{A}'$. *In general*

$$(\mathfrak{R} \cap \mathfrak{R}')' = \mathfrak{R} \cup \mathfrak{R}' \supseteqq (\mathfrak{R} \cup \mathfrak{R}')' \supseteqq \mathfrak{R} \cap \mathfrak{R}'.$$

2. Averaging operations, definition and preliminaries

DEFINITION 2.1. *An averaging operation from* $\mathfrak{A}$ *onto the subalgebra* $\mathfrak{R}$ *is any function* $\mathcal{C}$ *on* $\mathfrak{A}$ *onto* $\mathfrak{R}$ *such that*

(i) $\mathcal{C}$ *is linear*;

(ii) $\mathcal{C}$ *is positive*: *if* $A \geqq 0$ *then* $\mathcal{C}A \geqq 0$;

(iii) $\mathcal{C}$ *is idempotent*: *for all* $A \in \mathfrak{R}$, $A = \mathcal{C}A$;

(iv) *for all* $A, B \in \mathfrak{A}$, $\mathcal{C}(A\mathcal{C}B) = (\mathcal{C}A)\mathcal{C}B$.

The set of all averaging operations from $\mathfrak{A}$ onto $\mathfrak{R}$ will be denoted $\mathbf{M}(\mathfrak{A}, \mathfrak{R})$. (It might be interesting to see the consequences of dropping (ii), which is independent of the other conditions; cf. [10].)

PROPOSITION 2.1. $\mathcal{C}(A^*) = (\mathcal{C}A)^*$. $\mathcal{C}((\mathcal{C}A)B) = (\mathcal{C}A)\mathcal{C}B$.

PROPOSITION 2.2. *If* $\mathcal{C} \in \mathbf{M}(\mathfrak{A}, \mathfrak{R})$, *then the restriction of* $\mathcal{C}$ *to subalgebra* $\mathfrak{B}$ *is in* $\mathbf{M}(\mathfrak{B}, \mathcal{C}\mathfrak{B})$. *If* $\mathcal{C}_1 \in \mathbf{M}(\mathfrak{A}, \mathfrak{B})$, $C_2 \in \mathbf{M}(\mathfrak{B}, \mathfrak{R})$, *and* $\mathcal{C}_2 \circ \mathcal{C}_1$ *is their composition* $((\mathcal{C}_2 \circ \mathcal{C}_1)A = \mathcal{C}_2(\mathcal{C}_1 A))$, *then* $\mathcal{C}_2 \circ \mathcal{C}_1 \in \mathbf{M}(\mathfrak{A}, \mathfrak{R})$. *If* $\mathcal{C}_2 \circ \mathcal{C}_1$ *is the identity, so are* $\mathcal{C}_1$, $\mathcal{C}_2$.

PROPOSITION 2.3. $\mathbf{M}(\mathfrak{A}, \mathfrak{R})$ *is convex. That is, if* $\mathcal{C}_1, \mathcal{C}_2 \in \mathbf{M}(\mathfrak{A}, \mathfrak{R})$, $0 \leqq \lambda \leqq 1$, *and for all* $A \in \mathfrak{A}$, $\mathcal{C}_3 A = (1 - \lambda)\mathcal{C}_1 A + \lambda \mathcal{C}_2 A$, *then*

$$\mathcal{C}_3 \in \mathbf{M}(\mathfrak{A}, \mathfrak{R}).$$

Proof of these facts can be supplied without difficulty by the reader.

PROPOSITION 2.4. *If* $A = A^*$, *then* $(\mathcal{C}A)^2 \leqq \mathcal{C}(A^2)$.

Proof. This is a familiar computation:

$$0 \leqq \mathcal{C}((A - \mathcal{C}A)^2) = \mathcal{C}(A^2) - \mathcal{C}(A\mathcal{C}A) - \mathcal{C}((\mathcal{C}A)A) + \mathcal{C}((\mathcal{C}A)^2)$$
$$= \mathcal{C}(A^2) - (\mathcal{C}A)^2.$$

Justifying it uses all the assumptions on $\mathcal{C}$.

PROPOSITION 2.5. *The image of* $\mathcal{R}'$ *under* $\mathcal{C}$ *is* $\mathcal{R} \cap \mathcal{R}'$.

The only assertion here which is not absolutely trivial, namely $\mathcal{C}\mathcal{R}' \subseteqq \mathcal{R}'$, is proved like this: If $A \in \mathcal{R}'$, $B \in \mathcal{R}$, then by (iv) and Proposition 2.1

$$(\mathcal{C}A)B = \mathcal{C}(AB) = \mathcal{C}(BA) = B\mathcal{C}A.$$

3. Pinching operations

What I call *pinching operations*[4] may be defined in several different ways. Let $\mathcal{B}$ be a subalgebra of $\mathcal{A}$ such that $\mathcal{B} \supseteqq \mathcal{B}'$. The pinching from $\mathcal{A}$ to $\mathcal{B}$ is a (unique) function from $\mathcal{A}$ onto $\mathcal{B}$, which will also be denoted by $\mathcal{B}$. Until the definitions following are proved equivalent they will be written as defining different functions $\mathcal{B}_k$.

DEFINITION 3.1. $\mathcal{B}_1 A = \sum_i P_i A P_i$, *where the* P_i *are all the minimal projections in* $\mathcal{B}'$.

DEFINITION 3.2. $\mathcal{B}_2 A$ *is that* $B \in \mathcal{B}$ *for which* $\| B - A \|_2$ *is minimum.* Here $\| G \|_2$ means $(\operatorname{tr} G^*G)^{1/2}$, the Frobenius norm [16, §5.4].

DEFINITION 3.3. $\mathcal{B}_3 A$ *is the average of the* XAX, *where* X *ranges over all symmetries in* $\mathcal{B}'$.

DEFINITION 3.4. $\mathcal{B}_4 A$ *is the average of the* U^*AU, *where* U *ranges over all unitaries in* $\mathcal{B}'$.

THEOREM 3.1. *If* $\mathcal{B} \supseteqq \mathcal{B}'$, *there is just one averaging operation from* $\mathcal{A}$ *onto* $\mathcal{B}$, *and it is given by any of the four preceding definitions.*

Proof. In Definition 3.2, any faithful representation of $\mathcal{A}$ by linear transformations of a finite-dimensional space may be used to give meaning to the trace, but for simplicity assume multiplicity 1. Under Frobenius norm, $\mathcal{A}$ becomes a hilbert space, $\mathcal{B}$ a linear subspace; so $\mathcal{B}_2 A$ is uniquely defined and is determined as that $B \in \mathcal{B}$ for which $B - A$ is orthogonal to every $G \in \mathcal{B}$.

Let the Q_i be the minimal projections of $\mathcal{A}'$. $\mathcal{A} = \sum_i \oplus\, Q_i \mathcal{A} Q_i$, a direct sum of full matrix algebras (factors). It is easy to see that in any of the above definitions A can be replaced by the AQ_i, and those elements can then be treated separately (use too Proposition 2.2). Therefore without loss of generality assume $\mathcal{A}$ is all operators on a hilbert n-space. Because $\mathcal{B} \supseteqq \mathcal{B}'$, $\mathcal{B} = \sum_i \oplus\, P_i \mathcal{A} P_i$, a direct sum of full matrix algebras; here the P_i are the minimal projections of $\mathcal{B}'$.

To show $\mathcal{B}_1 = \mathcal{B}_2$, use a coordinate system in the underlying space whose first n_1 basis vectors are in the range of P_1, whose next n_2 basis vectors are in the range of P_2, $\cdots$ (n_i is the dimensionality of the range of P_i; $\sum n_i = n$). $G \in \mathcal{B}$ if and only if its matrix is in block form: an $n_1 \times n_1$

[4] Because of their analogy to compressions [7].

block followed by an $n_2 \times n_2$ block $\cdots$ along the main diagonal, and zeros elsewhere. To get $\mathcal{B}_1 A$ from A one merely replaces all entries of the matrix of A which lie outside this block pattern by zeros, leaving the other entries of A unchanged. $\mathcal{B}_1 A \in \mathcal{B}$, and $\mathcal{B}_1 A - A$ is evidently orthogonal to all $G \in \mathcal{B}$. Hence $\mathcal{B}_1 A = \mathcal{B}_2 A$.

$\mathcal{B}_3$ is linear, so it is enough to consider its action on A such that, for some $i, j, A = P_i A P_j$. First take $i \neq j$. In the finite abelian group of symmetries X in $\mathcal{B}'$, $2P_i - 1$ generates a subgroup. For any X,

$$XAX + X(2P_i - 1)A(2P_i - 1)X = XP_i AP_j X + XP_i A(-P_j)X = 0.$$

That is, the sum of XAX is zero over any coset of the subgroup, hence over the group. On the other hand, if $i = j$, for all X,

$$XAX = XP_i AP_i X = (\pm P_i)A(\pm P_i) = A.$$

In short, $\mathcal{B}_3(P_i AP_j) = \delta_{ij} P_i AP_j$, which agrees with $\mathcal{B}_1$. $\mathcal{B}_4$ is just the same idea, though integrals replace sums.

Therefore $\mathcal{B}_1 = \mathcal{B}_2 = \mathcal{B}_3 = \mathcal{B}_4$. Denote the operation simply $\mathcal{B}$. Is it an averaging?

It is evidently linear onto $\mathcal{B}$, and equal to the identity on $\mathcal{B}$. As to positivity, if $A \geqq 0$ then every summand XAX in Definition 3.3 is $\geqq 0$. Let us check (iv) of Definition 2.1. By linearity in A, it is enough again to take $A = P_i AP_j$; by linearity in B, to take $\mathcal{B}B = P_k BP_k$. Then

$$(\mathcal{B}A)\mathcal{B}B = \delta_{ij} P_i AP_j(P_k BP_k),$$
$$\mathcal{B}(A\mathcal{B}B) = \delta_{jk} \mathcal{B}(P_i ABP_k) = \delta_{jk} \delta_{ik} P_i ABP_k,$$

and the two are equal.

Finally, let $\mathcal{C}$ be any averaging operation from $\mathcal{A}$ onto $\mathcal{B}$. For any projection $P \in \mathcal{B}'$,

$$(*) \qquad \mathcal{C}(PAP + \tilde{P}A\tilde{P}) = \mathcal{C}A \qquad (\text{all } A \in \mathcal{A}).$$

Indeed, $P, \tilde{P} \in \mathcal{B}$, so by (iv)

$$\mathcal{C}(PAP + \tilde{P}A\tilde{P}) = P(\mathcal{C}A)P + \tilde{P}(\mathcal{C}A)\tilde{P} = P\mathcal{C}A + \tilde{P}\mathcal{C}A = \mathcal{C}A.$$

But each symmetry X in $\mathcal{B}'$ is of the form $P - \tilde{P}$, $P \in \mathcal{B}'$, so (*) implies $\mathcal{C}(XAX) = \mathcal{C}A$, $\mathcal{C}\mathcal{B}A = \mathcal{C}A$. Also $\mathcal{C}\mathcal{B}A = \mathcal{B}A$ just because $\mathcal{C}$ satisfies (iii.) Hence $\mathcal{C} = \mathcal{B}$, and the asserted uniqueness holds.

4. Structure of averaging operations

The set $\mathbf{M}(\mathcal{A}, \mathcal{R})$ of averaging operations from $\mathcal{A}$ onto $\mathcal{R}$ is convex (Proposition 2.3), and if $\mathcal{R} \supseteqq \mathcal{R}'$, it consists of a single element, a pinching (Theorem 3.1). It will now be shown how to express any averaging operation $\mathcal{C}$ as the composition of a pinching $\mathcal{B}$ followed by a convex combination of homomorphisms (cf. Proposition 2.2). Indeed, the conclusion of Theorem 4.2 is a slightly stronger assertion.

THEOREM 4.1. *If* $\mathcal{R}' = \mathcal{A}'$, *then the extreme points of* $\mathbf{M}(\mathcal{A}, \mathcal{R})$ *are exactly the homomorphisms.*[5]

THEOREM 4.2. *Given* $\mathcal{C} \in \mathbf{M}(\mathcal{A}, \mathcal{R})$, *there exists* $\mathcal{B} \subseteqq \mathcal{A}$ *such that*

$$\mathcal{B}' = \mathcal{R}' \cap \mathcal{B} \quad \textit{and} \quad \mathcal{C} \circ \mathcal{B} = \mathcal{C}.$$

The extreme points of $\mathbf{M}(\mathcal{B}, \mathcal{R})$ *are exactly the homomorphisms.*[5]

Before proving the theorems, here are a few remarks and examples.

For any subalgebra $\mathcal{R}$ of $\mathcal{A}$, $\mathbf{M}(\mathcal{A}, \mathcal{R})$ is not empty. This will be so evident, given the proof of Theorem 4.2, that no details need be given.

Since the condition $\mathcal{B}' = \mathcal{R}' \cap \mathcal{B}$ implies $\mathcal{B} \supseteqq \mathcal{B}'$, the implicit assumption in Theorem 4.2, that subalgebra $\mathcal{B}$ defines a pinching, is justified.

Furthermore $\mathcal{B}' = \mathcal{R}' \cap \mathcal{B}$ implies $\mathcal{B}' \subseteqq \mathcal{R}'$; hence $\mathcal{B}'' \supseteqq \mathcal{R}'' \supseteqq \mathcal{R}$. But by Proposition 1.1, $\mathcal{B} \supseteqq \mathcal{B}'$ implies $\mathcal{B} = \mathcal{B}''$; so $\mathcal{B} \supseteqq \mathcal{R}$ in Theorem 4.2.

The second sentence in Theorem 4.2 is a consequence of the first. For $\mathcal{B}' = \mathcal{R}' \cap \mathcal{B}$ is exactly the condition for Theorem 4.1 to be applicable to $\mathbf{M}(\mathcal{B}, \mathcal{R})$.

$\mathcal{C}$ does not always determine $\mathcal{B}$ uniquely. This will emerge in the proof of Theorem 4.2; the detailed explanation, since it seems distinctly less interesting, is exiled to §5.

Example 4.1. Let $\mathcal{A}$ be all operators on 2-dimensional $\mathcal{H}$, and let $\mathcal{R}$ be the constants. For fixed orthogonal unit vectors x_1, $x_2 \in \mathcal{H}$, and $\lambda \in [0, 1]$, let $\mathcal{C}A = (1 - \lambda)(Ax_1, x_1) + \lambda(Ax_2, x_2)$. $\mathcal{B}$ must be the algebra generated by $\{x_1, x_1\}$, the projection on $[x_1]$. That is, expressing A in matrix form $A = ((A_{ij}))$ with respect to basis vectors x_1, x_2, $\mathcal{B}A = \operatorname{diag}(A_{11}, A_{22})$, and $\mathcal{C}A = \operatorname{diag}((1 - \lambda)A_{11} + \lambda A_{22}, (1 - \lambda)A_{11} + \lambda A_{22})$. The restriction of $\mathcal{C}$ to $\mathcal{B}$ is a homomorphism if and only if λ is 0 or 1. This example illustrates the distinction between $\mathbf{M}(\mathcal{A}, \mathcal{R})$ and $\mathbf{M}(\mathcal{B}, \mathcal{R})$: here $\mathbf{M}(\mathcal{B}, \mathcal{R})$ is a line segment, but $\mathbf{M}(\mathcal{A}, \mathcal{R})$, the space of states of $\mathcal{A}$, is a closed ball in 3 real dimensions.

Example 4.2. This is the same as the preceding example, except that $\mathcal{C}A = \frac{1}{2}(Ax_1, x_1) + \frac{1}{2}(Ax_2, x_2)$, where unit vectors x_1 and x_2 are neither collinear nor orthogonal. (Without loss of generality, $0 < (x_1, x_2) < 1$.) Now $\mathcal{C}$ is expressed as a convex combination of averagings each of which (as already observed) is a composition of a pinching and a homomorphism. But two different pinchings are involved because $\{x_1, x_1\}$ and $\{x_2, x_2\}$ generate different subalgebras. The theorems say that it could be expressed using one common pinching. The reader is invited to verify this, in the coordinate system whose first basis vector (nonnormalized) is $x_1 + y_1$.

Example 4.3. More generally, let $n = m\mu$, and choose an orthogonal basis $\{x_{k\kappa}\}$, $k = 1, \cdots, m$, $\kappa = 1, \cdots, \mu$. Let $\mathcal{A}$ again be all operators, that is, all matrices $((A_{kl,\kappa\lambda}))$. Let $\mathcal{R}$ be those of the form $A_{kl,\kappa\lambda} = \delta_{kl} F_{\kappa\lambda}$.

[5] Of course there will ordinarily also be homomorphisms onto $\mathcal{R}$ which are not averagings at all.

For some $\mathcal{C}$, it may be possible to choose $\mathcal{B}$ to be those matrices of the form $A_{kl,\kappa\lambda} = \delta_{kl} F_{k,\kappa\lambda}$. The reader may repeat in this context the considerations of the preceding examples, in which μ was 1.

The examples are simple, and I now proceed to show that nothing much less simple can arise.

Proof of Theorem 4.1. This imitates the corresponding proof [17] for states on a commutative real algebra. One half of the theorem will be stated as a separate lemma because it does not use the hypothesis $\mathcal{A}' = \mathcal{R}'$.

LEMMA. *Every homomorphism in* $\mathbf{M}(\mathcal{A}, \mathcal{R})$ *is an extreme point.*[5]

Proof. Suppose homomorphism $\mathcal{C} = (1 - \lambda)\mathcal{C}_1 + \lambda\mathcal{C}_2$ for $\mathcal{C}_i \in \mathbf{M}(\mathcal{A}, \mathcal{R})$, $0 < \lambda < 1$. For $A \in \mathcal{A}$ hermitian, Proposition 2.4 gives

$$\mathcal{C}(A^2) = (1 - \lambda)\mathcal{C}_1(A^2) + \lambda\mathcal{C}_2(A^2) \geqq (1 - \lambda)(\mathcal{C}_1 A)^2 + \lambda(\mathcal{C}_2 A)^2.$$

But $\mathcal{C}$ is a homomorphism, so this is also equal to

$$(\mathcal{C}A)^2 = (1 - \lambda)^2(\mathcal{C}_1 A)^2 + 2\lambda(1 - \lambda)(\mathcal{C}_1 A)\mathcal{C}_2 A + \lambda^2(\mathcal{C}_2 A)^2.$$

Subtracting gives $0 \geqq \lambda(1 - \lambda)(\mathcal{C}_1 A - \mathcal{C}_2 A)^2$. Therefore $\mathcal{C}_1$ and $\mathcal{C}_2$ must agree on all hermitian members of $\mathcal{A}$; therefore they must be the same.

For the rest of the proof of Theorem 4.1, return to the stated hypothesis $\mathcal{R}' = \mathcal{A}'$.

Let $\mathcal{D} = \mathcal{R} \cup \mathcal{R}'$, the subalgebra generated by $\mathcal{R}$ and $\mathcal{R}'$. By Proposition 1.1, $\mathcal{D}' = (\mathcal{R} \cap \mathcal{R}') \cup \mathcal{A}'$; therefore $\mathcal{R}' = \mathcal{A}'$ implies $\mathcal{D}' = \mathcal{A}'$. Hence

$$\mathcal{D}'' = \mathcal{A}'' = \mathcal{A};$$

but $\mathcal{D} = \mathcal{D}''$, again by Proposition 1.1, so $\mathcal{D} = \mathcal{A}$. This implies that $\mathcal{A}$ is spanned linearly by elements of the form $A = RS$, $R \in \mathcal{R}$, $S \in \mathcal{R}'$. With this notation, $\mathcal{C}(RS) = R\mathcal{C}S$, of course.

Now let $\mathcal{C}$ be an averaging but not a homomorphism: there exist A, B for which $\mathcal{C}(AB) \neq (\mathcal{C}A)\mathcal{C}B$. In the rest of this paragraph, it will be shown that several special assumptions on A entail no loss in generality. By linearity in A, assume $A = RS$, $R \in \mathcal{R}$, $S \in \mathcal{R}'$. Then

$$R\mathcal{C}(SB) \neq R(\mathcal{C}S)\mathcal{C}B;$$

so assume $A \in \mathcal{R}' = \mathcal{A}'$. Since A can be replaced by $A - \mathcal{C}A$ here, and $\mathcal{C}A \in \mathcal{R}'$ by Proposition 2.5, assume $\mathcal{C}A = 0$; the failure of $\mathcal{C}$ to be multiplicative is now expressed by $\mathcal{C}(AB) \neq 0$. Write $A = H_1 + iH_2$, where $H_i \in \mathcal{R}'$ are hermitian; by Proposition 2.1, $\mathcal{C}H_i = 0$, but of course at least one $\mathcal{C}(H_i B) \neq 0$; so assume A hermitian. Since positive multiples are irrelevant, assume $\| A \| < 1$.

Define, for all $G \in \mathcal{A}$, $\mathcal{C}_+ G = \mathcal{C}((1 + A)G)$, $\mathcal{C}_- G = \mathcal{C}((1 - A)G)$. These are evidently linear functions on $\mathcal{A}$ into $\mathcal{R}$. Since $1 \pm A$ are positive operators in $\mathcal{A}'$, $G \geqq 0$ implies $(1 \pm A)G \geqq 0$, hence $\mathcal{C}_\pm G \geqq 0$. To complete the verification that $\mathcal{C}_\pm$ satisfy Definition 2.1, use the fact that $\mathcal{C}$ does, together with $\mathcal{C}A = 0$. Now to conclude the proof: $\mathcal{C} = \frac{1}{2}\mathcal{C}_+ + \frac{1}{2}\mathcal{C}_-$, a convex combination of two averagings from $\mathcal{A}$ onto $\mathcal{R}$, and $\mathcal{C}_+ \neq \mathcal{C}_-$ because $\mathcal{C}(AB) \neq 0$; therefore $\mathcal{C}$ is not extreme.

Proof of Theorem 4.2. Represent $\mathcal{A}$ as before as a subalgebra of the algebra $\mathfrak{M}$ of all operators on n-dimensional $\mathcal{H}$. It will be convenient *temporarily* to let $\mathcal{R}'$ mean the algebra of all $A \in \mathfrak{M}$ commuting with every member of $\mathcal{R}$, so that what we have been calling $\mathcal{R}'$ will temporarily be called $\mathcal{R}' \cap \mathcal{A}$, and similarly for other subalgebras of $\mathfrak{M}$. Thus the assumption that the representation of $\mathcal{A}$ has uniform multiplicity 1 is expressed by $\mathcal{A} \supseteqq \mathcal{A}'$. Let $\mathcal{C}$, $\mathcal{R}$ be as in the hypothesis of the theorem.

For any $x, y \in \mathcal{H}$, define $\{x, y\} \in \mathfrak{M}$ by $\{x, y\}z = (z, y)x$ for $z \in \mathcal{H}$. Now consider (cf. [4, I §3, Exercice 6]) tr $\mathcal{C} \circ \mathcal{A}\{x, y\}$ as a complex-valued function of x, y, $\mathcal{A}$ being the pinching from $\mathfrak{M}$ to $\mathcal{A}$. It is an everywhere-defined, positive sesquilinear form; hence

$$\operatorname{tr} \mathcal{C} \circ \mathcal{A}\{x, y\} = (Fx, y)$$

identically, for a determined positive $F \in \mathfrak{M}$.

The almost evident fact that $F \in \mathcal{A}$ may be proved explicitly as follows. Let the Q_i be the minimal projections of $\mathcal{A}'$, so that

$$\mathcal{A}\{x, y\} = \sum Q_i\{x, y\}Q_i = \sum \{Q_i x, Q_i y\}.$$

Then for all $x, y \in \mathcal{H}$

$$((\mathcal{A}F)x, y) = \sum (FQ_i x, Q_i y) = \operatorname{tr} \sum \mathcal{C} \circ \mathcal{A}\{Q_i x, Q_i y\}$$
$$= \operatorname{tr} \mathcal{C} \circ \mathcal{A} \circ \mathcal{A}\{x, y\} = (Fx, y).$$

There is more to the proof that $F \in \mathcal{R}'$: For any $A \in \mathcal{R}$,

$$(FAx, y) = \operatorname{tr} \mathcal{C} \circ \mathcal{A}\{Ax, y\} = \operatorname{tr} \mathcal{C} \circ \mathcal{A}(A\{x, y\}) = \operatorname{tr} (A\mathcal{C} \circ \mathcal{A}\{x, y\}),$$

$$(AFx, y) = \operatorname{tr} \mathcal{C} \circ \mathcal{A}\{x, A^*y\} = \operatorname{tr} \mathcal{C} \circ \mathcal{A}(\{x, y\}A) = \operatorname{tr} ((\mathcal{C} \circ \mathcal{A}\{x, y\})A),$$

and these are equal.

Now (knowing that $F \in \mathcal{R}' \cap \mathcal{A}$) define $\mathfrak{F}$ as some arbitrary maximal abelian subalgebra of $\mathcal{R}' \cap \mathcal{A}$ such that $F \in \mathfrak{F}$; define $\mathcal{B} = \mathfrak{F}' \cap \mathcal{A}$. This $\mathcal{B}$ will be shown to satisfy the requirements of the theorem: namely,

$$\mathcal{B}' \cap \mathcal{A} = \mathcal{B} \cap \mathcal{R}' \quad \text{and} \quad \mathcal{C} \circ \mathcal{B} = \mathcal{C}.$$

The first of these does not rely on the special properties of F. By the definition of $\mathfrak{F}$, $\mathfrak{F} = \mathfrak{F}' \cap \mathcal{R}' \cap \mathcal{A}$. Each factor on the right is $\supseteqq \mathcal{A}'$, so $\mathfrak{F} \supseteqq \mathcal{A}'$; therefore $\mathfrak{F}' \subseteqq \mathcal{A}$, $\mathcal{B} = \mathfrak{F}'$, $\mathcal{B}' \cap \mathcal{A} = \mathfrak{F} = \mathcal{B} \cap \mathcal{R}'$.

To prove $\mathcal{C} \circ \mathcal{B} = \mathcal{C}$ on $\mathcal{A}$, it would be more than enough to prove

$$\mathcal{C} \circ \mathcal{B} = \mathcal{C} \circ \mathcal{A}$$

on $\mathfrak{M}$. For this it would be enough to prove that every projection $P \in \mathcal{B}'$ has the property $\mathcal{C} \circ \mathcal{A}(PAP + \tilde{P}A\tilde{P}) = \mathcal{C} \circ \mathcal{A}A$ for all $A \in \mathfrak{M}$. For this it would be more than enough to prove that for every projection $P \in \mathcal{R}'$ commuting with F, $\mathcal{C} \circ \mathcal{A}(PA\tilde{P}) = 0$ for all $A \in \mathfrak{M}$. This I will do. But in this formulation there is no loss in generality in assuming $\mathcal{A} = \mathfrak{M}$, which will simplify the notation (now $\mathcal{C} \circ \mathcal{A} = \mathcal{C}$, $\mathcal{R}' \cap \mathcal{A} = \mathcal{R}'$, etc.).

Let the R_i be the minimal projections in $\mathcal{R} \cap \mathcal{R}'$. For any A,

$$\mathcal{C}A = \mathcal{C}(\sum R_i A R_j) = \sum R_i(\mathcal{C}A)R_j$$
$$= \sum R_i\, \mathcal{C}A = \sum R_i(\mathcal{C}A)R_i = \sum \mathcal{C}(R_i A R_i).$$

Clearly the task at hand is simply this: assuming $A = R_i A R_i$ for some i, and $A = PA\tilde{P}$ with P as above; to prove $\mathcal{C}A = 0$.

We may assume that $A = RS$, with $R = R_i R \in \mathcal{R}$, $S = R_i S \in \mathcal{R}'$. We may assume $R \geqq 0$ (premultiplying A by unitary $U \in \mathcal{R}$ affects nothing, not even $A = PA\tilde{P}$, because $P \in \mathcal{R}'$). Now $\mathcal{C}A = R\mathcal{C}S$; but since

$$\mathcal{C}S = R_i(\mathcal{C}S)R_i \in \mathcal{R} \cap \mathcal{R}'$$

(Proposition 2.5) and R_i is minimal, $\mathcal{C}A$ is a numerical multiple of $R \geqq 0$. Accordingly, if $\mathcal{C}A \neq 0$, tr $\mathcal{C}A \neq 0$.

For any B, tr $\mathcal{C}B$ = tr (FB), because this equation is linear in B and is true whenever $B = \{x, y\}$. Since P commutes with F,

$$\text{tr}\, \mathcal{C}A = \text{tr}\,(FA) = \text{tr}\,(FPA\tilde{P}) = \text{tr}\,(PFA\tilde{P}),$$

which is evidently zero. It follows that $\mathcal{C}A = 0$, as promised.

5. Nonuniqueness

In the proof just finished, F is determined by its definition, but (except in the case where F has simple spectrum) $\mathcal{B}'$ is not.

Example 5.1. Let $\mathcal{A}$ be all operators on 3-dimensional $\mathcal{H}$, and let $\mathcal{R}$ be the constants. Let $\mathcal{C}A = (Ax_1, x_1)$, for all A and for fixed unit vector x_1. Then F is the projection $\{x_1, x_1\}$. $\mathcal{B} = \mathcal{B}'$ must be the algebra generated by $\{x_1, x_1\}$, $\{x_2, x_2\}$, $\{x_3, x_3\}$, where x_1, x_2, x_3 form an orthogonal basis. But there are many ways to choose x_2, x_3.

One might suspect that uniqueness could fail even worse than this: that there could exist $\mathcal{B}'$ satisfying the conclusion of Theorem 4.2 but not obtained by the construction used in its proof. This does not happen, as I will show in the present section.

Any $\mathcal{B}'$ satisfying Theorem 4.2 is a maximal abelian subalgebra of $\mathcal{R}'$ each of whose projections P has property

$$(*) \qquad \mathcal{C}A = \mathcal{C}(PAP + \tilde{P}A\tilde{P}) \qquad (\text{all } A \in \mathcal{A}).$$

Any maximal abelian subalgebra of $\mathcal{R}'$ whose elements all commute with F, is obtainable as $\mathcal{B}'$ in the construction above. So the following is more information than is needed to fulfill my promise.

PROPOSITION 5.1. *Projection $P \in \mathcal{A}$ has property (*) if and only if it is the* *sum of two orthogonal projections, of which one is annihilated by* $\mathcal{C}$, *the other is in* $\mathcal{R}'$, *and both commute with* F.

Proof. Observe that if commuting projections P, Q both have (*), so do all Boolean functions of P, Q. The only part of this observation that takes any proving is that $P \cap Q = PQ$ has (*), and that is a simple computation: Because P and Q have (*),

$$\mathcal{C}A = \mathcal{C}(PAP + \tilde{P}A\tilde{P}) = \mathcal{C}(QPAPQ + Q\tilde{P}A\tilde{P}Q + \tilde{Q}PAP\tilde{Q} + \tilde{Q}\tilde{P}A\tilde{P}\tilde{Q})$$

for all A; if A is replaced by $PQAPQ + (PQ)^{\sim}A(PQ)^{\sim}$ here, each of the four terms on the right is unchanged, and hence so is the sum.

Therefore to justify "if" in the proposition it is enough to prove property (*) first for $P \in \mathcal{R}'$ commuting with F (this was done in the course of proving Theorem 4.2), then for P such that $\mathcal{C}P = 0$. These latter P can be treated easily without using F at all, but still more easily as follows.

LEMMA. *Let* $P_0 \in \mathcal{R}'$ *be the projection on the nullspace of* F. *For projection* $P \in \mathcal{A}$, *the following are equivalent*:

(i) $P \leqq P_0$;

(ii) $\mathcal{C}P = 0$;

(iii) *for all* $A \in \mathcal{A}$, $\mathcal{C}(PAP) = 0$.

Proof. Since $\mathcal{C}P \geqq 0$, clearly $\mathcal{C}P = 0$ is equivalent to $\operatorname{tr} \mathcal{C}P = 0$. But $\operatorname{tr} \mathcal{C}P = \operatorname{tr} FP = \operatorname{tr} PFP$, $PFP \geqq 0$; so $\operatorname{tr} \mathcal{C}P = 0$ if and only if $PFP = 0$, $P \leqq P_0$. This proves (i) equivalent to (ii). (ii) is a special case of (iii). To prove (iii) from (ii) it is enough, by linearity in A, to take $A \geqq 0$; but then

$$0 \leqq \mathcal{C}(PAP) \leqq \mathcal{C}(P \| A \|) = \| A \| \mathcal{C}P = 0.$$

The lemma is proved.

Now P_0 has (*) simply because $P_0 \in \mathcal{R}'$, $P_0 \leftrightarrow F$. Take any projection P such that $\mathcal{C}P = 0$. Because P_0 has (*),

$$\mathcal{C}(PA\tilde{P}) = \mathcal{C}(P_0 PA\tilde{P}P_0) + \mathcal{C}(\tilde{P}_0 PA\tilde{P}\tilde{P}_0);$$

the first term is zero by (iii) of the lemma, the second because $P \leqq P_0$. Likewise $\mathcal{C}(\tilde{P}AP) = 0$. Therefore P has (*).

The "if" half of the proposition is now taken care of; begin the interesting half.

In general, for projection P *not* to commute with hermitian A, is equivalent to the nonvanishing of

$$PA - AP = PA\tilde{P} - \tilde{P}AP = PA\tilde{P} - (PA\tilde{P})^*,$$

hence to the nonvanishing of $PA\tilde{P}$, hence to the nonvanishing of

$$(PA\tilde{P})(PA\tilde{P})^* = PA\tilde{P}AP \geqq 0.$$

I apply this criterion to any P having (*) and prove it commutes with F. By (*), $\mathcal{C}(\tilde{P}FP) = 0$. Hence

$$0 = \operatorname{tr} \mathcal{C}(\tilde{P}FP) = \operatorname{tr}(F\tilde{P}FP) = \operatorname{tr}(PF\tilde{P}FP),$$

$PF\tilde{P}FP = 0$, $P \leftrightarrow F$.

Therefore such P commute with P_0, and may be written as the sum of $P \cap P_0 = PP_0$ and $P \cap \tilde{P}_0 = P\tilde{P}_0$. These are orthogonal projections commuting with F, and the first is annihilated by $\mathcal{C}$. If we knew the second was in $\mathcal{R}'$, we would have finished the proof of the proposition. Accordingly we may consider, in the rest of the proof, only projections $P \leqq \tilde{P}_0$.

Any such P has the property that, if $A \geqq 0$ and $\mathcal{C}(PAP) = 0$, then $PAP = 0$. For otherwise there would be some nonzero projection Q and number $a > 0$ such that $PAP \geqq aQ$. Because $0 \leqq \mathcal{C}Q \leqq a^{-1}\mathcal{C}(PAP) = 0$, by the lemma $Q \leqq P_0$. But also $0 \neq Q \leqq P \leqq \tilde{P}_0$, a contradiction.

If this P has (*), I now prove it is in $\mathcal{R}'$. The method is to show, for any projection $R \in \mathcal{R}$, that P and R satisfy the criterion for commutativity used earlier. Clearly

$$R = \mathcal{C}R = \mathcal{C}(PRP + \tilde{P}R\tilde{P}) = \mathcal{C}(PR^2P + \tilde{P}R^2\tilde{P}).$$

But also

$$R = R^2 = R\mathcal{C}(PRP + \tilde{P}R\tilde{P}) = \mathcal{C}(RPRP + R\tilde{P}R\tilde{P})$$
$$= \mathcal{C}(PRPRP + \tilde{P}R\tilde{P}R\tilde{P}).$$

Subtracting these two expressions for R gives

$$0 = \mathcal{C}(PR\tilde{P}RP + \tilde{P}RPR\tilde{P}).$$

Both terms are $\geqq 0$; hence $\mathcal{C}(PR\tilde{P}RP) = 0$, and by the preceding paragraph $PR\tilde{P}RP = 0$. This proves $P \leftrightarrow R$.

6. Effect of pinching on the spectrum

Now shift the emphasis. Instead of looking at a single averaging operation acting on all operators in an algebra, fix one operator, and consider its images under various averaging operations. Consider first what seems the most interesting version: only pinchings are allowed.

All of the theorems of this section are restatements or easy consequences of known results (the most central of which are quoted as Theorem 6.0). Accordingly I give more references to the literature than proofs.

A matrix $S = ((S_{ij}))_{i,j=1,\cdots,n}$ is called *doubly stochastic* provided $S_{ij} \geqq 0$ and $\sum_i S_{ij} = \sum_j S_{ij} = 1$.

THEOREM 6.0. *Let $\alpha = (\alpha_1, \cdots, \alpha_n)$, $\beta = (\beta_1, \cdots, \beta_n)$ be n-tuples of real numbers. The following are equivalent*:

(i) *$\beta = S\alpha$, with S doubly stochastic*;

(ii) *β is a convex combination of permutations of α*;

(iii) *there exists a hermitian matrix with eigenvalues $\alpha_1, \cdots, \alpha_n$ and diagonal entries $\beta_1, \cdots, \beta_n$.*

These facts are all found in a paper by A. Horn [9, §2], who contributed the hardest part: namely the fact, which I will call Horn's Converse Theorem,[6] that (iii) is implied by the other conditions. The remaining implications are mostly drawn from [8, pp. 46–49], though Birkhoff's Theorem [9, Theorem 2] is helpful.

This immediately answers the question what spectrum a pinching of a given operator can have (cf. [5, Theorem 3]).

THEOREM 6.1. *Let hermitian A have eigenvalues $\alpha_1, \cdots, \alpha_n$. Numbers $\beta_1, \cdots, \beta_n$ are the eigenvalues of some pinching $\mathcal{B}A$ of A if and only if $\beta = S\alpha$, with S doubly stochastic.*

Proof. Assume $\beta_1, \cdots, \beta_n$ are the eigenvalues of $\mathcal{B}A = \sum_i P_i A P_i$ as in Definition 3.1. In a coordinate system in which the commuting hermitian matrices $\mathcal{B}A, P_1, P_2, \cdots$ are all diagonal, A (which need not be diagonal) has diagonal elements $\beta_1, \cdots, \beta_n$. Now use Theorem 6.0, (iii) implies (i).

Conversely, let $\beta = S\alpha$ with S doubly stochastic. The hermitian matrix whose existence is guaranteed by Horn's Converse Theorem belongs to our A in a suitable coordinate system. Then the required subalgebra $\mathcal{B}$ can be that of all diagonal matrices.

THEOREM 6.2. *Let A be hermitian. Let $\mathbf{K}(A)$ be the convex hull of the set of all unitary equivalents of A. Let $\mathbf{B}(A)$ be the set of all pinchings of unitary equivalents of A. Then $\mathbf{K}(A) = \mathbf{B}(A)$.*

Proof. That $\mathbf{B}(A) \subseteqq \mathbf{K}(A)$ follows from Definition 3.3 or 3.4, but it will be useful to know a stronger assertion. Let $B = \mathcal{B}A \in \mathbf{B}(A)$. We may (by unitary invariance and the proof of Theorem 6.1) assume without loss of generality that the coordinate system is such that $\mathcal{B}$ is exactly all diagonal matrices. Now it is true that B can be expressed as a convex combination of unitary equivalents A_i of A *which belong to* $\mathcal{B}$. The proof [3] involves Theorem 6.0, (i) implies (ii).

To complete the proof of Theorem 6.2, it must be shown that if

$$B = \sum \mu_i A_i ,$$

with $\mu_i \geqq 0, \sum \mu_i = 1$, and each A_i a unitary equivalent of A, then B is a pinching of a unitary equivalent of A. Choose subalgebra $\mathcal{B}$ so

$$B \in \mathcal{B} = \mathcal{B}';$$

then $B = \sum \mu_i \mathcal{B}A_i$. By the preceding paragraph, each $\mathcal{B}A_i$ is a convex combination of unitary equivalents of A lying in $\mathcal{B}$; this means that in

$$B = \sum \mu_i A_i$$

it may be assumed (by changing the notation) that all A_i were in $\mathcal{B}$ to begin with. By Theorem 6.0, (ii) implies (iii), we conclude $B \in \mathbf{B}(A)$.

---

[6] [9, Theorem 5]. A. J. Hoffman has an alternative proof (unpublished) using the theorem of Wielandt, Fan, and Pall [6, Theorem 1]. See also [12].

Thus Theorem 6.2 follows easily from Horn's Converse Theorem. In the other direction, if an independent proof could be found for Theorem 6.2, this would provide a new derivation of Horn's Converse Theorem

For any symmetric function f of n real variables and any hermitian A acting on n-space, let $f(A)$ mean the value of f when one substitutes for its arguments the eigenvalues of A.

THEOREM 6.3. *In order that $f(\mathcal{B}A) \leqq f(A)$ for all hermitian A and all pinchings $\mathcal{B}$, it is necessary and sufficient that f be Schur-convex. The inequality holds in particular for any convex symmetric f.*

On Schur-convex functions, see [15]. To say f is Schur-convex means that $\beta = S\alpha$, with S doubly stochastic, implies $f(\beta) \leqq f(\alpha)$. That every convex symmetric f is Schur-convex follows from Theorem 6.0, (i) implies (ii); cf. [3], [15, §24].

Theorem 6.3 is an immediate corollary of Theorem 6.1. I point out that it is in a sense weaker, for necessity can be proved without any form of Horn's Converse Theorem. Let $f(\mathcal{B}A) \leqq f(A)$ identically in A and $\mathcal{B}$, and let $\beta = S\alpha$, with S doubly stochastic. Since [8, pp. 46–49]

$$\beta = T^{(1)} \cdots T^{(k)}\alpha,$$

where each $T^{(i)}$ is doubly stochastic with at most two off-diagonal entries, we may assume S is of this sort. Hence we may assume $n = 2$. But for this case necessity in Theorem 6.3 is trivial.

THEOREM 6.4. (H. A. Dye [4, I §4, Exercice 2]) *If $\mathcal{B}'$ has m minimal projections, and $A \geqq 0$, then $A \leqq m\mathcal{B}A$.*

This suggests the following problem: Find restrictions on the spectrum of $\mathcal{B}A$, sharper than those implied by Theorem 6.1 (or 6.2), which follow from assuming a bound m on the number of minimal projections in $\mathcal{B}'$ $(1 < m < n)$. The known results are far from complete.

7. Effect of averaging on the spectrum

Now combine the results of §§4, 6.

Notation. $\sigma_k(\alpha_1, \cdots, \alpha_n)$, for real α_i and $k = 1, \cdots, n$, is the maximum of $\alpha_{i_1} + \cdots + \alpha_{i_k}$ over all choices of k integers i_j from among $\{1, \cdots, n\}$. (The minimum is therefore

$$-\sigma_k(-\alpha_1, \cdots, -\alpha_n) = \left(\sum_i \alpha_i\right) - \sigma_{n-k}(\alpha_1, \cdots, \alpha_n).)$$

This gives meaning to $\sigma_k(A)$ for hermitian A.

THEOREM 7.1. *Let hermitian A have eigenvalues $\alpha_1, \cdots, \alpha_n$. Numbers $\beta_1, \cdots, \beta_n$ are the eigenvalues of some averaging $\mathcal{C}A$ of A if and only if*

$$\sigma_k(A) \geqq \beta_{i_1} + \cdots + \beta_{i_k} \geqq -\sigma_k(-A)$$

whenever the β_{i_j} are distinct.

The following criterion will be helpful.

THEOREM 7.2. *Let real numbers $\alpha_1, \cdots, \alpha_n$ be given. In order that $\gamma_1, \cdots, \gamma_m$ be a subset of some n-tuple $(\gamma_1, \cdots, \gamma_n)$ such that $\gamma = S\alpha$ with S doubly stochastic, it is necessary and sufficient that*

$$\sigma_k(\alpha_1, \cdots, \alpha_n) \geqq \gamma_{i_1} + \cdots + \gamma_{i_k} \geqq -\sigma_k(-\alpha_1, \cdots, -\alpha_n)$$

for every choice of $\{i_1, \cdots, i_k\}$, $k = 1, \cdots, m$.

Two different special cases of this are known: the case $\gamma_1 = \cdots = \gamma_m$ is mentioned at the end of [10]; the case $m = n$ is a well-known theorem [8, Theorem 46]. I point out a consequence of the last-mentioned theorem:

Omitting the requirement of distinctness in the statement of Theorem 7.1, and replacing "averaging" by "pinching", converts it into another true statement, indeed into one equivalent to Theorem 6.1.

Proof of Theorem 7.2. Necessity is equivalent to necessity in the cited special case (and is easy), so I discuss only sufficiency. The idea of the proof is due to R. Radó [8, Theorem 75].

When $m = 1$ there is of course no difficulty: either $n = 1$, $\gamma_1 = \alpha_1$; or else we may number the α_i so that $\alpha_1 \geqq \gamma_1 \geqq \alpha_2$, in which case the requirements are clearly met by defining

$$\gamma_2 = \alpha_1 + \alpha_2 - \gamma_1 \quad \text{and} \quad \gamma_i = \alpha_i \quad \text{for } i = 3, \cdots, n.$$

Proceed by induction on m (letting n take care of itself). Let c be the minimum of the nonnegative numbers

$$D_k = \sigma_k(\alpha_1, \cdots, \alpha_n) - \sigma_k(\gamma_1, \cdots, \gamma_m),$$

$$d_k = -\sigma_k(-\alpha_1, \cdots, -\alpha_n) + \sigma_k(-\gamma_1, \cdots, -\gamma_m)$$

as k runs from 1 to $m - 1$. By symmetry, assume the minimum is attained by one of the D_k, say $D_h = c$.

As a notational convenience, assume now $\alpha_1 \geqq \cdots \geqq \alpha_n$, $\gamma_1 \geqq \cdots \geqq \gamma_m$.

$$c + \gamma_1 + \gamma_2 + \cdots + \gamma_h = \alpha_1 + \cdots + \alpha_h = \sigma_h(\alpha_1, \cdots, \alpha_n),$$

by the definitions of c and h. Likewise it is easy to verify that (i) the numbers $\gamma_1 + c, \gamma_2, \cdots, \gamma_h$ satisfy the conditions of the theorem relative to $\alpha_1, \cdots, \alpha_h$; and (ii) so do $\gamma_{h+1}, \cdots, \gamma_{m-1}, \gamma_m - c$ relative to $\alpha_{h+1}, \cdots, \alpha_n$. Remember $h < m$. By the inductive hypothesis, there exists a doubly stochastic matrix taking $(\alpha_1, \cdots, \alpha_h)$ to $(\gamma_1 + c, \cdots, \gamma_h)$, and another taking $(\alpha_{h+1}, \cdots, \alpha_n)$ to $(\gamma_{h+1}, \cdots, \gamma_{m-1}, \gamma_m - c, \gamma_{m+1}, \cdots, \gamma_n)$, where the new γ_i have been suitably chosen. Their direct sum is a doubly stochastic matrix R taking $(\alpha_1, \cdots, \alpha_n)$ to

$$(\gamma_1 + c, \gamma_2, \cdots, \gamma_{m-1}, \gamma_m - c, \gamma_{m+1}, \cdots, \gamma_n).$$

The latter n-tuple is taken to $(\gamma_1, \gamma_2, \cdots, \gamma_m, \cdots, \gamma_n)$ by the doubly stochastic matrix Q which has no nonzero off-diagonal entries other than

$$Q_{1m} = Q_{m1} = c/(2c + \gamma_1 - \gamma_m).$$

Then $S = QR$ is the matrix the theorem asks for.

Proof of Theorem 7.1. Assume $\mathcal{C}A$ with eigenvalues $\beta_1, \cdots, \beta_n$. By Theorem 4.2, $\mathcal{C}$ is the composition of a pinching followed by a convex combination of homomorphisms. By Theorem 6.1, as reformulated above with the aid of Theorem 7.2, the pinching can not increase any $\sigma_k(A)$ (or decrease any $-\sigma_k(-A)$). The rest of the proof of the inequalities in Theorem 7.1 need not be written out. The crucial fact, a special case of [3], is this: If $B = \sum_j \mu_j B^{(j)}$ is a convex combination of hermitian matrices, then

$$\sigma_k(B) \leqq \sum_j \mu_j \sigma_k(B^{(j)}).$$

Now for the converse! Assume the inequalities satisfied. Let the number of distinct β_i be m. $1 \leqq m \leqq n$; and choose the notation so that, for suitable n_j, $1 = n_1 < \cdots < n_m \leqq n$,

$$\beta_{n_j} = \cdots = \beta_{n_{j+1}-1} > \beta_{n_{j+1}}, \quad \cdots, \quad \beta_{n_m} = \cdots = \beta_n .$$

Let $\gamma_{n_j} = \beta_{n_j}$, with γ_i remaining undefined for the remaining $n - m$ values of i. That is, $\gamma_{n_1}, \cdots, \gamma_{n_m}$ are the distinct numbers among the β_i.

The inequalities relating $\alpha_1, \cdots, \alpha_n$ to $\beta_1, \cdots, \beta_n$ can now be expressed by saying that $\gamma_{n_1}, \cdots, \gamma_{n_m}$ satisfy the condition of Theorem 7.2 relative to $\alpha_1, \cdots, \alpha_n$. Hence values for the remaining γ_i can be chosen so that $\gamma = S\alpha$ for some doubly stochastic S. By Theorem 6.1 (or rather 6.0), A with eigenvalues $\alpha_1, \cdots, \alpha_n$ can be taken to $\mathcal{B}A$ with eigenvalues $\gamma_1, \cdots, \gamma_n$ by some pinching $\mathcal{B}$, of which we can assume $\mathcal{B} = \mathcal{B}'$. Choose orthonormal basis vectors x_i, $(\mathcal{B}A)x_i = \gamma_i x_i$. Subalgebra $\mathcal{B}$ is all diagonal matrices.

I have to define averaging $\mathcal{C}$ on $\mathcal{B}$. For any $B \in \mathcal{B}$, derive the diagonal entries of $\mathcal{C}B$ from those of B by

$$(\mathcal{C}B)_{ii} = B_{n_j n_j} \qquad (\text{for } n_j \leqq i < n_{j+1}).$$

This is an averaging, indeed a homomorphism. Clearly it gives the desired relation

$$(\mathcal{C} \circ \mathcal{B}A)_{ii} = \beta_i \qquad (i = 1, \cdots, n).$$

Here is the analogue of Theorem 6.3.

THEOREM 7.3. *Let $n \geqq 4$. In order that $f(\mathcal{C}A) \leqq f(A)$ for all hermitian A and all averagings $\mathcal{C}$, it is necessary and sufficient that f be of the form*

$$f(\alpha_1, \cdots, \alpha_n) = \max_{i,j} g(\alpha_i, \alpha_j),$$

for some symmetric $g(\gamma_1, \gamma_2)$ which, for $\gamma_1 \geqq \gamma_2$, is increasing in γ_1 and decreasing in γ_2.

By virtue of the properties of g, automatically

$$f(\alpha_1, \cdots, \alpha_n) = g(\max_i \alpha_i, \min_i \alpha_i).$$

It is also automatic that g is Schur-convex (hence f likewise). The monotonicity of g need not be strict.

Proof. If f has the stated form, then $f(\mathcal{C}A) \leqq f(A)$ is an easy consequence of the following trivial special case of Theorem 7.1:

$$\sigma_1(A) \geqq \sigma_1(\mathcal{C}A), \quad -\sigma_1(-\mathcal{C}A) \geqq -\sigma_1(-A).$$

Necessity is also not hard. It requires showing, for any $\alpha_1 \geqq \cdots \geqq \alpha_n$ and $\beta_1 \geqq \cdots \geqq \beta_n$ such that $\alpha_1 \geqq \beta_1$, $\beta_n \geqq \alpha_n$, $\alpha_1 > \alpha_n$, that

$$f(\beta_1, \cdots, \beta_n) \leqq f(\alpha_1, \cdots, \alpha_n).$$

I will do this by introducing an n-tuple $(\gamma_1, \cdots, \gamma_n)$. It is uniquely defined by the requirements that (for suitable $h = 0, 1, \cdots, n$)

$$\alpha_1 = \gamma_1 = \cdots = \gamma_{h-1} \geqq \gamma_h > \gamma_{h+1} = \cdots = \gamma_n = \alpha_n,$$
$$\textstyle\sum_{i=1}^n \gamma_i = \sum_{i=1}^n \beta_i.$$

Then in the sequence of n-tuples

$$(\alpha_1, \alpha_2, \cdots, \alpha_n), \quad (\alpha_1, \alpha_1, \alpha_n, \cdots, \alpha_n), \quad (\gamma_1, \cdots, \gamma_n), \quad (\beta_1, \cdots, \beta_n),$$

each satisfies the condition of Theorem 6.1 relative to the one before it, as one verifies at once. (This uses the hypothesis $n \geqq 4$.) Hence the value of f at each n-tuple is not greater than at the one before. This proves the theorem.

The cases $n = 2, 3$ are genuinely different. For example, for $n = 3$, f might be the symmetric function given when $\alpha_1 \geqq \alpha_2 \geqq \alpha_3$ by

$$f(\alpha_1, \alpha_2, \alpha_3) = 3(\alpha_1 - \alpha_3) - |\alpha_1 - 2\alpha_2 + \alpha_3|.$$

However, even for $n = 2$, some f which satisfy $f(\mathcal{C}A) \leqq f(A)$ when $\mathcal{C}$ is a pinching fail to satisfy it in general. For example, $f(\alpha_1, \alpha_2) = \alpha_1 + \alpha_2$.

Comparing the theorems of this section and §6, one might try to find corresponding theorems for a class of averagings containing the class of pinchings: namely, those $\mathcal{C}$ such that tr $\mathcal{C}A$ = tr A for all A (so that $F = 1$ in Theorem 4.2). Consider such a $\mathcal{C}$. Not only does $\mathcal{C}$ not have to be a pinching; in general there will be some A such that for no pinching $\mathcal{B}$ is $\mathcal{B}A = \mathcal{C}A$. Nevertheless it is easy to see that in their effect on the spectrum they are the same. That is, in Theorems 6.1, 6.2, 6.3, pinchings may be replaced by trace-preserving averagings throughout.

References

1. G. Birkhoff, *Lattice theory*, Amer. Math. Soc. Colloquium Publications, vol. 25, rev. ed., 1948.
2. ———, *Moyennes des fonctions bornées*, Algèbre et Théorie des Nombres, Colloques

Internationaux du Centre National de la Recherche Scientifique, no. 24, pp. 143–153, Paris, 1950.

3. C. DAVIS, *All convex invariant functions of hermitian matrices*, Arch. Math., vol. 8 (1957), pp. 276–278.
4. J. DIXMIER, *Les algèbres d'opérateurs dans l'espace hilbertien (Algèbres de von Neumann)*, Paris, 1957.
5. K. FAN, *Inequalities for eigenvalues of Hermitian matrices*, Contributions to the solution of systems of linear equations and the determination of eigenvalues, ed. Olga Taussky, Washington, 1954, pp. 131–139.
6. K. FAN AND G. PALL, *Imbedding conditions for Hermitian and normal matrices*, Canadian J. Math., vol. 9 (1957), pp. 298–304.
7. P. R. HALMOS, *Normal dilations and extensions of operators*, Summa Brasil. Math., vol. 2 (1950), pp. 125–134.
8. G. H. HARDY, J. E. LITTLEWOOD, AND G. PÓLYA, *Inequalities*, Cambridge, The University Press, 1934.
9. A. HORN, *Doubly stochastic matrices and the diagonal of a rotation matrix*, Amer. J. Math., vol. 76 (1954), pp. 620–630.
10. J. L. KELLEY, *Averaging operators on* $C_\infty(X)$, Illinois J. Math., vol. 2 (1958), pp. 214–223.
11. M. MARCUS, *A note on values of a quadratic form*, J. Washington Acad. Sci., vol. 47 (1957), pp. 97–99.
12. L. MIRSKY, *Matrices with prescribed characteristic roots and diagonal elements*, J. London Math. Soc., vol. 33 (1958), pp. 14–21.
13. M. NAKAMURA AND T. TURUMARU, *Expectations in an operator algebra*, Tôhoku Math. J. (2), vol. 6 (1954), pp. 182–188.
14. M. NAKAMURA AND H. UMEGAKI, *On a proposition of von Neumann*, Kōdai Math. Sem. Rep., vol. 8 (1956), pp. 142–144.
15. A. OSTROWSKI, *Sur quelques applications des fonctions convexes et concaves au sens de I. Schur*, J. Math. Pures Appl. (9), vol. 31 (1952), pp. 253–292.
16. R. SCHATTEN, *A theory of cross-spaces*, Princeton, 1950.
17. J. TATE, *On the relation between extremal points of convex sets and homomorphisms of algebras*, Comm. Pure Appl. Math., vol. 4 (1951), pp. 31–32.
18. J. VON NEUMANN, *On rings of operators. III*, Ann. of Math. (2), vol. 41 (1940), pp. 94–161.

THE INSTITUTE FOR ADVANCED STUDY
PRINCETON, NEW JERSEY