TWO SUBSPACES

BY

P. R. HALMOS

In the study of pairs of subspaces M and N in a Hilbert space H there are four thoroughly uninteresting cases, the ones in which both M and N are either 0 or H. In the most general case H is the direct sum of five subspaces:

$$M \cap N, \quad M \cap N^{\perp}, \quad M^{\perp} \cap N, \quad M^{\perp} \cap N^{\perp},$$

and the rest. The parts of M and N in the first four are "thoroughly uninteresting". In "the rest", the orthogonal complement of the span of the first four, M and N are in *generic position* ("position p" in [2]), in the sense that all four of the special intersections listed above are equal to 0. The purpose of this paper is to use graphs of linear transformations to represent pairs of subspaces in generic position. The results arose in the study of the invariant subspace lattices of operators and promise to be useful there. A more immediate by-product, described below, is a reasonably transparent new proof of Dixmier's theorem on the unitary equivalence of pairs of subspaces. Even aside from such external applications, however, the results answer at least one natural question and may be considered to be of geometric interest in their own right. Specialization to the finite-dimensional case makes neither the conclusions more obvious nor the proofs substantially simpler.

Axis-graph. Suppose that T is a closed but not necessarily bounded linear transformation on a dense subset of a Hilbert space K, with zero kernel and dense range. Write $H = K \oplus K$, let M be the "horizontal axis" consisting of all vectors of the form $\langle f, 0 \rangle$ in H, and let N be the graph of T, i.e., the set of all vectors of the form $\langle f, Tf \rangle$ in H. Assertion: M and N are in generic position. The first step of the proof is to show that $M \cap N = 0$. Indeed, how can an $\langle f, 0 \rangle$ be equal to a $\langle g, Tg \rangle$? Answer: only if $Tg = 0$, whence $g = 0$ (because ker $T = 0$), and therefore $f = 0$. To prove the rest of the assertion, it is necessary to know $M^{\perp}$ (trivial: all $\langle 0, f \rangle$) and $N^{\perp}$ (easy and standard computation: all $\langle -T^*f, f \rangle$). From this it is easy to deduce that $M^{\perp} \cap N^{\perp} = 0$: since ran T is dense, it follows that ker $T^* = 0$, and the proof just given applies again. The equations $M \cap N^{\perp} = 0$ and $M^{\perp} \cap N = 0$ are trivial.

The construction of the preceding paragraph is not new; it has been used, for instance, to exhibit pairs of subspaces whose vector sum is different from their closed span [4, p. 110], [5, p. 26]. The first result of the present paper is that this way of constructing pairs of subspaces in generic position is, to within unitary equivalence, the only way. To say that a pair $\langle M_1, N_1 \rangle$ of subspaces is unitarily

equivalent to a pair $\langle M_2, N_2 \rangle$ means, of course, that there exists a unitary operator U such that $UM_1 = M_2$ and $UN_1 = N_2$. In case M_1 and N_1 are in one Hilbert space H_1 and M_2 and N_2 in another Hilbert space H_2, then the requirement that U be unitary is to be interpreted to mean that U is an isometry from H_1 onto H_2.

THEOREM 1. *If M and N are subspaces in generic position in a Hilbert space H, then there exists a Hilbert space K, and there exists a closed linear transformation T on K with zero kernel and dense range, such that the pair $\langle K \oplus 0, \text{graph } T \rangle$ is unitarily equivalent to the pair $\langle M, N \rangle$.*

Proof. Let P be the projection with range M. Assertion: the restriction $P|N$ of P to N maps N one-to-one onto a linear manifold dense in M. Suppose, indeed, that $Pg = 0$ for some g in N. It follows that $g \in M^\perp \cap N$, and hence (generic position) that $g = 0$: this proves that the kernel of $P|N$ is 0. To prove that PN is dense in M, suppose that $f \in M$ and $f \perp PN$. This means that if $g \in N$, then $0 = (f, Pg) = (Pf, g) = (f, g)$, so that $f \in M \cap N^\perp$. It follows (generic position) that $f = 0$; the proof of the assertion is complete.

The existence of a transformation with the properties just proved for $P|N$ implies [5, p. 27] that M and N have the same dimension. Since M and $M^\perp$ on the one hand and N and $N^\perp$ on the other hand enter the hypotheses with perfect symmetry, it follows that all four of these subspaces have the same dimension. Since this applies to M and $M^\perp$ in particular, there exists an isometric mapping I from M onto $M^\perp$.

Everything is now prepared for the necessary definitions. Write

$$K = M,$$

define T on the dense subset PN of M by

$$TPg = I^{-1}(1-P)g \qquad (g \in N),$$

and, if both f and g are in $K\ (=M)$, write

$$U\langle f, g \rangle = f + Ig.$$

The definition of T may look artificial at first, but it is not. Here is what it says. To represent a vector g in N in the form $\langle f, Tf \rangle$, recall that the components of g with respect to the decomposition $H = M \oplus M^\perp$ are Pg and $(1-P)g$, and, therefore, let f be Pg and define T so that Tf is $(1-P)g$. Since T is to be a transformation on $K\ (=M)$, the last part does not quite make sense; the closest Tf can come to being $(1-P)g$ is to be $I^{-1}(1-P)g$, the vector in M that is identified with the vector $(1-P)g$ in M. That the definition of T is unambiguous is implied by the one-to-one character of $P|N$, proved above.

It remains to prove that K, T, and U have the required properties. Since I^{-1} is an isometry, the assertion that $\ker T = 0$ comes down to this: the restriction to N of the projection with range $M^\perp$ has kernel 0. That assertion is one of eight (go

from M, or $M^\perp$, to N, or $N^\perp$, or back); it differs from what was shown in the first paragraph of this proof in notation only. The density of ran T in K is proved similarly.

Next, why is T closed? It is to be proved that if $f_n = Pg_n$ (with $g_n \in N$), $f_n \to f$, and $I^{-1}(1-P)g_n \to h$, then $f = Pg$ for some g in N, and $h = I^{-1}(1-P)g$. Apply I to $I^{-1}(1-P)g_n \to h$ and infer $(1-P)g_n \to Ih$. Combine this with $Pg_n\ (=f_n) \to f$ to get

$$g_n = Pg_n + (1-P)g_n \to f + Ih.$$

In view of this, the obviously indicated thing to do is to write $g = f + Ih$. Since $g_n \in N$ and N is closed, it follows that $g \in N$. Since $I^{-1}(1-P)g_n \in M$ and M is closed, it follows that $h \in M$; hence $Ih \in M^\perp$ and $(1-P)Ih = Ih$. Since, finally, $f \in M$, it follows that

$$(1-P)g = (1-P)(f+Ih) = (1-P)Ih = Ih,$$

and hence that $I^{-1}(1-P)g = h$. Consequence: T is closed.

Since I maps M onto $M^\perp$, the transformation U maps $K \oplus K$ onto H. The isometric character of U follows from the computation:

$$\|U\langle f, g\rangle\|^2 = \|f+Ig\|^2 = \|f\|^2 + \|Ig\|^2 = \|f\|^2 + \|g\|^2 = \|\langle f, g\rangle\|^2.$$

It is trivial that U maps $K \oplus 0$ onto M; how does U map the graph of T? The graph of T is the same as the set of all $\langle Pg, I^{-1}(1-P)g\rangle$, with g in N; the image of $\langle Pg, I^{-1}(1-P)g\rangle$ under U is $Pg + II^{-1}(1-P)g = g$, and therefore the image of the graph is exactly N. The proof of the theorem is complete.

COROLLARY. *The transformation T can be chosen selfadjoint and positive, and, if it is so chosen, then it is unique, in the sense that if the pair $\langle K \oplus 0, \text{graph } T_1\rangle$ is unitarily equivalent to the pair $\langle K \oplus 0, \text{graph } T_2\rangle$, then T_1 is unitarily equivalent to T_2.*

Proof. To make T selfadjoint, consider its polar decomposition $T = WA$, where W is a partial isometry and A is selfadjoint and positive [3, p. 1249]. The conditions on T (zero kernel, dense range) imply that the partially isometric factor W is unitary. Apply the unitary operator $1 \oplus W^*$ to both axis and graph; the axis remains invariant, and the graph becomes the graph of the positive selfadjoint transformation A.

To prove uniqueness, observe first that if a unitary operator on $K \oplus K$ maps the axis $K \oplus 0$ onto itself, then it is a direct sum $U_1 \oplus U_2$ of two unitary operators on K. (A subspace mapped onto itself by a unitary operator is invariant under both the operator and its inverse, and therefore reduces the operator.) The assumption that $U_1 \oplus U_2$ maps graph T_1 onto graph T_2 implies that if $f \in \text{dom } T_1$, then $U_1 f \in \text{dom } T_2$ and $T_2 U_1 f = U_2 T_1 f$; in other words $T_2 U_1 = U_2 T_1$. By adjunction $U_1^* T_2 = T_1 U_2^*$ (because T_1 and T_2 are selfadjoint); by multiplication $T_1^2 = T_1 U_2^* U_2 T_1 = U_1^* T_2^2 U_1$, and therefore $T_1 = U_1^* T_2 U_1$ (because T_1 and T_2 are positive).

Projection matrix. It is an elementary exercise in analytic geometry to calculate the matrix of a projection of rank 1 acting on a space of dimension 2. The projection whose range is a line of inclination θ turns out to be

$$\begin{pmatrix} \cos^2 \theta & \cos \theta \sin \theta \\ \cos \theta \sin \theta & \sin^2 \theta \end{pmatrix}.$$

The graph of a linear transformation on a Hilbert space is very much like a line in a plane. Since the elements of the graph of T are ordered pairs $\langle f, Tf \rangle$, it follows, purely formally, that the ratio of the second coordinate to the first is always T; the transformation T plays the role of the slope of the line, i.e., of the tangent of the inclination θ. The following result is the operator version of the elementary exercise in analytic geometry mentioned above. The boundedness of *sin* and *cos* (as opposed to the unboundedness of *tan*) are reflected in that only bounded linear transformations need to be mentioned.

THEOREM 2. *If M and N are subspaces in generic position in a Hilbert space H, with respective projections P and Q, then there exists a Hilbert space K, and there exist positive contractions S and C on K, with $S^2 + C^2 = 1$ and* ker $S =$ ker $C = 0$, *such that P and Q are unitarily equivalent to*

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad and \quad \begin{pmatrix} C^2 & CS \\ CS & S^2 \end{pmatrix}$$

respectively.

Proof. Identify H with a direct sum $K \oplus K$ just as in the proof of Theorem 1. In this identification M and $M^\perp$ are the axes $K \oplus 0$ and $0 \oplus K$, and therefore

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

Since Q is a positive contraction, it follows that

$$Q = \begin{pmatrix} D & E \\ E^* & F \end{pmatrix},$$

where D $(=PQP|M)$ and F $(=(1-P)Q(1-P)|M^\perp)$ are positive contractions. Since $E = PQ(1-P)|M^\perp$ and $E^* = (1-P)QP|M$, it follows easily, just as in the first paragraph of the proof of Theorem 1, that ker $E =$ ker $E^* = 0$. This implies that in the polar decomposition $E = WA$ of E, the partially isometric factor W is unitary. Transform both P and Q by

$$\begin{pmatrix} W & 0 \\ 0 & 1 \end{pmatrix}$$

(i.e., multiply by

$$\begin{pmatrix} W^* & 0 \\ 0 & 1 \end{pmatrix}$$

on the left and

$$\begin{pmatrix} W & 0 \\ 0 & 1 \end{pmatrix}$$

on the right); the former remains invariant and the latter transforms to the same form as before except that now the off-diagonal terms are equal and positive. Assume therefore, with no loss of generality, that they were that way in the first place. (The identification of H with $K \oplus K$ was achieved by writing $K = M$ and choosing an arbitrary isometric correspondence between M and $M^\perp$. In these terms the change just made amounts to choosing a different, more convenient isometric correspondence.)

Let C and S be the positive square roots of D and F respectively, so that

$$Q = \begin{pmatrix} C^2 & E \\ E & S^2 \end{pmatrix},$$

where C, S, and E are positive contractions. The idempotence of Q says exactly that

$$C^2 - C^4 = S^2 - S^4 = E^2 \quad \text{and} \quad C^2 E + E S^2 = E.$$

Since, by the first two equations, C and S commute with E, the last equation says that $E(C^2 + S^2 - 1) = 0$. Since $\ker E = 0$, it follows that $S^2 = 1 - C^2$; since $E^2 = C^2 - C^4 = C^2(1 - C^2) = C^2 S^2$, therefore $E = CS$. The last assertion, together with $\ker E = 0$, implies that $\ker C = \ker S = 0$; the proof is complete.

Some of the known theory of pairs of subspaces in generic position can be recaptured from Theorem 2. Consider, as a sample, the problem of determining the commutant of the pair of projections P and Q. The commutant of P alone is easy to compute: it consists of all matrices of the form

$$\begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix}.$$

When does such a matrix commute with

$$\begin{pmatrix} C^2 & CS \\ CS & S^2 \end{pmatrix}?$$

Answer: if and only if both X and Y commute with C^2 (and hence with C, and S^2 and S), and also $CSX = CSY$. Since $\ker CS = 0$, the last condition implies $X = Y$. Conclusion: the simultaneous commutant of P and Q consists of all

$$\begin{pmatrix} X & 0 \\ 0 & X \end{pmatrix}$$

where X commutes with C.

The mere statement that the commutant of $\{P, Q\}$ is nontrivial is trivial: $(P+Q-1)^2$ commutes with both P and Q. N.B.:

$$(P+Q-1)^2 = \begin{pmatrix} C^2 & 0 \\ 0 & C^2 \end{pmatrix}.$$

It follows easily that the von Neumann algebra generated by two projections (on a Hilbert space of dimension greater than 2) is never the algebra of all operators; the cases in which the ranges are not in generic position are easily disposed of. Equivalently: if P and Q are projections on a Hilbert space H of dimension greater than 2, then there exists a nontrivial subspace of H invariant under both P and Q. (Note that since P and Q are Hermitian, invariance is the same as reduction.) Results such as these are almost explicitly contained in the earlier work on pairs of subspaces; cf. in particular [1] and [2].

Reflected graphs.   Theorems 1 and 2 were proved separately above, but there are other (not necessarily simpler) approaches to the theory. The same results can be obtained by proving only one of them directly (either one) and then deriving the other one from it.

The way to get Theorem 1 from Theorem 2, for instance, is to note first that the range of

$$\begin{pmatrix} C^2 & CS \\ CS & S^2 \end{pmatrix}$$

consists of all vectors of the form $\langle Cf, Sf \rangle$. (It really consists of all vectors of the form $\langle C(Cf+Sg), S(Cf+Sg) \rangle$. Since, however, the positive operator $C+S$ is invertible, because the function $x \to x+(1-x^2)^{1/2}$ is bounded away from zero in $[0, 1]$, the simpler description is valid.) To recognize that set as a graph, it is necessary to define a linear transformation T such that $TC=S$. (Formally: $\tan\theta = \sin\theta/\cos\theta$.) There is no conceptual difficulty here. The operator C has an inverse (not necessarily bounded), and the functional calculus for selfadjoint transformations can be invoked to justify the definition $T=C^{-1}(1-C^2)^{1/2}$.

To go in the other direction, from Theorem 1 to Theorem 2, it is necessary first to express in terms of T the projection whose range is the graph of T. Formally this too is an exercise in analytic geometry; the projection whose range is a line of slope $\tan\theta$ turns out to be

$$\begin{pmatrix} (1+\tan^2\theta)^{-1} & \tan\theta(1+\tan^2\theta)^{-1} \\ \tan\theta(1+\tan^2\theta)^{-1} & \tan^2\theta(1+\tan^2\theta)^{-1} \end{pmatrix}.$$

(Alternatively, obtain this matrix from the one in terms of sines and cosines by trigonometric identities.) It is easy to derive the precise version of the formalism from the basic definitions; for an explicit treatment see [6]. Once the projection Q is recognized as

$$\begin{pmatrix} (1+T^2)^{-1} & T(1+T^2)^{-1} \\ T(1+T^2)^{-1} & T^2(1+T^2)^{-1} \end{pmatrix},$$

then the functional calculus for selfadjoint transformations can be invoked again; the result is that $C^2 = (1 + T^2)^{-1}$ and $S^2 = T^2(1 + T^2)^{-1}$, and all the conditions of Theorem 2 are satisfied.

The purpose of this section is to exploit the trigonometric analogy once more. The underlying geometric fact is that the diagram
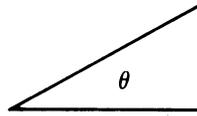
when rotated downward through $\theta/2$ becomes
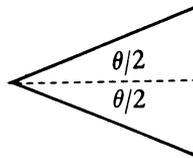
The operator version of this geometric fact is the following somewhat surprising assertion.

THEOREM 3. *If M and N are subspaces in generic position in a Hilbert space H, then there exists a Hilbert space K, and there exists a positive contraction T_0 on K, with* $\ker T_0 = \ker (1 - T_0) = 0$, *such that the pair $\langle M, N \rangle$ is unitarily equivalent to the pair $\langle \text{graph } T_0, \text{graph } (-T_0) \rangle$.*

Proof. The rotation that maps a line of inclination θ onto one of inclination $\theta/2$ and, at the same time, maps the horizontal axis onto the line of inclination $-\theta/2$ has the matrix

$$\begin{pmatrix} \cos(\theta/2) & \sin(\theta/2) \\ -\sin(\theta/2) & \cos(\theta/2) \end{pmatrix}.$$

This suggests that Theorem 3 could be derived from Theorem 1 by the operator analogue of that half-angle rotation. Formally this program is a simple computation, and precisely it is a straightforward argument about selfadjoint transformations. Since, however, the argument involves unbounded transformations, and since the same result can be achieved with bounded operators only, it is preferable to base the proof on Theorem 2. The boundedness of T_0 is, by the way, one of the mildly surprising features of Theorem 3. The heuristic reason for it is that if T corresponds to $\tan \theta$, $0 < \theta < \pi/2$, then T_0 corresponds to $\tan(\theta/2)$. As θ varies over the domain $(0, \pi/2)$, its tangent varies unboundedly, but, since $\theta/2$ varies over $(0, \pi/4)$, the tangent of the half-angle remains bounded.

By virtue of Theorem 2 there is no loss of generality in assuming that H is already represented as $K \oplus K$ in such a way that M is the axis $K \oplus 0$ and N is the range of the projection

$$Q = \begin{pmatrix} C^2 & CS \\ CS & S^2 \end{pmatrix}.$$

The problem is to find T_0 so that $\langle K \oplus 0, \operatorname{ran} Q \rangle$ is unitarily equivalent to $\langle \operatorname{graph} T_0, \operatorname{graph}(-T_0) \rangle$. Since the half-angle formula for the tangent is

$$\tan \frac{\theta}{2} = \left(\frac{1 - \cos \theta}{1 + \cos \theta} \right)^{1/2},$$

it is formally natural to write

$$T_0 = ((1 - C)(1 + C)^{-1})^{1/2}.$$

This makes sense. Since C is positive, $1 + C$ is invertible, and since C is a positive contraction, the same is true of $1 - C$. That T_0 and $1 - T_0$ have zero kernel is a consequence of the same properties of $1 - C$ and C, respectively.

The natural candidate for T_0 has been found; the next problem is to find a unitary operator that transforms M and N the desired way. The formalism suggests that too; since

$$\cos \frac{\theta}{2} = \left(\frac{1 + \cos \theta}{2} \right)^{1/2} \quad \text{and} \quad \sin \frac{\theta}{2} = \left(\frac{1 - \cos \theta}{2} \right)^{1/2},$$

it is natural to write

$$U = \begin{pmatrix} C_0 & S_0 \\ -S_0 & C_0 \end{pmatrix},$$

where

$$C_0 = \left(\frac{1 + C}{2} \right)^{1/2} \quad \text{and} \quad S_0 = \left(\frac{1 - C}{2} \right)^{1/2}.$$

All this makes sense; the verification that U is unitary is trivial.

The rest of the proof is computational: it verifies that as f varies over K, the images

$$\begin{pmatrix} C_0 & S_0 \\ -S_0 & C_0 \end{pmatrix} \begin{pmatrix} Cf \\ Sf \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} C_0 & S_0 \\ -S_0 & C_0 \end{pmatrix} \begin{pmatrix} f \\ 0 \end{pmatrix}$$

fill out exactly the graph of T_0 and the graph of $-T_0$ respectively. It is to be verified, in other words, that
   (i) $\operatorname{ran}(C_0 C + S_0 S) = K$,
   (ii) $T_0(C_0 C + S_0 S) = -S_0 C + C_0 S$,
   (iii) $\operatorname{ran} C_0 = K$,
   (iv) $(-T_0)C_0 = -S_0$.
For (i): in fact $C_0 C + S_0 S = C_0$; since C_0 is invertible, this settles (i) and (iii). The verifications of (ii) and (iv) are mechanical.

COROLLARY. *If M and N are subspaces in generic position in a Hilbert space H, with respective projections P and Q, then a complete set of unitary invariants of the pair $\langle M, N \rangle$ is the unitary equivalence class of the Hermitian operator $P + Q$.*

Proof. The statement means that $\langle M_1, N_1 \rangle$ is unitarily equivalent to $\langle M_2, N_2 \rangle$ if and only if $P_1 + Q_1$ is unitarily equivalent to $P_2 + Q_2$. Since the problem of

unitary equivalence for Hermitian operators is in principle solved, this is a solution of the problem of unitary equivalence for pairs of subspaces. The result is due to Dixmier [2].

To prove the corollary, use Theorem 3 to justify the assumption that M and N are the graphs of T_0 and $-T_0$ respectively. In that case the matrices of the projections P and Q are given by

$$P = \begin{pmatrix} (1+T_0^2)^{-1} & T_0(1+T_0^2)^{-1} \\ T_0(1+T_0^2)^{-1} & T_0^2(1+T_0^2)^{-1} \end{pmatrix}$$

and

$$Q = \begin{pmatrix} (1+T_0^2)^{-1} & -T_0(1+T_0^2)^{-1} \\ -T_0(1+T_0^2)^{-1} & T_0^2(1+T_0^2)^{-1} \end{pmatrix},$$

or, equivalently, by

$$P = \begin{pmatrix} C_0^2 & C_0 S_0 \\ C_0 S_0 & S_0^2 \end{pmatrix} \quad \text{and} \quad Q = \begin{pmatrix} C_0^2 & -C_0 S_0 \\ -C_0 S_0 & S_0^2 \end{pmatrix}.$$

The rest of the proof is the same as Dixmier's. In view of the form of $P+Q$, it is natural to consider the Hermitian operator

$$R = P + Q - 1 = \begin{pmatrix} C & 0 \\ 0 & -C \end{pmatrix}.$$

Since the positive and negative parts of a Hermitian operator are unitary invariants of it, it follows that the unitary equivalence class of R uniquely determines that of C. (Here it is important that ker $C=0$.) Since C, in turn, uniquely determines C_0, and thence P and Q, it follows that the unitary equivalence class of R uniquely determines that of the pair P, Q. The opposite direction is trivial; the proof of the corollary is complete. (The difference between $P+Q$ and R is a technical triviality. For R the parts that matter are the ones above and below 0, whereas for $P+Q$ they are the ones above and below 1.)

REFERENCES

1. Arlen Brown, *The unitary equivalence of binormal operators*, Amer. J. Math. **76** (1954), 414–434.

2. J. Dixmier, *Position relative de deux variétés linéaires fermées dans un espace de Hilbert*, Rev. Sci. **86** (1948), 387–399.

3. N. Dunford and J. T. Schwartz, *Linear operators*, Vol. II, Interscience, New York, 1963.

4. P. R. Halmos, *Introduction to Hilbert space and the theory of spectral multiplicity*, Chelsea, New York, 1951.

5. ———, *A Hilbert space problem book*, Van Nostrand, Princeton, N. J., 1967.

6. M. H. Stone, *On unbounded operators in Hilbert space*, J. Indian Math. Soc. **15** (1951), 155–192.

UNIVERSITY OF HAWAII,
HONOLULU, HAWAII