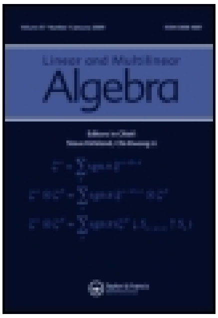

Linear and Multilinear Algebra
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/glma20

Bad products of good matrices

Paul R. Halmos [a]
[a] Department of Mathematics, Santa Clara University, Santa Clara, CA 95053
Version of record first published: 30 May 2007.

To cite this article: Paul R. Halmos (1991): Bad products of good matrices, Linear and Multilinear Algebra, 29:1, 1-20

To link to this article: http://dx.doi.org/10.1080/03081089108818053

Linear and Multilinear Algebra, 1991, Vol. 29, pp. 1–20
Reprints available directly from the publisher
Photocopying permitted by license only

Printed in the United States of America

This paper is the first in a series of invited expository articles by distinguished contributors. The editors anticipate that the entire series will appear in the next two volumes.

Bad Products of Good Matrices

PAUL R. HALMOS
Department of Mathematics, Santa Clara University, Santa Clara, CA 95053

(*Received May 18, 1990*)

1. SYMMETRIC

A matrix (a square matrix) is called *diagonal* if every entry off the main diagonal is zero; in symbols, $A = (a_{ij})$ is diagonal if $a_{ij} = 0$ whenever $i \neq j$. A matrix is called *symmetric* if its entries remain unchanged when reflected through the diagonal; in symbols, A is symmetric if $a_{ij} = a_{ji}$ for all i and j. The product of two diagonal matrices is always diagonal. Is the product of two symmetric matrices always symmetric?

The question is preposterous—the product of two symmetric matrices is extremely unlikely to be symmetric. Write down at random almost any two symmetric matrices, say for instance $\left(\begin{smallmatrix}1 & 2\\ 2 & 3\end{smallmatrix}\right)$ and $\left(\begin{smallmatrix}4 & 5\\ 5 & 6\end{smallmatrix}\right)$, and multiply them; the result is $\left(\begin{smallmatrix}14 & 17\\ 23 & 28\end{smallmatrix}\right)$, which is not symmetric. Once you have the courage, it's no trouble at all to construct much simpler examples. One easy example is $\left(\begin{smallmatrix}1 & 0\\ 0 & 0\end{smallmatrix}\right)$ and $\left(\begin{smallmatrix}0 & 1\\ 1 & 0\end{smallmatrix}\right)$ with product $\left(\begin{smallmatrix}0 & 1\\ 0 & 0\end{smallmatrix}\right)$. Note that, for both examples, the product in one order is not the same as the product in the other order; in other words the pairs of matrices are not commutative. (**Exercise 1**: is commutativity either a necessary or a sufficient condition that the product of two symmetric matrices be symmetric?)

Granted that the product of two symmetric matrices can fail to be symmetric, it makes sense to ask which matrices can be such products—which non-symmetric matrices are products of symmetric ones? The question belongs to a large class of interesting ones that are often non-trivial, questions that ask which bad matrices are products of good ones. Could

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

for instance, possibly be written as a product of two symmetric matrices? And if the answer happens to be yes, if that non-symmetric matrix is not bad enough to be a counterexample, how about something thoroughly non-symmetric, something like

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 4 & 5 & 6 \end{pmatrix}$$

—could that be a product of two symmetric ones?

The equation

$$\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix} \cdot \begin{pmatrix} -1 & -1 & -1 \\ -1 & -1 & 0 \\ -1 & 0 & 0 \end{pmatrix}$$

may or may not be considered shocking, but in any event it does answer the first of the two questions raised above—a mildly bad matrix is the product of two good ones. The frightful equation

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 1 & -1 & \frac{1}{5} \\ -1 & 2 & \frac{-28}{15} \\ \frac{1}{5} & \frac{-28}{15} & \frac{127}{45} \end{pmatrix} \cdot \begin{pmatrix} 2908 & 3303 & 1980 \\ 3303 & 3753 & 2250 \\ 1980 & 2250 & 1350 \end{pmatrix}$$

is likely to be considered frightening by most people, and it answers the second question: a thoroughly bad matrix can also be the product of two symmetric good ones. The factoring is far from unique; another possibility (discovered by Sheldon Axler) is

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} \frac{25}{22} & \frac{-25}{22} & \frac{5}{22} \\ \frac{-25}{22} & \frac{409}{418} & \frac{15}{418} \\ \frac{5}{22} & \frac{15}{418} & \frac{-25}{418} \end{pmatrix} \begin{pmatrix} \frac{103}{5} & 25 & \frac{132}{5} \\ 25 & 31 & 30 \\ \frac{132}{5} & 30 & 18 \end{pmatrix}.$$

How are such examples arrived at? What are the general facts; how can all possible examples be characterized?

The principal fact was known to Frobenius in 1910 [2], and it is easy to state: *every* matrix is a product of two symmetric matrices. The proof has two parts, cogitation and calculation. The cogitation needed is simple, but it relies on a deep theorem, the deepest theorem of linear algebra. The calculation is simple, but inspired—it seems to pull a rabbit out of a hat, and invites the question of just how a mere mortal could have thought to look for that rabbit.

The calculation involves the consideration of matrices of the form

$$K = \begin{pmatrix} 0 & 0 & 0 & a \\ 1 & 0 & 0 & b \\ 0 & 1 & 0 & c \\ 0 & 0 & 1 & d \end{pmatrix},$$

or, to be more accurate, of their $n \times n$ generalizations (consisting of all 0's except a subdiagonal of 1's and arbitrary entries in the last column). The 4×4 case is computationally absolutely typical and has the virtue of demanding less notation. The rabbit out of the hat is the matrix

$$L = \begin{pmatrix} b & c & d & -1 \\ c & d & -1 & 0 \\ d & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}$$

formed by using the entries in the last column of K, all except the very first, in the first $n-1$ backward diagonals, and putting -1's in the main backward diagonal. The product

$$KL = \begin{pmatrix} 0 & 0 & 0 & a \\ 1 & 0 & 0 & b \\ 0 & 1 & 0 & c \\ 0 & 0 & 1 & d \end{pmatrix} \cdot \begin{pmatrix} b & c & d & -1 \\ c & d & -1 & 0 \\ d & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}.$$

is not obvious at a glance, but, once it is formed, all it takes is a blink to see that the result

$$KL = \begin{pmatrix} -a & 0 & 0 & 0 \\ 0 & c & d & -1 \\ 0 & d & -1 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} = E$$

is symmetric. Even a casual glance at L is enough to prove that L is invertible. (**Exercise 2**: the invertibility of L implies that the transpose L' of L is invertible also and, moreover, $(L')^{-1} = (L^{-1})'$: the transpose of the inverse of every invertible matrix is equal to the inverse of its transpose. Recall: the transpose $A' = (a'_{ij})$ of a matrix $A = (a_{ij})$ is defined by $a'_{ij} = a_{ji}$.) Consequence: $F = L^{-1}$ can be formed; since the inverse of a symmetric matrix is symmetric it follows that

$$K = EL^{-1} = EF$$

is the product of two symmetric matrices.

Matrices of the form K are called *companion matrices*, and they are what the deepest theorem of linear algebra is about. The word "companion" refers to a relation between matrices and polynomials. The relation is that the entries of the last column of a companion matrix can be used as the coefficients of a polynomial, and, in reverse, the coefficients of a polynomial can be used as the last column of a companion matrix. More precisely, the polynomial associated with the companion matrix K above is

$$(-1)^4(\lambda^4 - (a + b\lambda + c\lambda^2 + d\lambda^3)).$$

The odd-looking coefficient $(-1)^4$ is a reminder: for $n \times n$ matrices it is to be replaced by $(-1)^n$. (**Exercise 3**: what is the relation between the companion polynomial and the characteristic polynomial of a companion matrix?)

Reminder: two matrices are called *similar* if they represent the same linear transformation with respect to two possibly different bases of the underlying vector space, or equivalently, two matrices A and B are called similar if there exists an invertible matrix S such that $B = S^{-1}AS$. (The matrix S determines the transformation that sends one of the bases onto the other.) The deepest theorem of linear algebra is the statement that every matrix is similar to its *rational canonical form*. Reference: [8, p. 360]. In detail: a matrix is in rational canonical form if it is a direct sum of companion matrices, or, in other words, if it is of the form

$$\begin{pmatrix} K_1 & 0 & \cdots & 0 \\ 0 & K_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & K_r \end{pmatrix},$$

where each K_j is a companion matrix (and where, moreover, the companions of the K_j's are related to one another by certain divisibility properties that are irrelevant to the present application).

Once every companion matrix is known to be a product of two symmetric matrices, then the rational canonical form implies that every matrix is similar to a product of two symmetric matrices. Does that imply that every matrix is actually equal to a product of two symmetric ones? The question is whether representability as a product of symmetric matrices is preserved by similarity. In other words: if $A = EF$, with E and F symmetric, and if $B = S^{-1}AS$, does it follow that there exist symmetric matrices G and H such that $Y = GH$?

There is a standard device for answering questions like that, and it's important to realize that it does not work here. The device is to insert SS^{-1} as an extra factor in a computation; since $SS^{-1} = 1$, the insertion does no harm, and, sometimes, it puts into evidence an unexpected property of the product being studied. If, for instance, $A = EF$ and $B = S^{-1}AS$, so that $B = S^{-1}EFS$, then

$$B = (S^{-1}ES)(S^{-1}FS),$$

so that $B = GH$ with $G = S^{-1}ES$ and $H = S^{-1}FS$. In other words, if A, a product of two factors, is similar to B, then B is the product of the factors obtained from those of A by transforming them with the same similarity. All that remains is to verify that $S^{-1}ES$ and $S^{-1}FS$ are just as symmetric as E and F—that's all that remains, but it just happens not to be universally true. In other words, a matrix similar to a symmetric one can perfectly well fail to be symmetric. Examples are easy to come by:

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

Since the transpose of a product of two matrices is the product of their transposes in reverse order ($(AB)' = B'A'$), what is true is that if A is symmetric, then $S'AS$ is also symmetric. (Proof: $(S'AS)' = S'A'S''$.) This remark makes possible a small modification of the standard device that fails so as to produce a device that works. The modified device is to choose the inserted form of the identity matrix so as to balance with transposes some of the factors occurring earlier and later. Example: insert $S'^{-1}S'$, so as to balance an earlier occurring S^{-1} and a later S. Special case at hand: if A is the product of two symmetric matrices, $A = EF$, and if $B = S^{-1}AS = S^{-1}EFS$, then

$$B = (S^{-1}ES'^{-1})(S'FS)$$

—and the latter factors are just as symmetric as E and F.

Conclusion: since every matrix is similar to one that is the product of two symmetric matrices, and since, by the modified device just described, every matrix similar to such a product is a product of the same kind, it follows that every matrix is a product of two symmetric ones—victory! Reference: Radjavi [12].

Comment: the proof is sufficiently "constructive" that it can be used in concrete numerical cases for explicit computation. Thus, for instance, symmetric factorings of

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 4 & 5 & 6 \end{pmatrix}$$

can be found by following the proof step by step. Some free parameters enter in the process, namely at the stage where a prescribed matrix has to be reduced to rational canonical form. There are many similarities that will do the work, and they can be chosen at random with nothing in mind except possibly an attempt to make the resulting numbers simpler than some other choices might make them.

2. HERMITIAN

What happens in the complex case? The discussion in the preceding section is valid for any field, and, in particular, not only for the field $\mathbb{R}$ of real numbers but also for the field $\mathbb{C}$ of complex numbers. For matrices over $\mathbb{C}$, however, symmetry is not the most useful concept. The field $\mathbb{C}$ has another structure, namely complex conjugation, and wisdom and experience warn that we ignore it at our peril. The role of the transpose of a complex matrix A should, accordingly, be played by the adjoint (conjugate transpose) of A defined by $A^* = \overline{A}'$, or, more explicitly, by the matrix A^* whose entries are $a^*_{ij} = \overline{a_{ji}}$. If that change is made, then the role of symmetric matrices is played by the Hermitian matrices, that is the matrices A for which $A = A^*$.

Is the product of two Hermitian matrices Hermitian? Since the Hermitian matrices that happen to have real entries are just the symmetric ones, the negative answer to that question is already known. All right, then: which matrices over $\mathbb{C}$ are products of Hermitian ones? At this point a new phenomenon enters the theory of factoring bad matrices into good ones, namely the consideration of determinants. The determinant of a Hermitian matrix is always real. (Possible proof: the determinant of every matrix over $\mathbb{C}$ is the product of its eigenvalues, and the eigenvalues of a Hermitian matrix are real.) Consequence: every product of Hermitian matrices has real determinant, and, therefore, a necessary condition for the Hermitian factorability of a matrix A is that its determinant be real. Is that condition sufficient as well? The question is not trivial; what follows is a more detailed search for other possible necessary conditions. Assume therefore that $A = EF$, with E and F Hermitian; what besides the realness of the determinant can be said about A?

If at least one of the factors, say E, is invertible, then another necessary condition is easy to discover: in that case

$$A = EF = E(FE)E^{-1} = EA^*E^{-1},$$

so that, surprise, A is similar to A^*. (**Exercise 4**: what if E is not invertible but F is? **Exercise 5**: can a matrix have real determinant but not be similar to its adjoint? What about the other direction: can a matrix be similar to its adjoint but not have a real determinant?) In case both the Hermitian factors of A are singular, a new technique enters.

Some of the techniques needed for the discussion of Hermitian factoring are the same as the ones used in the symmetric case. Note, in particular, that

$$S^{-1}AS = S^{-1}EFS = (S^{-1}ES^{*-1})(S^*FS).$$

Since the factors on the right are Hermitian along with E and F, it follows that the property of being a product of two Hermitian matrices is preserved by similarity (although the property of being Hermitian is not so preserved.)

Since

$$AE = EFE = EA^*,$$

it follows inductively that

$$A^nE = EA^{*n}$$

for every positive integer n. Since every matrix is similar to a direct sum of a nilpotent one and an invertible one [8, p. 113, Theorem 1], the rest of this discussion assumes, with no loss of generality, that the prescribed matrix A is already of the form $\begin{pmatrix} J & 0 \\ 0 & K \end{pmatrix}$ where J is nilpotent and K is invertible. Write E and F as matrices with respect to the same decomposition of the space:

$$E = \begin{pmatrix} L & M \\ M^* & N \end{pmatrix}, \qquad F = \begin{pmatrix} P & Q \\ Q^* & R \end{pmatrix}.$$

It follows that

$$\begin{pmatrix} J^n & 0 \\ 0 & K^n \end{pmatrix}\begin{pmatrix} L & M \\ M^* & N \end{pmatrix} = \begin{pmatrix} L & M \\ M^* & N \end{pmatrix}\begin{pmatrix} J^{*n} & 0 \\ 0 & K^{*n} \end{pmatrix},$$

and hence that

$$K^n M^* = M^* J^{*n}$$

for every positive integer n. Since $J^{*n} = 0$ for n sufficiently large, and since K is invertible, it follows that $M = 0$, and hence that

$$\begin{pmatrix} J & 0 \\ 0 & K \end{pmatrix} = \begin{pmatrix} L & 0 \\ 0 & N \end{pmatrix} \begin{pmatrix} P & Q \\ Q^* & R \end{pmatrix}$$

and

$$\begin{pmatrix} J & 0 \\ 0 & K \end{pmatrix} \begin{pmatrix} L & 0 \\ 0 & N \end{pmatrix} = \begin{pmatrix} L & 0 \\ 0 & N \end{pmatrix} \begin{pmatrix} J^* & 0 \\ 0 & K^* \end{pmatrix}.$$

The first of these equations implies that $K = NR$. The invertibility of K implies then that N and R are invertible also, and then the second equation implies that

$$N^{-1} K N = K^*.$$

At this point a powerful canonical form theorem is needed again; the pertinent fact is that every nilpotent matrix is similar to its adjoint. (Either the rational canonical form, or its complex special case, the Jordan form, will do the trick. Reference: [8, p. 114.] The point is that the adjoint of a nilpotent matrix is exactly as nilpotent as the matrix itself, and, therefore, has the same Jordan form; two matrices with the same Jordan form are similar.) Consequence: there exists an invertible matrix H such that

$$H^{-1} J H = J^*.$$

Since the matrix

$$S = \begin{pmatrix} H & 0 \\ 0 & N \end{pmatrix}$$

is invertible, and

$$S^{-1} A S = A^*,$$

it follows A and A^* are similar.

The conclusion so arrived at is that if A is a product of two Hermitian matrices (which may or may not be invertible), then A is similar to A^*. In other words, the similarity of A and A^* is another necessary condition for Hermitian factorability—the possession of a real determinant is not enough.

Very well then: for A to be a product of two Hermitian matrices, it is necessary that there exist an invertible matrix S such that

$$AS = SA^*,$$

and hence (form adjoints) such that

$$AS^* = S^* A^*.$$

Take any complex number u, multiply the first of these equations by u and the second by $\overline{u}$, add, and get

$$A(uS + \overline{u}S^*) = (uS + \overline{u}S^*)A^*.$$

Write $E_u = uS + \bar{u}S^*$, and note that E_u is Hermitian and, since

$$AE_u = E_u A^*,$$

the product AE_u is Hermitian. If, in particular, $|u| = 1$, then

$$E_u = \left(SS^{*-1} + \frac{\bar{u}}{u}\right) uS^*.$$

Choose u so that $SS^{*-1} + \bar{u}/u$ is invertible; since the spectrum (the set of all eigenvalues) of SS^{*-1} is finite, that can always be done. It follows that E_u is invertible (and Hermitian along with AE_u.) Since

$$A = (AE_u)E_u^{-1},$$

it follows that A is the product of two Hermitian matrices (and, moreover, so that at least one of the two is invertible). (Reference: [10].) Summary: a matrix A is the product of two Hermitian matrices if and only if it is similar to its adjoint.

Is that the end of the story? For a hint to the answer consider the product

$$A = \begin{pmatrix} i & 0 \\ 0 & -2i \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

The eigenvalues of A are i and $-2i$; the adjoint of A is

$$A^* = \begin{pmatrix} -i & 0 \\ 0 & 2i \end{pmatrix}$$

with eigenvalues $-i$ and $2i$. The two sets of eigenvalues are not the same, which implies that A and A^* cannot be similar, and therefore A is not the product of two Hermitian matrices. Since, however, A *is* the product of Hermitian matrices, three of them, there is more to the story than the necessary and sufficient condition derived above tells. The restriction to two factors is the right thing to do when the answer is always yes (as it was for symmetric factors). If, however, two factors are enough sometimes, and at other times three are needed, and, for all we know, four, or five, or any number, then the question must be reformulated. Which matrices are products of three, or four, or any finite number of Hermitian matrices?

One of the necessary conditions mentioned in the preceding discussion, namely real determinant, is obviously still necessary. It is pleasant to be able to report that that condition, by itself, is sufficient for the general case. That is: *a complex matrix is a product of a finite number of Hermitian matrices if and only if its determinant is real, and, if that condition is satisfied, then it is the product of (not more than) four Hermitian matrices*. That's Radjavi's theorem [12], and what follows is a presentation of its proof.

The "only if" is trivial: since the determinant of a matrix is the product of its eigenvalues, every Hermitian matrix has real determinant, and, therefore, so does every finite product of Hermitian operators. The "if" can be treated by proving directly that every matrix with real determinant is the product of four Hermitian ones.

Observe to begin with that if $A = EFGH$, with E, F, G, and H Hermitian, and if S is an arbitrary invertible matrix, then

$$S^{-1}AS = (S^{-1}ES^{*-1})(S^*FS)(S^{-1}GS^{*-1})(S^*HS),$$

and therefore every matrix similar to a product of four Hermitians is a product of four Hermitians. Consequence: it is sufficient to prove that every matrix in rational canonical form, with real determinant, is a product of four Hermitian ones.

Begin with an easy special case, $A = \left(\begin{smallmatrix} 0 & a \\ 1 & b \end{smallmatrix}\right)$ (rational canonical form); note that $\det A = -a$. Motivation for next step: the principal entry a is easier to deal with if it is in a diagonal position, and that can be achieved by permuting the columns. Write $X = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ and calculate $AX^{-1} (= AX) = \left(\begin{smallmatrix} 0 & a \\ 1 & b \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} a & 0 \\ b & 1 \end{smallmatrix}\right)$. Since X is already Hermitian, it remains to prove that the product AX^{-1} is a product of three (or fewer) Hermitian matrices, and it is (more than) sufficient to prove that AX^{-1} is similar to a real matrix (because then it is similar to a product of two real symmetric matrices, and such matrices are Hermitian). Indeed: a similarity of AX^{-1} with a real matrix can be achieved via a diagonal matrix $\left(\begin{smallmatrix} \alpha & 0 \\ 0 & \beta \end{smallmatrix}\right)$. The proof is a small computation:

$$\begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}^{-1} \begin{pmatrix} a & 0 \\ b & 1 \end{pmatrix} \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} = \begin{pmatrix} a & 0 \\ \frac{\alpha}{\beta} b & 1 \end{pmatrix}.$$

All that is needed is to choose α and β so as to make $(\alpha/\beta)b$ real, and that is obviously always possible.

The next most complicated case is 3×3; look at

$$A = \begin{pmatrix} 0 & 0 & a \\ 1 & 0 & b \\ 0 & 1 & c \end{pmatrix}.$$

This time the permutation of columns is given by

$$X = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

for which

$$X^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

This time X is not Hermitian, but it is real, and, therefore, it is the product of two Hermitian matrices. Since

$$AX^{-1} = \begin{pmatrix} 0 & 0 & a \\ 1 & 0 & b \\ 0 & 1 & c \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} a & 0 & 0 \\ b & 1 & 0 \\ c & 0 & 1 \end{pmatrix},$$

with determinant a, and since X^{-1} is the product of two real symmetric matrices, it is, once again, sufficient to prove that AX^{-1} is similar to a real matrix. Once again such a similarity can be effected by a diagonal matrix:

$$\begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{pmatrix}^{-1} \begin{pmatrix} a & 0 & 0 \\ b & 1 & 0 \\ c & 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{pmatrix} = \begin{pmatrix} a & 0 & 0 \\ \frac{\alpha}{\beta}b & 1 & 0 \\ \frac{\alpha}{\gamma}c & 0 & 1 \end{pmatrix}.$$

What has to be verified is that by appropriate choices of α, β, and γ the latter matrix can be made real, and that is easy.

Neither of the preceding special cases is typical: they fairly represent what can happen with companion matrices, but the most general case is that of a direct sum of companion matrices, and that needs one more look. A totally typical case is the direct sum

$$A = \begin{pmatrix} 0 & a & 0 & 0 & 0 \\ 1 & b & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & c \\ 0 & 0 & 1 & 0 & d \\ 0 & 0 & 0 & 1 & e \end{pmatrix}$$

with determinant $-$ ac. In this case the appropriate permutation matrix is

$$X = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix},$$

with

$$X^{-1} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

As above, the thing to do is to compute AX^{-1} and prove that it is similar to a real matrix. Once the program is clear, the steps are routine, and, as before, the desired similarity can be achieved by a diagonal matrix. Since

$$AX^{-1} = \begin{pmatrix} 0 & a & 0 & 0 & 0 \\ 1 & b & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & c \\ 0 & 0 & 1 & 0 & d \\ 0 & 0 & 0 & 1 & e \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & a & 0 & 0 \\ 0 & 1 & b & 0 & 0 \\ c & 0 & 0 & 0 & 0 \\ d & 0 & 0 & 1 & 0 \\ e & 0 & 0 & 0 & 1 \end{pmatrix},$$

the matrix to compute is one of the form

$$\begin{pmatrix} \alpha & 0 & 0 & 0 & 0 \\ 0 & \beta & 0 & 0 & 0 \\ 0 & 0 & \gamma & 0 & 0 \\ 0 & 0 & 0 & \delta & 0 \\ 0 & 0 & 0 & 0 & \mathcal{E} \end{pmatrix}^{-1} \begin{pmatrix} 0 & 0 & a & 0 & 0 \\ 0 & 1 & b & 0 & 0 \\ c & 0 & 0 & 0 & 0 \\ d & 0 & 0 & 1 & 0 \\ e & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha & 0 & 0 & 0 & 0 \\ 0 & \beta & 0 & 0 & 0 \\ 0 & 0 & \gamma & 0 & 0 \\ 0 & 0 & 0 & \delta & 0 \\ 0 & 0 & 0 & 0 & \mathcal{E} \end{pmatrix},$$

which turns out to be

$$\begin{pmatrix} 0 & 0 & \frac{\gamma}{\alpha}a & 0 & 0 \\ 0 & 1 & \frac{\beta}{\gamma}b & 0 & 0 \\ \frac{\alpha}{\gamma}c & 0 & 0 & 0 & 0 \\ \frac{\alpha}{\delta}d & 0 & 0 & 1 & 0 \\ \frac{\alpha}{\mathcal{E}}e & 0 & 0 & 0 & 1 \end{pmatrix}.$$

The problem of making all the entries real is easily solved: just write $a = c = 1$, and then choose b and d and e so as to make each of $(1/\beta)b$ and $(1/\delta)d$ and $(1/\mathcal{E})e$ real.

3. POSITIVE

One of the most important and most useful subsets of the set of Hermitian matrices is the set of positive ones. "Positve" is used here to mean what cautious classical pedantry describes as "non-negative semi-definite". The classical definition involves the quadratic form associated with a matrix and demands that all the values of that form be non-negative real numbers. According to an equivalent definition E is positive if it has a Hermitian square root (that is, there exists a Hermitian F such that $E = F^2$); or, also equivalently, if there exists a matrix X such that $E = X^*X$.

Positive matrices are, in particular, Hermitian; they are distinguished within the class of Hermitian matrices by having only positive (meaning non-negative) eigenvalues. Positive matrices are in many respects excellent analogues of positive numbers but such "non-commutative numbers" tend to behave in unpredictable ways.

Which matrices are products of positive ones? Since the determinant of a positive matrix is not only real but in fact positive, an obvious necessary condition that A be the product of a finite number of positive matrices is that A have positive determinant. Could that possibly be enough? Consider, as an extreme case, the matrix

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

It has positive determinant, yes, but it looks as negative as any matrix can be; most people would regard a representation of it as a product of a finite number of positive matrices surprising if not shocking. Here is the surprise:

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 5 & 7 \\ 7 & 10 \end{pmatrix} \begin{pmatrix} \frac{13}{2} & -5 \\ -5 & 4 \end{pmatrix} \begin{pmatrix} 8 & 5 \\ 5 & \frac{13}{4} \end{pmatrix} \begin{pmatrix} \frac{25}{8} & -\frac{11}{2} \\ -\frac{11}{2} & 10 \end{pmatrix}.$$

A large amount of work was done on products of positive matrices by Ballantine (see [1] and its predecessors). The most striking outcome of the work is the statement that *if* $\det A > 0$, *then* A *is the product of five positive matrices*. Note: this is the *strictly* positive case. The theory has been extended to the case $\det A \geq 0$ (see in particular [16]). The simplest proof of Ballantine's principal result was given by Sourour [13]; here is how it goes.

Begin with almost, but not quite, the most general case: assume that $\det A > 0$ and A is not a scalar. Assertion: in that case A is the product of four positive matrices. There is no loss of generality in assuming that $\det A = 1$; if the theorem is known for that case, then any other positive value of the determinant can be absorbed by any one of the positive factors. At this point a great loss of generality will be permitted: I'll restrict attention to the case of 2×2 matrices only. The general case is more complicated, but not conceptually so; it goes by induction on the dimension, and the main idea is the same as in the 2×2 case.

Since A is not a scalar, there exists a non-zero vector x that is not an eigenvector of A, and, in particular, such that if $y = Ax - x$, then $y \neq 0$. It follows that x and y are linearly independent—for if $\alpha x + \beta(Ax - x) = 0$, then $\beta Ax = (\beta - \alpha)x$. Since x is not an eigenvector, it follows that β must be 0, and since $x \neq 0$, it follows that $\beta - \alpha = 0$, which proves the asserted linear independence.

The set $\{x, y\}$ is a basis for the 2-dimensional space at hand. Since $Ax = x + y$, the first column of the matrix of A with respect to that basis is $\binom{1}{1}$. In view of the assumption that $\det A = 1$, it follows that the matrix of A is of the form $\left(\begin{smallmatrix} 1 & \xi \\ 1 & \xi+1 \end{smallmatrix}\right)$, and therefore that A can be factored as follows:

$$\begin{pmatrix} 1 & \xi \\ 1 & \xi + 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & 0 \\ \frac{1}{2} & 2 \end{pmatrix} \begin{pmatrix} 2 & 2\xi \\ 0 & \frac{1}{2} \end{pmatrix}.$$

Write

$$B = \begin{pmatrix} \frac{1}{2} & 0 \\ \frac{1}{2} & 2 \end{pmatrix} \quad \text{and} \quad C = \begin{pmatrix} 2 & 2\xi \\ 0 & \frac{1}{2} \end{pmatrix}.$$

Since each of these two matrices has distinct eigenvalues, each of them is similar to a diagonal matrix. That is: $B = S^{-1}ES$ and $C = T^{-1}FT$, with

$$E = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 2 \end{pmatrix} \quad \text{and} \quad F = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}.$$

It follows that

$$B = (S^{-1}S^{*-1})(S^*ES)$$

and

$$C = (T^{-1}T^{*-1})(T^*FT),$$

and each of the two factors in each of these two representations is positive. That concludes the proof.

Comment: if $\det A > 0$ and A is a positive scalar, then A is obviously the product of four positive matrices, because $A = A \cdot 1 \cdot 1 \cdot 1$. Conclusion: if $\det A > 0$, and A is not a negative scalar, then A is a product of four positive matrices.

Corollary: if $\det A > 0$, then A is a product of five positive matrices. Proof: in view of the principal theorem, the only case that needs to be considered is the one in which A is a negative scalar, and that case reduces to the case $A = -1$. In that case, let P be a positive matrix that is not a scalar, and note that $A = (-P^{-1})P$. Apply the principal theorem to the factor $-P^{-1}$ to represent it as a product of four positives—since P itself plays the role of the fifth positive factor, the proof is complete.

The result about $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ is, of course, a special case of what was just proved. Here too, just as for symmetric factorization, the proof is constructive, and concrete factorings can be found by following the proof step by step.

4. INFINITE

Do infinite matrices make sense in this context? Answer: with extreme caution only. The trouble is that infinite matrices are not easy to multiply; infinite series enter the formula for the entries of the product, and difficulties with convergence abound.

Questions about factoring matrices into symmetric ones might be asked in the infinite case, but I don't know of any work at all along those lines.

Questions about Hermitian and positive factoring are of a completely different character. Such words make sense in the theory of operators on Hilbert spaces, and the case of finite-dimensional Hilbert spaces is in effect what the preceding sections treat. With a little reformulation they can be asked in the infinite case, and it turns out that some of the answers remain the same, others change, and many are unknown.

Probably the best known Hilbert space is the sequence space ℓ^2 consisting of all infinite sequences $x = \{x_1, x_2, x_3, \ldots\}$ such that $\sum_{n=1}^{\infty} |x_n|^2 < \infty$. For such sequences a natural inner product is defined by $(x,y) = \sum_{n=1}^{\infty} x_n \bar{y}_n$, and the corresponding natural norm is defined by $\|x\|^2 = (x,x)$. The well-behaved linear transformations (operators) A on ℓ^2 are the bounded ones, meaning the ones for which there exists a positive constant c with the property that $\|Ax\| \le c\|x\|$ for all x; the smallest such c is the norm, $\|A\|$ of A. (There is an equivalent and simpler definition of boundedness, which, however is less useful; according to that definition A is bounded if it maps ℓ^2 into ℓ^2.) A (bounded) operator A always has an adjoint A^* characterized by the identity $(Ax, y) = (x, A^*y)$ for all x and y. The operator A is Hermitian in case $A = A^*$; it is positive in case it has a Hermitian square root, or, equivalently, in case it is of the form X^*X for some operator X.

A comment about the relations between operators and matrices is in order. Every operator A corresponds to a matrix (a_{ij}) this way: if $x = \{x_n\}$ and if $y = Ax = \{y_n\}$, then $y_i = \sum_{j=1}^{\infty} a_{ij}y_j$. If, moreover, A and B are operators, with corresponding matrices (a_{ij}) and (b_{ij}), then linear combinations of the operators correspond to linear combinations of their matrices, the adjoint A^* of an operator corresponds to the conjugate transpose matrix $(\overline{a_{ji}})$, and, most importantly, the product AB of operators corresponds to the matrix product, that is to the matrix

$$\left(\sum_{k=1}^{\infty} a_{ik}b_{kj}\right).$$

Everything converges that should converge; the correspondence between operators and matrices is perfect.

Well, almost perfect. Every word in the preceding paragraph applies equally well to the finite and infinite-dimensional case, but in the finite-dimensional case more is true, namely that not only does every operator determine a matrix, but, conversely, every matrix is determined by an operator. In the infinite-dimensional case the latter part is not true. The boundedness condition is quite severe. The matrices determined by the (bounded) operators might well be called bounded matrices (and usually they are called just that), but not every matrix is bounded. The matrix that has every entry off the main diagonal equal to 0 and the entries on the main diagonal equal to 1 (the identity matrix) is bounded, but the matrix that has *every* entry equal to 1 is not, and neither is the matrix that has 0's off the main diagonal and the entry n in the nth position of the diagonal. The correspondence is perfect between bounded operators and bounded matrices, and that's as far as the good news goes.

If an operator is the product of two Hermitian operators of which one is invertible, then it is similar to its adjoint; the simple argument that proved that assertion in the finite case is purely algebraic and works just as well in the infinite case. If the invertibility condition is not satisfied, things change.

Consider the Hilbert space ℓ^2 and on it the Hermitian operator

$$E = \operatorname{diag}\left(1, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{3}}, \dots\right)$$

defined by

$$E(x_1, x_2, x_3, \dots) = \left(x_1, \frac{1}{\sqrt{2}}x_2, \frac{1}{\sqrt{3}}x_3, \dots\right).$$

To define F, first define the projection P onto the 1-dimensional space spanned by the vector $y = (1, \frac{1}{2}, \frac{1}{3}, \ldots)$, so that

$$Px = \frac{(x,y)y}{\|y\|^2}$$

for every x in ℓ^2, and define F by $F = 1 - P$.

If $A = EF = E(1-P) = E - EP$, then

$$Ay = Ey - EPy = Ey - Ey = 0,$$

so that $y \in \ker A$. (The abbreviation "ker" stands for "kernel" or "null space"; $\ker A$ is the set of all those vectors x for which $Ax = 0$.) As a matter of fact $\ker A$ is exactly the span of y. Indeed, if $x \perp y$, then $Ax = E(1-P)x = Ex$, which can never be 0 (unless $x = 0$), because (obviously) $\ker E = 0$.

What about $\ker A^* = \ker FE = \ker(E - PE)$? If $Ex = PEx$, then

$$Ex \in \operatorname{ran} E \cap \operatorname{ran} P.$$

(Here "ran" stands for "range", or, in many other people's usage, "image".) Since every vector in $\operatorname{ran} P$ is a multiple of y, and vice versa, the question becomes: which multiples of y are in $\operatorname{ran} E$? Since a non-zero multiple of y is in $\operatorname{ran} E$ if and only if $y \in \operatorname{ran} E$, the question becomes: does y belong to $\operatorname{ran} E$? Since E is one-to-one, the answer is easy to see: if $y = Ez$, with

$$z = (z_1, z_2, z_3, \ldots),$$

then $1/n = 1/(\sqrt{n})z_n$ for $n = 1, 2, 3, \ldots$, so that $z_n = 1/\sqrt{n}$—which is impossible—the sequence with those terms does not belong to ℓ^2. Consequence: $\ker A^* = 0$.

Conclusion: since $\ker A^* = 0$ and $\ker A \neq 0$, it is not true that A and A^* are similar—the finite-dimensional statement can be false in infinite-dimensional spaces. (There is some technical merit in knowing that an example of this sort can be constructed even if it is demanded that $\dim \ker A = \dim \ker A^*$—that was done by Gray [3]. Recall that if A is the product of two Hermitian operators of which at least one is invertible, then there is an easy algebraic argument that works equally well in the finite and infinite-dimensional cases.)

What about the converse: if A is similar to A^*, does it follow that A is the product of two Hermitian operators? The answer is not known. Conjecturally, the answer is yes, but it is not known even in the invertible case.

So much, for the present, about Hermitian factors; what can be said about positive factors? Not everything is known, but some things are; typical among them is the following statement about the invertible case. An invertible operator is the product of two positive operators (which then are necessarily invertible) if and only if it is similar to a positive operator (which then is necessarily invertible). The proof uses a well-known classical technique; it runs as follows. Assume that $A = EF$, with A invertible, and E and F positive (and then necessarily invertible). The assumption implies that, since

$$\left(\sqrt{E}\right)^{-1} EF \left(\sqrt{E}\right) = \left(\sqrt{E}\right) F \left(\sqrt{E}\right),$$

and since the left side of this equation is similar to A and the right side is positive, the operator A must indeed be similar to a positive operator. Conversely, if $A = X^{-1}CX$, where X is, of course, invertible, and C is positive, then

$$A = (X^{-1}X^{-1*})(X^*CX)$$

is the product of two positive operators. (The technique is the one used to prove the familiar finite-dimensional statement that a product of two positive matrices must have positive eigenvalues. In fact the technique proves that such a product is similar to a positive matrix.)

5. NORMAL

The best known and most natural generalizations of Hermitian operators are the normal ones. Recall that an operator A is normal if $A^*A = AA^*$. In the finite-dimensional case the so-called spectral theorem says that a matrix is normal if and only if it is unitarily equivalent (meaning unitarily similar) to a diagonal matrix—in other words, the normal matrices are exactly the (unitarily) diagonalizable ones. A Hermitian operator is trivially normal (every operator commutes with itself), and so is a unitary operator (every invertible operator commutes with its inverse, and, by definition, an operator U is unitary if and only if U^* is the inverse of U).

Which operators are products of normal ones? In the finite-dimensional case, the use of the best known and probably most important factorization theorem makes the answer to the question turn out to be trivial. The pertinent factorization theorem is the so-called polar decomposition. The first statement of polar decomposition that most of us learn is that every linear transformation on a finite-dimensional *real* vector space is the product of a rotation-or-reflexion and a dilatation; more elegantly, every finite matrix is the product of an orthogonal matrix and a positive one. According to the complex version, every finite matrix is the product of a unitary matrix and a positive one—that is the polar decomposition. (Caution: it is not asserted that the representation is unique, and sometimes it is not. A more detailed description of the polar decomposition exists, [5, p. 170] which yields a unique result. For present purposes the additional detailes are irrelevant.) Since unitary matrices and positive matrices are normal, it follows that *every matrix is the product of two normal ones*, and that's all that deserves to be said in the finite-dimensional case.

In the infinite case things change again. Consider the so-called *unilateral shift*, call it S, defined on ℓ^2 by

$$S(x_1, x_2, x_3, \ldots) = (0, x_1, x_2, x_3, \ldots).$$

The adjoint S^* of the unilateral shift S is given by

$$S^*(x_1, x_2, x_3, \ldots) = (x_2, x_3, x_4, \ldots).$$

An obvious and easy computation shows that $S^*S = 1$ (which makes S look unitary), but $SS^* \neq 1$ (which implies that S is not invertible, and, therefore, not unitary). The point of mentioning S and some of its properties here is that *the unilateral shift is not the product of a finite number of normal operators.*

The principal tool in the proof is the observation that if a normal operator has a one-sided inverse, then it has an inverse. (Proof: for every operator, left invertibility is the same as boundedness from below, which means the existence of a positive constant c such that $\|Ax\| \geq c\|x\|$ for all x. Boundedness from below for a normal operator is the same as boundedness from below for its adjoint.) Suppose indeed that

$$S = E_1 \cdots E_n,$$

where $E_1 \cdots E_n$ are normal. Since

$$S^* = E_n^* \cdots E_1^*,$$

it follows that

$$E_n^* \cdots E_1^* E_1 \cdots E_n = 1,$$

and hence that E_n is left invertible. In view of the preceding comments this implies that E_n is invertible, and therefore so is E_n^*. Invertible operators can be peeled off either end of a product without altering its invertibility character. It follows by an obvious inductive repetition of the argument that each of the E's is invertible, and so therefore is S. This is a contradiction, and the proof is complete.

Granted that not every operator is a product of normal ones, it makes sense to ask again, and in fact it is almost obligatory to ask, which ones are? As the discussion above shows, this is an infinite-dimensional question; in the finite-dimensional case it is so easy that it almost shouldn't be asked. The proof that every (finite) matrix is the product of two normal ones depended on the polar decomposition. There is a good and usable infinite version of the polar decomposition theorem, but it is not identical in statement with the finite-dimensional one, and, consequently, it does not have the same implications. In the infinite case it is simply not true that every operator is the product of a unitary operator and a positive one.

The fact is that the polar decomposition theorem, even for finite-dimensional spaces, is more delicate than the one usually stated in linear algebra books. The truth is that every operator (on finite or infinite-dimensional spaces) is a product UP, where U is a partial isometry and P is positive. (A partial isometry is like a unitary operator in that it often preserves distances, but it is unlike a unitary one in that it doesn't always do so. Precisely: to say that U is a partial isometry means that it is an isometry on the orthogonal complement of its kernel.)

If the prescribed operator is invertible, then the polar representation is unique, and the partially isometric factor is unitary. In general, however, the polar representation of an operator, as here described, is not necessarily unique. The most degenerate, and seemingly unimportant, case of non-uniqueness is visible in the representation of 0. In that case the "positive" factor P is the operator 0, and the unitary factor U can be any unitary operator whatever. (The phenomenon is visible even in the 1-dimensional case, that is, the classical polar form of a complex number. It is always true that $z = \rho e^{i\theta}$, and ρ is uniquely determined as $|z|$, but if $z = 0$, then θ can be changed to an arbitrary real number without destroying the equation.) The statement of the polar decomposition theorem can be sharpened as follows: every operator is a product UP, where U is a partial isometry and P is positive, and where, moreover, $\ker U = \ker P$; such a representation is uniquely determined by the given operator. If the underlying space is finite-dimensional, then every partial isometry can be "enlarged" to a unitary operator (in general in many ways), and that's what makes a representation as the product of two normal operators possible. In the infinite case such an enlargement is not always possible.

What can go wrong in the infinite case (as the unilateral shift shows) is that the initial space of a partial isometry (that is, the orthogonal complement of the kernel) does not sufficiently resemble the final space (that is, the range). For a partial isometry to have a unitary enlargement it is clearly necessary (and sufficient) that the orthogonal complements of its initial space and its final space have the same dimension, and in a finite-dimensional space that is always true. In an infinite-dimensional space, however, that condition can fail—that is the sense in which the initial space of a partial isometry may fail sufficiently to resemble the final space. If U is a partial isometry, the orthogonal complement of the initial space of U is $\ker U$, by definition of initial space, and the orthogonal complement of $\operatorname{ran} U$ is $\ker U^*$. In other words, a necessary and sufficient condition for U to possess a unitary enlargement is that $\dim \ker U = \dim \ker U^*$.

If T is an operator with polar decomposition UP, then $\ker U = \ker T (= \ker P)$, and, moreover, $\ker U^* = \ker T^*$. (Reason: $U^*f = 0 \Rightarrow UU^*f = 0 \Rightarrow U^*f \in \ker U \Rightarrow U^*f \in \ker P \Rightarrow PU^*f = 0 \Rightarrow T^*f = 0$; since $T^* = PU^*$, the reverse implication is trivial.) Consequence: if T is an operator such that $\dim \ker T = \dim \ker T^*$, then the partially isometric factor U of the polar decomposition $T = UP$ has the property that $\dim \ker U = \dim \ker U^*$, hence U has a unitary enlargement, hence T is a product of a unitary operator and a positive one, and hence, finally, T is a product of two normal operators. Thus a sufficient condition for T to be a product of *two* normal operators is that $\dim \ker T = \dim \ker T^*$.

It turns out that the condition is not quite necessary, but a careful study by Wu [15] reveals all the secrets. The major fact is that if T is a product of finitely many normal operators, and if T is a "good" operator, meaning in this case that $\operatorname{ran} T$ is closed, then $\dim \ker T = \dim \ker T^*$. Surprise: for good operators, being a finite product of normal operators implies being a product of two (but, of course, as the unilateral shift still stands by to show, a good operator may well fail to be a finite product of normal ones). The proof is not trivial, but it uses easily available standard techniques.

What then can be said about bad operators, the ones whose range is not closed? Surprising, and definitely non-trivial answer: bad operators are always finite products of normal operators, and, in fact, three are always enough. Corollary: every finite product of normal operators is a product of not more than three normal operators.

6. INFINITE, HERMITIAN

The problem of representing operators on an infinite-dimensional space as products of *two* Hermitian operators was discussed above; what can be said about finite products with possibly many factors? Since every Hermitian operator is normal, Hermitian product representability is more restrictive than normal product representability; if an operator is not a product of normal ones, then, all the more, it cannot be a product of Hermitian ones. All right then: which products of normal operators can be products of Hermitian ones? The question is essentially the same as this one: which normal operators are products of Hermitian ones? In the finite-dimensional case the answer turned out to be the ones whose determinant is real—but in the infinite case determinant makes no sense and a different answer has to be looked for.

For normal operators the polar decomposition is better behaved than in general; if A is normal, then it is always possible to write $A = UP$ with U unitary and P positive. An obviously pertinent question, therefore, is this: which unitary operators are products of Hermitian ones? That question was raised and answered a long time ago in [7]. An operator is called a *symmetry* if it is both unitary and Hermitian; those two properties imply that the operator is an involution also (that is: if $A = A^*$ and $A^* = A^{-1}$, then $A^2 = 1$, and, in fact, any two of these three properties imply the third). The relevant result about unitary operators is that every unitary operator is the product of two shifts, and every shift is the product of two symmetries (and, moreover, those symmetries can be chosen so that they have both $+1$ and -1 as eigenvalues of infinite multiplicity). Definition: an operator T is a shift if it is unitary and if there exists a sequence $\{H_n\}$ of equidimensional subspaces, $n = 0, \pm 1, \pm 2, \ldots$, such that $TH_n = H_{n+1}$ for all n. The main algebraic idea behind the proof that every shift is the product of two symmetries is that the product GF of the two mappings $F : n \mapsto -n$ and $G = n \mapsto 1 - n$ of $\mathbf{Z}$ onto itself is the mapping $n \mapsto n + 1$—that is, that a translation can be expressed as the product of two reflections, or, in parallel language, that a shift can be expressed as the product of two symmetries.

Once that's done, the question about factoring a unitary operator into Hermitian ones is solved: every unitary operator is the product of four symmetries, and, therefore, every normal operator is the product of five Hermitian ones. That result is, in fact, not the best possible one: Radjavi [12] has proved that every normal operator is the product of four Hermitian ones.

The results above imply that in the infinite-dimensional case the products of normal operators and the products of Hermitian operators and the products of positive operators are all the same class. As for how many are needed: 3 normal factors, or 6 Hermitian ones, or 18 positive ones are enough, but it is not known whether these numbers are the best possible.

7. THE END?

Yes, that's the end as far as this report is concerned, but it is not the end of the subject. Any time you specify a class of matrices (operators), you can ask which matrices are finite products of the specified ones. Some such questions are interesting and some are not; some such questions are solved and some are not; and examples exist of all four intersections of these two classifications. The factorization problem has been studied and at least partially solved for involutions [4], for symmetries

[7], for partial isometries [9], for nilpotent matrices [14], and other classes, and the subject has many unanswered (and in some cases not yet asked) questions.

One factorization problem that does not seem to have received attention is the one for *unitarily diagonalizable* matrices. That phrase is quite a mouthful. I wish there were a better word for it, but there isn't. In order to avoid having to say and write the whole thing over and over again, I shall temporarily use the artificial word *diagonable*. Formal definition: a linear transformation A (an operator) is diagonable if there exists an orthonormal basis for the space consisting of eigenvectors of A. In the matrix spirit: a matrix A is diagonable if there exists a unitary matrix U such that U^*AU is diagonal. Question: which matrices are products of diagonable ones?

In the finite-dimensional case the question degenerates. Since every matrix is a product of a unitary one and a positive one, and since all finite unitary and all finite positive matrices are diagonable, it follows that every finite matrix is a product of two diagonable matrices, and nothing more can be said.

The question becomes interesting (and unsolved) in the infinite-dimensional case. The answer to some small special cases can be derived from known facts, but the general case is open. Since a diagonable matrix is automatically normal, the problem is a refinement of the problem of normal factorization. Not every matrix is a finite product of normal ones, and, therefore, not every matrix is a finite product of diagonable ones. In effect the question reduces to this: which normal matrices are finite products of diagonable matrices? The known facts about symmetries imply that every unitary matrix is such a product. What about Hermitian matrices? Is every Hermitian matrix a finite product of diagonable ones?

Consider the special case of the "position operator" on $L^2(0,1)$. That is, the Hilbert space consists of all square-integrable functions on the unit interval, and the operator M is the one that sends each function f onto the function $g = Mf$ defined by $g(t) = tf(t)$. Is M a finite product of diagonable operators? No one knows.

References

[1] C. S. Ballantine, Products of positive definite matrices. IV, *Linear Algebra Appl.* **3** (1970), 79–114.

[2] G. Frobenius, Über die mit einer Matrix vertauschbaren Matrizen, *Sitzungsber. Preuss, Akad. Wiss* (1910), 3–15.

[3] L. J. Gray, Products of Hermitian operators, *Proc. A.M.S.* **59** (1976), 123–126.

[4] W. H. Gustafson, P. R. Halmos and H. Radjavi, Products of involutions, *Linear Algebra Appl.* **13** (1976), 157–162.

[5] P. R. Halmos, Finite-dimensional vector spaces, Springer, New York (1974).

[6] P. R. Halmos, Hilbert space problem book, Springer, New York (1974).

[7] P. R. Halmos and S. Kakutani, Products of symmetries, *Bull. A.M.S.* **64** (1958), 77–78.

[8] T. W. Hungerford, Algebra, Springer, New York (1974).

[9] K-H. Kuo and P.-Y. Wu, Factorization of matrices into partial isometries, *Proc. A.M.S.* **105** (1989), 263–272.

[10] H. Radjavi and J. P. Williams, Products of self-adjoint operators, *Michigan Math. J.* **16** (1969), 177–185.

[11] H. Radjavi, On self-adjoint factorization of operators, *Can. J. Math.* **21** (1969), 1421–1426.

[12] H. Radjavi, Products of Hermitian matrices and symmetries, *Proc. A.M.S.* **21** (1969), 369–372.

[13] A. R. Sourour, A factorization theorem for matrices, *Linear Multilin. Alg.* **19** (1986), 141–147.

[14] P.-Y. Wu, Products of nilpotent matrices, *Linear Algebra Appl.* **96** (1987), 227–232.

[15] P.-Y. Wu, Products of normal operators, *Canadian Journal of Mathematics* **40** (1988), 1322–1330.

[16] P.-Y. Wu, Products of positive semidefinite matrices, *Linear Algebra Appl.* **111** (1988), 53–61.

[17] P.-Y. Wu, The operator factorization problems, *Linear Algebra Appl.* **117** (1989), 35–63.