NUMERICAL RADII OF ZERO-ONE MATRICES

Kenneth R. Davidson and John A. R. Holbrook

1. Introduction. Recently V. Müller (see [6]) has constructed a striking example related to the following long-standing problem: What is the best constant C such that $w(TS) \leq Cw(T)\|S\|$ for all commuting operators T and S? The setting here is complex Hilbert space and $\|S\|$ denotes the usual operator norm. The numerical radius $w(T)$ may be defined as $\sup\{|(Tu, u)|: \|u\| = 1\}$. Müller's example shows that $C > 1.01$, in spite of various results that appeared to support the conjecture $C = 1$ (for further details see §3).

A study of Müller's example led us to related operators that have the following advantages. They are simpler to describe and can be set in a lower-dimensional space (dimension 9 seems to be the best we can do; the example of Müller lives in a 12-dimensional space). The computations can be carried out directly and the relevant numerical radii identified (Müller relies on a computer check of some aspects of his example). The constant C is more closely constrained by our examples (we shall see that $C \geq 1/\cos(\pi/9) > 1.064$; the best upper bound known for C appears to be $C < 1.169$). Müller refers to an approximation result of Holbrook (see [5]) to establish that $w(TS)$ may exceed $w(T)\|S\|$ even when T is a polynomial in S; in some of our examples this feature is built in. Finally, we can adapt our examples to settle a related problem about ρ-dilations (see §3).

We end this introduction by describing explicitly one of our examples. Let $S = S_9$, the shift on the Hilbert space of dimension 9. Let $T = S^3 + S^7$. Then, of course, $\|S\| = 1$, and we shall see that $w(T) = \cos(\pi/10)$ while $w(TS) = 1$, so that $C \geq 1/\cos(\pi/10)$ (> 1.05).

2. Zero-one matrices. Given an $n \times n$ matrix M whose entries are zeroes or ones, we form the incidence graph $G(M)$ with vertices labelled 1 through n and edges joining exactly those pairs (i, j) such that $m_{ij} = 1$.

PROPOSITION 1. *Suppose that an $n \times n$ zero-one matrix M has zero diagonal and, for each $1 \leq k \leq n$, no more than two ones in the cross-shaped region X_k defined as the union of the kth row and the kth column. Then*

$$w(M) = \cos(\pi/(L+1)),$$

where L is the number of vertices in the longest chain in the incidence graph $G(M)$.

REMARKS. The condition on the X_k in the statement of our proposition is a way of saying that $G(M)$ has no vertices of valence greater than 2, that is, that no more than two edges meet at a given vertex. As a result, $G(M)$ breaks into

Received November 24, 1987. Revision received February 12, 1988.
Research supported in part by NSERC of Canada.
Michigan Math. J. 35 (1988).

connected components that are either cycles or chains (with two free ends). We consider a cycle to be a chain of infinite length L, interpreting the formula of the proposition to say that $w(M)=1$. If i and j are distinct and $m_{ij}=m_{ji}=1$ then $G(M)$ has two edges joining the vertices i and j, forming a special sort of short cycle. Finally, if $M=0$ then each vertex in $G(M)$ is isolated and $L=1$.

Proof. Since the entries of M are nonnegative, $\max\{|(Mu,u)|: \|u\|=1\}$ is attained as (Mu,u) for some $u=(x_1,\dots,x_n)$, where each $x_i\geq 0$. In this case

$$(Mu,u)=\sum_{m_{ij}=1} x_i x_j$$

and this sum breaks into portions corresponding to the connected components of $G(M)$. As we have remarked, these components are either cycles or chains. Let us denote by $Cy_1, Cy_2, \dots$ the portions of the sum (Mu,u) corresponding to the cycles in $G(M)$, and by Vy_k the set of vertices belonging to the kth cycle. Let Ch_k and Vh_k refer to the chains of $G(M)$ in the same fashion. Then

$$(Mu,u)=\sum_k Cy_k+\sum_k Ch_k=\sum_k\left\{\left(\sum_{i\in Vy_k} x_i^2\right)Sy_k\right\}+\sum_k\left\{\left(\sum_{i\in Vh_k} x_i^2\right)Sh_k\right\},$$

where $Sy_k=Cy_k/(\sum_{i\in Vy_k} x_i^2)$ and Sh_k is similarly defined for the chains.

Since $\sum x_i^2=\|u\|^2=1$, we see that (Mu,u) is a convex combination of the Sy_k and Sh_k and that $w(M)$ is the maximum among these.

To maximize the individual Sy_k and Sh_k is a familiar problem. To simplify the notation, let us assume that the cycle corresponding to Sy_k involves the vertices $1,2,\dots,q$. Then

$$\max Sy_k=\max\left\{\sum_1^q x_{i-1}x_i : \sum_1^q x_i^2=1\right\},$$

where we interpret x_0 as x_q. The sum we are maximizing is certainly no greater than $\sum_1^q (x_{i-1}^2+x_i^2)/2$ and under our constraints this is 1; on the other hand, value 1 is attained by the choice $x_i=1/\sqrt{q}$ $(1\leq i\leq q)$. Thus $w(M)=1$ if there are any cycles. We treat the chain term Sh_k similarly. Supposing that this chain involves the vertices $1,2,\dots,q$, we have

$$\max Sh_k=\max\left\{\sum_2^q x_{i-1}x_i : \sum_1^q x_i^2=1\right\}.$$

There are several approaches to this computation. One way is to regard this as $w(S_q)$, where S_q denotes the shift on q-dimensional space represented as a $q\times q$ matrix with ones on the superdiagonal and zeroes elsewhere. However, $w(S_q)=w(\operatorname{Re} S_q)=$ maximum eigenvalue of the symmetric matrix $\operatorname{Re} S_q$. This is a tridiagonal matrix with zero diagonal and $\frac{1}{2}$ along the two neighboring diagonals, and it is easy to verify that a simple recurrence relation holds for its characteristic polynomial $P_q(x)$:

$$P_q(x)=xP_{q-1}(x)-\tfrac{1}{4}P_{q-2}(x),$$

with $P_1(x)=x$ and $P_0(x)=1$. This relation identifies $P_q(x)$ as a type of Chebyshev polynomial; in fact $P_q(x)=(\sin(q+1)\theta)/(2^q\sin\theta)$ where $\cos\theta=x$, because these functions satisfy the same recursion. Thus $P_q(x)$ has q distinct roots in $(-1,1)$ corresponding to $\theta = k\pi/(q+1)$ $(k=1,2,\ldots,q)$, and the largest eigenvalue is $x=\cos(\pi/(q+1))$.

Since the largest of these values occurs for the chain where q is maximal (i.e., when $q=L$), the formula of our proposition is verified. □

The following special case of Proposition 1 was noted (with a different proof) by Allen in [1, §3.5].

COROLLARY 2. *If S_n denotes the shift on n-dimensional Hilbert space, then*

$$w(S_n^m) = \cos(\pi/([(n-1)/m]+2))$$

for all positive integers n and m (here $[x]$ denotes the integer part of x).

Proof. Evidently the chains in $G(S_n^m)$ are of the form

$$i-(i+m)-(i+2m)-\cdots$$

and the longest of these begins with 1 and ends with $1+km$, where k is maximal subject to $1+km \le n$; that is, $k=[(n-1)/m]$. The number of vertices in this longest chain is $L=k+1$. □

REMARK. The same considerations show that S_n^m is unitarily equivalent to an orthogonal sum of shifts of smaller dimension, the largest dimension being the L of the last proof. Hence it is no surprise that $w(S_n^m)=w(S_L)$. In the general case treated in the proposition, M is not usually unitarily equivalent to such a sum of shifts. Nevertheless, an alternate proof of the proposition might be based on the unitary equivalence of $\mathrm{Re}(M)$ with the real part of such a sum.

COROLLARY 3. *There exist commuting 9-dimensional operators T and S such that $w(TS) > 1.05w(T)\|S\|$. In fact, with $S=S_9$ (the shift on a Hilbert space of dimension 9) and $T=S^3+S^7$ (a polynomial in S!), we have $\|S\|=1$, $w(T)=\cos(\pi/10)$, and $w(TS)=1$, so that we need only compute $1/\cos(\pi/10)=1.0514\ldots$.*

Proof. It is easy to check that both T and TS have matrices to which we may apply Proposition 1. In $G(TS)$ we find a cycle (involving vertices 1, 5, and 9); hence $w(TS)=1$. On the other hand, $G(T)$ consists of a single chain involving all the vertices; that is, $L=9$. □

Many variations on this theme are possible. The following is somewhat more closely related to Müller's construction.

COROLLARY 4. *Consider the 9-dimensional operators defined via 3×3 blocks as follows:*

$$S=\begin{bmatrix} S_3 & 0 & 0 \\ 0 & S_3 & 0 \\ 0 & 0 & S_3 \end{bmatrix} \quad \textit{and} \quad T=\begin{bmatrix} 0 & I_3 & S_3 \\ 0 & 0 & I_3 \\ 0 & 0 & 0 \end{bmatrix},$$

where I_3 is the 3×3 identity matrix. Then S and T commute, $\|S\|=1$, $w(T)=\cos(\pi/10)$, and $w(TS)=1$ $(>1.05w(T)\|S\|)$.

Proof. Clearly $\|S_3\| = \|S\| = 1$, and it is easy to check that $TS = ST$ and that $G(TS)$ has a cycle involving 1, 5, and 9. Proposition 1 applies to show that $w(TS) = 1$. Moreover, the T of this corollary is the same as that of Corollary 3, so that again $w(T) = \cos(\pi/10)$. □

One can tinker with incidence matrices to examine many such pairs T and S, hoping to obtain better lower bounds on C. It appears that the following corollary gives the best estimate among those based on operators that can be analyzed using Proposition 1.

COROLLARY 5. *Let* $S = S_{16}$, *the* 16-*dimensional shift operator, and let* $T = S^4 + S^{14}$. *Then* $\|S\| = 1$, $w(T) = \cos(\pi/9)$, *and* $w(TS) = 1$. *Hence the constant* C (*as defined in* §1) *is no less than* $1/\cos(\pi/9) = 1.064\ldots$.

Proof. We may apply Proposition 1 to T and TS. The graph $G(TS)$ includes a cycle (involving vertices 1, 6, 11, and 16) so that $w(TS) = 1$. Examining $G(T)$, we find that it falls into two chains of length 8; that is, $L = 8$ for $G(T)$. □

REMARK. There is no reason to expect that zero-one matrices give the best lower bounds for C. In fact, by replacing T in Corollary 5 by $S^4 + aS^{14}$, where a is an appropriate positive constant, we can obtain slightly larger ratios. Numerical studies suggest that $a = 1.22$ is close to the best choice for this purpose and that the corresponding estimate is roughly $C > 1.066$.

3. Implications for dilation theory. The problem of evaluating C was initially suggested by dilation theory. Such estimates as $C < 0.5(2 + 2\sqrt{3})^{1/2}$ (< 1.169), due to Okubo and Ando [7] (who attribute part of their argument to M. J. Crabb), and various results about special classes of commuting operators came in response to this stimulus. We shall briefly review the relevant parts of dilation theory below; for more detail the reader is referred to Sz.-Nagy–Foiaş [8, Chap. 1].

Recall that an operator S on Hilbert space H is a contraction (i.e., $\|S\| \le 1$) exactly when S has a Nagy dilation—that is, a unitary operator U on some Hilbert space containing H such that

$$S^n = P_H U^n|_H \quad (n \ge 1)$$

(here P_H is the orthogonal projection from the larger "dilation space" onto H). On the other hand, an operator T on H is a "numerical contraction" (i.e., $w(T) \le 1$) exactly when T has a Berger dilation—that is, a unitary operator V on some larger space such that

$$T^n = 2P_H V^n|_H \quad (n \ge 1).$$

It is natural to ask, for commuting T and S, whether these representations may be achieved "simultaneously"; that is, whether there exist unitary U and V commuting on a larger space such that, for all integers $n, m \ge 0$,

$$S^n T^m = \begin{cases} P_H U^n|_H & \text{if } m = 0, \\ 2P_H U^n V^m|_H & \text{if } m > 0. \end{cases} \tag{1}$$

This would mean in particular that ST has a Berger dilation UV, so that $w(ST) \leq 1$. In Holbrook [3] it was noted that such a simultaneous dilation is possible when S and T double commute (i.e., $ST = TS$ and $S^*T = TS^*$). Ando's dilation theorem for two commuting contractions (see [2]) does not require double commutativity and raises the possibility of a similar simultaneous dilation theorem for a numerical contraction T simply commuting with a contraction S. We have seen, however, that this would imply that $w(ST) \leq 1$ whenever $\|S\| \leq 1$, $w(T) \leq 1$, and $ST = TS$. By homogeneity the constant C of §1 would be 1; thus the examples of §2 (or Müller's example) yield the following.

PROPOSITION 2. *There exists a commuting pair of (finite-dimensional) operators S, T such that S has a Nagy dilation and T has a Berger dilation but the pair has no simultaneous dilation as in* (1).

Sz.-Nagy and Foiaş introduced the classes C_ρ of operators T having ρ-dilations: unitary V on a space containing H such that

$$T^n = \rho P_H V^n|_H \quad (n \geq 1).$$

The earlier history of these classes is presented in [8, §I.11]. In [4] (see also Williams [9]), Holbrook develops the properties of the corresponding (homogeneous) operator radii defined by

$$w_\rho(T) \leq 1 \quad \text{iff } T \in C_\rho.$$

We shall say that a contraction S on H and an operator $T \in C_\rho$ commuting with S have a simultaneous $(1, \rho)$-dilation if there are commuting unitaries U, V on a larger Hilbert space such that

$$S^n T^m = \begin{cases} P_H U^n|_H & \text{if } m = 0, \\ \rho P_H U^n V^m|_H & \text{if } m > 0. \end{cases} \tag{2}$$

Proposition 2 says that the Ando dilation theorem ((1, 1)-dilations) does not extend to the (1, 2) case. The following proposition says that failure occurs for $(1, \rho)$-dilations with any $\rho > 1$.

PROPOSITION 3. *For any $\rho > 1$, there are commuting operators S and T such that $w_\rho(ST) > w_\rho(T)\|S\|$. Consequently there are commuting S' and T' such that S' is a contraction and T' has a ρ-dilation but the pair has no simultaneous* $(1, \rho)$-*dilation (as in* (2)). *In particular, if $S = S_9$ and $T = S^3 + (\rho - 1)S^7$, we have $w_\rho(T) < 1$ while $w_\rho(ST) = 1$.*

Proof. In general (see [8, §I.11]), $w_\rho(T) < 1$ if and only if

$$\rho I + \sum_1^\infty \{(e^{i\theta}T)^n + (e^{i\theta}T)^{*n}\}$$

is positive definite for every θ. Writing our particular T in the form

$$T = \begin{bmatrix} 0 & I & (\rho-1)J \\ 0 & 0 & I \\ 0 & 0 & 0 \end{bmatrix},$$

where I is the 3×3 identity matrix and $J=S_3$, we must verify that

$$\begin{bmatrix} \rho I & e^{i\theta}I & (\rho-1)e^{i\theta}J+e^{2i\theta}I \\ e^{-i\theta}I & \rho I & e^{i\theta}I \\ (\rho-1)e^{-i\theta}J^*+e^{-2i\theta}I & e^{-i\theta}I & \rho I \end{bmatrix}$$

is positive definite. This matrix is unitarily equivalent (via $\mathrm{diag}\{1, e^{i\theta}, e^{2i\theta}\}$) to

$$\begin{bmatrix} \rho I & I & I+(\rho-1)e^{-i\theta}J \\ I & \rho I & I \\ I+(\rho-1)e^{i\theta}J^* & I & \rho I \end{bmatrix}$$

so that we need to show that $[I]+(\rho-1)(I_9-Q)>0$, where

$$[I]=\begin{bmatrix} I & I & I \\ I & I & I \\ I & I & I \end{bmatrix} \quad \text{and} \quad Q=\begin{bmatrix} 0 & 0 & -e^{-i\theta}J \\ 0 & 0 & 0 \\ -e^{i\theta}J^* & 0 & 0 \end{bmatrix}.$$

Now $([I]h, h)=\|x+y+z\|^2$ (≥ 0), where

$$h=\begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

Also, Q is a (self-adjoint) contraction, so $(I_9-Q)\geq 0$. This makes it clear that $[I]+(\rho-1)(I_9-Q)\geq 0$. To show that the inequality is strict we must examine those h for which $([I]h, h)=0$ and $((I_9-Q)h, h)=0$. This requires $\|x+y+z\|=0$ and

$$\|x\|^2+\|y\|^2+\|z\|^2=-2\,\mathrm{Re}(e^{i\theta}(x, Jz));$$

because this last quantity is no greater than $\|x\|^2+\|z\|^2$ we must have $y=0$ and $\|x+z\|=0$, so that $-2\,\mathrm{Re}(e^{i\theta}(x, Jz))=2\,\mathrm{Re}(e^{i\theta}(z, Jz))$. This can only be $2\|z\|^2$ if $Jz=e^{i\theta}z$, which forces $z=0$ and hence $h=0$.

To show that $w_\rho(ST)=1$ we may first note that $ST=S^4+(\rho-1)S^8$ and that S_9^4 is unitarily equivalent to $S_3\oplus S_2\oplus S_2\oplus S_2$. Thus ST is unitarily equivalent to $(J+(\rho-1)J^2)\oplus S_2\oplus S_2\oplus S_2$, and we need only check that $w_\rho(J+(\rho-1)J^2)=1$. Using the general criterion described at the beginning of this proof, we see that we must check that the 3×3 matrix

$$\begin{bmatrix} \rho & e^{i\theta} & (\rho-1)e^{i\theta}+e^{2i\theta} \\ e^{-i\theta} & \rho & e^{i\theta} \\ (\rho-1)e^{-i\theta}+e^{-2i\theta} & e^{-i\theta} & \rho \end{bmatrix}$$

is nonnegative for every θ. Since $\rho>1$, the first two principal minors are certainly positive, and it is a matter of computing the 3×3 determinant. This turns out to be $2(\rho-1)^2(1-\mathrm{Re}(e^{i\theta}))$, a nonnegative quantity that vanishes at $\theta=0$. Thus $w_\rho(J+(\rho-1)J^2)=1$. $\square$

REFERENCES

1. L. Allen, *Analytic inequalities for finite-dimensional operators,* M.Sc. thesis, University of Guelph, 1973.
2. T. Ando, *On a pair of commutative contractions,* Acta Sci. Math. (Szeged) 24 (1963), 88–90.
3. J. A. R. Holbrook, *Multiplicative properties of the numerical radius in operator theory,* J. Reine Angew. Math. 237 (1969), 166–174.
4. ———, *On the power-bounded operators of Sz.-Nagy and Foiaş,* Acta Sci. Math. (Szeged) 29 (1968), 299–310.
5. ———, *Polynomials in a matrix and its commutant,* Linear Algebra Appl. 48 (1982), 426–432.
6. V. Müller, *The numerical radius of a commuting product,* Michigan Math. J. 35 (1988), 255–260.
7. K. Okubo and T. Ando, *Operator radii of commuting products,* Proc. Amer. Math. Soc. 56 (1976), 203–210.
8. B. Sz.-Nagy and C. Foiaş, *Harmonic analysis of operators on Hilbert space,* North-Holland, Amsterdam, 1970.
9. J. P. Williams, *Schwarz norms for operators,* Pacific J. Math. 24 (1968), 181–188.

Department of Pure Mathematics
University of Waterloo
Waterloo, Ontario
Canada N2L 3G1

Department of Mathematics and Statistics
University of Guelph
Guelph, Ontario
Canada N1G 2W1