Unitary Invariance and Spectral Variation

Rajendra Bhatia
Indian Statistical Institute
New Delhi-110016, India

and

John A. R. Holbrook
Department of Mathematics
University of Guelph
Guelph, Ontario N1G 2W1, Canada

Dedicated to Chandler Davis on the occasion of his sixtieth birthday.

Submitted by Richard A. Brualdi

ABSTRACT

We call a norm on operators or matrices weakly unitarily invariant if its value at operator A is not changed by replacing A by U^*AU, provided only that U is unitary. This class includes such norms as the numerical radius. We extend to all such norms an inequality that bounds the spectral variation when a normal operator A is replaced by another normal B in terms of the arclength of any normal path from A to B, computed using the norm in question. Related results treat the local metric geometry of the "manifold" of normal operators. We introduce a representation for weakly unitarily invariant matrix norms in terms of function norms over the unit ball, and identify this correspondence explicitly in certain cases.

---

1. INTRODUCTION

In the first part of this paper we study a class of norms on complex matrices. A norm τ from this class is characterized by the invariance property

$$\tau(A) = \tau(UAU^*) \tag{1.1}$$

*Research supported in part by NSERC of Canada under Operating Grant A8745.

52 Vanderbilt Ave., New York, NY 10017 0024-3795/87/$3.50

for all matrices A and all unitary matrices U. We call such a norm weakly unitarily invariant (wui).

If the norm satisfies the stronger requirement

$$\tau(A) = \tau(UAV) \tag{1.2}$$

for all matrices A and all unitary matrices U and V, we shall say in this paper that τ is strongly unitarily invariant (sui). Such norms have most often been called "unitarily invariant" in the literature. We need the adverbs to distinguish these two classes here. The class of sui norms is properly contained in the class of wui norms.

In Section 2 we consider several examples of such norms (some "classical," some more exotic), various ways of generating them, and a few of their properties. We prove a theorem which characterizes wui norms in terms of certain function norms.

In the latter part of the paper we continue our study of the spectral variation of matrices, building upon our earlier papers [1], [5]. The main result can be described as follows. In [1] it was shown that if $N(t)$, $0 \leqslant t \leqslant 1$, is a piecewise smooth path in the space of normal matrices, then for any sui norm τ, the τ-distance between the eigenvalues of $N(0)$ and $N(1)$ is bounded above by the τ-length of the curve $N(t)$. (See Section 3 for precise definitions.) Several known spectral variation results were seen to follow from this theorem. We show in Section 5 that this "path inequality" can be extended to the class of all wui norms.

In [5], along with our study of the "short normal path" geometry of the unitary matrices, we derived the path inequality, for the operator norm only, using a different technique that removed the smoothness restriction and some other technical conditions imposed on the path in [1]. In Section 5 we shall carry out a similar analysis for the entire class of wui norms. In this approach some differential geometry of Section 3 is replaced by analysis. This not only allows us to work without the technical restrictions mentioned above; it also brings out some interesting local metric properties of the set of normal matrices.

In Section 4 we point out connections between our work and some results of Halmos and Bouldin on approximating a normal operator by another with restricted spectrum.

2. WEAKLY UNITARILY INVARIANT (WUI) NORMS

Let us first fix some notation. We denote by $\mathbf{B}(\mathbf{C}^n)$ the space of linear operators on the vector space $\mathbf{C}^n$. Unless otherwise specified, the space $\mathbf{C}^n$ will be assumed to be equipped with the usual Euclidean inner product and norm, i.e., $\mathbf{C}^n$ is the n-dimensional l_2 space. Occasionally we shall need to think of it as an l_p space. An operator A will be identified with its matrix with respect to the standard Cartesian basis for $\mathbf{C}^n$. The space of n by n matrices will be denoted by $\mathrm{M}(n)$, and the group of unitary matrices by $\mathrm{U}(n)$. We shall drop the parenthetical n when there is no danger of confusion.

We shall consider several different norms on $\mathrm{M}(n)$. The following notation will be used for them:

$\|A\|$ will always denote the operator bound norm of the operator A on the Hilbert space l_2.

The norm symbol with a subscript will denote some other norm. Thus for example the Frobenius norm is defined as

$$\|A\|_F = (\operatorname{tr} A^*A)^{1/2} = \Big(\sum_{i,j} |a_{ij}|^2\Big)^{1/2}, \tag{2.1}$$

where A is the matrix with entries a_{ij}.

The symbol m will denote any arbitrary norm on $\mathrm{M}(n)$. The symbol μ will be reserved for strongly unitarily invariant (sui) norms [see the relation (1.2)], while τ will denote any weakly unitarily invariant (wui) norm [i.e., one satisfying (1.1)]. Note that the norms $\|A\|$ and $\|A\|_F$ are both sui.

A detailed study of sui norms was made by von Neumann [17] and Schatten [16]. These norms have been used frequently in the study of theoretical and computational problems. See, in particular, the books by Gohberg and Krein [10], Hewitt and Ross [13], and Marshall and Olkin [15].

Numerical analysts have often used certain matrix norms which are easy to compute but which are not unitarily invariant. For example, we can define the maximal row sum norm, the maximal column sum norm, the maximal entry norm, and the total sum norm of a matrix A as, respectively,

$$\|A\|_{\text{row}} = \max_i \sum_j |a_{ij}|, \tag{2.2}$$

$$\|A\|_{\text{col}} = \max_j \sum_i |a_{ij}|, \tag{2.3}$$

$$\|A\|_{\max} = \max_{i,j} |a_{ij}|, \tag{2.4}$$

$$\|A\|_{\text{tot}} = \sum_{i,j} |a_{ij}|. \tag{2.5}$$

Each of these norms is easy to compute from the entries of the matrix; none of them is a wui norm. The following facts are easy to verify:

(i) $\|A\|_{\text{row}}$ is the norm of A as an operator from l_∞ to l_∞,
(ii) $\|A\|_{\text{col}}$ is the norm of A as an operator from l_1 to l_1,
(iii) $\|A\|_{\max}$ is the norm of A as an operator from l_1 to l_∞,
(iv) $\|A\|_{\text{tot}}$ is the norm of A as an operator from l_∞ to l_1.

It is only recently that the family of wui norms has been studied in some detail. A restricted class of such norms has been analysed by Fong and Holbrook [8]; some special examples of such norms have been studied in detail by Fong, Radjavi, and Rosenthal [9]. In the following subsections we look at wui norms from a somewhat different perspective.

2A. *Generation of WUI Norms*

We give below some methods of constructing wui norms which are not of the better-known sui class.

(a) The numerical radius of an operator A on a Hilbert space H is defined as

$$w(A) = \sup\{|\langle Ax, x\rangle| : x \in \mathbf{H},\ \|x\| = 1\}. \tag{2.6}$$

This defines a norm on the space $\mathbf{M}(n) = \mathbf{B}(\mathbf{C}^n)$ that is wui but not sui. There is a whole family of such norms associated with "ρ-dilations" in a Hilbert space. These give rise to norms $w_\rho(A)$ each of which is wui. A detailed study of such norms has been made in [8]. We should mention here that each of these norms $w_\rho(A)$ lies between $w(A)$ and $\|A\|$. In particular, if A is a Hermitian operator, then all these norms coincide with $\|A\|$.

(b) The pointwise sum or maximum of any wui seminorm and wui norm is again a wui norm. Thus, for example, the norm

$$\tau(A) = \|A\| + |\operatorname{tr} A| \tag{2.7}$$

is a simple example of a norm that is wui but not sui. We shall give a more comprehensive description of such norms later.

(c) Let $m(A)$ be any norm on M. Define two induced norms as follows:

$$m_{\mathbf{U}}(A) = \max_{U \in \mathbf{U}} m(UAU^*), \tag{2.8}$$

$$m'_{\mathbf{U}}(A) = \int_{\mathbf{U}} m(UAU^*)\, dU, \tag{2.9}$$

where the integration in (2.9) is with respect to any invariant probability measure on U. If the norm $m(A)$ is wui, then $m_{\mathbf{U}}(A) = m'_{\mathbf{U}}(A) = m(A)$. However, interesting new wui norms are obtained by this procedure if $m(A)$ is not wui. Thus, each of the norms (2.2)–(2.5) induces a pair of wui norms.

Let us look at some examples of norms generated by the procedures above. Given an operator A, define

$$q(A) = \sup\{|\langle Ax, y\rangle| : \|x\| = \|y\| = 1, \langle x, y\rangle = 0\}. \tag{2.10}$$

Then $q(A)$ is a wui seminorm that is not sui. This can be combined with any wui norm according to procedure (b) above. A specially interesting example is the norm

$$\|A\|_{\max,\mathbf{U}} = \max(w(A), q(A)), \tag{2.11}$$

which is also the norm generated by $\|A\|_{\max}$ [defined in (2.4)] via the procedure (2.8). Note that

$$\|A\|_{\max,\mathbf{U}} = \max_{\{e_1,\dots,e_n\}} \max_{i,j} |\langle Ae_i, e_j\rangle|,$$

where $\{e_1,\dots,e_n\}$ varies over all choices of orthonormal bases in $\mathbf{C}^n$. This is one of the norms studied in some detail in [9].

We should point out that sui norms can be generated by a procedure analogous to (c) above, starting from wui or arbitrary norms. Specifically, if $\tau(A)$ is a wui norm, we may simply use the definitions

$$\tau'(A) = \max_{U \in \mathbf{U}} \tau(UA), \tag{2.12}$$

$$\tau''(A) = \int_{\mathbf{U}} \tau(UA)\, dU \tag{2.13}$$

to obtain sui norms.

For example, the numerical radius $w(A)$ is a wui norm that leads by the procedure (2.12) to the operator norm

$$\begin{aligned} w'(A) &= \max_{U \in \mathbf{U}} \max_{x : \|x\| = 1} |\langle UAx, x\rangle| \\ &= \max_{x, y : \|x\| = \|y\| = 1} |\langle Ax, y\rangle| \\ &= \|A\|. \end{aligned}$$

In the same spirit, consider the wui norm

$$\tau_k(A) = \max_{(x_1,\dots,x_k)} \sum_{j=1}^{k} |\langle Ax_j, x_j\rangle|, \tag{2.14}$$

where the maximum is taken over all choices of orthonormal k-tuples $\{x_2,\dots,x_k\}$ in $\mathbf{C}^n$. Then

$$\tau_k'(A) = \max_{U\in\mathbf{U}} \max_{(x_1,\dots,x_k)} \sum_{j=1}^{k} |\langle UAx_j, x_j\rangle|$$

$$= \sum_{j=1}^{k} s_j(A) = \|A\|_k,$$

the kth Ky Fan norm of A [here the $s_j(A)$ are the singular values of A, written in decreasing order].

This procedure sometimes produces a (sui) norm from a (wui) seminorm. For example, the wui seminorm $|\mathrm{tr}\, A|$ leads to the trace class norm:

$$\max_{U\in\mathbf{U}} |\mathrm{tr}\, UA| = \max_{u\in\mathbf{U}} \left| \sum_{j=1}^{n} \langle UAx_j, x_j\rangle \right|$$

$$= \sum_{j=1}^{n} s_j(A) = \|A\|_{\mathrm{tr}}.$$

(The "maximal principles" used in making the above assertions may be found in Gohberg and Krein [10].)

Finally let us note that we could go from an arbitrary matrix norm m to a wui norm via (2.8) and then to a sui norm via (2.12). This would be the same norm as the one defined by

$$\tilde{m}(A) = \max_{U,V\in\mathbf{U}} m(UAV). \tag{2.15}$$

Similarly, applying the procedures (2.9) and (2.13) successively to m would lead to the same norm as is defined by

$$\hat{m}(A) = \int\int m(UAV)\,dU\,dV. \qquad (2.16)$$

2B. *A Characterization of WUI Norms*

We shall now identify a source which generates all wui norms.

Fix n, and let S be the unit sphere in $\mathbf{C}^n$, consisting of all unit vectors u. Let $C(S)$ be the linear space of all complex continuous functions on S. The group U acts naturally on S and hence on $C(S)$. We say that a norm Φ on $C(S)$ is a unitarily invariant function norm if it is invariant under this action. That is,

$$\Phi(f \circ U) = \Phi(f) \qquad (2.17)$$

for all $f \in C(S)$ and $U \in \mathbf{U}$.

Let du indicate the normalized Lebesgue measure on S. Since du is invariant under rotations, the familiar p-norms

$$\Phi_p(f) = \|f\|_p = \left(\int |f(u)|^p\,du\right)^{1/p} \qquad (1 \leqslant p < \infty), \qquad (2.18)$$

$$\Phi_\infty(f) = \|f\|_\infty = \max_{u \in S} |f(u)| \qquad (2.19)$$

provide natural examples of unitarily invariant function norms.

Given an operator A on $\mathbf{C}^n$, define the function f_A in $C(S)$ by

$$f_A(u) = \langle Au, u \rangle, \qquad (2.20)$$

and for Φ as above let Φ' be defined by

$$\Phi'(A) = \Phi(f_A). \qquad (2.21)$$

The map taking A to f_A is linear with trivial kernel, so Φ' is certainly a norm on $\mathbf{B}(\mathbf{C}^n)$. Furthermore, the unitary invariance of Φ implies that Φ' is wui:

$$\Phi'(U^*AU) = \Phi(f_{U^*AU}) = \Phi(f_A \circ U) = \Phi(f_A) = \Phi'(A).$$

We shall show that every wui norm on $\mathbf{M}$ arises in this way from some unitarily invariant function norm on $C(S)$. In some cases we are able to identify a correspondence between "natural" Φ and "natural" matrix norms, but interesting problems remain in this area.

Let $F = \{f_A : A \in \mathbf{M}\}$; this is a finite-dimensional subspace of $C(S)$. Given a wui norm τ, define Φ_0 on F by

$$\Phi_0(f_A) = \tau(A). \tag{2.22}$$

Since the map from A to f_A is linear and invertible (on F), Φ_0 is a norm on F, and it is clear that the wui property of τ means that Φ_0 is unitarily invariant in the sense that

$$\Phi_0(f \circ U) = \Phi_0(f) \qquad \text{for all} \quad f \in F.$$

If we can extend Φ_0 to a unitarily invariant function norm Φ defined on all of $C(S)$, we shall have $\tau = \Phi'$. Such an extension may be obtained by Hahn-Banach argument; here are the details.

Consider $C(S)$ as a Banach space with the uniform norm $\|f\|_\infty$. The finite-dimensional subspace F now has two norms Φ_0 and $\|\cdot\|_\infty$, and there exist constants $0 < \alpha \leqslant \beta < \infty$ such that $\alpha\|f\|_\infty \leqslant \Phi_0(f) \leqslant \beta\|f\|_\infty$ for all $f \in F$. Let G be the set of all linear functionals g on F such that $|g(f)| \leqslant \Phi_0(f)$ for all $f \in F$. Then, for every $f \in F$, we have $\Phi_0(f) = \sup\{|g(f)|: g \in G\}$. Note that $|g(f)| \leqslant \beta\|f\|_\infty (g \in G,\ f \in F)$. By the Hahn-Banach theorem each $g \in G$ has a linear extension $\tilde{g}$ to $C(S)$ such that $|\tilde{g}(f)| \leqslant \beta\|f\|_\infty$ $[f \in C(S)]$. Given $f \in C(S)$, let $\theta(f) = \sup\{|\tilde{g}(f)|: g \in G\}$. This defines a seminorm on $C(S)$ that coincides with Φ_0 on F. Replace $\theta(f)$ with $\max(\theta(f), \alpha\|f\|_\infty)$, if necessary, to obtain a norm with the same property. Finally define Φ on $C(S)$ by

$$\Phi(f) = \sup\{\theta(f \circ U) : U \in \mathbf{U}\}.$$

By construction Φ is a unitarily invariant function norm extending Φ_0 from F to $C(S)$. It is not clear when such an extension is unique. We have proved the following.

Theorem 2.1. *A norm τ on $\mathbf{M}$ is wui exactly when there is some unitarily invariant function norm Φ on $C(S)$ with $\tau = \Phi'$ (as defined by the relations* (2.20) *and* (2.21)*).*

REMARK. Since the value of f_A is constant on the classes $[u] = \{e^{i\theta}u: \theta \in \mathbf{R}\}$, one might replace S in the above discussion by the complex projective space $\mathbf{C}P^{n-1}$ to avoid a certain redundancy. At present, we do not see any advantage to this representation.

It is now natural to ask: what are the wui matrix norms Φ_p' corresponding to the unitarily invariant function norms defined by (2.18) and (2.19)? It is clear that

$$\Phi_\infty'(A) = w(A), \tag{2.23}$$

the numerical radius of A. We shall also identify Φ_2'. Notice that this is an inner product norm on $\mathbf{M}$, arising from the inner product

$$\langle X, Y\rangle_2 = \int_S f_X(u)(f_Y(u))^* \, du \tag{2.24}$$

(here we use the notation z^* for the complex conjugate of a complex number z). It is perhaps not surprising that the norm turns out to be related to the Frobenius norm.

We first identify all wui sesquilinear forms on $\mathbf{M}$. Let $\langle X, Y\rangle$ be such a form. By the Riesz representation theorem there exists a linear operator Γ on $\mathbf{M}$ such that

$$\langle X, Y\rangle = \langle X, \Gamma(Y)\rangle_F,$$

where $\langle X, Y\rangle_F$ stands for the Frobenius inner product $\operatorname{tr} Y^*X$. Since our form is wui, we have for each unitary U

$$\begin{aligned}\langle X, \Gamma(Y)\rangle_F &= \langle X, Y\rangle = \langle U^*XU, U^*YU\rangle \\ &= \langle U^*XU, \Gamma(U^*YU)\rangle_F = \langle X, U\Gamma(U^*YU)U^*\rangle_F.\end{aligned}$$

It follows that the linear map Γ satisfies

$$\Gamma(U^*YU) = U^*\Gamma(Y)U \tag{2.25}$$

for all $Y \in \mathbf{M}$ and $U \in \mathbf{U}$.

Our next proposition characterizes all such Γ. The proof presented here is a close relative of an argument suggested by Ken Davidson.

PROPOSITION 2.2. *Let Γ be a linear map from* M *to itself that satisfies the relation* (2.25). *Then Γ must have the form*

$$\Gamma(Y) = aY + b(\operatorname{tr} Y)I$$

for some scalars a and b.

Proof. Each contraction on $\mathbf{C}^n$ is a convex combination of unitary operators (see [11, Problem 136], for example). Using this fact and the spectral theorem, we have one way to see that every operator Y is a linear combination of rank one projections.

Let $\{e_1, \dots, e_n\}$ be the standard basis in $\mathbf{C}^n$, and let E_{11} be the projection onto the span of e_1. Let U be any unitary of the form $1 \oplus U_1$, leaving e_1 fixed. Such a U commutes with E_{11}, so that

$$\Gamma(E_{11}) = \Gamma(U^* E_{11} U) = U^* \Gamma(E_{11}) U,$$

i.e., U commutes with $\Gamma(E_{11})$ as well. Hence, using for example the fact cited at the beginning of this proof, $\Gamma(E_{11})$ commutes with every operator of the form $1 \oplus R$, where R is any operator on the space span$(e_2, \dots, e_n)$. It follows that $\Gamma(E_{11})$ must be of the form $t \oplus sI_1$, where $t, s \in C$ and I_1 is the identity operator on span$(e_2, \dots, e_n)$. Note that $\operatorname{tr} E_{11} = 1$, so that

$$\Gamma(E_{11}) = aE_{11} + b(\operatorname{tr} E_{11})I$$

with $b = s$ and $a = t - s$. Now if E is any other rank one projection, there exists a unitary U such that $E = U^* E_{11} U$. Hence

$$\Gamma(E) = U^* \Gamma(E_{11}) U = U^* [aE_{11} + b(\operatorname{tr} E_{11})I] U$$
$$= aE + b(\operatorname{tr} E)I.$$

Thus the same constants a and b work for each E, and by the first paragraph of this proof they work for any operator Y.

COROLLARY 2.3. *Any wui sesquilinear form $\langle X, Y \rangle$ on $\mathrm{M}(n)$ is a linear combination of $\langle X, Y \rangle_F$ and $\operatorname{tr} X (\operatorname{tr} Y)^*$. If $\langle X, Y \rangle$ is an inner product, then*

$$\langle X, Y \rangle = a \langle X, Y \rangle_F + b \operatorname{tr} X (\operatorname{tr} Y)^*$$

where a, b are real constants, $a \geq 0$ and $b \geq -a/n$.

Proof. By the proposition and the discussion preceding it there exist complex constants a, b such that

$$\langle X,Y\rangle = \langle X, aY + b(\operatorname{tr} Y)I\rangle_F$$
$$= a^*\langle X,Y\rangle_F + b^* \operatorname{tr} X\,(\operatorname{tr} Y)^*.$$

If $\langle X,Y\rangle$ is an inner product, a and b must be real with $a \geqslant 0$. Also note that

$$0 \leqslant \langle X,X\rangle = a\|X\|_F^2 + b|\operatorname{tr} X|^2$$

and

$$|\operatorname{tr} X|^2 = |\langle X, I\rangle_F|^2 \leqslant n\|X\|_F^2.$$

Hence $b \geqslant -a/n$.

With the help of this corollary we identify Φ_2'.

THEOREM 2.4. *Let Φ_2' be the wui norm on* M(n) *associated with the natural L^2-norm on $C(S)$. Then*

$$\Phi_2'(A) = \left(\frac{\|A\|_F^2 + |\operatorname{tr} A|^2}{n+n^2}\right)^{1/2}.$$

Proof. For the inner product of (2.24) we must have

$$\langle X,Y\rangle_2 = a\langle X,Y\rangle_F + b\operatorname{tr} X\,(\operatorname{tr} Y)^* \tag{2.26}$$

for some real a and b. Since the measure in (2.24) is normalized, $\langle I, I\rangle_2 = 1$, so that

$$an + bn^2 = 1. \tag{2.27}$$

Now consider the operators E_{ij} defined by $E_{ij}e_j = e_i$ and $E_{ij}e_k = 0$ $(k \neq j)$ (i.e., E_{ij} is the "matrix unit" whose sole nonzero entry is a 1 in the ijth place). Note that if $A = E_{ij}$ and u is any unit vector, then

$$f_A(u) = \langle\langle u, e_j\rangle e_i, u\rangle = \langle u, e_j\rangle(\langle u, e_i\rangle)^*.$$

In this way one sees that

$$\langle E_{ii}, E_{jj}\rangle_2 = \langle E_{ij}, E_{ij}\rangle_2 \tag{2.28}$$

for all i, j. Choose $i \neq j$, and use (2.26), (2.28) to see that $a = b$. Together with (2.27) this implies that $a = b = 1/(n + n^2)$.

We also identify the trace in terms of the functional representation.

PROPOSITION 2.5. *For any A in* $\mathrm{M}(n)$, $\operatorname{tr} A = n\int_S f_A(u)\,du$.

Proof. By linearity, it is enough to verify the relation for each of the matrix units E_{ij} (see the proof above). First let us check that $\int f_A = 0$ when $A = E_{ij}$ with $i \neq j$. Referring again to the last proof, we see that the integral is $\int\langle u, e_j\rangle\langle e_i, u\rangle\,du$. But there is a unitary U such that $Ue_j = e_j$ and $Ue_i = -e_i$, so that the invariance of du under U implies that

$$\int\langle u, e_j\rangle\langle e_i, u\rangle\,du = -\int\langle u, e_j\rangle\langle e_i, u\rangle\,du = 0.$$

Since $\operatorname{tr} E_{ii} = 1$, it only remains to check that $\int|\langle u, e_i\rangle|^2\,du = 1/n$. There are unitary U taking e_i to e_j; hence all the $\int|\langle u, e_i\rangle|^2\,du$ are equal. Since their sum is $\int 1\,du$, i.e. 1, each must indeed be $1/n$.

The proposition allows us to reinterpret Theorem 2.4 in various ways. For example, we have the following "probabilistic" interpretation of the Frobenius norm.

PROPOSITION 2.6. *For any A in* $\mathrm{M}(n)$,

$$\|A\|_F^2 = nE\left(|f_A|^2\right) + n^2\,\mathrm{Var}(f_A),$$

where the expectation E *and variance* Var *are computed with respect to the unitarily invariant probability measure du.*

Proof. Compute, using Theorem 2.4, Proposition 2.5, and the standard relation

$$\mathrm{Var}(f) = \int\left|f - \int f\right|^2 = \int|f|^2 - \left|\int f\right|^2.$$

2C. The Pinching Inequality

Let $P_1, P_2, \dots, P_r$ be a complete family of mutually orthogonal projections in $\mathbf{C}^n$. These projections induce a pinching operator π on $\mathbf{B}(\mathbf{C}^n)$ defined by

$$\pi(A) = \sum_i P_i A P_i. \tag{2.29}$$

Properties of this operator have been studied by Davis [7] and Gohberg and Krein [10], who call it the diagonal-cell operator. In an appropriate basis for $\mathbf{C}^n$, the effect of a pinching is to replace A by a block diagonal matrix $\pi(A)$ consisting of r diagonal blocks whose sizes are the ranks of the projections P_i. This matrix is obtained from A by replacing the entries outside those diagonal blocks by zeros. A pinching induced by a family of r projections as in (2.29) will be called an r-pinching.

For $j = 1, 2, \dots, r-1$, put $Q_j = P_1 + \cdots + P_j$, and define a 2-pinching π_j by

$$\pi_j(A) = Q_j A Q_j + (I - Q_j) A (I - Q_j).$$

It is easy to see that the pinching (2.29) can be expressed as

$$\pi(A) = \pi_{r-1} \cdots \pi_2 \pi_1(A). \tag{2.30}$$

Thus an r-pinching can be obtained by successively applying 2-pinchings. One more diagonal cell is pinched off at each stage.

THEOREM 2.7. *Every wui norm is diminished by a pinching. That is, if* τ *is a wui norm and* π *a pinching on* $\mathbf{B}(\mathbf{C}^n)$, *then*

$$\tau(\pi(A)) \leq \tau(A) \qquad \text{for all} \quad A \in \mathbf{B}(\mathbf{C}^n). \tag{2.31}$$

Proof. Because of the decomposition (2.30) it suffices to prove (2.31) when π is a 2-pinching. In this case we can write, in an appropriate basis, the block forms

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \qquad \pi(A) = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}.$$

Let U be the unitary operator

$$U = \begin{bmatrix} I_1 & 0 \\ 0 & -I_2 \end{bmatrix}.$$

where I_1 and I_2 are the identity operators of the appropriate sizes. It is easily seen that $\pi(A) = \frac{1}{2}(A + UAU^*)$. Since τ is a wui norm, the inequality (2.31) follows directly from this representation.

Gohberg and Krein [10, p. 82] prove the "pinching inequality" (2.31) for the class of sui norms. Their proof rests on certain properties of those norms that are not available in our more general setting.

3. THE PATH INEQUALITY FOR SPECTRAL VARIATION

In this section we shall show that some inequalities for the distance between eigenvalues of normal matrices that were established in [1] for sui norms can be extended to the class of wui norms. At the same time we shall rectify a false step in the proof of the main theorem in [1].

For a matrix A, let Eig A denote the unordered n-tuple consisting of the eigenvalues of A, and let $D(A)$ be a diagonal matrix whose diagonal entries are the elements of Eig A. Given a wui norm τ, define the τ-spectral distance between A and B by

$$\tau(\operatorname{Eig} A, \operatorname{Eig} B) = \min \tau(D(A) - WD(B)W^*), \tag{3.1}$$

where the minimum is taken over all permutation matrices W. Because τ is wui, this definition does not depend on the choice of $D(A)$ and defines a pseudometric on the space $\mathbf{M}$.

If A is Hermitian, let $D\downarrow(A)$ denote the diagonal matrix whose diagonal entries are the eigenvalues of A arranged in decreasing order. It follows from a famous theorem of Lidskii [14] that, if A, B are any two Hermitian matrices and μ is any sui norm, then

$$\mu(\operatorname{Eig} A, \operatorname{Eig} B) = \mu(D\downarrow(A) - D\downarrow(B)) \tag{3.2}$$

and, further,

$$\mu(\operatorname{Eig} A, \operatorname{Eig} B) \leqslant \mu(A - B). \tag{3.3}$$

In [1] a new proof of (3.3) was given and it was shown that this inequality holds, more generally, when A, B, and $A - B$ are normal matrices. The standard method of proving that an inequality of the type $\mu(A) \leqslant \mu(B)$ holds simultaneously for all sui norms is to invoke a theorem of Ky Fan saying that it is only necessary to check the inequality for each of the n Ky Fan norms.

[See the discussion after (2.14) for definitions.] In [1] estimates for $\mu(\text{Eig}\, A, \text{Eig}\, B)$ were first obtained for the class of Ky Fan norms, and then it was claimed that they can be extended to all sui norms by the above reasoning. This last argument is quite wrong. Implicit here is the assumption that in the definition of $\mu(\text{Eig}\, A, \text{Eig}\, B)$ given by (3.1) the minimum would be attained at the same permutation matrix W for all μ. By (3.2) this is indeed the case when A, B are both Hermitian, but for arbitrary normal matrices this is no longer true. [This has been recently observed by Ando and Bhatia (unpublished).] However, this step in the proof can be avoided. The method used in [1] then not only gives all the theorems proved there, but leads also to a stronger result.

THEOREM 3.1. *Let τ be any wui norm on* $\mathbf{M}$. *Let A_0, A_1 be any two normal matrices, and let $A:[0,1] \to \mathbf{M}$ be any piecewise C^1 curve such that $A(0) = A_0$, $A(1) = A_1$, and $A(t)$ is a normal matrix for each t. Then*

$$\tau(\text{Eig}\, A_0, \text{Eig}\, A_1) \leqslant \int_0^1 \tau(A'(t))\, dt, \tag{3.4}$$

where $A'(t)$ is the derivative of $A(t)$.

Proof. All the essential details are the same as in [1]. We shall only sketch the proof and point out how s.u. invariance in [1] can be replaced by w.u. invariance.

We recall from [1] the basic decomposition

$$\mathbf{M} = T_A O_A \oplus Z_A, \tag{3.5}$$

valid for every normal matrix A, where $T_A O_A$ is the tangent space at A to the similarity orbit O_A of the matrix A, and Z_A is the commutant of A in $\mathbf{M}$.

If A and B commute, then we can find a $U \in \mathbf{U}$ such that UAU^* and UBU^* are both upper triangular. The diagonal entries of these triangular matrices are the eigenvalues of A and B, respectively. Thus the pinching inequality (2.31) implies

$$\tau(\text{Eig}\, A, \text{Eig}\, B) \leqslant \tau(A - B) \qquad \text{if} \quad [A, B] = 0. \tag{3.6}$$

Also, we have, trivially,

$$\tau(\text{Eig}\, A, \text{Eig}\, B) = 0 \qquad \text{if} \quad B \in O_A. \tag{3.7}$$

Now we estimate the spectral variation along the two components in (3.5) separately using (3.6) and (3.7). For the given path $A(t)$ consider the decomposition

$$\mathbf{M} = T_{A(t)}O_{A(t)} \oplus Z_{A(t)}$$

for each t. Let $P_t^{(1)}$ and $P_t^{(2)}$ be the complementary projection operators in $\mathbf{M}$ corresponding to the above direct sum decomposition. Arguing as in [1] we obtain

$$\tau(\mathrm{Eig}\, A_0, \mathrm{Eig}\, A_1) \leqslant \int \tau\left(P_t^{(2)}A'(t)\right) dt.$$

We now claim that

$$\tau\left(P_t^{(2)}B\right) \leqslant \tau(B) \qquad \text{for all} \quad B \in \mathbf{M}.$$

Since τ is wui and $A(t)$ normal, we may assume without loss of generality that $A(t)$ is diagonal. Then $Z_{A(t)}$ consists of block diagonal matrices, and $P_t^{(2)}(B)$ is the pinching of B by the spectral projections of $A(t)$. Thus our claim follows from the pinching inequality. This proves the theorem.

We must warn the reader that we have left out some technical details here, since they are the same as in [1]. Notably, the idea works first in a "generic" case and is then extended.

As a consequence we have the following result.

THEOREM 3.2. *If A_0, A_1 are normal matrices such that $A_1 - A_0$ is also normal, then for every wui norm τ we have*

$$\tau(\mathrm{Eig}\, A_0, \mathrm{Eig}\, A_1) \leqslant \tau(A_0 - A_1).$$

Note that the conditions of the theorem are surely met if A_0, A_1 are Hermitian. The famous Lidskii inequality is thus included in Theorem 3.2.

Notice that the right-hand side of (3.4) is the length of the path $A(t)$ in the norm τ. In Section 5 we shall prove the path inequality using a different method that also works for nondifferentiable paths.

4. ON A THEOREM OF HALMOS AND BOULDIN

In [11] Halmos considered the following problem. Let K be a closed nonempty set in $\mathbf{C}$. Let $\mathbf{N}$ denote the set of all normal operators acting in some fixed Hilbert space $\mathbf{H}$, and let $\mathbf{N}(K)$ be those elements of $\mathbf{N}$ whose spectrum is contained in K. Then, given an element A of N, which is the element of $\mathbf{N}(K)$ closest to A? He answered this question when the distance between operators is measured by the operator norm. Bouldin [6] showed that the same best approximant works for the distance as measured by any Schatten p-norm with $p \geqslant 2$. In this section we point out the connection between these results and the problem of spectral variation in finite dimensions.

The Halmos-Bouldin result can be described as follows. A map $F:\mathbf{C} \to K$ is called here a retraction onto K if it maps each point of $\mathbf{C}$ to a point in K which is as close as possible [i.e., $|z - F(z)| \leqslant |z - w|$ for all $w \in K$]. For every nonempty closed set there is a Borel measurable F with the above roperty. The Halmos-Bouldin theorem states that for every $A \in \mathbf{N}$ and $V \in \mathbf{N}(K)$

$$\|A - F(A)\|_p \leqslant \|A - N\|_p \tag{4.1}$$

for all Schatten p-norms $\|\cdot\|_p$ with $p \geqslant 2$. The $p = \infty$ case treats the operator norm and was proved by Halmos. Bouldin established the case $2 \leqslant p < \infty$. [The statement (4.1) has to be interpreted to mean that if there exists $N \in \mathbf{N}(K)$ such that $A - N$ is in the Schatten p-class, then $F(A)$ also has this property and (4.1) holds. In finite dimensions this qualification is redundant.]

We next point out some consequences for spectral variation.

Proposition 4.1. *Let $A, B \in \mathbf{N}(n)$, the set of n by n normal matrices. Let $\alpha_1,\dots,\alpha_n$ and $\beta_1,\dots,\beta_n$ be the respective eigenvalues of A and B. Suppose there is a permutation σ such that*

$$|\alpha_i - \beta_{\sigma(i)}| \leqslant |\alpha_i - \beta_{\sigma(j)}| \tag{4.2}$$

for all i, j. Then for every Schatten p-norm, $p \geqslant 2$, we have

$$\|(\operatorname{Eig} A, \operatorname{Eig} B)\|_p \leqslant \|A - B\|_p. \tag{4.3}$$

Proof. Let $K = \{\beta_1, \dots, \beta_n\}$. Then (4.2) lets us define a retraction F onto K by the relation $F(\alpha_i) = \beta_{\sigma(i)}$. Thus, by (4.1),

$$\left(\sum_{i=1}^{n} |\alpha_i - \beta_{\sigma(i)}|^p \right)^{1/p} \leqslant \|A - B\|_p,$$

which implies (4.3).

We should remark that the condition (4.2) is very stringent, so that Proposition 4.1 fails to cover many situations where (4.3) is nevertheless known to hold. Consider for example the Hermitian case where $\{\alpha_1, \alpha_2\} = \{1, 10\}$ and $\{\beta_1, \beta_2\} = \{9, 12\}$.

However, for A and B sufficiently close we can always appeal to Proposition 4.1 in the following sense.

PROPOSITION 4.2. *Let B be a normal matrix with eigenvalues $\beta_1, \dots, \beta_n$. Let δ be the minimum of $|\beta_i - \beta_j|$ with $\beta_i \neq \beta_j$, and let A be any normal matrix such that $\|A - B\| \leqslant \delta/6$. Then the eigenvalues of A and B satisfy* (4.2) *and hence* (4.3).

Proof. It follows from Theorem 5.1 in [3] that there exist a universal constant c and a permutation σ such that

$$|\alpha_i - \beta_{\sigma(i)}| \leqslant c\|A - B\| \qquad \text{for all } i.$$

It has been shown in [2] that this constant c is smaller than 3, so that our hypothesis implies that $|\alpha_i - \beta_{\sigma(i)}| < \delta/2$ for all i. We could not have $|\alpha_i - \beta_{\sigma(j)}| < |\alpha_i - \beta_{\sigma(i)}|$, for then $0 < |\beta_{\sigma(j)} - \beta_{\sigma(i)}| < \delta$.

With the last proposition we can derive the path inequality for spectral variation as in [5]. Let $N(t)$, $0 \leqslant t \leqslant 1$, be a continuous path in $\mathbf{N}$. Given any wui norm τ, define the τ-length of this path by

$$L_\tau(N(\cdot)) = \sup\left\{ \sum \tau(N(t_{k+1}) - N(t_k)) : 0 = t_0 < t_1 < \cdots < t_m = 1 \right\}. \tag{4.4}$$

Using Proposition 4.2 and following the argument of Theorem 2.2 in [5], we have

$$\|(\text{Eig}\, N(0), \text{Eig}\, N(1))\|_p \leqslant L_{p\text{-norm}}(N(\cdot)) \tag{4.5}$$

for every Schatten p-norm with $p \geqslant 2$. The method of [5] will be reviewed in detail in the course of the next section. For the present we have an alternate proof of Theorem 3.1 for general paths but within a special class of norms.

We end this section by pointing out that the $F(A)$ of Halmos and Bouldin is not a best approximant to A from $N(K)$ for the Schatten p-norms with $1 \leqslant p < 2$. In fact, no function of A can be such an approximant.

To see this let $K = \{i, -i\}$, and let

$$N = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \qquad A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Then $N \in \mathrm{N}(K)$ and $\|A - N\|_p = 2$ for all Schatten p-norms. If $F(A)$ is any function of A with spectrum contained in K, then in some orthonormal basis, A can be written as a diagonal matrix with entries 1 and -1, while $F(A)$ is diagonal with entries in $\{i, -i\}$. Hence $\|A - F(A)\|_p$ is $2^{1/2+1/p}$, which is larger than 2 for $p < 2$.

Note that in the example above both A and N were unitary; even in this restricted setting $F(A)$ is not a best approximant to A for $p > 2$. A "worse" example can be found in [3].

5. LOCAL METRIC GEOMETRY OF N AND THE PATH INEQUALITY

Here we establish some results on the behavior of spectral variation and normal paths in the neighborhood of a given normal operator. This will tell us about the local metric geometry of the normal manifold N and will allow us to extend our path inequality to the broadest context.

PROPOSITION 5.1. *For a fixed normal N_0 let δ be half the minimum distance between distinct eigenvalues of N_0. Then there exists a finite M (depending only on δ and the dimension n of the space) such that any normal N_1 with $\|N_1 - N_0\| \leqslant \delta/3$ can be represented as $N_1 = UN_2U^*$, where N_2 is a (normal) operator commuting with N_0, and U is unitary with $\|I - U\| \leqslant M\|N_1 - N_0\|$.*

Proof. With respect to an appropriate basis we may write N_0 as a diagonal matrix $D_0 = \oplus \alpha_k I_k$, where the α_k are distinct eigenvalues and the I_k are identity submatrices of the appropriate dimensions. Now δ will be half the minimum distance between distinct α_k's. By the arguments of Proposition 4.2, our condition on N_1 means that the eigenvalues of N_1 can be grouped into diagonal blocks B_k of the same dimensions as the I_k in such a way that every eigenvalue in B_k is within δ of α_k. By our choice of δ we have ensured that

$$\left\|(B_j - \alpha_k)^{-1}\right\| \leqslant 1/\delta \qquad \text{whenever} \quad j \neq k. \tag{5.1}$$

For the appropriate unitary matrix W we have N_1 represented by WD_2W^*, where $D_2 = \oplus B_k$. If the corresponding block form of W is $[W_{kj}]$, we have

$$\left\|\left[W_{kj}(B_j - \alpha_k)\right]\right\| = \|WD_2 - D_0W\| = \|WD_2W^* - D_0\| = \|N_1 - N_0\|. \tag{5.2}$$

From (5.1) and (5.2) it follows that $\|W_{kj}\| \leqslant \|N_1 - N_0\|/\delta$ whenever $j \neq k$. We conclude that for some M' (depending only on δ and the dimension of the space) $\|W - \oplus W_{kk}\| \leqslant M'\|N_1 - N_0\|$ whenever $\|N_1 - N_0\| \leqslant \delta/3$.

Now let W_k be the unitary part (in the polar decomposition) of W_{kk}, so that $W_{kk} = |W_{kk}|W_k$. Since $\|W_{kk} - W_k\| = \|\,|W_{kk}| - I_k\| \leqslant \|\,|W_{kk}|^2 - I_k\|$, we have $\|\oplus W_{kk} - \oplus W_k\| \leqslant \|(\oplus W_{kk})^* \oplus W_{kk} - I\|$. Setting $V = \oplus W_k$ and $X = \oplus W_{kk}$, we have $\|X - V\| \leqslant \|X^*X - W^*W\|$ and $\|W - X\| \leqslant M'\|N_1 - N_0\|$. Noting that the pinching inequality ensures that $\|X\| \leqslant 1$, we have

$$\begin{aligned}\|W - V\| &\leqslant \|W - X\| + \|X - V\| \leqslant \|W - X\| + \|X^*X - W^*W\| \\ &\leqslant \|W - X\| + \|(X^* - W^*)X\| + \|W^*(X - W)\| \\ &\leqslant 3\|W - X\| \leqslant 3M'\|N_1 - N_0\|.\end{aligned}$$

With $U = WV^*$ and $M = 3M'$, the inequality of the lemma is satisfied. Furthermore, if N_2 is the operator represented by VD_2V^*, we have $N_1 = UN_2U^*$; N_2 certainly commutes with N_0, since it has the form $\oplus W_kB_k(W_k)^*$.

Proposition 5.2. *For fixed wui norm τ, normal N_0, and $\epsilon > 0$,*

$$\tau(\operatorname{Eig} N_1, \operatorname{Eig} N_0) \leqslant (1 + \epsilon)\tau(N_1 - N_0)$$

whenever N_1 is normal and sufficiently close to N_0.

Proof. Let U and N_2 be as in Proposition 5.1. Since N_0 and N_2 commute, we may represent them simultaneously by diagonal matrices C_0 and C_2. Let $S = U - I$, so that $U = I + S$, $U^* = I - S + S^2U^*$, and

$$N_0 - N_1 = C_0 - C_2 + C_2S - SC_2 - UC_2S^2U^* + SC_2S.$$

It follows that

$$\tau(C_0 - C_2 + C_2S - SC_2) \leqslant \tau(N_0 - N_1) + \tau\big(UC_2S^2U^* - SC_2S\big). \quad (5.3)$$

Since τ is equivalent to the operator norm, there is some constant r such that all ratios $\tau(T)/\|T\|$ lie in $[1/r, r]$. Thus the second term on the right of (5.3) is bounded by $2r\|C_2\|\,\|S\|^2$. The diagonal of $C_2S - SC_2$ is empty, so that the pinching inequality ensures that

$$\tau(C_0 - C_2) \leqslant \tau(C_0 - C_2 + C_2S - SC_2). \quad (5.4)$$

Thus (5.3) implies that

$$\tau(\operatorname{Eig} N_1, \operatorname{Eig} N_0) \leqslant \tau(N_1 - N_0) + 2r\|N_1\|\{rM\tau(N_1 - N_0)\}^2,$$

where M is the constant in Proposition (5.1). If N_1 is (also) so close to N_0 that $2r^3M^2\|N_1\|\tau(N_1 - N_0) < \epsilon$, the inequality of our proposition is satisfied.

Remark 5.3. As noted in Proposition 4.2, Bouldin's theorem implies that if τ is any of the Schatten p-norms with $p \geqslant 2$, then Proposition 5.2 remains valid with $\epsilon = 0$. Nevertheless, we must retain the condition $\epsilon > 0$ in general. To see this consider the unitary operators

$$U(t) = \begin{bmatrix} 0 & 1 \\ e^{it} & 0 \end{bmatrix}.$$

It is evident that $\sigma(U(t)) = \{e^{it/2}, -e^{it/2}\}$. Clearly, then, if $w(T)$ denotes the numerical radius of T and $0 \leqslant t \leqslant \pi$, we have

$$w(\operatorname{Eig} U(t), \operatorname{Eig} U(0)) = |1 - e^{it/2}| = 2\sin(t/4).$$

On the other hand, $w(U(t) - U(0))$ is easily seen to be $|e^{it} - 1|/2$, i.e.,

$\sin(t/2)$. Although $U(t)$ approaches $U(0)$ as t tends to 0, the ratio

$$\frac{w(\operatorname{Eig} U(t), \operatorname{Eig} U(0))}{w(U(t)-U(0))} = \sec\left(\frac{t}{4}\right)$$

exceeds 1. The same ratio is obtained for the trace norm, so that we must generally retain the condition $\epsilon > 0$ in Proposition 5.2 even for sui norms.

We are now in a position to prove the most general path inequality by a modification of the method used in [5].

THEOREM 5.4. *For any wui norm τ and any (continuous) normal path $N(\cdot)$ (defined on* [0, 1]), *we have the path inequality*

$$\tau(\operatorname{Eig} N(0), \operatorname{Eig} N(1)) \leqslant L_\tau(N(\cdot)).$$

Proof. It is clearly enough to prove the inequality

$$\tau(\operatorname{Eig} N(0), \operatorname{Eig} N(1)) \leqslant (1+\epsilon) L_\tau(N(\cdot)),$$

for each $\epsilon > 0$. Let $N_t(\cdot)$ denote the part of the path defined on $[0, t]$, and let $G = \{t : \tau(\operatorname{Eig} N(t), \operatorname{Eig} N(0)) \leqslant (1+\epsilon) L_\tau(N_t(\cdot))\}$; we must show that 1 is in G. Let $g = \sup G$. By the continuity of $N(\cdot)$ and of spectra, g belongs to G. If g were less than 1, we could find, in view of Proposition 5.2 [applied with $N_0 = N(g)$], some $t > g$ such that $\tau(\operatorname{Eig} N(t), \operatorname{Eig} N(g))$ is not more than $(1+\epsilon)\tau(N(t) - N(g))$. But then we would have, since the spectral distance is a pseudometric,

$$\begin{aligned}
&\tau(\operatorname{Eig} N(t), \operatorname{Eig} N(0)) \\
&\quad\leqslant \tau(\operatorname{Eig} N(g), \operatorname{Eig} N(0)) + \tau(\operatorname{Eig} N(t), \operatorname{Eig} N(g)) \\
&\quad\leqslant (1+\epsilon)\left\{L_\tau(N_g(\cdot)) + \tau(N(t) - N(g))\right\} \\
&\quad\leqslant (1+\epsilon) L_\tau(N_t(\cdot)),
\end{aligned}$$

i.e., $t \in G$, a contradiction.

In [5] we showed that for the operator norm some parts of the normal manifold N are metrically flat. These were called "plains" in [5]. Here we shall say that a subset Y of normal operators is τ-flat if every pair of operators N_0, N_1 in Y can be joined by a path $N(\cdot)$ lying in Y that is τ-short, i.e., such that $L_\tau(N(\cdot)) = \tau(N_1 - N_0)$. In [5] we saw that the set of scalar multiples of unitary operators, CU, is $\|\cdot\|$-flat, but it is known that the normal manifold as a whole is not $\|\cdot\|$-flat for dimension $\geqslant 3$.

For other norms, flatness it is still more rare. For the Frobenius norm, for example, the geometry is Euclidean, so that normals N_0 and N_1 can lie in a flat subset only if the line segment joining them consists of normal operators; this is equivalent to requiring that $N_1 - N_0$ be normal. With this norm, then, not even CU is flat; the examples of Remark 5.3 also illustrate this phenomenon.

From Proposition 5.1 we can nevertheless conclude that the whole normal manifold is "locally asymptotically τ-flat" for any wui τ. More precisely, we have the following proposition.

PROPOSITION 5.5. *For a fixed wui norm τ, normal N_0, and $\epsilon > 0$, there is a normal path $N(\cdot)$ from N_0 to N_1 such that $L_\tau(N(\cdot)) \leqslant (1+\epsilon)\tau(N_1 - N_0)$ whenever N_1 is normal and sufficiently close to N_0.*

Proof. Every unitary matrix U can be expressed as $U = e^K$, where K is a skew-Hermitian matrix with $\|K\| \leqslant (\pi/2)\|I - U\|$ (see [3]). Hence Proposition 5.1 implies that there is a constant M such that if N_1 is a normal sufficiently close to N_0 then we may represent N_0, N_1 simultaneously by D_0, $e^K D_1 e^{-K}$, where D_0, D_1 are diagonal and K is skew-Hermitian with $\|K\| \leqslant M\|N_1 - N_0\|$.

For $0 \leqslant t \leqslant 1$, let $D(t) = (1-t)D_0 + tD_1$, and let $N(t) = e^{tK}D(t)e^{-tK}$. Then $N(\cdot)$ is a path in N. Now

$$\begin{aligned} L_\tau(N(\cdot)) &= \int_0^1 \tau(N'(t))\,dt \\ &= \int_0^1 \tau\left(e^{tK}[KD(t) + D'(t) - D(t)K]e^{-tK}\right)dt. \end{aligned}$$

Note that

$$\begin{aligned} &KD(t) + D'(t) - D(t)K \\ &\quad = (1-t)(D_1 - D_0 + KD_0 - D_0K) + t(D_1 - D_0 + KD_1 - D_1K). \end{aligned}$$

Since τ is a wui norm, this gives

$$\begin{aligned} L_\tau(N(\cdot)) \leqslant \int_0^1 \{&(1-t)\tau(D_1 - D_0 + KD_0 - D_0K) \\ &+ t\tau(D_1 - D_0 + KD_1 - D_1K)\}\,dt. \qquad (5.5) \end{aligned}$$

By the usual power series expansion

$$D_1 - D_0 + KD_0 - D_0K = D_1 - e^{-K}D_0e^K + \text{HOT}$$

where HOT stands for higher order terms that involve at least two factors of K. Routine estimates with the operator norm, the bound $\|K\| \leqslant M\|N_1 - N_0\|$, and the equivalence of τ with $\|\cdot\|$ show that there is a constant M' such that $\tau(\text{HOT}) \leqslant M'\tau^2(N_1 - N_0)$. Thus

$$\begin{aligned}\tau(D_1 - D_0 + KD_0 - D_0K) &\leqslant \tau\left(D_1 - e^{-K}D_0e^K\right) + M'\tau^2(N_1 - N_0)\\ &= \{1 + M'\tau(N_1 - N_0)\}\,\tau(N_1 - N_0).\end{aligned}$$

The second term in the integrand on the right side of (5.5) can be estimated similarly. Together, these estimates show that

$$L_\tau(N(\cdot)) \leqslant \{1 + M'\tau(N_1 - N_0)\}\,\tau(N_1 - N_0).$$

If we require N_1 to be so close to N_0 that we also have $M'\tau(N_1 - N_0) \leqslant \epsilon$, then the inequality of our proposition is assured.

REMARK 5.6. In general, we must retain the condition $\epsilon > 0$ in the proposition above. Indeed, if the proposition is true with $\epsilon = 0$, then, by Theorem 5.4, we have $\tau(\operatorname{Eig} N_0, \operatorname{Eig} N_1) \leqslant \tau(N_0 - N_1)$ for all normal N_1 sufficiently close to N_0. As we noted in Remark 5.3, this is not generally the case.

Finally we note that local spectral perturbation is especially well behaved for the operator norm. The following proposition may be compared with Proposition 2.1 in [5].

PROPOSITION 5.7. *Let N_0 be a fixed normal matrix. Then for every matrix T (normal or not) in a certain small neighborhood of N_0 we have*

$$\|(\operatorname{Eig} N_0, \operatorname{Eig} T)\| \leqslant \|N_0 - T\|.$$

Proof. Let $\alpha_1, \ldots, \alpha_n$ be the eigenvalues of N_0, and let δ be half the minimum distance between distinct α_k. By general results on spectral perturbation (see e.g. [4]), if T is sufficiently close to N_0, there is a matching to the α_k with the eigenvalues β_k of T such that $|\beta_k - \alpha_k| < \delta$ for all k. By our choice of δ, it is clear that $|\beta_k - \alpha_j|$ is smallest when $j = k$. Choosing u to be a unit eigenvector of T belonging to β_k and an orthonormal basis u_j such that $N_0 u_j = \alpha_j u_j$, we have $u = \Sigma\langle u, u_j\rangle u_j$, so that $1 = \Sigma|\langle u, u_j\rangle|^2$ and

$$(T - N_0)u = \sum_j (\beta_k - \alpha_j)\langle u, u_j\rangle u_j.$$

It follows that

$$|\beta_k - \alpha_k|^2 \leqslant \sum_j |\beta_k - \alpha_j|^2 |\langle u, u_j\rangle|^2 = \|(T - N_0)u\|^2 \leqslant \|T - N_0\|^2.$$

Since this holds for each k, we have in fact $\|(\mathrm{Eig}\, N_0, \mathrm{Eig}\, T)\| \leqslant \|T - N_0\|$.

The authors wish to thank NSERC of Canada for its support (via Operating Grant A8745) of their research in this area.

REFERENCES

1 R. Bhatia, Analysis of spectral variation and some inequalities, *Trans. Amer. Math. Soc.* 272:323–331 (1982).

2 R. Bhatia, C. Davis, and P. Koosis, An extremal problem in Fourier analysis with applications to operator theory, to appear.

3 R. Bhatia, C. Davis, and A. McIntosh, Perturbation of spectral subspaces and solution of linear operator equations, *Linear Algebra Appl.* 52–53:45–67 (1983).

4 R. Bhatia and S. Friedland, Variation of Grassman powers and spectra, *Linear Algebra Appl.* 40:1–18 (1981).

5 R. Bhatia and J. A. R. Holbrook, Short normal paths and spectral variation, *Proc. Amer. Math. Soc.* 94:377–382 (1985).

6 R. Bouldin, Best approximation of a normal operator in the Schatten p-norm, *Proc. Amer. Math. Soc.* 80:277–282 (1980).

7 C. Davis, Various averaging operations onto subalgebras, *Illinois J. Math.* 3:538–553 (1959).

8 C.-K. Fong and J. A. R. Holbrook, Unitarily invariant operator norms, *Canad. J. Math.* 35:274–299 (1983).

9 C.-K. Fong, H. Radjavi, and P. Rosenthal, Norms for matrices and operators, to appear.

10 I. C. Gohberg and M. G. Krein, *Introduction to the Theory of Linear Nonselfadjoint Operators*, Amer. Math. Soc., 1969.

11 P. R. Halmos, *A Hilbert Space Problem Book*, 2nd ed., Springer, 1982.

12 P. R. Halmos, Spectral approximants of normal operators, *Proc. Edinburgh Math. Soc.* 19:51–58 (1974).

13 E. Hewitt and K. A. Ross, *Abstract Harmonic Analysis*, Vol. II, Springer, 1970.
14 V. B. Lidskii, On the eigenvalues of sums and products of symmetric matrices (in Russian), *Dokl. Akad. Nauk SSSR* 75:769–772 (1950).
15 A. W. Marshall and I. Olkin, *Inequalities*, Academic, 1979.
16 R. Schatten, *Norm Ideals of Completely Continuous Operators*, Springer, 1960.
17 J. von Neumann, Some matrix inequalities and metrization of matric space, *Tomsk Univ. Rev.* 1:286–299 (1937); also in *Collected Works*, Vol. IV, Pergamon, 1961, pp. 205–214.

Received 7 August 1986; revised 6 November 1986