PACIFIC JOURNAL OF MATHEMATICS
Vol. 38, No. 2, 1971

NORMS OF DERIVATIONS ON $\mathscr{L}(\mathfrak{X})$

B. E. JOHNSON

If $\mathfrak{X}$ is a real or complex Banach space and $\mathscr{L}(\mathfrak{X})$ is the algebra of bounded linear endomorphisms of $\mathfrak{X}$ then each element T of $\mathscr{L}(\mathfrak{X})$ defines an operator D_T on $\mathscr{L}(\mathfrak{X})$ by $D_T(A) = AT - TA$. Clearly $\|D_T\| \leqq 2\ \inf_\lambda\ \|T + \lambda I\|$ and Stampfli has shown that when $\mathfrak{X}$ is a complex Hilbert space equality holds. In this paper it is shown, by methods which apply to a large class of uniformly convex spaces, that this formula for $\|D_T\|$ is false in l^p and $L^p(0, 1)$ $1 < p < \infty$, $p \neq 2$. For L^1 spaces the formula is true in the real case but not in the complex case when the space has dimension 3 or more.

Stampfli's results appear in [1] stated for complex Hilbert space but the same proofs yield the corresponding result for real spaces.

Throughout this paper $\boldsymbol{K}$ will denote either $\boldsymbol{R}$ or $\boldsymbol{C}$. We begin by describing the construction of an operator T of rank 1 with $\|D_T\| < d_T = 2\inf_{\lambda \in K} \|T + \lambda I\|$. The reason that this fails in Hilbert space is precisely because for an ellipse, conjugacy is a symmetric relation on the set of diameters; more precisely if x, y are two points on the unit ball then y is parallel to the tangent plane at x if and only if x is parallel to the tangent plane at y.

DEFINITION 1. Let $x \in \mathfrak{X}$, $\|x\| = 1$. The unit ball $\mathfrak{X}_1$ is *uniformly convex* at x if whenever $\{y_n\}$ is a sequence with $\|y_n\| \leqq 1$, $\|x + y_n\| \to 2$ then $y_n \to x$.

PROPOSITION 2. *Let $\mathfrak{X}$ be a normed space over $\boldsymbol{K}$ and let $x, y \in \mathfrak{X}$ with the following properties*

(i) *$\|x\| = 1$ and there is $f \in \mathfrak{X}^*$ with $\|f\| = 1$ and such that if $\{x_n\}$ is a sequence with $\|x_n\| \leqq 1, f(x_n) \to 1$ then $x_n \to x$.*

(ii) *$\|y\| = 1$ and the unit ball $\mathfrak{X}_1$ is uniformly convex at y.*

(iii) *For some $\lambda \in \boldsymbol{K}$, $\|x + \lambda y\| < 1$.*

(iv) *For all λ in $\boldsymbol{K}$, $\|y + \lambda x\| \geqq 1$.*

Define $T \in \mathscr{L}(\mathfrak{X})$ by $Tz = f(z)y$. Then $2\|T\| = d_T > \|D_T\|$.

Proof. $\|T + \lambda I\| \geqq \|(T + \lambda I)x\| = \|y + \lambda x\| \geqq \|y\| = 1$ by (iv) and $\|T\| = 1$ so $d_T = 2$. Suppose $\|D_T\| = 2$ and choose sequences $\{A_n\}$ from $\mathscr{L}(\mathfrak{X})$ and $\{x_n\}$ from $\mathfrak{X}$ with $\|A_n\| = 1 = \|x_n\|$ and $\|D_T(A_n)x_n\| \to 2$. As $\|TA_nx_n\| \leqq 1$, $\|A_nTx_n\| \leqq 1$ we have $\|TA_nx_n\| \to 1$, $\|A_nTx_n\| \to 1$

and hence $||A_n x_n|| \to 1$, $||Tx_n|| \to 1$. This shows $|f(x_n)| \to 1$ and so, replacing x_n by $w_n x_n$ if necessary where $\{w_n\}$ is a sequence of elements of $\boldsymbol{K}$ with $|w_n| = 1$, we may assume $f(x_n) \to 1$. Condition (i) now implies $x_n \to x$ and hence $Tx_n \to y$. In the same way $||TA_n x_n|| \to 1$ implies $|f(A_n x_n)| \to 1$ and replacing A_n by $w'_n A_n$ if necessary we can assume $f(A_n x_n) \to 1$ from which we see $A_n x_n \to x$, $TA_n x_n \to y$. As $||A_n|| \leqq 1$ we have $A_n Tx_n - A_n y \to 0$ and so $||A_n Tx_n - TA_n x_n|| \to 2$ implies $||A_n y - y|| \to 2$. Condition (ii) now shows $A_n y \to -y$ so that $A_n(x + \lambda y) \to x - \lambda y$. However if λ satisfies condition (iii) then $||x - \lambda y|| > 1$, as otherwise $2 = 2\,||x|| \leqq ||x + \lambda y|| + ||x - \lambda y|| < 2$, and so $\lim ||A_n(x + \lambda y)|| = ||x - \lambda y|| > 1$ which is impossible because $||A_n(x + \lambda y)|| \leqq ||A_n||\,||x + \lambda y|| < 1$.

PROPOSITION 3. *Let $\mathfrak{X}$ be a uniformly convex Banach space, $x, y \in \mathfrak{X}$, $f, g \in \mathfrak{X}^*$ with $||x|| = ||y|| = ||f|| = ||g|| = f(x) = g(y) = 1$, $g(x) = 0$, $f(y) \neq 0$ and suppose f is the only element h of $\mathfrak{X}^*$ with $||h|| = h(x) = 1$. Then x, y, f satisfy the conditions of Proposition 2.*

Proof. (i) If $||x_n|| \leqq 1, f(x_n) \to 1$ then $f(x + x_n) \to 2$ and as $||x + x_n|| \leqq 2$, $||f|| = 1$ we have $||x + x_n|| \to 2$ so $x_n \to x$ by uniform convexity.

(ii) is clearly part of the present hypotheses.

(iii) x and y are linearly independent as $g(x) = 0$, $g(y) = 1$, $x \neq 0$. If $||x + \lambda y|| \geqq 1$ for all $\lambda \in \boldsymbol{K}$ then $\alpha x + \beta y \mapsto \alpha$ is a norm one linear functional on the space spanned by x and y and so has an extension h in $\mathfrak{X}^*$ with $||h|| = 1$, $h(x) = 1$ but $h \neq f$ because $h(y) = 0 \neq f(y)$.

(iv) As $g(y + \lambda x) = g(y) = 1$, for all λ in $\boldsymbol{K}$ and $||g|| = 1$ we have $||y + \lambda x|| \geqq 1$ for all λ in $\boldsymbol{K}$.

COROLLARY 4. *If $1 < p < 2$ or $2 < p < \infty$ and $\mathfrak{X} = l^p(0, \infty)$ or $\mathfrak{X} = L^p(-1, +1)$ is the corresponding $\boldsymbol{K}$ Banach space of $\boldsymbol{K}$ valued functions then there is $T \in \mathscr{L}(\mathfrak{X})$ with $||D_T|| \neq d_T$.*

Proof. The spaces are uniformly convex and at each point z of $\mathfrak{X}$ with $||z|| = 1$ the element h of $\mathfrak{X}^*$ with $h(z) = 1 = ||h||$ is unique. Thus the construction in Proposition 2 applies once we find two suitable points x, y and these exist in such abundance that we can take anything but multiples of characteristic functions for x. First of all we give the construction in the two dimensional space $l^p(1, 2)$.

If $x = (x_1, x_2)$, $x_1 > 0$, $x_2 > 0$, $x_1^p + x_2^p = 1$ then $f(z) = x_1^{p-1} z_1 + x_2^{p-1} z_2$ so y can be taken as $\alpha(x_2^{p-1}, -x_1^{p-1})$ where $\alpha^{-p} = x_1^{p(p-1)} + x_2^{p(p-1)}$ and

$g(z) = \alpha^{p-1}(x_2^{(p-1)^2} z_1 - x_1^{(p-1)^2} z_2)$. Then $g(x) = \alpha^{p-1}(x_1 x_2^{(p-1)^2} - x_1^{(p-1)^2} x_2)$ which will be zero if and only if $x_1 = x_2$. Thus taking say $x = 3^{-1/p}(2^{1/p}, 1)$ and y, f, g as above the result is shown in $l^p(1, 2)$.

As $l^p(0, \infty)$ and $L^p(-1, +1)$ each contain subspaces isometric with $l^p(1, 2)$ we can construct x, y, f, g in this subspace and then extend f and g to $\mathfrak{X}$ using the Hahn-Banach theorem.

In order to prove the results for spaces of measures we establish the equation $d_T = \|D_T\|$ for finite dimensional l^1 spaces.

PROPOSITION 5. *Let n be a positive integer and $\mathfrak{X}$ be the real Banach space $\boldsymbol{R}^n$ with norm $\|x\| = \Sigma |x_i|$. Let $T \in \mathscr{L}(\mathfrak{X})$. Then $\|D_T\| = 2 \inf_{\lambda \in \boldsymbol{R}} \|T + \lambda I\|$.*

Proof. Suppose T is given by the matrix a_{ij} in the standard basis $e_1, e_2, \cdots, e_n$. We have $\|T\| = \sup_j \sum_i |a_{ij}|$. Suppose $\sum_i |a_{ij}| = \|T\|$ for $j = 1, \cdots, m$ but not for $j > m$. The condition $\|T\| = \frac{1}{2} d_T$ is equivalent to saying that 0 is in the convex hull of $a_{11}, \cdots, a_{mm}$ since if 0 does not lie in this convex hull then either $|a_{jj} + \lambda| < |a_{jj}|$ for $j = 1, \cdots, m$ and small positive λ or for small negative λ and so there are small values of λ with $\|T + \lambda I\| < \|T\|$ whereas if 0 lies in this hull and $\lambda \neq 0$ there is j with $1 \leqq j \leqq m$ and $|a_{jj} + \lambda| > |a_{jj}|$ so that $\|T + \lambda I\| > \|T\|$.

It is clearly sufficient to prove the result when $\|T\| = \frac{1}{2} d_T$. First of all consider the case $m \geqq 2$ and suppose $a_{11} \geqq 0 \geqq a_{22}$. Let $A \in \mathscr{L}(\mathfrak{X})$ be an operator of the form $Ae_1 = e_2$, $Ae_2 = \pm e_1$, $Ae_i = \pm e_i$ $i = 3, \cdots, n$. Clearly $\|A\| = 1$ and

$$\begin{aligned} \|D_T(A)e_1\| &= \|ATe_1 - Te_2\| \\ &= |\pm a_{21} - a_{12}| + |a_{11} - a_{22}| + \sum_{i=3}^{n} |\pm a_{i1} - a_{i2}| \\ &= \sum_{i=1}^{n} |a_{i1}| + \sum_{i=1}^{n} |a_{i2}| \\ &= 2\|T\| \end{aligned}$$

for a suitable choice of signs of the Ae_i since each sign to be chosen corresponds to exactly one term $|\pm a_{i1} - a_{i2}|$.

If $m = 1$ then $a_{11} = 0$ because 0 lies in the convex hull of $a_{11}, \cdots, a_{mm}$, and we define A by $Ae_1 = e_1$, $Ae_j = -e_j$ $j = 2, \cdots, n$ which gives $\|A\| = 1$ and $ATe_1 = -Te_1$ so that

$$\|D_T(A)e_1\| = \|ATe_1 - TAe_1\| = 2\|Te_1\| = 2\|T\| .$$

PROPOSITION 6. *Let Ω be a compact topological space and $\mathfrak{X}$ a closed linear subspace of the* (real) *Banach space of real valued measures on Ω with the property that if $\mu \in \mathfrak{X}$ then every measure* *absolutely continuous with respect to* μ *is in* $\mathfrak{X}$. *Let* $T \in \mathscr{L}(\mathfrak{X})$. *Then* $\|D_T\| = 2 \inf_{\lambda \in \boldsymbol{R}} \|T + \lambda I\|$.

Proof. We may assume $d_T = 2\|T\|$. Let $\varepsilon > 0$. For each $\nu > 0$ in $\mathfrak{X}$ let $E_\nu(\mu)$ be the part of $\mu \in \mathfrak{X}$ which is absolutely continuous with respect to ν. The E_ν form a system of commuting idempotents of norm 1 and $E_\nu E_{\nu'} = E_\nu$ if $\nu' > \nu$, so that $\|E_\nu S E_\nu\|$, where the elements ν are directed by the usual ordering of measures, is a monotonic direct net. It is easy to see that $\|E_\nu S E_\nu\| \to \|S\|$. Thus applying Dini's theorem to the functions $\lambda \mapsto \|E_\nu(T + \lambda I)E_\nu\|$ we can find $\nu \in \mathfrak{X}, \nu > 0$ with $\|E_\nu(T + \lambda I)E_\nu\| > \|T + \lambda I\| - \varepsilon \geqq \|T\| - \varepsilon$ for $|\lambda| \leqq 2\|T\|$.

For each dissection $\Delta = (\Omega_1, \cdots, \Omega_n)$ of Ω into disjoint measurable sets of positive ν measure we define

$$P_\Delta(\mu) = (E_\nu\mu(\Omega_1), \cdots, E_\nu\mu(\Omega_n))$$
$$Q_\Delta(\xi) = (\textstyle\sum c_i \xi_i \nu(\Omega_i)^{-1})\nu$$

where $\mu \in \mathfrak{X}$, $\xi \in \mathbf{R}^n$, $P_\Delta; \mathfrak{X} \to \boldsymbol{R}^n$, $Q_\Delta: \boldsymbol{R}^n \to \mathfrak{X}$ and c_i is the characteristic function of Ω_i. Directing the dissections in the usual way it is easy to see that for each $S \in \mathscr{L}(\mathfrak{X})$ $\|P_\Delta E_\nu S E_\nu Q_\Delta\|$, where $\boldsymbol{R}^n$ has the l^1 norm, is a monotonic directed set with limit $\|E_\nu S E_\nu\|$. Applying Dini's theorem again we see that there is a dissection Δ with

$$(*) \qquad \|P_\Delta E_\nu(T + \lambda I)E_\nu Q_\Delta\| > \|T\| - \varepsilon$$

for all $|\lambda| \leqq 2\|T\|$. For convenience we now denote $E_\nu, P_\Delta, Q_\Delta$ by E, P, Q. As these operators have norm 1 we see that inequality (*) holds for all values of λ. As $PE = P$, $EQ = Q$, $PEQ = PQ =$ identity on $\boldsymbol{R}^n$, $(*)$ shows that $d_{PTQ} \geqq 2(\|T\| - \varepsilon)$. By proposition 5 there is $A \in \mathscr{L}(\boldsymbol{R}^n)$ with $\|D_{PTQ}(A)\| = d_{PTQ}$, $\|A\| = 1$. As Q is an isometry and P maps the unit ball of $\mathfrak{X}$ onto that of $\boldsymbol{R}^n$ we have

$$\begin{aligned} d_{PTQ} &= \|QD_{PTQ}(A)P\| \\ &= \|QAPTQP - QPTQAP\| \\ &= \|QPD_T(QAP)QP\| \\ &\leqq \|D_T(QAP)\| . \end{aligned}$$

As $\|QAP\| = 1$ we have $\|D_T\| \geqq d_{PTQ} \geqq 2(\|T\| - \varepsilon)$ for each $\varepsilon > 0$ and the result follows.

In the complex space $l^1(1, 2)$ Proposition 5 is true and the proof is similar to that for the real case. However the result is false in higher dimensions for complex spaces, e.g., in $l^1(1, 2, 3)$ let T be the linear transformation given by the matrix

$$\begin{matrix} 1 & -\omega & -\omega^2 \\ 1 & \omega & -\omega^2 \\ 1 & \omega & \omega^2 \end{matrix}$$

where $\omega^3 = 1, \omega \neq 1$. The situation is similar to that at the beginning of the proof of Proposition 5 with $m = n = 3$ and the argument given there shows that because 0 is a convex combination of diagonal entries we have $\inf_{\lambda \in C} \|T + \lambda I\| = \|T\| = 3$. If $\|x\| = 1, \|A\| = 1$ and $\|D_T(A)x\| = 6$ then $\|Tx\| = 3$ and since $|x_1 \pm \omega x_2 \pm \omega^2 x_3| \leqq 1$ we see that $|x_1 - \omega x_2 - \omega^2 x_3| = |x_1 + \omega x_2 - \omega^2 x_3| = |x_1 + \omega x_2 + \omega^2 x_3| = |x_1| + |x_2| + |x_3|$ which occurs only if two of x_1, x_2, x_3 are 0. Multiplying by a complex number of absolute value 1, if necessary, we can assume $x = e_1$ or e_2 or e_3. In the same way $Ax = e_1$ or e_2 or e_3. If $x = e_1 = Ax$ then

$$\begin{aligned} \|D_T(A)e_1\| &= \|e_1 + Ae_2 + Ae_3 - e_1 - e_2 - e_3\| \\ &= \|Ae_2 + Ae_3 - e_2 - e_3\| \\ &\leqq 4 \end{aligned}$$

and if $x = e_1, Ax = e_2$ then

$$\begin{aligned} \|D_T(A)e_1\| &= \|e_2 + Ae_2 + Ae_3 + \omega e_1 - \omega e_2 - \omega e_3\| \\ &= \|(1 - \omega)e_2 + Ae_2 + Ae_3 - \omega e_1 - \omega e_3\| \\ &\leqq \sqrt{3} + 4 . \end{aligned}$$

The other four possibilities give similar results and so we cannot in fact have $\|D_T\| = 6$.

A similar construction in the complex spaces $l^1(1, n)$, $l^1(0, \infty)$, $L^1(0, 1)$, $M(0, 1)$ shows that Proposition 6 is false in these spaces too.

Reference

1. J. G. Stampfli, *On the norm of a derivation*, Pacific J. Math., **33** (1970), 737-747.

Received November 2, 1970. The author gratefully acknowledges financial support from the National Science Foundation Grant GP-21193.

YALE UNIVERSITY