A quantitative version of the commutator theorem for zero trace matrices *

William B. Johnson†, Narutaka Ozawa‡, Gideon Schechtman§

Abstract

Let A be a $m \times m$ complex matrix with zero trace and let $\varepsilon > 0$. Then there are $m \times m$ matrices B and C such that $A = [B, C]$ and $\|B\|\|C\| \leq K_\varepsilon m^\varepsilon \|A\|$ where K_ε depends only on ε. Moreover, the matrix B can be taken to be normal.

1 Introduction

It is well known that a complex $m \times m$ matrix A is a commutator (i.e., there are matrices B and C of the same dimensions as A such that $A = [B, C] = BC - CB$) if and only if A has zero trace. In such a situation clearly $\|A\| \leq 2\|B\|\|C\|$ where $\|D\|$ denotes the norm of D as an operator from ℓ_2^m to itself.

Is it true that the converse holds? That is, if A has zero trace are there $m \times m$ matrices B and C such that $A = [B, C]$ and $\|B\|\|C\| \leq K\|A\|$ for some absolute constant K?

Here we provide a weaker estimate: The above holds for $K = K_\varepsilon m^\varepsilon$ for every $\varepsilon > 0$ where K_ε depends only on ε. Moreover, the matrix B can be taken to be normal.

*AMS subject classification: 47B47, 15A60 Key words: commutators, zero trace, norm of matrices

†Supported in part by NSF DMS-1001321 and U.S.-Israel Binational Science Foundation

‡Supported in part by JSPS

§Supported in part by U.S.-Israel Binational Science Foundation. Participant NSF Workshop in Analysis and Probability, Texas A&M University

The proof will be presented in the next section. It is self contained except for two facts. The first is a relatively easy result of Rosenblum [1] which gives a solution for X of the matrix equation $A = SX - XT$ where all matrices are square and S and T have separated spectra in the sense that there is a domain D, whose boundary is a simple curve, which contains the spectrum of S and is disjoint from the spectrum of T. The solution then is:

$$X = \frac{1}{2\pi\imath} \int_{\partial D} (zI - S)^{-1} A (zI - T)^{-1} dz.$$

The second fact is a heavy theorem of Bourgain and Tzafriri [2] related to restricted invertibility of matrices and to the Kadison–Singer conjecture. It is stated as Theorem 1 in the sequel.

After two of us were led to this problem while considering classification problems for commutators in spaces of operators on Banach spaces, one of us raised the problem discussed here on MathOverFlow.net [3]. Although the MO discussion did not produce a solution to the problem, it did put the authors in contact with one another and the discussion itself contains some useful tidbits.

2 The main result

Given $0 < \varepsilon < 1$, define a sequence of sets Λ_n inductively: Λ_1 is the set of 4 points $\{\pm 1 \pm \imath 1\}$ and

$$\Lambda_n = \frac{1-\varepsilon}{2}\Lambda_{n-1} + \{\pm\frac{1+\varepsilon}{2} \pm \imath\frac{1+\varepsilon}{2}\}.$$

Note that Λ_n is a subset of the square $[-1, 1] \times [-\imath, \imath]$ of cardinality 4^n and that it consists of a disjoint union of 4 sets each of which is a translate of $\frac{1-\varepsilon}{2}\Lambda_{n-1}$ and for each two of them their projection on either the real or imaginary axis is 2ε separated.

As we shall discuss below, every square matrix with zero trace is unitarily equivalent to a matrix with zero diagonal. It is thus enough to consider such matrices A. In our main result the matrix B can then be chosen to be a diagonal matrix. This is the reason for the definitions of μ and λ below. We do not know if one can get better results with more general B.

Given a $4^n \times 4^n$ matrix A with zero diagonal denote by $\mu(A)$ the smallest number μ such that there is a diagonal matrix B with diagonal elements

exactly the points of Λ_n and a $4^n \times 4^n$ matrix C such that $A = [B, C] = BC - CB$ and $\|C\| \leq \mu$. Note that since A has zero diagonal, for each diagonal matrix B with distinct diagonal entries $\{b_i\}$ such a matrix C exist and its non diagonal entries are uniquely defined by $c_{ij} = a_{ij}/(b_i - b_j)$. Put also $\mu(4^n) = \max \mu(A)$ where the max ranges over all zero diagonal $4^n \times 4^n$ matrices of norm one.

Similarly, for m not necessarily of the form 4^n, we denote by $\lambda(A)$ the smallest number λ such that there is a diagonal matrix B with diagonal elements in $[-1, 1] \times [-\imath, \imath]$ and a $m \times m$ matrix C such that $A = [B, C] = BC - CB$ and $\|C\| \leq \lambda$. Put $\lambda(m) = \max \lambda(A)$ where the max ranges over all $m \times m$ matrices of zero diagonal and norm one.

Given a $m \times m$, $m = 4^n$, matrix A write it as a 4×4 block matrix with blocks of size $4^{n-1} \times 4^{n-1}$

$$\begin{pmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ A_{21} & A_{22} & A_{23} & A_{24} \\ A_{31} & A_{32} & A_{33} & A_{34} \\ A_{41} & A_{42} & A_{43} & A_{44} \end{pmatrix}$$

Claim 1

$$\mu(A) \leq \frac{2}{1-\varepsilon} \max_{1 \leq i \leq 4} \mu(A_{ii}) + \frac{6\|A\|}{\varepsilon^2}.$$

In particular

$$\mu(4^n) \leq \frac{2}{1-\varepsilon} \mu(4^{n-1}) + \frac{6}{\varepsilon^2}.$$

Also,

$$\lambda(A) \leq \frac{2}{1-\varepsilon} \max_{1 \leq i \leq 4} \lambda(A_{ii}) + \frac{6\|A\|}{\varepsilon^2} \quad \textit{and} \quad \lambda(4^n) \leq \frac{2}{1-\varepsilon} \lambda(4^{n-1}) + \frac{6}{\varepsilon^2}. \quad (1)$$

Proof: Let B_{ii} be diagonal matrices with diagonal entries in Λ_{n-1} and C_{ii} $4^{n-1} \times 4^{n-1}$ matrices with $A_{ii} = [B_{ii}, C_{ii}]$ and $\|C_{ii}\| = \mu(A_{ii})$. Let

$$\{B'_{ii}\}_{i=1}^4 = \left\{ \frac{1-\varepsilon}{2} B_{\frac{a+1}{2}+b+2, \frac{a+1}{2}+b+2} + \left(a\frac{1+\varepsilon}{2} + \imath b \frac{1+\varepsilon}{2} \right) I_{4^{n-1}} \right\}_{a,b=\pm 1}$$

(the order doesn't matter), and, for $i \neq j$, let C'_{ij} be defined (uniquely) by

$$A_{ij} = B'_{ii} C'_{ij} - C'_{ij} B'_{jj}.$$

Then by the result mentioned in the Introduction (see[1] or [4]),

$$C'_{ij} = \frac{1}{2\pi\imath} \int_{\partial D_{ij}} (zI - B'_{ii})^{-1} A_{i,j} (zI - B'_{jj})^{-1} dz$$

where D_{ij} is the boundary curve of any domain containing the spectrum of B'_{ii} and disjoint from the spectrum of B'_{jj}. Since we can easily find such a curve of distance at least ε from the spectra of B'_{ii} and B'_{jj} and of length $4 + 4\varepsilon < 8$ we get that $\|C'_{ij}\| < \frac{2}{\varepsilon^2}\|A_{ij}\|$.

Let $C'_{ii} = \frac{2}{1-\varepsilon} C_{ii}$ and set

$$B = \begin{pmatrix} B'_{11} & 0 & 0 & 0 \\ 0 & B'_{22} & 0 & 0 \\ 0 & 0 & B'_{33} & 0 \\ 0 & 0 & 0 & B'_{44} \end{pmatrix}$$

and

$$C = (C'_{ij})_{i,j=1,2,3,4}.$$

Then

$$\|C\| \leq \frac{2}{1-\varepsilon} \max_{i,i} \mu(A_{ii}) + \frac{6}{\varepsilon^2}\|A\|.$$

This gives the claim for μ and the proof for λ is almost identical. ∎

In the proof of the main theorem we shall use the parameter λ. The reason we also included μ here is that the matrices B in the proof for the property of μ depend only on ε and not on the matrices A. Optimizing over ε we get

Corollary 1 *(i) For each m there is a $m \times m$ diagonal matrix B with spectrum in the square $[-1,1] \times [-\imath, \imath]$ such that for each $m \times m$ matrix A with diagonal zero there is a $m \times m$ matrix C with norm at most $O((\log m)^3 \sqrt{m})\|A\|$ such that $A = [B, C]$.*

(ii) For each $m = 4^n$ there is a subset Λ_m of $[-1,1] \times [-\imath, \imath]$ such that any trace zero $m \times m$ matrix A there is a normal matrix B with spectrum Λ_m and a matrix C with norm at most $O((\log m)^3 \sqrt{m})\|A\|$ such that $A = [B, C]$.

Proof: For each $0 < \varepsilon < 1$, m of the form 4^n, and an $m \times m$ matrix A with norm 1 and zero diagonal, Claim 1 gives, as long as $\frac{6}{\varepsilon^2} \leq \frac{2\varepsilon}{1-\varepsilon}\mu(m/4)$, that

$$\mu(m) \leq 2\frac{1+\varepsilon}{1-\varepsilon}\mu(m/4).$$

Let k be the largest natural number smaller than $\log_4 m$ such that $\frac{6}{\varepsilon^2} \leq \frac{2\varepsilon}{1-\varepsilon}\mu(m/4^k)$. (If no such k exists take $k = \log_4 m$ and change the argument below a bit, getting a better estimate.) Then

$$\begin{aligned}\mu(m) &\leq (2\frac{1+\varepsilon}{1-\varepsilon})^k \mu(m4^{-k}) \leq (2\frac{1+\varepsilon}{1-\varepsilon})^k (\frac{2}{1-\varepsilon}\mu(m4^{-(k+1)}) + \frac{6}{\varepsilon^2}) \\ &\leq (2\frac{1+\varepsilon}{1-\varepsilon})^k (\frac{6}{\varepsilon^3} + \frac{6}{\varepsilon^2}) \leq \frac{12}{\varepsilon^3}(2\frac{1+\varepsilon}{1-\varepsilon})^k.\end{aligned}$$

For $\varepsilon = \frac{1}{k}$ we get

$$\mu(m) \leq 12k^3 2^k (1 + \frac{3}{k})^k.$$

Since k is at most $\log_4 m$ we get (i) to get (ii) use the fact (see e.g. [5] or [6]) that any trace zero matrix is unitarily equivalent to a matrix with zero diagonal. ∎

Remark 1 The power $1/2$ of m in the first part of Corollary 1 can't be lowered. Indeed, if B is any $m \times m$ diagonal matrix with spectrum in $[-1, 1] \times [-\imath, \imath]$ then there are $i \neq j$ in $\{1, 2, \cdots, m\}$ with $|i - j| \leq \sqrt{8/m}$. If A is the $m \times m$ matrix with 1 in the i, j place and zero elsewhere and $A = [B, C]$, then it is easy to see that the absolute value of the i, j entry of C is at least $\sqrt{m/8}$.

Note that the constant $\frac{2}{1-\varepsilon}$ in (1) is what leads to the power $1/2$ of m in the Corollary above. If we could replace it with $\frac{1}{1-\varepsilon}$ we could eliminate the power of m altogether and be left with only a log factor. The next Claim is a step in this direction. The Claim, which has a proof similar to the previous one, shows that if a zero diagonal $2m \times 2m$ matrix A has its two $m \times m$ central submatrices having substantially different λ values and the smaller one is substantially larger than the norm of the matrix, then $\lambda(A)$ is, up to a multiplicative constant close to 1, basically the same as the larger of these two values. This will be used in the proof of the main theorem.

Claim 2 *Let*

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

be a $2m \times 2m$ matrix with zero diagonal where the A_{ij} are all $m \times m$ matrices. Assume also that $\lambda(A_{ii}) \leq c_i$ where $c_1/c_2 < 1/4$. Then

$$\lambda(A) \leq (1 + K((c_1/c_2)^{1/2} + \|A\|/c_1))c_2$$

For some absolute constant $K > 0$.

Proof: Write $A_{ii} = B_{ii}C_{ii} - C_{ii}B_{ii}$, $i = 1, 2$ where the B_{ii} are diagonal matrices with spectrum in $[-1, 1] \times [-\imath, \imath]$ and $\|C_{ii}\| = \lambda(A_{ii}) \leq c_i$. Assume also that $c_1 < c_2$. For any $1/2 > \delta \geq c_1/c_2$ put

$$B'_{11} = (-1 + \delta)I + \delta B_{11}, \qquad B'_{22} = 2\delta I + (1 - 2\delta)B_{22}$$

and

$$C'_{11} = \delta^{-1}C_{11}, \qquad C'_{22} = (1 - 2\delta)^{-1}C_{22}.$$

Then $A_{ii} = B'_{ii}C'_{ii} - C'_{ii}B'_{ii}$ and the B'_{ii}-s are diagonal matrices with spectrum in $[-1, 1] \times [-\imath, \imath]$. Moreover, the spectrum of B'_{11} lies to the left of the vertical line $\Re z = -1 + 2\delta$ and that of B'_{22} to the right of the vertical line $\Re z = -1 + 4\delta$. Also

$$\max_{i=1,2} \|C'_{ii}\| \leq \max\{\delta^{-1}c_1, (1 - 2\delta)^{-1}c_2\} = \frac{c_2}{1 - 2\delta}.$$

Define C'_{ij}, $i \neq j \in 1, 2$, by

$$A_{ij} = B'_{ii}C'_{ij} - C'_{ij}B'_{jj}$$

then, by the same argument as in the proof of Claim 1, using Rosenblum's result, $\|C_{ij}\| \leq K\|A\|/\delta^2$ for some universal K. Define

$$B' = \begin{pmatrix} B'_{11} & 0 \\ 0 & B'_{22} \end{pmatrix} \quad \textit{and} \quad C' = \begin{pmatrix} C'_{11} & C'_{12} \\ C'_{21} & C'_{22} \end{pmatrix}$$

then $A = B'C' - C'B'$, B is a diagonal matrix with spectrum in $[-1, 1] \times [-\imath, \imath]$ and

$$\|C'\| \leq \frac{c_2}{1 - 2\delta} + \frac{K\|A\|}{\delta^2}.$$

Taking $\delta = (c_1/c_2)^{1/2}$ we get that

$$\lambda(A) \leq (1 + K((c_1/c_2)^{1/2} + \|A\|/c_1))c_2$$

for some absolute constant K (which, a careful examination of the proof shows, can be taken to be $4/\pi$). ∎

We next recall a theorem of Bourgain and Tzafriri [2].

Theorem 1 *[2]. For some absolute constant $K > 0$, if A is a $m \times m$ matrix with zero diagonal then for all $\varepsilon > 0$ there is a central (i.e., whose diagonal is a subset of the diagonal of A) submatrix A' of dimension $\lfloor \varepsilon^2 m \times \varepsilon^2 m \rfloor$ whose norm is at most $K\varepsilon\|A\|$.*
Consequently, If A is a norm one $2 \cdot 4^n \times 2 \cdot 4^n$ matrix with zero diagonal then for all $l \leq n$ there are 4^l disjoint subsets σ_i of $1, 2, \ldots, 2 \cdot 4^n$ each of size 4^{n-l} such that all the submatrices corresponding to the entries in $\sigma_i \times \sigma_i$ have norm at most $K2^{-l}$.

Theorem 2 *(i) For each $\varepsilon > 0$ there is a constant K_ε such that for all m*

$$\lambda(m) \leq K_\varepsilon m^\varepsilon.$$

(ii) For each $\varepsilon > 0$ there is a constant K_ε such that for all m and every $m \times m$ zero trace matrix A there is a normal matrix B with spectrum in $[-1, 1] \times [-\imath, \imath]$ and a matrix C with norm at most $K_\varepsilon m^\varepsilon \|A\|$ such that $A = [B, C]$.

Proof: Let A be a $2 \cdot 4^n \times 2 \cdot 4^n$ matrix with zero diagonal and norm one. Let $1 \leq l \leq n$ and let A' be the $4^n \times 4^n$ submatrix corresponding to the entries in $\cup_{i=1}^{4^l} \sigma_i \times \cup_{i=1}^{4^l} \sigma_i$ where σ_i are given by Theorem 1. Let A_{ii}^l denote the submatrix corresponding to the entries in $\sigma_i \times \sigma_i$, $i = 1, 2, \ldots, 4^l$. Divide $1, 2, \ldots, 4^l$ into 4^{l-1} disjoint sets each a union of 4 σ_i-s and let A_{ii}^{l-1}, $i = 1, 2, \ldots, 4^{l-1}$, denote the $4^{n-l+1} \times 4^{n-l+1}$ submatrices corresponding to the entries corresponding to these sets. Continue in this manner to define A_{ii}^s, $i = 1, 2, \ldots, 4^s$ for each $s = 0, 1, 2, \ldots, l$ where for $s \geq 1$ A_{ii}^s is a $4^{n-s} \times 4^{n-s}$ submatrix of one of the A_{jj}^{s-1}. Note that $A' = A_{11}^0$.

Now, By Claim 1 for each $\varepsilon > 0$,

$$\begin{aligned}
\lambda(A') &\leq \frac{2}{1-\varepsilon} \max_{1 \leq i \leq 4} \lambda(A_{ii}^1) + \frac{6}{\varepsilon^2} \\
&\leq \left(\frac{2}{1-\varepsilon}\right)^2 \max_{1 \leq i \leq 16} \lambda(A_{ii}^2) + \left(\frac{2}{1-\varepsilon} + 1\right) \frac{6}{\varepsilon^2} \\
&\leq \quad \cdots\cdots \\
&\leq \left(\frac{2}{1-\varepsilon}\right)^{l-1} \max_{1 \leq i \leq 4^{l-1}} \lambda(A_{ii}^{l-1}) + \left(\left(\frac{2}{1-\varepsilon}\right)^{l-2} + \cdots + \frac{2}{1-\varepsilon} + 1\right) \frac{6}{\varepsilon^2} \\
&\leq \left(\frac{2}{1-\varepsilon}\right)^{l} \lambda(4^{n-l}) K 2^{-l} + \left(\left(\frac{2}{1-\varepsilon}\right)^{l-1} + \cdots + \frac{2}{1-\varepsilon} + 1\right) \frac{6}{\varepsilon^2},
\end{aligned}$$

where the last step is the place we use Theorem 1. Now use Corollary 1 to get that for some absolute constants K (not necessarily the same in each row)

$$\begin{aligned}\lambda(A') \leq & K\left(\frac{1}{1-\varepsilon}\right)^l \lambda(4^{n-l}) + l\left(\frac{2}{1-\varepsilon}\right)^{l-1}\frac{6}{\varepsilon^2} \\ \leq & K\left(\frac{1}{1-\varepsilon}\right)^l (n-l)^3 2^{n-l} + l\left(\frac{2}{1-\varepsilon}\right)^{l-1}\frac{6}{\varepsilon^2}.\end{aligned} \tag{2}$$

For $\varepsilon = 1/l$ we get

$$\lambda(A') \leq K((n-l)^3 2^{n-l} + l^3 2^l)$$

and taking $l = n/2$ gives

$$\lambda(A') \leq Kn^3 2^{n/2} = K(\log m)^3 m^{1/4}. \tag{3}$$

We managed to reduce the power of m in the bound on $\lambda(A)$ from $m^{1/2}$ to $m^{1/4}$ but only for a large submatrix. Next we are going to utilize Claim 2 to get a similar bound for the whole matrix. Let $\sigma^c = \{1, 2, \cdots, 2 \cdot 4^n\} \setminus \cup_{i=1}^{4^l}\sigma_i$ and let A'' be the submatrix of A with entries in $\sigma^c \times \sigma^c$. Put $c_1 = K(\log m)^3 m^{1/4}$ and $c_2 = \max\{K(\log m)^7 m^{1/4}, \lambda(A'')\}$. Then A, $A_{11} = A'$ and $A_{22} = A''$ satisfy the assumptions of Claim 2 with c_1, c_2. Consequently,

$$\lambda(A) \leq (1 + K(\log m)^{-2}) \max\{K(\log m)^7 m^{1/4}, \lambda(A'')\}$$

where we continue to use K to denote a universal constant, possibly different in different occurrences, and for $m = 4^n$, $n \geq 1$,

$$\lambda(2m) \leq (1 + K(\log m)^{-2}) \max\{K(\log m)^7 m^{1/4}, \lambda(m)\}.$$

Repeating the argument again reducing from matrices of size $4^{n+1} \times 4^{n+1}$ to ones of size $2 \cdot 4^n \times 2 \cdot 4^n$ and combining with the above we get, for $m = 4^n$,

$$\lambda(4m) \leq (1 + K(\log m)^{-2}) \max\{K(\log m)^7 m^{1/4}, \lambda(m)\}.$$

Let $k \leq m$ be the largest power of 4 such that $\lambda(k) \leq K(\log_4 k)^7 k^{1/4}$. Then

$$\lambda(4m) \leq \left(\prod_{s=\log_4 k+1}^{\log_4 m} (1 + Ks^{-2})\right) K(\log k)^7 k^{1/4}.$$

For some other absolute constant K this last quantity is at most $K(\log m)^7 m^{1/4}$. We thus improved the previous bound on $\lambda(m)$ (for $m = 4^n$) to

$$\lambda(m) \leq K(\log m)^7 m^{1/4}$$

for some absolute K.

Repeating the argument one can improve the bound further: Go back to (2) and plug this new bound to get

$$\lambda(A') \leq K \left(\frac{1}{1-\varepsilon}\right)^l (n-l)^7 2^{(n-l)/2} + l \left(\frac{2}{1-\varepsilon}\right)^{l-1} \frac{6}{\varepsilon^2}.$$

For $\varepsilon = 1/l$ we get

$$\lambda(A') \leq K((n-l)^7 2^{(n-l)/2} + l^3 2^l)$$

and taking $l = n/3$ gives

$$\lambda(A') \leq K n^7 2^{n/3} = K(\log m)^7 m^{1/6}.$$

replacing (3) with this new estimate and following the rest of the argument above leads to

$$\lambda(m) \leq K(\log m)^{11} m^{1/6}.$$

Iterating, this leads to a bounds of the form:

$$\lambda(m) \leq K_k(\log m)^{4k-1} m^{1/2k} \tag{4}$$

for every $m = 4^n$ and every positive integer k, where K_k depends only on k. This gives the statement of the theorem for m being a power of 4. For a general $m \times m$ zero diagonal matrix A, complete it to a $4^n \times 4^n$ matrix A' where $4^{n-1} < m \leq 4^n$ by adding zero entries and keeping A supported on $\{1, 2, \cdots, m\} \times \{1, 2, \cdots, m\}$. Apply the theorem to A' and note that the fact that B is diagonal implies that we can assume that C has non zero entries only in $\{1, 2, \cdots, m\} \times \{1, 2, \cdots, m\}$. This proves the first part of the theorem. The second follows from the fact that any trace zero matrix is unitarily equivalent to a zero diagonal matrix. ∎

3 Concluding remarks

1. Recall that the paving conjecture states that for every $\varepsilon > 0$ there is a positive integer $n(\varepsilon)$ such that any norm one zero diagonal matrix has a paving of length at most $n(\varepsilon)$ and norm at most ε. By a paving of A we mean a block diagonal submatrix of A whose diagonal is the same as that of A. The length of a paving is the number of blocks. Anderson [7] showed that this conjecture is equivalent to the Kadison–Singer conjecture [8] on the extension of pure states. For s recent expository paper on these conjectures see [9].

It is clear from the proof above that if the paving conjecture holds with the right parameters than the proof can be simplified and the main result strengthened to get a polylog estimate on $\lambda(m)$. We next show that the reverse holds in a very strong sense. In particular if $\lambda(m)$ is bounded independently of m then the paving conjecture holds.

Claim 3 *Assume $A = [B, C]$ with B a $m \times m$ diagonal matrix with spectrum in $[-1, 1] \times [-\imath, \imath]$ and C an $m \times m$ matrix. Then for every $0 < \varepsilon < 1$ A has a paving of length $\lfloor \frac{2}{\varepsilon} \rfloor^2$ and norm $\sqrt{2}\varepsilon\|C\|$.*

Proof: Partition $[-1, 1]$ into $\lfloor \frac{2}{\varepsilon} \rfloor$ disjoint intervals I_i of length at most ε each. Let $B(i, j)$ be the central (diagonal) submatrix of B whose diagonal entries are in $I_i \times \imath I_j$, let $A(i, j)$ and $C(i, j)$ be the central submatrices of A and C respectively with the same support as $B(i, j)$. $A(i, j)$, $i, j = 1, 2, \cdots, \lfloor \frac{2}{\varepsilon} \rfloor$. $A(i, j)$, $i, j = 1, 2, \cdots, \lfloor \frac{2}{\varepsilon} \rfloor$, is a paving of A and it is enough to prove that $\|A(i, j)\| \leq \sqrt{2}\varepsilon\|C\|$.

Clearly $A(i, j) = [B(i, j), C(i, j)]$. Pick i, j, let b be the center of the square $I_i \times \imath I_j$ and note that $bI - B(i, j)$ (with I the identity matrix of the same dimensions as $B(i, j)$) is a diagonal matrix with entries of absolute value at most $\varepsilon/\sqrt{2}$. Therefore

$$\|A(i, j)\| = \|(B(i, j) - bI)C(i, j) - C(i, j)(B(i, j) - bI)\| \leq \sqrt{2}\varepsilon\|C\|.$$

■

2. A more careful examination of the proof of Theorem 2 shows that the constant we get in (4) is

$$\lambda(m) \leq K^k (\log m)^{4k-1} m^{1/2k}$$

for some absolute constant K. Optimizing over k gives

$$\lambda(m) \leq m^{K(\log\log m/\log m)^{1/2}}$$

for some absolute K.

3. It is quite easy to see that $1/2$ is also the best constant for K in the second paragraph of the introduction (assume A has zero diagonal and take B to be diagonal with diagonal elements $1/2$ and $-1/2$). It also follows that $\lambda(2) = 1/2\sqrt{2}$. We did not try to compute the best constants for other small values of the dimension.

4. Although the problem we discuss seems basic enough not to need further motivation, we would like to indicate one. If any trace zero matrix A could be written as $A = [B, C]$ with $\|B\|\|C\| \leq K\|A\|$ for a universal K, then we would get a simple characterization of the commutators in an important class of II_1 factors, the Wright factors; an element there would be a commutator if and only if it has zero trace. See [10] for this and related matters.

5. One can ask similar questions to the one addressed here for norms other than the operator norm. In particular, what is the (order of) the best constant in

$$\|B\|\|C\|_{HS} \leq K\|A\|_{HS}$$

where A ranges over all trace zero $m \times m$ matrices, $A = BC - CB$ and $\|\cdot\|_{HS}$ denotes the Hilbert–Schmidt norm?

We checked that a proof with a similar idea but much simpler gives $K = O((\log m)^{5/2})$.

References

[1] Rosenblum M (1956) On the operator equation $BX - XA = Q$. Duke Math. J. 23: 263–269.

[2] Bourgain J, Tzafriri L (1991) On a problem of Kadison and Singer. J. Reine Angew. Math. 420: 1–43.

[3] http://mathoverflow.net/questions/27345.

[4] Marcoux L-W (2006) Sums of small number of commutators. J. Operator Theory 56 no. 1: 111–142.

[5] Fillmore P-A (1969) On similarity and the diagonal of a matrix. Amer. Math. Monthly 76: 167–169.

[6] Halmos P-R (1982) A Hilbert space problem book. Second edition. Graduate Texts in Mathematics, 19. Encyclopedia of Mathematics and its Applications, 17. Springer-Verlag, New York-Berlin.

[7] Anderson J (1979) Extensions, restrictions and representations of states on C^*-algebras. Trans. AMS 249: 303-329.

[8] Kadison R, Singer I (1959) Extensions of pure states. American Jour. Math. 81: 383-400.

[9] Casazza P-G, Edidin D (2007) Equivalents of the Kadison-Singer problem. Function spaces, 123-142, Contemp. Math. 435, Amer. Math. Soc., Providence, RI.

[10] Dykema K, Skripka A, On single commutators in II_1-factors. http://arxiv.org/abs/1007.1701

W.B. Johnson
Department of Mathematics
Texas A&M University
College Station, TX 77843 U.S.A.
johnson@math.tamu.edu

N. Ozawa
Research Institute for Mathematical Sciences
Kyoto University
Kyoto 606-8502, Japan
narutaka@kurims.kyoto-u.ac.jp

G. Schechtman
Department of Mathematics
Weizmann Institute of Science
Rehovot, Israel
gideon@weizmann.ac.il