A CONSTRUCTIVE PROOF OF THE DERIVATION THEOREM

SOUMYASHANT NAYAK

Abstract. We give a constructive proof of the well-known result that every derivation on $M_n(\mathbb{C})$ is inner, taking an approach based on fixed point theory. The proof yields an explicit formula for the matrix implementing the inner derivation (modulo translation by scalar matrices) by employing a standard averaging technique over the unitary group of $M_n(\mathbb{C})$. As a corollary, it follows that the matrix may be chosen with norm less than or equal to the norm of the derivation.

Keywords: Derivations, Fixed point theory, Haar measure
MSC2010 subject classification: 16W25, 17B40, 47B47

1. Introduction

For a positive integer n, let $M_n(\mathbb{C})$ denote the set of $n \times n$ complex matrices. A derivation $\delta : M_n(\mathbb{C}) \to M_n(\mathbb{C})$ is a $\mathbb{C}$-linear transformation on $M_n(\mathbb{C})$ satisfying the Leibnitz rule, that is, $\delta(AB) = A\delta(B)+\delta(A)B$ for all $A, B \in M_n(\mathbb{C})$. The study of derivations is motivated by the fact that they are infinitesimal generators of (continuous) one-parameter groups of automorphisms, which arise in the study of time evolution of physical systems. Since the motivation on our part is one of mathematical curiosity and pedagogy, we limit our scope to the case of matrix algebras. For a fixed matrix $Z \in M_n(\mathbb{C})$, it is easy to verify that the linear transformation on $M_n(\mathbb{C})$ given by $A \mapsto ZA - AZ =: [Z, A]$ is a derivation. Such derivations are called *inner derivations* and we say that the matrix Z implements the inner derivation. An automorphism $\Phi : M_n(\mathbb{C}) \to M_n(\mathbb{C})$ is said to be an inner automorphism if there is an invertible matrix P in $M_n(\mathbb{C})$ such that $\Phi(A) = PAP^{-1}$ for all $A \in M_n(\mathbb{C})$, that is, the automorphism is 'spatially' implemented. An inner derivation generates a one-parameter group of inner automorphisms.

In this note, we present a constructive proof of the fact that all derivations on $M_n(\mathbb{C})$ are inner (henceforth called the Derivation Theorem). For a derivation δ, we find an explicit formula for a matrix implementing δ as an inner derivation. As an application, we use the formula to show that it is possible to choose such a matrix with norm less than or equal to the norm of the derivation. We note that the strategy also works for a proof of the Derivation Theorem for $M_n(\mathbb{R})$.

2. Preliminaries

Definition 2.1. For a locally compact topological group G, let Σ denote a σ-algebra of subsets of G containing the Borel subsets. A Radon measure $\mu : \Sigma \to \mathbb{R}$ is said to be a left-invariant *Haar measure* on G if $\mu(gS) = \mu(S)$ for all $g \in G$ and μ-measurable subsets S of G.

It was first shown by Haar that a left-invariant Radon measure exists on any locally compact topological group and is unique upto multiplication by a positive constant. This generalizes the notion of Lebesgue measure on the Euclidean spaces $\mathbb{R}^n$. Our interest in the Haar measure stems from the averaging procedure it allows via integration over the group which, loosely speaking, results in functions or other mathematical objects that are invariant under action of the group.

Since the Haar measure is finite for a compact group, we may normalize it to a probability measure. On an n-dimensional Lie group, the measure induced by a left-invariant n-form is a

Haar measure. For our discussion, the pertinent groups are the unitary group $\mathscr{U}(n)$ of $M_n(\mathbb{C})$, and the orthogonal group $\mathscr{O}(n)$ of $M_n(\mathbb{R})$, both of which are compact Lie groups.

Lemma 2.2. *Every matrix in $M_n(\mathbb{C})$ can be written as a $\mathbb{C}$-linear combination of four unitary matrices in $M_n(\mathbb{C})$.*

Proof. For a matrix A, without loss of generality we may assume that $\|A\| \le 1$ as $A = \|A\|(\frac{A}{\|A\|})$. Let $\Re(A) := \frac{A+A^*}{2}, \Im(A) := \frac{A-A^*}{2i}$, denote the real and imaginary parts, respectively, of A, so that $A = \Re(A) + i\,\Im(A)$. The matrices $\Re(A)$ and $\Im(A)$ are Hermitian, and have matrix norm less than or equal to 1. Using the spectral theorem and the continuous functional calculus for a Hermitian matrix H with $\|H\| \le 1$ (and thus, $-I \le H \le I$), we note that $\phi_1(H) := H + i\sqrt{I - H^2}$, and $\phi_2(H) := H - i\sqrt{I - H^2}$, are unitary matrices. Thus the decomposition,

$$A = \frac{1}{2}\phi_1(\Re(A)) + \frac{1}{2}\phi_2(\Re(A)) + \frac{i}{2}\phi_1(\Im(A)) + \frac{i}{2}\phi_2(\Im(A)),$$

shows that A can be written as a $\mathbb{C}$-linear combination of four unitary matrices. $\square$

Lemma 2.3. *Every matrix in $M_n(\mathbb{R})$ can be written as an $\mathbb{R}$-linear combination of finitely many orthogonal matrices in $M_n(\mathbb{R})$.*

Proof. For $n = 1$, the assertion holds trivially and hence we assume that $n \ge 2$. Note that we may represent $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ as an $\mathbb{R}$-linear combination of five orthogonal matrices in $M_2(\mathbb{R})$ as follows,

$$\frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{1}{2\sqrt{2}}\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} + \frac{1}{2\sqrt{2}}\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} - \frac{1}{2\sqrt{2}}\begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} - \frac{1}{2\sqrt{2}}\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

Thus the matrix unit $\mathcal{E}_{11}$ in $M_n(\mathbb{R})$, with 1 in the $(1,1)^{\text{th}}$ position and 0's elsewhere, can be represented as an $\mathbb{R}$-linear combination of five orthogonal matrices in $M_n(\mathbb{R})$ using the canonical embedding $\mathscr{O}(2) \hookrightarrow \mathscr{O}(n)$ given by

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \mapsto \left[\begin{array}{cc|c} a_{11} & a_{12} & \multirow{2}{*}{$\mathbf{0}$} \\ a_{21} & a_{22} & \\ \hline \multicolumn{2}{c|}{\mathbf{0}} & \mathbf{I}_{n-2} \end{array}\right].$$

Recall that permutation matrices are orthogonal matrices. By multiplying suitable permutation matrices to $\mathcal{E}_{11}$ (on the left or the right), we conclude that any matrix unit $\mathcal{E}_{ij}, 1 \le i, j \le n$, may be represented as a linear combination of five orthogonal matrices in $M_n(\mathbb{R})$. Since the n^2 matrix units $\mathbb{R}$-linearly generate $M_n(\mathbb{R})$, we note that every matrix in $M_n(\mathbb{R})$ can be written as an $\mathbb{R}$-linear combination of $5n^2$ orthogonal matrices in $M_n(\mathbb{R})$.

$\square$

3. Proof of the Derivation Theorem

Theorem 3.1. *Let δ be a derivation on $M_n(\mathbb{C})$. Then there exists a matrix $Z \in M_n(\mathbb{C})$ such that $\delta(A) = ZA - AZ =: [Z, A]$ for all $A \in M_n(\mathbb{C})$. Furthermore, Z is unique upto translation by a scalar matrix. Let $\mathscr{U}(n)$ denote the compact group of unitary matrices in $M_n(\mathbb{C})$, with the left-invariant Haar probability measure on $\mathscr{U}(n)$ denoted by dU. We have the following formula,*

$$Z = \alpha I + \int_{\mathscr{U}(n)} \delta(U)U^{-1}\, dU,$$

for some $\alpha \in \mathbb{C}$.

Proof. Define a family of affine linear maps on $M_n(\mathbb{C})$, indexed by $\mathscr{U}(n)$, in the following manner:

$$\psi_U(X) := UXU^{-1} + \delta(U)U^{-1}, \text{ for } U \in \mathscr{U}(n).$$

For $U, V \in \mathscr{U}(n)$ and $X \in M_n(\mathbb{C})$, we observe that

$$\begin{aligned}
\psi_U \circ \psi_V(X) &= U\psi_V(X)U^{-1} + \delta(U)U^{-1} \\
&= U\big(VXV^{-1} + \delta(V)V^{-1}\big)U^{-1} + \delta(U)U^{-1} \\
&= (UV)X(UV)^{-1} + \big(U\delta(V) + \delta(U)V\big)(UV)^{-1} \\
&= (UV)X(UV)^{-1} + \big(\delta(UV)\big)(UV)^{-1} \\
&= \psi_{UV}(X)
\end{aligned}$$

Thus for any $U, V \in \mathscr{U}(n)$, we have $\psi_U \circ \psi_V = \psi_{UV}$.

For a fixed matrix $X_0 \in M_n(\mathbb{C})$, note that $\psi_U(X_0)$ is a continuous function in U on the compact group $\mathscr{U}(n)$. Let us define

$$Z := \int_{\mathscr{U}(n)} \psi_U(X_0)\, dU.$$

Keeping in mind the affine linearity of ψ_V and the left-invariance of the Haar probability measure on $\mathscr{U}(n)$, for every unitary matrix $V \in \mathscr{U}(n)$ we have,

$$\begin{aligned}
\psi_V(Z) = \psi_V\Big(\int_{\mathscr{U}(n)} \psi_U(X_0)\, dU\Big) &= \int_{\mathscr{U}(n)} \psi_V \circ \psi_U(X_0)\, dU \\
= \int_{\mathscr{U}(n)} \psi_{VU}(X_0)\, dU &= \int_{\mathscr{U}(n)} \psi_U(X_0)\, dU \\
= Z.
\end{aligned}$$

Thus for all $U \in \mathscr{U}(n)$, we observe that $\psi_U(Z) = UZU^{-1} + \delta(U)U^{-1} = Z$, which implies that $\delta(U) = ZU - UZ$. Using Lemma 2.2 and the linearity of δ, we conclude that $\delta(A) = ZA - AZ = [Z, A]$ for all $A \in M_n(\mathbb{C})$. For an explicit formula for Z in terms of δ, we choose $X_0 = 0$ and note that

$$Z = \int_{\mathscr{U}(n)} \delta(U)U^{-1}\, dU.$$

If matrices Z_1, Z_2 both implement the inner derivation δ, then $[Z_1, A] = [Z_2, A]$ for all $A \in M_n(\mathbb{C})$, which implies that $Z_1 - Z_2$ is in the center of $M_n(\mathbb{C})$. Thus $Z_1 = \alpha I + Z_2$ for some $\alpha \in \mathbb{C}$. $\square$

Remark 3.2. Note that for a unitary matrix $V \in \mathscr{U}(n)$ and $X_0 \in M_n(\mathbb{C})$, we have

$$\begin{aligned}
V\Big(\int_{\mathscr{U}(n)} UX_0U^{-1}\, dU\Big)V^{-1} &= \int_{\mathscr{U}(n)} (VU)X_0(VU)^{-1}\, dU \\
&= \int_{\mathscr{U}(n)} UX_0U^{-1}\, dU.
\end{aligned}$$

Thus $\int_{\mathscr{U}(n)} UX_0U^{-1}\, dU$ commutes with every unitary matrix and by Lemma 2.2, it must lie in the center of $M_n(\mathbb{C})$. This explains why different choices of X_0 perturb Z only by scalar matrices.

Let $\|\cdot\|$ be a norm on $M_n(\mathbb{C})$ which makes it a normed algebra ($\|AB\| \le \|A\|\|B\|$ for $A, B \in M_n(\mathbb{C})$, $\|I\| = 1$) such that $\|U\| \le 1$ for every unitary matrix $U \in M_n(\mathbb{C})$. Examples of such a norm include the Frobenius norm,

$$\|A\| = \sqrt{\frac{1}{n}\mathrm{tr}(A^*A)},$$

and the usual matrix norm,

$$\|A\| = \sup_{x \in \mathbb{C}^n \setminus \{\mathbf{0}\}} \frac{\|Ax\|_2}{\|x\|_2}.$$

For a derivation δ on $M_n(\mathbb{C})$, let $\|\delta\|$ denote the usual operator norm of δ relative to the norm $\|\cdot\|$ (with properties as described above) on $M_n(\mathbb{C})$. If $\delta(A) = ZA - AZ$ for all $A \in M_n(\mathbb{C})$, then $\|\delta(A)\| \leq \|ZA\| + \|AZ\| \leq 2\|Z\|\|A\|$, and hence

$$\frac{\|\delta\|}{2} \leq \|Z\|.$$

Corollary 3.3. *Let δ be a derivation on $M_n(\mathbb{C})$ $(n \in \mathbb{N})$. Then we can choose $Z \in M_n(\mathbb{C})$ such that $\delta(A) = [Z, A]$ for all $A \in M_n(\mathbb{C})$ and $\|Z\| \leq \|\delta\|$.*

Proof. From Theorem 3.1, choosing

$$Z = \int_{\mathscr{U}(n)} \delta(U)U^{-1}\, dU,$$

we see that

$$\|Z\| \leq \int_{\mathscr{U}(n)} \|\delta(U)U^{-1}\|\, dU \leq \|\delta\| \Big(\int_{\mathscr{U}(n)} 1\, dU \Big) = \|\delta\|.$$

□

Theorem 3.4. *Let δ be a derivation on $M_n(\mathbb{R})$. Then there exists a matrix $Z \in M_n(\mathbb{R})$ such that $\delta(A) = ZA - AZ =: [Z, A]$ for all $A \in M_n(\mathbb{R})$. Furthermore, Z is unique upto translation by a scalar matrix. Let $\mathscr{O}(n)$ denote the compact group of orthogonal matrices in $M_n(\mathbb{R})$, with the left-invariant Haar probability measure on $\mathscr{O}(n)$ denoted by dO. We have the following formula,*

$$Z = \alpha I + \int_{\mathscr{O}(n)} \delta(O)O^{-1}\, dO,$$

for some $\alpha \in \mathbb{R}$.

Proof. The proof is almost identical to the one for Theorem 3.1 using the affine $\mathbb{R}$-linear maps $\psi_O(X) = OXO^{-1} + \delta(O)O^{-1}$, indexed by $\mathscr{O}(n)$, and using Lemma 2.3. □

4. Concluding Remarks

There are generalizations of the main result discussed in several directions. For instance, the Kadison-Sakai theorem states that all derivations of a von Neumann algebra are inner (cf. [3], [4]). Our proof is inspired by Sakai's approach which uses the Markov-Kakutani fixed point theorem in an essential way (see [5, Lemma 2.5.1]). The astute reader may have noticed that the proof of Theorem 3.1 works equally well for $M_{n_1}(\mathbb{C}) \oplus \cdots \oplus M_{n_k}(\mathbb{C})$ (for positive integers $n_1, \cdots, n_k$) by considering its unitary group, and for $M_{n_1}(\mathbb{R}) \oplus \cdots \oplus M_{n_k}(\mathbb{R})$ by considering its orthogonal group.

The theory of derivations has been studied extensively in abstract algebraic settings. The approach by Hochschild (cf. [1], [2]) involves building a cohomology theory for associative algebras in which the first cohomology group corresponds to the space of outer derivations. In the context of the Derivation Theorem, one studies associative algebras that have trivial first Hochschild cohomology group. Since all (K-linear) derivations on a field K are trivial and Morita equivalent rings have isomorphic Hochschild cohomology, all derivations on $M_n(K)$ are inner. Note that this is a 'global' approach studying the space of derivations in contrast with our 'local' approach in which we directly work with the derivation under consideration.

For an account of the Hochschild cohomology of algebras, the reader may refer to [8, Chapter 9]. A survey of the (continuous) Hochschild cohomology theory of von Neumann algebras can be found in [6], [7]. The reader may also benefit from the exposition on derivations in the context of operator algebras in [5, Chapter 3-4].

References

[1] G. Hochschild. On the cohomology groups of an associative algebra. *Ann. of Math. (2)*, 46:58–67, 1945.

[2] G. Hochschild. On the cohomology theory for associative algebras. *Ann. of Math. (2)*, 47:568–579, 1946.

[3] Richard V. Kadison. Derivations of operator algebras. *Ann. of Math. (2)*, 83:280–293, 1966.

[4] Shôichirô Sakai. Derivations of W^*-algebras. *Ann. of Math. (2)*, 83:273–279, 1966.

[5] Shôichirô Sakai. *Operator algebras in dynamical systems*, volume 41 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1991. The theory of unbounded derivations in C^*-algebras.

[6] Allan M. Sinclair and Roger R. Smith. *Hochschild cohomology of von Neumann algebras*, volume 203 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1995.

[7] Allan M. Sinclair and Roger R. Smith. A survey of Hochschild cohomology for von Neumann algebras. In *Operator algebras, quantization, and noncommutative geometry*, volume 365 of *Contemp. Math.*, pages 383–400. Amer. Math. Soc., Providence, RI, 2004.

[8] Charles A. Weibel. *An introduction to homological algebra*, volume 38 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1994.

Smilow Center for Translational Research, University of Pennsylvania, Philadelphia, PA 19104

E-mail address: `nsoum@upenn.edu`

URL: `https://nsoum.github.io/`