Diagonal Norm Hermitian Matrices

HANS SCHNEIDER AND R. E. L. TURNER*
University of Wisconsin-Madison
Madison, Wisconsin

ABSTRACT

If ν is a norm on $\mathbb{C}^n$, let $H(\nu)$ denote the set of all norm-Hermitians in $\mathbb{C}^{nn}$. Let S be a subset of the set of real diagonal matrices D. Then there exists a norm ν such that $S = H(\nu)$ (or $S = H(\nu) \cap D$) if and only if S contains the identity and S is a subspace of D with a basis consisting of rational vectors. As a corollary, it is shown that, for a diagonable matrix h with distinct eigenvalues $\lambda_1, \ldots, \lambda_r$, $r \leqslant n$, there is a norm ν such that $h \in H(\nu)$, but $h^s \notin H(\nu)$, for some integer s, if and only if $\lambda_2 - \lambda_1, \ldots, \lambda_r - \lambda_1$ are linearly dependent over the rationals. It is also shown that the set of all norms ν, for which $H(\nu)$ consists of all real multiples of the identity, is an open, dense subset, in a natural metric, of the set of all norms.

INTRODUCTION

For a norm ν on $\mathbb{C}^n$, the complex n-tuples, an $(n \times n)$ matrix h is called norm-Hermitian if the numerical range of h with respect to ν is real. (For a precise definition see Section 1 and the beginning of Section 3.) An unsolved problem in this area is

(1) Given a norm ν on $\mathbb{C}^n$, characterize the set $H(\nu)$ of norm-Hermitian matrices.

An alternative, easier problem is

* The research of the first author was supported by NSF Grant GP 17815, and by the Mathematics Research Center under Contract No. DA-31-124-ARO-D-402. The research of the second author was supported by NSF Grant GP 32833.

(2) Characterize all subsets S of $\mathbb{C}^{nn}$, the set of $(n \times n)$ complex matrices, such that there exists a norm ν on $\mathbb{C}^n$ for which $S = H(\nu)$.

In an earlier paper [14] (cf. Theorem (6.2)), we solved (1) under the additional hypothesis that ν is absolute. (The norm ν is absolute if $\nu(x)$ depends only on the absolute values of the coordinates of x in $\mathbb{C}^n$.) Implicit in [14] is the solution of problem

(2a) Characterize all S in $\mathbb{C}^{nn}$ such that there exists an absolute norm ν on $\mathbb{C}^n$ for which $S = H(\nu)$.

The result is: There exists an absolute norm ν with $S = H(\nu)$ if and only if there exists an equivalence relation $\equiv$ on the set of integers $\{1, 2, \ldots, n\}$ for which

$$S = \{h \in \mathbb{C}^{nn} | h_{ij} = \bar{h}_{ji}, \text{ for } i \equiv j \text{ and } h_{ij} = 0 \text{ otherwise}\}.$$

In the present paper we take a step toward a solution of problem (2). Here, in Sections 3 and 4 we deal with sets S contained in the set D of real diagonal matrices. We solve two problems. The first is

(2d): Characterize all subsets S in D such that there exists a norm ν for which $S = H(\nu) \cap D$.

The solution is stated in Theorem 3.3. The second problem is

(2c): Characterize all S in D such that there exists a norm ν for which $S = H(\nu)$.

For a solution of (2c) see Theorem 4.54. In describing S we identify the space D of real diagonal matrices with $\mathbb{R}^n$, the real n-tuples. For each of (2d) and (2c) S is characterized by being a subspace of $\mathbb{R}^n$ which contains the identity and which has a basis of rational vectors. By a rational vector we mean a vector $(\alpha_1, \ldots, \alpha_n)$ where each α_i is a rational number. In resolving (2d), one could omit the proof of Theorem 3.2, and use instead Theorem 4.7. However, we include Theorem 3.2, both to show that the more complicated Theorem 4.7 is not required to resolve (2d) and also to shed light on the proof of Theorem 4.7. The construction of the norm in Theorem 3.2 was motivated by an example due to M. J. Crabb [6]; [4, p. 57]. The more elaborate norm (4.3) used in the proof of Theorem

4.7 is apparently required to deal with the off-diagonal elements of a norm-Hermitian matrix without disturbing the desired class of diagonal elements.

A crucial ingredient in the characterization of diagonal norm-Hermitian matrices is a theorem on inhomogeneous diophantine approximation. The archetypal theorem is due to Kronecker, but for our purposes we require the form given in Theorem IV of the appendix. There are several versions of the approximation theorem in the literature (Perron [11], Koksma [9], Cassels [5]) but we have been unable to find the version we require. We feel it is informative to show, as we have done in the appendix, how the different forms of the approximation theorem can be derived from Theorem 63, p. 153 of [11], 1st Edition (or Theorem 64, p. 159 of [11], 4th Edition). In fact, Theorem I of the appendix is a restatement of Perron's Theorem.

As a consequence of our characterization of diagonal norm-Hermitian matrices, we are able to shed light on Problem 4, p. 128, of Bonsall and Duncan [4], which concerns norm-Hermitian elements whose powers are not all norm-Hermitian. Near the end of Section 3, in Corollaries 3.5 and 3.6, we give conditions on the eigenvalues of a diagonable matrix h which are necessary and sufficient for the existence of a norm with respect to which some, but not all, powers of h are norm-Hermitian. In Corollary 3.7, we show that a norm is absolute if and only if there is diagonal norm-Hermitian matrix $\operatorname{diag}(d_1, \ldots, d_n)$ where $d_2 - d_1, \ldots, d_n - d_1$ are linearly independent over the rationals.

In Section 5, we examine the set of norms which allow only the identity and its real multiples as norm-Hermitian elements. It is shown that almost all norms are of this type. More precisely, we introduce a metric in the space of norms on $\mathbb{C}^n$ and show that the set of norms allowing only real multiples of the identity as norm-Hermitians constitutes an open, dense set.

1. NORMS AND DUALITY

We will be concerned with the vector space of n-tuples of complex numbers, $\mathbb{C}^n$, over the field $\mathbb{C}$.

DEFINITIONS 1.1. *A* semi-norm *on $\mathbb{C}$ is a function σ from $\mathbb{C}^n$ to the non-negative real numbers $\mathbb{R}^+$ satisfying*

$$(1) \qquad \sigma(x+y) \leqslant \sigma(x) + \sigma(y); \qquad x, y \in \mathbb{C}^n,$$

$$\sigma(\xi x) = |\xi|\sigma(x); \qquad \xi \in \mathbb{C}, \qquad x \in \mathbb{C}^n, \tag{2}$$

where $|\xi|$ *denotes the absolute value of* ξ. *A* norm *is a semi-norm satisfying the additional condition*

$$\sigma(x) = 0 \Rightarrow x = 0. \tag{3}$$

We denote the usual Euclidean norm on $\mathbb{C}^n$ by χ and use the fact that any semi-norm on $\mathbb{C}^n$ is continuous with respect to the Euclidean norm topology.

LEMMA 1.2. *Let* $\sigma_\alpha, \alpha \in \mathscr{A}$, *be an indexed family of semi-norms and suppose there exists a norm* ν *on* $\mathbb{C}^n$ *such that* $\sigma_\alpha(x) \leqslant \nu(x)$ *for all* $\alpha \in \mathscr{A}$ *and all* $x \in \mathbb{C}^n$. *Then* σ *defined by*

$$\sigma(x) = \sup_{\alpha \in \mathscr{A}} \sigma_\alpha(x)$$

is a semi-norm. If σ_{α_0} *is a norm for some* $\alpha_0 \in \mathscr{A}$, *then* σ *is a norm.*

Proof. Straightforward.

The dual space of $\mathbb{C}^n$, that is, the space of linear functionals on $\mathbb{C}^n$, can be identified with $\mathbb{C}^n$ and if y is a linear functional, its value at x is denoted by $\langle y, x\rangle$. We assume $\langle y, x\rangle$ is conjugate linear in y. If ν is a norm on $\mathbb{C}^n$, then the dual norm ν^D, on linear functionals, is defined by

$$\nu^D(y) = \sup_{x \neq 0} \frac{|\langle y, x\rangle|}{\nu(x)}.$$

If x, y are in $\mathbb{C}^n$ and

$$1 = \langle y, x\rangle = \nu^D(y)\nu(x)$$

we say that y is dual to x and indicate this relationship by writing $y||x$. It is well known that for each $x \in \mathbb{C}^n$, $x \neq 0$, there is at least one y such that $y||x$ and for each $y \neq 0$, there is at least one x such that $y||x$ (e.g., [1]).

LEMMA 1.3. *Let* η *be a norm on* $\mathbb{C}^n$ *and let* σ *be a semi-norm. Let*

$$\nu(x) = \sup(\eta(x), \sigma(x)).$$

Suppose $1 = \sigma(x_0) > \eta(x_0)$ *and that a vector* x' *satisfies* $\langle x', x_0\rangle = 1$ *and* $|\langle x', x\rangle| \leqslant \sigma(x)$ *for all* x. *Then* $x'||x_0$ *with respect to* ν.

Proof. Since $\langle x', x_0\rangle = \nu(x_0)$, it suffices to show that $|\langle x', x\rangle| \leqslant \nu(x)$ for all x. However, $|\langle x', x\rangle| \leqslant \sigma(x) \leqslant \nu(x)$ for $x \in \mathbb{C}^n$.

DEFINITION 1.4. *Suppose ν is a semi-norm and is defined by*

$$\nu(x) = \sup_{\alpha\in\mathscr{A}} \sigma_\alpha(x),$$

where the σ_α are semi-norms. Suppose that for a given $x_0 \in \mathbb{C}^n$ and $\beta \in \mathscr{A}$,

$$\sigma_\beta(x_0) > \sup_{\substack{\alpha\in\mathscr{A}\\ \alpha\neq\beta}} \sigma_\alpha(x_0).$$

Then we say that σ_β is active *(with respect to ν) at x_0.*

LEMMA 1.5. *Let ν be the semi-norm in Definition 1.4. If σ_β is active at x_0, then it is active in a Euclidean neighborhood of x_0.*

Proof. Since $\nu(x) \leqslant k\chi(x)$ for some constant $k > 0$, $\sigma_\alpha(x) \leqslant k\chi(x)$ for each α and, by Lemma 1.2,

$$\nu_1(x) = \sup_{\substack{\alpha\in\mathscr{A}\\ \alpha\neq\beta}} \sigma_\alpha(x)$$

is a semi-norm. Then $\nu(x) = \sup(\sigma_\beta(x), \nu_1(x))$ and $\sigma_\beta(x_0) - \nu_1(x_0) = \varepsilon > 0$. Since both σ_β and ν, are continuous with respect to χ there exists a χ-neighborhood V of x_0 such that $\sigma_\beta(x) - \nu_1(x) > \varepsilon/2$ for $x \in V$. Hence, σ_β is active in V.

Let $e_1, e_2, \ldots, e_n$ be a basis for $\mathbb{C}^n$. Then there exists a unique basis $e_1', \ldots, e_n'$ for the dual space, satisfying $\langle e_i', e_j\rangle = \delta_{ij}$, the Kronecker delta. The set $\{e_i'\}$ is the algebraic dual basis to $\{e_i\}$.

DEFINITION 1.6. *Let ν be a norm on $\mathbb{C}^n$. If the basis $\{e_i\}$ and its algebraic dual $\{e_j'\}$ satisfy $e_i' \| e_i$ with respect to ν and $\nu(e_i) = 1$ for $i = 1, 2, \ldots, n$, we call $\{e_i\}$ a* double-dual *basis.*

We will use the result:

THEOREM 1.7 (e.g. Schneider [13]). *For any norm ν on $\mathbb{C}^n$ there exists a double-dual basis with respect to ν.*

If $e_1, \ldots, e_n$ is any basis for $\mathbb{C}^n$ then each x is uniquely represented as $x = \sum_{i=1}^n \alpha_i e_i$. The quantity $|x|_p = (\sum_{i=1}^n |\alpha_i|^p)^{1/p}$ defines a norm on $\mathbb{C}^n$ for $1 \leqslant p < \infty$ and as usual we set $|x|_\infty = \sup_{1\leqslant i\leqslant n}|\alpha_i|$.

LEMMA 1.8 (cf. Schneider [13]). *Let* $e_1, \ldots, e_n$ *be a double-dual basis with respect to a norm* ν. *Then with respect to this basis*

$$|x|_\infty \leqslant \nu(x).$$

Proof. Suppose $x = \sum_{i=1}^n \alpha_i e_i$ and $|x|_\infty = |\alpha_k|$. Then

$$|x|_\infty = |\langle e_k', x\rangle| \leqslant \nu^D(e_k')\nu(x) = \nu(x).$$

DEFINITION 1.9. *Let* E *be a set. A* semi-metric *on* E *is a function* d *from* $E \times E$ *into* R^+ *satisfying*

$$d(x, x) = 0, \tag{1}$$

$$d(x, y) = d(y, x), \tag{2}$$

$$d(x, z) \leqslant d(x, y) + d(y, z), \tag{3}$$

for all $x, y, z \in E$. *If also*

$$d(x, y) = 0 \quad \textit{implies} \quad x = y \tag{4}$$

then (as usual) d *is a* metric *on* E.

The distance function used in the following lemma is similar to ones that have been used in other contexts (e.g. G. Birkhoff [2]).

LEMMA 1.10. *Let* $N = N(n)$ *be the set of all norms on* $\mathbb{C}^n$. *Then*
(1) *The function defined by*

$$d(\rho, \nu) = \log\left(\sup_{x\neq 0} \frac{\rho(x)}{\nu(x)} \cdot \sup_{y\neq 0} \frac{\nu(y)}{\rho(y)}\right) \tag{1.11}$$

is a semi-metric on N.
(2) $d(\rho, \nu) = 0$ *if and only if there is a* $c > 0$ *such that* $\rho = c\nu$.

Proof. Clearly $d(\rho, \rho) = 0$ and $d(\rho, \nu) = d(\nu, \rho)$ for all $\rho, \nu \in N$. To prove the triangle inequality, suppose also that $\sigma \in N$. Then

$$\begin{aligned} d(\rho, \nu) &= \log\left(\sup_{x\neq 0} \frac{\rho(x)}{\nu(x)} \cdot \sup_{y\neq 0} \frac{\nu(y)}{\rho(y)}\right) \\ &= \log\left(\sup_{x\neq 0} \frac{\rho(x)}{\sigma(x)} \cdot \frac{\sigma(x)}{\nu(x)} \cdot \sup_{y\neq 0} \frac{\nu(y)}{\sigma(y)} \cdot \frac{\sigma(y)}{\rho(y)}\right) \end{aligned}$$

$$\leqslant \log\left(\sup_{x\neq 0}\frac{\rho(x)}{\sigma(x)}\cdot\sup_{w\neq 0}\frac{\sigma(w)}{\nu(w)}\cdot\sup_{y\neq 0}\frac{\nu(y)}{\sigma(y)}\cdot\sup_{z\neq 0}\frac{\sigma(z)}{\rho(z)}\right)$$

$$= d(\rho,\sigma) + d(\sigma,\nu).$$

If $\rho = c\nu$ where $c > 0$, then obviously $d(\rho,\nu) = \log(c\cdot 1/c) = 0$. Conversely, suppose $d(\rho,\nu) = 0$ and let

$$c = \inf_{x\neq 0}\frac{\rho(x)}{\nu(x)}.$$

This infinum will be achieved at a point x_0 on the unit Euclidean sphere. Then

$$\sup\frac{\nu(y)}{\rho(y)} \geqslant \frac{\nu(x_0)}{\rho(x_0)} = c^{-1},$$

so to have $d(\rho,\nu) = 0$ one must have

$$\sup\frac{\rho(x)}{\nu(x)} \leqslant c,$$

yielding $\rho(x) = c\nu(x)$ for all x.

THEOREM 1.12. *Let e be a nonzero element of $\mathbb{C}^n$ with $\chi(e) = 1$ and let N_1 be the set of all norms ρ on $\mathbb{C}^n$ such that $\rho(e) = 1$. Then the function d defined in* (1.11) *is a metric on N_1 and N_1 is complete with respect to d.*

Proof. Let $d(\rho,\nu) = 0$ where ρ, ν are in N_1. Then by Lemma 1.10 (part 2), $\rho = c\nu$ for some $c > 0$. But $\rho(e) = \nu(e) = 1$ so $\rho = \nu$. Hence d is a metric on N_1.

Let ρ_r be a Cauchy sequence of norms in N_1. The Cauchy property implies there is an $M > 0$ such that $d(\rho_r,\chi) \leqslant M$ for all r, implying that $\rho_r(x) \leqslant e^M$ for all x on the sphere

$$S = \{x\in\mathbb{C}^n \mid \chi(x) = 1\}.$$

Further, for $x_1, x_2 \in S$,

$$\frac{|\rho_r(x_1) - \rho_r(x_2)|}{\chi(x_1 - x_2)} \leqslant \frac{\rho_r(x_1 - x_2)}{\chi(x_1 - x_2)} \leqslant e^M$$

independently of r, making ρ_r a uniformly-bounded, equicontinuous set of functions on S. By the Arzela-Ascoli Theorem [8, p. 266], the sequence

ρ_r, restricted to S, is precompact in the space of continuous real valued functions on S and has a subsequence $\tilde{\rho}_r$ converging uniformly to a function $\tilde{\rho}$. Since ρ_r is a Cauchy sequence, the whole sequence converges to $\tilde{\rho}$. Since $d(\rho_r, \chi) \leqslant M$ for all r, restricting to S we see that the limit $\tilde{\rho}$ must be positive on S. We extend $\tilde{\rho}$ to ρ defined on $\mathbb{C}^n$ setting

$$\rho(x) = \begin{cases} \chi(x)\tilde{\rho}\left(\dfrac{x}{\chi(x)}\right), & x \neq 0 \\ 0 \quad , & x = 0. \end{cases}$$

One readily verifies that this homogeneous extension is a norm in N_1 and that $d(\rho_r, \rho) \to 0$ as $r \to \infty$.

2. RATIONAL BASES

Let $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$ denote the integers, the rational numbers, and the real numbers, respectively. By $\mathbb{Z}^n$, $\mathbb{Q}^n$, and $\mathbb{R}^n$ we denote the modules (linear spaces in the case of $\mathbb{Q}$ or $\mathbb{R}$) of n-tuples with values from $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$, respectively. If $\alpha = (\alpha_1, \ldots, \alpha_n)$ belongs to $\mathbb{Z}^n$, $\mathbb{Q}^n$, or $\mathbb{R}^n$, then we set

$$|\alpha|_\infty = \max_{1 \leqslant i \leqslant n} |\alpha_i|,$$
$$\|\alpha\| = \max_{1 \leqslant i \leqslant n} |\alpha_i(\text{mod } 1)|, \tag{2.1}$$

where $|\alpha_i|$ denotes the absolute value of α_i and $\alpha_i(\text{mod } 1)$ is the number in the interval $(-\frac{1}{2}, \frac{1}{2}]$ which is congruent to α_i modulo 1.

DEFINITION 2.2. *We say a subspace $S \subset \mathbb{R}^n$ has a* rational basis *if S consists of all real linear combinations of a set of vectors $\{v_1, \ldots, v_k\}$ from $\mathbb{Q}^n$.*

If S is a subset of $\mathbb{R}^n$ we let $\mathrm{sp}(S)$ denote the subspace consisting of real linear combinations of elements from S. We are interested in the largest subspace contained in the set of vectors that can be approximated "modulo 1" by vectors in $\mathrm{sp}(S)$. Accordingly, we introduce the following:

DEFINITION 2.3. *Let S be a non-empty subset of $\mathbb{R}^n$ and α a vector in $\mathbb{R}^n$. Then $\alpha \in \mathrm{App}(S)$ if and only if for each $\varepsilon > 0$ and each real t, there is a $\lambda \in \mathrm{sp}(S)$ such that $\|\lambda - t\alpha\| < \varepsilon$.*

It is readily seen that $\mathrm{App}(S)$ is a subspace of $\mathbb{R}^n$ containing S. To obtain another characterization of $\mathrm{App}(S)$ we introduce the polar of a set S; that is,

$$S^{\perp} = \{\beta \in \mathbb{R}^n | \langle \beta, s \rangle = 0 \quad \text{for all } s \in S\},$$

where $\langle \beta, s \rangle$ denotes the standard scalar product in $\mathbb{R}^n$. Even though S need not be a subspace, the set $S^{\perp}$ will be a subspace and is easily seen to have the property that if $S_1 \subseteq S_2$ then $S_1{}^{\perp} \supseteq S_2{}^{\perp}$.

THEOREM 2.4. $\text{App}(S) = (S^{\perp} \cap \mathbb{Q}^n)^{\perp}$.

Proof. We appeal to Theorem IV of the appendix. Let $w_1, \ldots, w_m$ be a basis for $\text{sp}(S)$ over $\mathbb{R}$ and let M be the $n \times m$ matrix having $w_1, \ldots, w_m$ as its columns. An arbitrary element of $\text{sp}(S)$ then has the form Mx for $x \in \mathbb{R}^m$, while $\beta \in \mathbb{R}^n$ will be in the polar of S if and only if $\beta M = 0$ as a row vector. Given $\alpha \in \mathbb{R}^n$, we have $\alpha \in \text{App}(S)$ if and only if A4 of Theorem IV (Appendix) holds. But A4 is equivalent to B4 of that theorem which states that if β is in $\mathbb{Q}^n$ and in $S^{\perp}$ then $\langle \beta, \alpha \rangle = 0$; that is, $\alpha \in (S^{\perp} \cap \mathbb{Q}^n)^{\perp}$.

REMARK. Since $S^{\perp} = (\text{sp}(S))^{\perp}$, $\text{App}(S) = \text{App}(\text{sp}(S))$.

LEMMA 2.5. *Let $T \subset \mathbb{R}^n$ be a set consisting of rational vectors (elements of $\mathbb{Q}^n$). Then $T^{\perp}$ has a rational basis.*

Proof. The subspace $\text{sp}(T) \subset \mathbb{R}^n$ has a basis $v_1, v_2, \ldots, v_q$ with each $v_i \in \mathbb{Q}^n$. A vector $\alpha = (\alpha_1 \cdots \alpha_n)$ is in $\text{sp}(T)^{\perp} = T^{\perp}$ if and only if it satisfies a matrix equation $M\alpha = 0$, where M is the $q \times n$ matrix having $v_1, \ldots, v_q$ as its rows. The matrix M contains a $q \times q$ submatrix with nonzero determinant and hence has a $k = n - q$ dimensional nullspace over $\mathbb{Q}$ or $\mathbb{R}$. If $w_1, \ldots, w_k$ from $\mathbb{Q}^n$ are a basis for the nullspace over $\mathbb{Q}$, then $w_1, \ldots, w_k$ are also a basis for the nullspace over $\mathbb{R}$.

LEMMA 2.6. *For any non-empty set $S \subset \mathbb{R}^n$,* $\text{App}(S)$ *has a rational basis.*

Proof. Let $S^{\perp} \cap \mathbb{Q}^n = T$ in Lemma 2.5.

LEMMA 2.7. *A subspace $S \subset \mathbb{R}^n$ has a rational basis if and only if* $S = \text{App}(S)$.

Proof. If $S = \text{App}(S)$, then from Lemma 2.6, S has a rational basis. Conversely, if S has a rational basis and has dimension q, the argument given in the proof of Lemma 2.5 shows that $S^\perp$ has a rational basis and has dimension $n - q$. Since $\mathrm{sp}(S^\perp \cap \mathbb{Q}^n) = S^\perp$, $\mathrm{App}(S) = (S^\perp \cap \mathbb{Q}^n)^\perp = S^{\perp\perp}$. But $S^{\perp\perp}$ has dimension q and contains S, so $\mathrm{App}(S) = S$.

COROLLARY 2.8. App *is a "closure" operation; that is,*

$$S \subset \mathrm{App}(S) \tag{1}$$

$$\mathrm{App}\,\mathrm{App}(S) = \mathrm{App}(S) \tag{2}$$

$$S \subset \tilde{S} \Rightarrow \mathrm{App}(S) \subset \mathrm{App}(\tilde{S}). \tag{3}$$

Proof. Item (1) is a consequence of $S \subset S^{\perp\perp} \subset (S^\perp \cap \mathbb{Q}^n)^\perp$. Item (2) follows from Lemmas 2.6 and 2.7. If $S \subset \tilde{S}$ then $\tilde{S}^\perp \subset S^\perp$ so $\tilde{S}^\perp \cap \mathbb{Q}^n \subset S^\perp \cap \mathbb{Q}^n$ and $\mathrm{App}(\tilde{S}) = (\tilde{S}^\perp \cap \mathbb{Q}^n)^\perp \supset (S^\perp \cap \mathbb{Q}^n)^\perp = \mathrm{App}(S)$.

COROLLARY 2.9. *If S is any set in $\mathbb{R}^n$ then*

$$\mathrm{App}(S) = \bigcap_{W \in \mathscr{S}} W$$

where $\mathscr{S}$ is the collection of subspaces W having rational bases and containing S.

Proof. If $W \in \mathscr{S}$, then $W = \mathrm{sp}(W) \supset \mathrm{sp}(S)$ so $W = \mathrm{App}(W) \supset \mathrm{App}(\mathrm{sp}\, S) = \mathrm{App}(S)$, using Lemma 2.7. Since $\mathrm{App}(S) \in \mathscr{S}$, the desired equality follows.

3. DIAGONAL NORM-HERMITIAN MATRICES

We denote the collection of $n \times n$ complex matrices by $\mathbb{C}^{nn}$. If ν is a norm on $\mathbb{C}^n$ then we say $h \in \mathbb{C}^{nn}$ is a *ν-Hermitian* matrix if $y \| x$ with respect to ν implies that $\langle y, hx\rangle$ is real. We let $H(\nu)$ denote the set of ν-Hermitian matrices. If D denotes the set of real diagonal matrices in $\mathbb{C}^{nn}$ we set $D(\nu) = H(\nu) \cap D$. In the sequel we will identify D with $\mathbb{R}^n$. Note that $I = (1, 1, \ldots, 1)$ is always in $D(\nu)$, and that $D(\nu)$ is a subspace of D.

THEOREM 3.1. *Let ν be a norm on $\mathbb{C}^n$ and let S be a non-empty subset of $D(\nu)$. Then* $\mathrm{App}(S) \subset D(\nu)$.

Proof. Suppose $\mu = (\mu_1, \ldots, \mu_n)$ is in $\mathrm{App}(S)$. Then for each $\varepsilon > 0$ and each $t \in R$ there is an element $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n) \in \mathrm{sp}(S)$ satisfying

$\|\lambda - t\mu\| < \varepsilon$ (cf. (2.1)). Since $\lambda \in D(\nu)$, the diagonal matrix

$$e^{2\pi i\lambda} = (e^{2\pi i\lambda_1}, \ldots, e^{2\pi i\lambda_n})$$

will be an isometry for the norm ν (cf. [4, p. 46]). Then for any $x \in \mathbb{C}^n$ with $\nu(x) = 1$,

$$\begin{aligned} |\nu(e^{2\pi i t\mu}x) - \nu(x)| &= |\nu(e^{2\pi i\lambda}e^{2\pi i(t\mu-\lambda)}x) - \nu(x)| \\ &= |\nu(e^{2\pi i(t\mu-\lambda)}x) - \nu(x)| \\ &\leqslant \nu((e^{2\pi i(t\mu-\lambda)} - I)x). \end{aligned}$$

A matrix map is continuously dependent on its entries and

$$e^{2\pi i(t\mu-\lambda)} = I \quad \text{if} \quad \|t\mu - \lambda\| = 0.$$

Hence, given $\delta > 0$, we can find $\lambda \in \operatorname{sp}(S)$ such that

$$\nu((e^{2\pi i(t\mu-\lambda)} - I)x) \leqslant \delta\nu(x).$$

Hence $|\nu(e^{2\pi i t\mu}x) - \nu(x)| \leqslant \delta\nu(x)$ and it follows that $e^{2\pi i t\mu}$ is an isometry for each real t. But then using Lemma 2, p. 46, of [4] again, we see that μ must be ν-Hermitian and hence in $D(\nu)$.

The following theorem is a counterpart to Theorem 3.1. A strengthened version is contained in Theorem 4.7, but we feel that the proof given here will help to illuminate the more complicated proof of 4.7.

THEOREM 3.2. *Let S be a subspace of $\mathbb{R}^n$, and suppose $I \in S$. Then there exists a norm ν on $\mathbb{C}^n$ such that* $D(\nu) = \operatorname{App}(S)$.

Proof. Let $\{e_i\}$ be a basis for $\mathbb{C}^n$ and let $\{e_i'\}$ be its algebraic dual basis. If $x = \sum_{i=1}^n \alpha_i e_i$ and $l = \sum_{i=1}^n e_i'$ then $\langle l, x\rangle = \sum_{i=1}^n \alpha_i$. If $\lambda = (\lambda_1, \ldots, \lambda_n)$ is in S we form the semi-norm $x \to |\langle l, e^{i\lambda}x\rangle|$, where $e^{i\lambda}x = \sum_{i=1}^n e^{i\lambda_i}\alpha_i e_i$. Such a semi-norm satisfies $|\langle l, e^{i\lambda}x\rangle| \leqslant \nu_1(x)$, where $\nu_1(x) = \sum_{i=1}^n |\alpha_i|$ and hence σ defined by

$$\sigma(x) = \sup_{\lambda\in S} |\langle l, e^{i\lambda}x\rangle|$$

is a semi-norm, by Lemma 1.2. Setting $\nu(x) = \nu_1(x) + \sigma(x)$ we see that ν is a norm.

If $\mu \in S$, then to show $\mu \in D(\nu)$ it suffices to show that $e^{i\mu t}$ is a ν-isometry for all real t (cf. [4, p. 46]); that is, that $\nu(e^{i\mu t}x) = \nu(x)$ for each

$x \in \mathbb{C}^n$ and all $t \in \mathbb{R}$. It is clear that $\nu_1(e^{i\mu t}x) = \nu_1(x)$ and since S is a subspace,

$$\sup_{\lambda \in S} |\langle l, e^{i(\lambda + t\mu)}x \rangle| = \sup_{\lambda \in S} |\langle l, e^{i\lambda}x \rangle|,$$

yielding $\sigma(e^{it\mu}x) = \sigma(x)$. It follows that $e^{it\mu}$ is isometric and hence $S \subset D(\nu)$.

We next show that $D(\nu) \subset \mathrm{App}(S)$.

Suppose that $\mu' \notin \mathrm{App}(S)$. Then there exists a $\delta > 0$ and a $t \in R$ such that for all $\lambda \in S$, $\|\lambda - t\mu'\| \geqslant \delta$, and clearly, such a t is non-zero. We let $t\mu' = \mu = (\mu_1, \ldots, \mu_n)$ and show that $\mu \notin D(\nu)$. Then it will follow that $\mu' \notin D(\nu)$. We let x be the vector $(e^{-2\pi i \mu_1}, \ldots, e^{-2\pi i \mu_n})$ and have

$$\begin{aligned} \nu(e^{2\pi i \mu}x) &= \nu_1(e^{2\pi i \mu}x) + \sup_{\lambda \in S} |\langle l, e^{i\lambda}e^{2\pi i \mu}x \rangle| \\ &= n + \sup_{\lambda \in S} \left| \sum_{i=1}^{n} e^{i\lambda_i} \right| \\ &= 2n. \end{aligned}$$

If $\nu(x) \neq 2n$, then $e^{2\pi i \mu}$ is not an isometry and μ cannot be ν-Hermitian. But

$$\begin{aligned} \nu(x) &= \nu_1(x) + \sup_{\lambda \in S} |\langle l, e^{2\pi i \lambda}x \rangle| \\ &= n + \sup_{\lambda \in S} \left| \sum_{i=1}^{n} e^{2\pi i(\lambda_i - \mu_i)} \right| \end{aligned}$$

and can equal $2n$ only if, by varying $\lambda \in S$, one can come arbitrarily close to making the complex numbers $e^{2\pi i(\lambda_i - \mu_i)}$ equal.

Since we may multiply a complex number by $e^{i\theta}$, θ real, without changing its modulus and since $(\theta, \theta, \ldots, \theta) \in S$ for each real θ, we need only consider those λ for which $\lambda_1 = \mu_1$. Hence

$$\nu(x) = n + \sup_{\substack{\lambda \in S \\ \lambda_1 = \mu_1}} \left| \sum_{i=1}^{n} e^{2\pi i(\lambda_i - \mu_i)} \right|$$

where we still have $\|\lambda - \mu\| \geqslant \delta > 0$. Since $|(\lambda_i - \mu_i) \bmod 1| \geqslant \delta$ for some $i \geqslant 2$, a simple estimate yields

$$\nu(x) \leqslant 2n - 2 + (2 + 2\cos 2\pi\delta)^{1/2} < 2n.$$

Thus μ is not ν-Hermitian.

Since $S \subset D(\nu) \subset \mathrm{App}(S)$, it follows from Lemma 2.7, Corollary 2.8, and Theorem 3.1 that $\mathrm{App}(S) \subset \mathrm{App}\, D(\nu) = D(\nu)$ and hence $\mathrm{App}(S) = D(\nu)$.

THEOREM 3.3. *Let S be a subset of D. Then there exists a norm ν on $\mathbb{C}^n$ such that $S = D(\nu)$ if and only if $I \in S$ and S is a subspace of D with a rational basis.*

Proof. Let $S = D(\nu)$. Clearly $I \in S$. By Theorem 3.1, $\mathrm{App}(S) \subset D(\nu) = S$ and hence $\mathrm{App}\, S = S$. It follows by Lemma 2.7, that S is a subspace with a rational basis. Conversely, if S contains I and is a subspace with a rational basis then, again by Lemma 2.7, $\mathrm{App}\, S = S$, and it follows that $S = D(\nu)$, for some norm ν on $\mathbb{C}^n$, by Theorem 3.2.

COROLLARY 3.4. *Suppose S is a subset of D and that $I \in S$. Define*

$$S_1 = \bigcap \{D(\nu) | \nu \text{ is a norm on } \mathbb{C}^n \text{ and } H(\nu) \supseteq S\}.$$

Then $\mathrm{App}(S) = S_1$.

Proof. Since $D(\nu)$ has a rational basis for each ν, $\mathrm{App}(S) \subset S_1$ follows from Corollary 2.9. But from Theorem 3.3 there exists a norm ν such that $\mathrm{App}(S) = D(\nu)$, so $\mathrm{App}(S) = S_1$.

A matrix $h \in \mathbb{C}^{nn}$ is called *diagonable* if there exists a non-singular $p \in C^{nn}$ such that $p^{-1}hp$ is a diagonal matrix.

COROLLARY 3.5. *Let $\lambda_1, \ldots, \lambda_r$, $1 \leqslant r \leqslant n$ be pairwise distinct real numbers. Let $h \in \mathbb{C}^{nn}$ be a diagonable matrix with spectrum $\{\lambda_1, \ldots, \lambda_r\}$. For each positive integer s and $\lambda \in \mathbb{R}$, put $v_s(\lambda) = (1, \lambda, \lambda^2, \ldots, \lambda^s) \in \mathbb{R}^{s+1}$. Then there exists a norm ν on $\mathbb{C}^n$ such that $h^m \in H(\nu)$ for $m = 1, 2, \ldots, s$, but $h^k \notin H(\nu)$ for some $k > s$ if and only if $v_s(\lambda_1), \ldots, v_s(\lambda_r)$ are linearly dependent over Q.*

Proof. If ν is a norm and the norm ν_p is defined by $\nu_p(x) = \nu(px)$, for $x \in \mathbb{C}^n$, then the numerical range of h with respect to ν equals the numerical range of $p^{-1}hp$ with respect to ν_p (cf. Nirschl and Schneider [10]). Hence we need consider only diagonal h. Let $\tilde{S} \subset \mathbb{R}^r$ be spanned by the r-tuples $\tilde{h}^m = (\lambda_1{}^m, \ldots, \lambda_r{}^m)$ for $m = 0, 1, \ldots, s$. Suppose there exists a non-zero vector $\beta \in \mathbb{Q}^r$ such that $\sum_{i=1}^r \beta_i v_s(\lambda_i) = 0$. Then $\beta \in (\tilde{S}^\perp \cap \mathbb{Q}^r)$, and it follows that $\mathrm{App}(\tilde{S}) \neq \mathbb{R}^r$. Since the numbers λ_i are distinct, the vectors $\tilde{h}^0, \tilde{h}^1, \ldots, \tilde{h}^{r-1}$ span R^r and hence for some $k > s$, $\tilde{h}^k \notin \mathrm{App}(\tilde{S})$.

Suppose that the eigenvalues $\lambda_1, \ldots, \lambda_r$ of h, occur with multiplicities $m_1, \ldots, m_r$, respectively. Let $S \subseteq \mathbb{R}^n$ be the span of $I, h, \ldots, h^s$ with s as above. For the same $k > s$, suppose $h^k \in \mathrm{App}(S)$. Then for each $\varepsilon > 0$ and each real t, there is a $\mu \in S$ such that $\|\mu - th^k\| < \varepsilon$. Since, in the vector $\mu - th^k$, the entries are repeated according to the multiplicities m_j, one sees that $\|\tilde{\mu} - t\tilde{h}^k\| < \varepsilon$ follows for an appropriate $\tilde{\mu} \in \tilde{S}$. This contradicts $\tilde{h}^k \notin \mathrm{App}(\tilde{S})$. Hence $h^k \notin \mathrm{App}(S)$. But by Theorem **3.3** there is a norm ν such that $\mathrm{App}(S) = D(\nu)$, implying $h^k \notin H(\nu)$, since $h^k \in D$.

Conversely, let a norm ν and an integer s be given. Suppose $h^1, h^2, \ldots, h^s \in H(\nu)$ and that $h^k \notin H(\nu)$ for some $k > s$. Then with $\tilde{S}$ as in the last section, $\tilde{h}^k$ is not approximable modulo 1 by elements of $\tilde{S}$; that is, $\mathrm{App}\, \tilde{S} \neq \mathbb{R}^r$. Hence $\tilde{S}^\perp \cap \mathbb{Q}^r \neq \{0\}$ so there is a $\beta \in \mathbb{Q}^r$, $\beta \neq 0$, such that $\sum_{i=1}^r \beta_i v_s(\lambda_i) = 0$.

COROLLARY **3.6.** *Let $\lambda_1, \ldots, \lambda_r$ and h be as in Corollary* **3.5**. *Then there exists a norm ν on $\mathbb{C}^n$ such that $h \in H(\nu)$ but $h^k \notin H(\nu)$ for some $k > 0$, if and only if $(\lambda_2 - \lambda_1, \ldots, \lambda_r - \lambda_1)$ are linearly dependent over $\mathbb{Q}$.*

Proof. One uses Corollary **3.5** in the case $s = 1$ observing that, in the notation of that corollary, $v_1(\lambda_1), \ldots, v_1(\lambda_r)$ are linearly dependent over $\mathbb{Q}$ if and only if $\lambda_2 - \lambda_1, \ldots, \lambda_r - \lambda_1$ are linearly dependent over $\mathbb{Q}$.

DEFINITION **3.7.** *A norm ν on $\mathbb{C}^n$ is* absolute *if for every $(\theta_1, \theta_2, \ldots, \theta_n) \ni \mathbb{R}^n$, the diagonal matrix $(e^{i\theta_1}, \ldots, e^{i\theta_n})$ is an isometry.*

COROLLARY **3.8.** *A norm ν is absolute if and only if there exists $d \in H(\nu)$, where d is a diagonal matrix $(d_1, d_2, \ldots, d_n)$ and $d_1 - d_2, \ldots, d_1 - d_n$ are linearly independent over $\mathbb{Q}$.*

Proof. If a norm ν is absolute then every real diagonal matrix is ν-Hermitian (cf. [13]) and one can choose the entries so that $d_1 - d_2, \ldots, d_1 - d_n$ are linearly independent over $\mathbb{Q}$. On the other hand, given the independence for a real diagonal matrix $d = (d_1, \ldots, d_n)$, there is no $\beta \in \mathbb{Q}^n$ satisfying $\langle \beta, I \rangle = 0$ and $\langle \beta, d \rangle = 0$. Hence $\mathrm{App}(\mathrm{sp}\{I, d\}) = \mathbb{R}^n$, which means every diagonal matrix is ν-Hermitian (cf. Theorem **3.1**). Then for any diagonal $\theta = (\theta_1, \ldots, \theta_n)$, $e^{i\theta}$ is an isometry.

In the remainder of this section we shall prove the converse of a theorem due to Nirschl and Schneider[1] [10]. For a given norm ν on $\mathbb{C}^n$ and

---

[1] We are grateful to B. D. Saunders for a suggestion which led to the proof of Theorem 3.11.

$a \in \mathbb{C}^{nn}$ let $V(a) = \{\langle y, ax\rangle | y \| x\}$ be the numerical range of a. We begin by restating [10, Theorem 3] (see first line of proof).

THEOREM 3.9. *Let ν be an absolute norm on $\mathbb{C}^n$, and let $a \in \mathbb{C}^{nn}$. If W is an open convex subset of $\mathbb{C}$ which contains the spectrum of a, then there exists a non-singular $s \in \mathbb{C}^{nn}$ such that $V(sas^{-1}) \subset W$.*

We first prove a lemma, the proof of which is modeled on the proof of Crabb, Duncan, and McGregor [7, Lemma 1.5].

LEMMA 3.10 (B. D. Saunders). *Let ν be a norm, let $a \in \mathbb{C}^{nn}$ have distinct eigenvalues. Let $W_1 \supset W_2 \supset \ldots$ be a sequence of closed bounded subsets of $\mathbb{C}$. If, for $k = 1, 2, \ldots$, there exists a non-singular matrix $s_k \in \mathbb{C}^{nn}$ such that $V(s_k a s_k^{-1}) \subset W_k$, then there exists a non-singular matrix s in $\mathbb{C}^{nn}$ such that $V(sas^{-1}) \subset \bigcap_{k=1}^{\infty} W_k$.*

Proof. For $c \in \mathbb{C}^{nn}$, let $v(c) = \sup\{|\lambda| \,|\lambda \in V(c)\}$. Then it is known (see Bohnenblust and Karlin [3] or [4, Theorem 1.4.1] that $\nu^0(c) \leqslant ev(c)$, where ν^0 is the operator norm on $\mathbb{C}^{nn}$ associated with ν. Since $V(s_k a s_k^{-1}) \subset W_k$, for $k = 1, 2, \ldots$, it follows that $\{\nu^0(s_k a s_k^{-1}) | k = 1, 2, \ldots\}$ is a bounded subset of $\mathbb{R}$. Hence the $b_k = s_k a s_k^{-1}$, $k = 1, 2, \ldots$, lie in some compact subset of $\mathbb{C}^{nn}$, and we may select a convergent subsequence of $b_1, b_2, \ldots$. Without loss of generality, we may assume that $b_1, b_2, \ldots$ converges to a matrix b. Let d be a diagonal matrix similar to a. Then, for $k = 1, 2, \ldots$, there exists $q_k \in \mathbb{C}^{nn}$ such that $q_k b_k q_k^{-1} = d$. We may choose the q_k so that $\nu^0(q_k) = 1$, $k = 1, 2, \ldots$, and then suppose that $q_1, q_2, \ldots$ converges to $q \in \mathbb{C}^{nn}$. Then $q \neq 0$ and $qb = dq$, and since the diagonal elements of d are distinct, it follows that q is non-singular. Hence b is similar to a, say $b = sas^{-1}$, for a non-singular $s \in \mathbb{C}^{nn}$. Let $y \| x$. Then

$$\langle y, sas^{-1}x\rangle = \lim_{k\to\infty}\langle y, s_k a s_k^{-1} x\rangle \in \bigcap_{k=1}^{\infty} W_k,$$

since $\bigcap_{k=1}^{\infty} W_k$ is closed. The lemma follows.

THEOREM 3.11. *Let ν be a norm on $\mathbb{C}^n$. Let h be a diagonable matrix in $\mathbb{C}^{nn}$ with real spectrum $\{d_1, \ldots, d_n\}$, where $d_1 - d_2, \ldots, d_1 - d_n$ are linearly independent over $\mathbb{Q}$. Then the following are equivalent:*

(i) *There is a non-singular $s \in \mathbb{C}^{nn}$ such that the norm ν_s is absolute, where ν_s is defined by $\nu_s(x) = \nu(sx)$, for $x \in \mathbb{C}^n$.*

(ii) *For each* $a \in \mathbb{C}^{nn}$ *and each open convex subset* W *of* $\mathbb{C}$ *containing the spectrum of* a *there is a non-singular* $s \in \mathbb{C}^{nn}$ *such that* $V_s(a) \subset W$, *where* $V_s(a)$ *is the numerical range of* a *for the norm* ν_s.

(iii) *For each open convex subset* W *of* $\mathbb{C}$ *which contains the spectrum of* h *there is a non-singular* $s \in \mathbb{C}^{nn}$ *such that* $V_s(h) \subset W$.

Proof. Since $V_s(a) = V(sas^{-1})$, it follows immediately from (**3**.9) that (i) implies (ii). It is trivial that (ii) implies (iii). Thus we need only prove that (iii) implies (i). So suppose that (iii) holds. Let X be the convex hull of $\{d_1, \ldots, d_n\}$ and let $W_k = \{\xi \in \mathbb{C} \mid |\xi - \omega| \leqslant 1/k, \text{ for some } \omega \in X\}$. By assumption, there is a non-singular $s_k \in \mathbb{C}^{nn}$ such that $V(s_k h s_k^{-1}) \subset W_k$. Since $d_1, \ldots, d_n$ are pairwise distinct, it follows by Lemma **3**.10 that there exists a non-singular $p \in \mathbb{C}^{nn}$ such that $V_p(h) = V(php^{-1}) \subset \bigcap_{k=1}^{\infty} W_k = X \subset \mathbb{R}$. Let s be a non-singular matrix in $\mathbb{C}^{nn}$ for which $s^{-1}php^{-1}s = d$ is a diagonal matrix. Then $V_s(d) = V(sds^{-1}) = V(php^{-1}) \subset \mathbb{R}$. Hence $d \in H(\nu_s)$ and so by Corollary **3**.8, the norm ν_s is absolute.

4. NORM HERMITIAN MATRICES

In this section we suppose we are given $S \subset D$ with $\text{App}(S) = S$ and $I \in S$, and construct a norm ν so that $H(\nu) = S$. In order to force all ν-Hermitian elements to be diagonal, but still preserve all elements of S as ν-Hermitian we require a more complicated construction than was used in the proof of Theorem **3**.2. We begin with a technical lemma that will be used several times.

LEMMA **4**.1. *Let* $\theta_1 < \theta_2 < \cdots < \theta_k$ *be real numbers and let* $h_1, h_2, \ldots, h_k$ *be complex. If the function*

$$f(t) = h_1 e^{it\theta_1} + \cdots + h_k e^{it\theta_k}$$

is real for real t *in some neighborhood of* $t = 0$, *then the following hold*:

(i) *if* $\theta_i = 0$ *for some* i, *then* h_i *is real*,

(ii) *if* $\theta_i = -\theta_j$ *for some pair of indices* i, j, *then* $h_i = \bar{h}_j$,

(iii) *for each* i, *if there does not exist a* j *with* $j \neq i$ *such that* $|\theta_j| = |\theta_i|$, *then* $h_i = 0$.

Proof. If, for a given i with $\theta_i > 0$, there is a j such that $\theta_i = -\theta_j$, then there is at most one such j and we have

$$h_i e^{it\theta_i} + h_j e^{it\theta_j} = (h_i - \bar{h}_j) e^{it\theta_i} + \bar{h}_j e^{-it\theta_j} + h_j e^{it\theta_j}.$$

Since $\bar{h}_j e^{-it\theta_j} + h_j e^{it\theta_j}$ is real for t real, we can incorporate it into $f(t)$ and consider a sum of exponentials in which the absolute values of all exponents are distinct. We shall show in this latter case that all coefficients are zero except for a possible constant term and show that the constant term is real. It will then follow that each difference $h_i - \bar{h}_j$, above, is zero, and all parts of the conclusion will follow.

Suppose then that we have a function

$$r(t) = g_0 + g_1 e^{it\phi_1} + \cdots + t_m e^{it\phi_m}$$

where the g_i are complex constants and the ϕ_i are real with all values $|\phi_i|$ distinct. Assuming $r(t)$ is real for t in a neighborhood of zero, the derivatives

$$r^{(k)}(t) = g_1 (i\phi_1)^k e^{it\phi_1} + \cdots + g_m (i\phi_m)^k e^{it\phi_m}$$

will be real for such t. Taking the imaginary part at $t = 0$ we obtain

$$\sum_{j=1}^{m} \operatorname{Im} g_j \phi_j{}^{4p} = 0; \qquad p = 1, 2, \ldots, m.$$

Since the determinant of the coefficients $\{\phi_j{}^{4p}\}$ is of Vandermonde type with distinct constants $\phi_j{}^4$, we must have $\operatorname{Im} g_j = 0$ for $j = 1, 2, \ldots, m$. Similarly,

$$\operatorname{Re} g_j (\phi_j)^{4p+1} = 0, \qquad p = 0, 1, \ldots, m-1,$$

yielding $\operatorname{Re} g_j = 0$, for $j = 1, 2, \ldots, m$. Then $g_j = 0$ for $j = 1, \ldots, m$ and $g_0 = r(t)$ which must then be real. The lemma is complete.

Let S be a subspace of $\mathbb{R}^n$ having a rational basis, with $I \in S$. Define a relation on the integers $\{1, 2, \ldots, n\}$ with respect to S, setting $i \sim j$ if, for every element $(\lambda_1, \ldots, \lambda_n)$ in S, $\lambda_i = \lambda_j$. The relation is easily seen to be an equivalence relation. For a given S suppose there are τ equivalence classes, denoted by $C_1, C_2, \ldots, C_\tau$. If C_m consists of a single integer we call C_m, and the integer it contains, a *singleton*. Otherwise we call C_m, and any integer it contains, *multiple*.

As before, we use $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ to denote a point in $\mathbb{R}^n$ and a diagonal matrix in $\mathbb{C}^{nn}$.

Suppose $\{e_i\}$ is the canonical basis for $\mathbb{C}^n$ and that $\{e_i'\}$ is the algebraic dual basis. Given an ε, $0 < \varepsilon < \frac{1}{4}$, we describe a collection of functionals L to be used in the construction of a norm. A functional $l = \sum_{i=1}^n \gamma_i e_i'$ belongs to L if and only if l satisfies:

$$\left.\begin{array}{l}\text{There exists } i_r \in C_r,\ r = 1, 2, \ldots, \tau, \text{ and, if } C_r \text{ is multiple,}\\ j_r \text{ in } C_r,\ j_r \neq i_r \text{ such that}\\ \quad (1)\ \text{Among the coefficients } \gamma_{i_r},\ r = 1, 2, \ldots, \tau \text{ precisely}\\ \text{one has the value 2 and the remaining } \tau - 1 \text{ coefficients have}\\ \text{the value 1.}\\ \quad (2)\ \text{If } C_r \text{ is multiple, } \gamma_{j_r}/\gamma_{i_r} = \pm\varepsilon \text{ and } \gamma_t = 0 \text{ for } t \in C_r,\\ t \neq i_r,\ t \neq j_r.\end{array}\right\} \quad (4.2)$$

Note that the numbers i_r, j_r may be different for different elements of L. We let l stand for an arbitrary functional from the collection L. Later we will distinguish among the functionals with subscripts. We now define

$$\nu(x) = \sup_{\substack{l \in L \\ \lambda \in S}} |\langle l, e^{i\lambda} x\rangle| \tag{4.3}$$

for each $x = \sum_{i=1}^n \alpha_i e_i$ in $\mathbb{C}^n$, where, as before, $e^{i\lambda}x = \sum_{j=1}^n e^{i\lambda_j}\alpha_j e_j$.

LEMMA 4.4. *ν defined by* (4.3) *is a norm on* $\mathbb{C}^n$.

Proof. For each pair $l \in L$ and $\lambda \in S$, the map $x \to |\langle l, e^{i\lambda}x\rangle|$ is a semi-norm and if $x = \sum_{i=1}^n \alpha_i e_i$ it is easily seen that

$$|\langle l, e^{i\lambda}x\rangle| \leqslant 2\sum_{i=1}^{n} |\alpha_i|.$$

Thus the semi-norms are uniformly bounded with respect to the l_1 norm. Using Lemma 1.2 we see that ν is a semi-norm, so one need only show that $\nu(x) = 0$ implies $x = 0$. For this purpose we shall show that there exists a $\lambda = (\lambda_1, \ldots, \lambda_n)$ from S such that $i \not\sim j$ implies $\lambda_i \neq \lambda_j$. For each pair i, j such that $i \not\sim j$, the set of $\lambda \in S$ for which $\lambda_i = \lambda_j$ is a proper subspace. Since S is not the union of a finite number of proper subspaces (cf. [12]), there exists a $\lambda = (\lambda_1, \ldots, \lambda_n) \in S$ which, after a renumbering of the basis vectors, satisfies

$$\left.\begin{array}{ll}(1) & i \not\sim j \Rightarrow \lambda_i \neq \lambda_j,\\ (2) & i \leqslant j \Rightarrow \lambda_i \leqslant \lambda_j.\end{array}\right\} \tag{4.5}$$

Let $\lambda = (\lambda_1, \ldots, \lambda_n)$ satisfy (4.5) and let $\mu_1, \mu_2, \ldots, \mu_\tau$ denote the distinct numbers occurring among the λ_i. Then for any $l \in L$, the quantity $\langle l, e^{it\lambda}x\rangle$ will be of the form $s(t) = e^{it\mu_1}\beta_1 + \cdots + e^{it\mu_\tau}\beta_\tau$ where $\beta_i = \langle l, P^i x\rangle$ and $P^i x$ is the projection $\sum_{j \in C_i}\langle e_j', x\rangle e_j$. If $\nu(x) = 0$, then $s(t) = 0$ for all real t and its derivatives at $t = 0$ will vanish. Then $\beta_1, \ldots, \beta_\tau$ will provide a solution of

$$\left.\begin{array}{l} \beta_1 + \beta_2 + \cdots + \beta_\tau = 0 \\ \mu_1\beta_1 + \mu_2\beta_2 + \cdots + \mu_\tau\beta_\tau = 0 \\ \vdots \\ \mu_1{}^{\tau-1}\beta_1 + \mu_2{}^{\tau-1}\beta_2 + \cdots + \mu_\tau{}^{\tau-1}\beta_\tau = 0. \end{array}\right\} \tag{4.6}$$

However, the coefficient matrix of (4.6) is nonsingular, being a Vandermonde matrix with distinct μ_i, and so $\langle l, P^i x\rangle = \beta_i = 0$ for each i. But the span of the functions in L is all of $\mathbb{C}^n$, so $P^i x = 0$ for each i, giving $x = 0$.

THEOREM 4.7. *Let S be a subspace of $\mathbb{R}^n$ and suppose $I \in S$. If the norm ν is defined by* (4.3) *then* $H(\nu) = \mathrm{App}(S)$.

Proof. Let h be a ν-Hermitian matrix. We examine what the Hermitian condition implies about the elements h_{pq}, of h, and begin with the case in which both p and q are multiple.

Case *mm*: $p \in C_r$, $q \in C_s$, both multiple. We assume $p \neq q$, but do not exclude $r = s$. We let λ be an element of S satisfying conditions (4.5). Let $l_1 \in L$ have $i_r = p$ and $i_s = q$ if $p \not\sim q$; and let $i_r = i_s = p$, $j_r = j_s = q$ if $p \sim q$. Assume $\gamma_{j_r}/\gamma_{i_r} = \gamma_{j_s}/\gamma_{i_s} = +\varepsilon$ and that $\gamma_{i_s} = 2$. Beyond these requirements, but maintaining condition 4.2, l_1 can be arbitrarily chosen.

Having $0 < \varepsilon < \frac{1}{4}$, we set $r_1 = (1 + \varepsilon^2)^{-1}$. Letting $J = \{s_1 | r_1 \leqslant s_1 < 1\}$ and setting $s_2 = \varepsilon^{-1}(1 - s_1)$ one has $s_1 + \varepsilon s_2 = 1$ for $s_1 \in J$. Further

$$\begin{aligned} \varepsilon s_1 + s_2 &< \varepsilon + \varepsilon^{-1}(1 - r_1) \\ &= \varepsilon + \varepsilon^{-1}[1 - (1 + \varepsilon^2)^{-1}] \\ &\leqslant \tfrac{1}{2}, \end{aligned}$$

and one then easily sees that for $s_1 \in J$, each of the quantities $\pm\, \varepsilon s_1 + s_2$, $s_1 - \varepsilon s_2$, s_1, $\pm\, \varepsilon s_1$, s_2, and $\pm\, \varepsilon s_2$ has absolute value less than 1. By the continuity of addition and multiplication there is a neighborhood U of J in the *complex* plane such that for $z_1 \in U$, z_2 exists such that

$$\left.\begin{array}{ll} (1) & z_1 + \varepsilon z_2 = 1, \\ (2) & \text{the numbers } \pm\, \varepsilon z_1 + z_2,\ z_1 - \varepsilon z_2,\ z_1,\ \pm\, \varepsilon z_1,\ z_2,\ \pm\, \varepsilon z_2 \\ & \text{each has modulus less than 1.} \end{array}\right\} \tag{4.8}$$

With l_1 as fixed above and λ satisfying (4.5) we define

$$y_t = e^{-it\lambda} l_1 \tag{4.9}$$

and

$$x_t = \delta_1 e^{-it\lambda_p}(z_1 e_{i_r} + z_2 e_{j_r}) + \delta_2 e^{-it\lambda_q}(z_1 e_{i_s} + z_2 e_{j_s}) \tag{4.10}$$

where t is real; $z_1 \in U$; z_1, z_2 satisfy (4.8); and δ_1, δ_2 satisfy

$$\left.\begin{array}{ll} (1) & \delta_2 > \delta_1 \geqslant 0, \\ (2) & \delta_1 + 2\delta_2 = 1, \\ (3) & \delta_1 = 0, \quad \text{if} \quad p \sim q. \end{array}\right\} \tag{4.11}$$

Suppose we choose any l' from L and any λ' from S. Then

$$|\langle l', e^{i\lambda'} x_t\rangle| \leqslant \delta_1 |\langle l', e^{i(\lambda'_p - t\lambda_p)}(z_1 e_{i_r} + z_2 e_{j_r})\rangle|$$
$$+ \delta_2 |\langle l', e^{i(\lambda'_q - t\lambda_q)}(z_1 e_{i_s} + z_2 e_{j_s})\rangle|.$$

Depending upon whether l' has a coefficient 2 associated with the class C_r, with the class C_s, or with neither class, the expression $|\langle l', e^{i\lambda'} x_t\rangle|$ will be bounded by a quantity $2\delta_1|a_1| + \delta_2|a_2|$, $\delta_1|a_1| + 2\delta_2|a_2|$, or $\delta_1|a_1| + \delta_2|a_2|$, respectively, where each number a_i ($i = 1, 2$) is either zero or one of the complex numbers listed in parts (1) and (2) of (4.8). In any case, using (4.11) we see that

$$|\langle l', e^{i\lambda'} x_t\rangle| \leqslant 1.$$

Thus, referring to (4.3), we conclude that $\nu(x_t) \leqslant 1$. However,

$$\langle y_t, x_t\rangle = \langle e^{-it\lambda} l_1, x_t\rangle = \langle l_1, e^{it\lambda} x_t\rangle$$
$$= \delta_1 + 2\delta_2 = 1,$$

so $\nu(x_t) = 1$ and $\nu^D(y_t) \geqslant 1$. Since, for arbitrary $z \in \mathbb{C}^n$,

$$|\langle y_t, z\rangle| = |\langle l_1, e^{it\lambda} z\rangle| \leqslant \sup_{\substack{l' \in L \\ \lambda' \in S}} |\langle l', e^{i\lambda'} z\rangle| = \nu(z),$$

we see that $\nu^D(y_t) = 1$ and that $y_t \| x_t$ with respect to ν for all real t. Consequently, $\langle y_t, hx_t\rangle$ is real for all t. We can write the scalar product as

$$\langle y_t, hx_t\rangle = \langle e^{-it\lambda}l_1, hx_t\rangle$$

$$= \langle l_1, e^{it\lambda}hx_t\rangle$$

$$= \langle l_1, \sum_{k=1}^{n} (e^{it(\lambda_k-\lambda_p)}\delta_1 t_{kr} + e^{it(\lambda_k-\lambda_q)}\delta_2 t_{ks})e_k\rangle, \tag{4.12}$$

where

$$t_{km} = h_{ki_m}z_1 + h_{kj_m}z_2, \qquad m = r \text{ or } s. \tag{4.13}$$

Before pursuing (4.12) further we obtain corresponding expressions for the remaining cases.

Case *ms*: $p \in C_r$ multiple; $q \in C_s$, singleton. Let l_1 have $i_r = p$, $\gamma_{j_r}/\gamma_{i_r} = +\varepsilon$ and $\gamma_{i_s} = 2$. Define y_t by (4.9) and let

$$x_t = \delta_1 e^{-it\lambda_p}(z_1 e_{i_r} + z_2 e_{j_r}) + \delta_2 e^{-it\lambda_q} e_q. \tag{4.14}$$

Again, $y_t||x_t$ and the expression (4.12) is real, where now

$$\left.\begin{aligned} t_{kr} &= h_{ki_r}z_1 + h_{kj_r}z_2, \\ t_{ks} &= h_{kq}. \end{aligned}\right\} \tag{4.15}$$

Case *sm*: $p \in C_r$, singleton; $q \in C_s$, multiple. Let $l_1 \in L$ have $i_s = q$, $\gamma_{j_s}/\gamma_{i_s} = +\varepsilon$ and $\gamma_{i_s} = 2$. Define y_t by (4.9) and let

$$x_t = \delta_1 e^{-it\lambda_p} e_p + \delta_2 e^{-it\lambda_q}(z_1 e_{i_s} + z_2 e_{j_s}). \tag{4.16}$$

Again, (4.12) is real with

$$\left.\begin{aligned} t_{kr} &= h_{kp}, \\ t_{ks} &= h_{ki_s}z_1 + h_{kj_s}z_2. \end{aligned}\right\} \tag{4.17}$$

Case *ss*: $p \in C_r$, $q \in C_s$, both singletons. We require $\gamma_{i_s} = 2$, define y_t by (4.9), and set

$$x_t = \delta_1 e^{-it\lambda_p} e_p + \delta_2 e^{-it\lambda_q} e_q. \tag{4.18}$$

One finds that (4.12) is real with

$$t_{km} = h_{km}, \qquad m = r \text{ or } s. \tag{4.19}$$

We now return to examine (4.12).

Let q be in C_1. The first necessary condition we obtain is that $h_{pq} = 0$ whenever $p \not\sim q$. For this purpose we take $\delta_1 = 0$ and maintain this assumption through formula (4.31).

In Case *mm*, (4.12) yields the real expression

$$\sum_{k=1}^{n} e^{it(\lambda_k - \lambda_q)} t_{ks} \langle l_1, e_k \rangle, \quad s = 1. \tag{4.20}$$

Since $q \in C_1$, the terms $\lambda_k - \lambda_q$ are all non-negative and for k in a given class C_r, $r > 1$, all of the terms $\lambda_k - \lambda_q$ are equal and positive. Moreover, if $k \in C_r$ and $k' \notin C_r$, then $|\lambda_{k'} - \lambda_q| \neq |\lambda_k - \lambda_q|$. By grouping the terms in (4.20) according to equivalence classes for the index k and applying Lemma 4.1, we can conclude that for $r > 1$,

$$\sum_{k \in C_r} t_{ks} \langle l_1, e_k \rangle = 0, \qquad s = 1, \tag{4.21}$$

or, evaluating $\langle l_1, e_k \rangle$,

$$t_{i_r s} + \varepsilon t_{j_r s} = 0, \qquad s = 1. \tag{4.22}$$

Were one to follow the same type of argument using a functional $l_2 \in L$, which differs from l_1 *only* in having ε replaced by $-\varepsilon$, then with $y_t = e^{-it\lambda} l_2$ and

$$x_t = \tfrac{1}{2} e^{-it\lambda_q} (z_1 e_{i_s} - z_2 e_{j_s}) \tag{4.23}$$

one would arrive at

$$t'_{i_r s} - \varepsilon t'_{j_r s} = 0, \qquad s = 1, \tag{4.24}$$

with

$$t_{ks}' = h_{k i_s} z_1 - h_{k j_s} z_2. \tag{4.25}$$

Adding (4.22) and (4.24) with the use of (4.13) and (4.25), one obtains

$$h_{i_r i_s} z_1 + \varepsilon h_{j_r j_s} z_2 = 0, \qquad s = 1, \tag{4.26}$$

or, noting (4.8),

$$(h_{i_r i_s} - h_{j_r j_s}) z_1 + h_{j_r j_s} = 0, \qquad s = 1. \tag{4.27}$$

Since z_1 can vary while maintaining conditions (4.8), Eq. (4.27) can hold only if $h_{pq} = h_{i_r i_1} = h_{j_r j_1} = 0$.

In Case *ms* we obtain (4.21) with t_{ks} as in (4.14), yielding

$$h_{i_r q} + \varepsilon h_{j_r q} = 0, \qquad q \in C_1. \tag{4.28}$$

Using an $l_2 \in L$ differing from l_1 *only* in having ε replaced by $-\varepsilon$, setting $y_t = e^{-it\lambda} l_2$, and replacing z_2 by $-z_2$ in (4.14), with $\delta_1 = 0$, one obtains

$$h_{i_r q} - \varepsilon h_{j_r q} = 0, \qquad q \in C_1 \tag{4.29}$$

or, from (4.28) and (4.29), $h_{pq} = h_{i_r q} = 0$.

In Case *sm*, (4.12) provides

$$h_{i_r i_1} z_1 + h_{i_r i_1} z_2 = 0 \tag{4.30}$$

from which $h_{pq} = h_{i_r i_1} = 0$ follows, using (4.8) and the variability of z_1.

In Case *ss*, (4.12) immediately yields $h_{pq} = 0$ for $p > 1$, with the aid of Lemma 4.1.

Thus, when $q \in C_1$, we have shown that

$$h_{pq} = 0 \tag{4.31}$$

when $p \not\sim q$.

If the matrix h is partitioned into blocks corresponding to the classes $C_1, C_2, \ldots, C_\tau$ one must then have all zero entries in the first column of blocks with the exception of the block in the upper left hand corner. We will later return to the diagonal blocks. First, however, we proceed to show that the other off-diagonal blocks must have entries equal to zero.

Let $1 \leqslant p, q \leqslant n$. We shall call the pair (p, q) *non-degenerate* if

$$k \not\sim p \Rightarrow |\lambda_k - \lambda_q| \neq |\lambda_p - \lambda_q|. \tag{4.32}$$

Otherwise we call the pair (p, q) *degenerate*. (As we have already observed (p, q) is non-degenerate if $q \in C_1$.) We proceed by induction. Suppose we have shown that for some $s \geqslant 2$, $q \in C_{s'}$, $s' \leqslant s - 1$ and $p \not\sim q$ imply $h_{pq} = 0$. Let q be in C_s. We shall show that $h_{pq} = 0$, for $p \not\sim q$. If (p, q) is non-degenerate, then the arguments accompanying Eqs. (4.19) through (4.31) hold without the restriction $s = 1$. Hence $h_{pq} = 0$ follows.

Now let (p, q) be degenerate. Suppose we have shown that $h_{pq} = 0$ for $p \in C_r$ with $r < s$. Let (p', q) be degenerate, where $p' \in C_{r'}$, $r' > s$. Then $|\lambda_{p'} - \lambda_q| = |\lambda_p - \lambda_q|$ for some $p \in C_r$, $r < s$. Again, the arguments accompanying Eqs. (4.19) through (4.31) (without the restriction $s = 1$) show that $h_{pq} - \bar{h}_{p'q} = 0$. But $h_{pq} = 0$, so $h_{p'q} = 0$. It suffices, then, to show $h_{pq} = 0$ for $p \in C_r$ with $r < s$.

Assume $p \in C_r$, $r < s$ and that $C_{r'}$ is the unique class such that $p' \in C_{r'}$ implies $\lambda_{p'} - \lambda_q = \lambda_q - \lambda_p$.

Suppose we are in the Case *mm*. We now use expressions (4.9) and (4.10) under conditions (4.8) and (4.11) with $\delta_1 > 0$. Since, by our inductive hypothesis, $h_{ki_r} = h_{kj_r} = 0$ for $k \notin C_r$, the expression (4.12) contains the terms

$$e^{it(\lambda_p - \lambda_q)}\delta_1(t_{i_r s} + \varepsilon t_{j_r s}) + e^{it(\lambda_{p'} - \lambda_q)}\delta_1(t_{i_{r'} s} + \varepsilon t_{j_{r'} s}) \tag{4.33}$$

together with terms having exponents *unequal* to $\pm\, it(\lambda_p - \lambda_q)$. The expression (4.12) is still real and Lemma 4.1 enables us to conclude that

$$(t_{i_r s} + \varepsilon t_{j_r s}) - \overline{(t_{i_{r'} s} + \varepsilon t_{j_{r'} s})} = 0, \tag{4.34}$$

where the bar denotes the complex conjugate.

Let l_3 be a functional from L having $\gamma_{i_r} = 2, \gamma_{j_r} = 2\varepsilon, \gamma_{i_s} = 1, \gamma_{j_s} = \varepsilon$, and having its remaining coefficients equal to those of l_1, introduced for the case *mm* at the beginning of the proof. Let $y_t = e^{it\lambda}l_3$ and let

$$x_t = \delta_2 e^{-it\lambda_p}(z_1 e_{i_r} + z_2 e_{j_r}) + \delta_1 e^{-it\lambda_q}(z_1 e_{i_s} + z_2 e_{j_s}), \tag{4.35}$$

which differs from (4.10) in having δ_1 and δ_2 interchanged. Repeating the type of argument given in the last paragraph, but using the quantities l_3, y_t, and x_t just defined, one finds $y_t \| x_t$, and the reality of $\langle y_t, hx_t\rangle$ yields

$$2(t_{i_r s} + \varepsilon t_{j_r s}) - \overline{(t_{i_{r'} s} + \varepsilon t_{j_{r'} s})} = 0, \tag{4.36}$$

with t_{ks} still given by (4.13). Together, (4.34) and (4.36) lead to

$$t_{i_r s} + \varepsilon t_{j_r s} = 0. \tag{4.37}$$

Assume l_2 (introduced after formula (4.22)) and l_4 differ from l_1 and l_3, respectively, *only* in having ε replaced by $-\varepsilon$. Then with $y_t = e^{-it\lambda}l_i$ $(i = 2, 4)$ and with an x_t differing from (4.10) and (4.35), respectively, *only* in having $-z_2$ where z_2 stands, one can repeat the steps above to obtain equations differing from (4.34) and (4.36), respectively, only in having $-\varepsilon$ where ε stands. From the new equations one obtains

$$t_{i_r s} - \varepsilon t_{j_r s} = 0 \tag{4.38}$$

in place of (4.37), which, together with (4.37) implies

$$t_{i_r s} = h_{i_r i_s} z_1 + h_{i_r j_s} z_2 = 0. \tag{4.39}$$

But we have already seen (cf. (4.30)) that an equation of the type (4.39) can hold only if $h_{i_r i_s} = 0$.

The argument just completed can be imitated in the Cases *ms*, *sm*, and *ss*. For Case *ms* we refer back to definitions (4.14), (4.15) and the accompanying discussion. Using (4.12) one obtains

$$t_{i_r s} + \varepsilon t_{j_r s} - \overline{t_{i_r' s}} = 0. \tag{4.40}$$

Use of an l_3 differing from l_1 *only* in having $\gamma_{i_r} = 2$ rather than $\gamma_{i_s} = 2$ together with an x_t, differing from (4.14) *only* in having δ_1, δ_2 interchanged, provides

$$2(t_{i_r s} + \varepsilon t_{j_r s}) - \overline{t_{i_r' s}} = 0. \tag{4.41}$$

Combining (4.40) and (4.41) we have $t_{i_r s} + t_{j_r s} = 0$. Next, one uses l_2 and l_4 differing from l_1 and l_3, respectively, *only* in having ε replaced by $-\varepsilon$. These, together with appropriate vectors y_t and x_t, supply the vanishing of $t_{i_r s} - \varepsilon t_{j_r s}$. Then $t_{i_r s} = 0$, leading as before to $h_{i_r i_s} = 0$.

For Case *sm* we use (4.16) and then the vector obtained by interchanging δ_1, δ_2 together with the appropriate l_1 and l_3, respectively, to obtain

$$t_{i_r s} - (t_{i_r' s} + \varepsilon t_{j_r' s}) = 0 = 2t_{i_r s} - (t_{i_r' s} + \varepsilon t_{j_r' s})$$

and hence $t_{i_r s} = h_{pq} = 0$. Finally, in the Case *ss* one uses functionals l_1 and l_3 to obtain $h_{pq} = 0$. This completes the argument showing that $q \in C_s$ and $p \not\sim q$ imply $h_{pq} = 0$. By induction the assertion holds for $1 \leqslant s \leqslant n$.

We now know that h can have non-zero entries only in blocks along the diagonal. The structure of a multiple block is obtained by further examining the Case *mm*, setting $i_r = i_s = p$ and $j_r = j_s = q$. We use (4.9) and (4.10) under conditions (4.8) and (4.11) with $\delta_1 = 0$ and $t = 0$. Now, the fact that (4.12) is real means that

$$(h_{pp} z_1 + h_{pq} z_2) + \varepsilon(h_{qp} z_1 + h_{qq} z_2) \tag{4.42}$$

must be real. Using l_2 (introduced earlier) and replacing z_2 by $-z_2$ in (4.9) one finds that

$$(h_{pp} z_1 - h_{pq} z_2) - \varepsilon(h_{qp} z_1 - h_{qq} z_2) \tag{4.43}$$

is real and, subtracting, that

$$h_{pq} z_2 + \varepsilon h_{qp} z_1 \tag{4.44}$$

is real. Since $z_1 + \varepsilon z_2 = 1$,

$$h_{pq} + (\varepsilon^2 h_{qp} - h_{pq})z_1 \tag{4.45}$$

is real. As z_1 can vary in a complex open set we conclude that h_{pq} is real and

$$h_{pq} = \varepsilon^2 h_{qp}. \tag{4.46}$$

But then

$$h_{pq} = \varepsilon^2 h_{qp} = \varepsilon^4 h_{pq} \tag{4.47}$$

and since $0 < \varepsilon < \frac{1}{4}$, $h_{pq} = 0$.

If we add (4.42) and (4.43) we get the real expression

$$h_{pp} z_1 + \varepsilon h_{qq} z_2 = z_1(h_{pp} - h_{qq}) + h_{qq}. \tag{4.48}$$

Again, as z_1 can vary, we see that h_{qq} is real and $h_{pp} = h_{qq}$.

If p is a singleton, and e_p is the unit vector having its pth coordinate equal to 1, then $\frac{1}{2}e_p$ and any $l \in L$ with $\gamma_p = 2$ are easily seen to be dual vectors, forcing $\frac{1}{2}\langle l, he_p\rangle = h_{pp}$ to be real.

We have shown that a necessary condition for h, represented by $\{h_{pq}\}$ to be ν-Hermitian is that its entries be real and

$$\left.\begin{aligned} h_{pq} &= 0 \quad \text{if} \quad p \neq q \\ h_{pp} &= h_{qq} \quad \text{if} \quad p \sim q. \end{aligned}\right\} \tag{4.49}$$

The elements $\lambda \in S$, of course, satisfy (4.49), a fact that will also follow from the inclusion $S \subset H(\nu)$. The inclusion, in turn, follows easily from the definition of the norm since

$$\begin{aligned} \nu(e^{it\lambda}x) &= \sup_{\substack{l\in L \\ \lambda'\in S}} |\langle l, e^{i\lambda'} e^{it\lambda} x\rangle| \\ &= \sup_{\substack{l\in L \\ \lambda''\in S}} |\langle l, e^{i\lambda''} x\rangle| \\ &= \nu(x), \end{aligned} \tag{4.50}$$

and any λ generating an isometric group must be ν-Hermitian. Since $S \subset H(\nu)$ and $H(\nu) = D(\nu)$ by (4.49), it follows from Theorem 3.1 that $\mathrm{App}(S) \subset H(\nu)$. It remains to show that $H(\nu) \subset \mathrm{App}(S)$. We note that if S generates just one equivalence class then by (4.49) $H(\nu)$ consists of real multiples of the identity and thus $H(\nu) \subset (S) \subset \mathrm{App}(S)$. Accordingly, we can assume that the number τ, of equivalence classes, is at least 2. We shall show that if $h' \notin \mathrm{App}(S)$ then $h' \notin H(\nu)$. We may restrict attention to an h' satisfying (4.49), since if (4.49) fails, then $h' \notin H(\nu)$. Thus let h' be the diagonal matrix $(h_1', \ldots, h_n')$ where $h_i' = h_j'$ if $i \sim j$. If $h' \notin \mathrm{App}(S)$, there exists a $\delta > 0$ and a $t \in R$ such that for any $\lambda \in S$,

$$\|\lambda - th'\| \geqslant \delta. \tag{4.51}$$

Since (4.51) cannot be satisfied with $t = 0$, if we show that $h = th' \notin H(\nu)$, it will follow that $h' \notin H(\nu)$.

From each equivalence class C_r choose an integer i_r and let

$$x = 2\tau e^{-2\pi i h_{i_1}} e_{i_1} + \sum_{2 \leqslant r \leqslant \tau} e^{-2\pi i h_{i_r}} e_{i_r}. \tag{4.52}$$

Then for any $\lambda \in S$,

$$e^{2\pi i \lambda} x = 2\tau e^{2\pi i (\lambda_{i_1} - h_{i_1})} e_{i_1} + \sum_{2 \leqslant r \leqslant \tau} e^{2\pi i (\lambda_{i_r} - h_{i_r})} e_{i_r} \tag{4.53}$$

and

$$\nu(x) = \sup_{\substack{l \in L \\ \lambda \in S}} |\langle l, e^{2\pi i \lambda} x \rangle|.$$

As was noted in the proof of Theorem 3.2, the norm of x will be unaffected by taking the supremum over those $\lambda \in S$ having $\lambda_{i_1} = h_{i_1}$.

If $l = \sum_{i=1}^n \gamma_i e_i'$ and $\gamma_{i_1} = 1$, then

$$|\langle l, e^{2\pi i \lambda} x \rangle| \leqslant 2\tau + 2 + \tau - 2 = 3\tau,$$

since $\gamma_{i_r} = 2$ for at most one class C_r with $r \neq 1$. However, if $\gamma_{i_1} = 2$, then from the triangle inequality,

$$|\langle l, e^{2\pi i \lambda} x \rangle| \geqslant 4\tau - (\tau - 1) = 3\tau + 1.$$

So the norm $\nu(x)$ can be determined by using only functionals in L with $\gamma_{i_1} = 2$. Suppose l is within this restricted class and that λ is an arbitrary element of S normalized so that $\lambda_{i_1} = h_{i_1}$. Since, in (4.53), $|(\lambda_{i_k} - h_{i_k}) \bmod 1| \geqslant \delta$ for some $k \geqslant 2$, $|\langle l, e^{2\pi i \lambda} x \rangle|$ has the form

$$|4\tau + r e^{2\pi i \theta} + a_3 + \cdots + a_\tau|,$$

where $0 \leqslant r \leqslant 1$, $\delta \leqslant |\theta| \leqslant \frac{1}{2}$ and $|a_i| \leqslant 1$ for $3 \leqslant i \leqslant \tau$. Hence, using the fact that $(\alpha^2 - \beta)^{1/2} \leqslant \alpha - \beta/(2\alpha)$ for positive α and β, we find

$$\begin{aligned} |\langle l, e^{2\pi i\lambda}x\rangle| &\leqslant (16\tau^2 + 8\tau\cos 2\pi\delta + 1)^{1/2} + \tau - 2 \\ &\leqslant [(4\tau+1)^2 - 8\tau(1-\cos 2\pi\delta)]^{1/2} + \tau - 2 \\ &\leqslant 4\tau + 1 - \frac{8\tau(1-\cos 2\pi\delta)}{8\tau+2} + \tau - 2 \\ &= 5\tau - 1 - \frac{8\tau(1-\cos 2\pi\delta)}{8\tau+2} \end{aligned}$$

and consequently $\nu(x)$ is *strictly* less than $5\tau - 1$.

If h were ν-Hermitian, then $\nu(e^{2\pi ih}x)$ would equal $\nu(x)$. However, $e^{2\pi ih}x = 2\tau e_{i_1} + \sum_{r=2}^{\tau} e_{i_r}$, so if $l \in L$ has $\gamma_{i_1} = 2$ and $\gamma_{i_r} = 1$ for $r \geqslant 2$, then $|\langle l, e^{2\pi ih}x\rangle| = 5\tau - 1$ from which it follows that $\nu(e^{2\pi ih}x) \geqslant 5\tau - 1$. Consequently, $h \notin H(\nu)$ and the proof of Theorem 4.7 is complete.

THEOREM 4.54. *Let S be a subset of D. Then there exists a norm ν on $\mathbb{C}^n$ such that $S = H(\nu)$ if and only if $I \in S$ and S is a subspace of D with a rational basis.*

Proof. The "only if" portion is the same as that of Theorem **3.3**. In the other direction, one obtains $S = \text{App } S$ and the desired norm is that given in Theorem 4.7.

While one does not need such an elaborate norm construction to obtain the following result, it follows immediately from Theorem 4.7.

COROLLARY 4.55. *For any positive $n \in \mathbb{Z}$. There exists a norm ν on $\mathbb{C}^n$ such that*

$$H(\nu) = \{h | h = \alpha I, \alpha \text{ real}\}.$$

Proof. In Theorem 4.7 let S be the span of I.

In the next section we will strengthen this last result.

5. NORMS WITH TRIVIAL HERMITIANS

Recalling that $N_1 = N_1(n)$ is a metric space of norms in $\mathbb{C}^n$, we can strengthen Corollary 4.55 as follows:

THEOREM 5.1. *For each positive $n \in \mathbb{Z}$, the norms in $N_1(n)$ which permit only real multiples of the identity as Hermitians are dense in $N_1(n)$.*

Proof. Since the result is clearly true for $n = 1$, we may suppose $n \geqslant 2$. Let η be a norm in $N(n)$. By Theorem 1.7 there is a basis $\{e_i\}$ for $\mathbb{C}^n$ which is double dual with respect to η, and we use this basis for the rest of the proof. Without loss of generality, we assume that the $\{e_i\}$ are the canonical unit vectors.

Let $\nu_\varepsilon = \frac{1}{2}\nu_\varepsilon'$ where ν_ε' denotes the norm (4.3) in the case that S is the span of I and $0 < \varepsilon < \frac{1}{4}$ is the number entering in the description of the space L of functionals. Since $\lambda \in S$ has the form $\alpha \cdot I$ for $\alpha \in R$, the expression $|\langle l, e^{i\lambda}x\rangle|$ equals $|\langle l, x\rangle|$ and thus

$$\nu_\varepsilon(x) = \tfrac{1}{2} \sup_{l \in L} |\langle l, x\rangle|. \tag{5.2}$$

As S gives rise to just one equivalence class C_1 consisting of the integers $\{1, 2, \ldots, n\}$, if $x = \sum_{i=1}^n \alpha_i e_i$ then

$$\begin{aligned} \nu_\varepsilon(x) &= \tfrac{1}{2} \sup_{\substack{1 \leqslant i,j \leqslant n \\ i \neq j}} |2\alpha_i \pm 2\varepsilon\alpha_j| \\ &= \sup_{\substack{1 \leqslant i,j \leqslant n \\ i \neq j}} |\alpha_i \pm \varepsilon\alpha_j|. \end{aligned} \tag{5.3}$$

If $|\cdot|_\infty$ denotes the l_∞ norm with respect to the basis $\{e_i\}$, then $\nu_\varepsilon(x) \leqslant (1 + \varepsilon)|x|_\infty$ with equality occurring for $x = e_1 + e_2$.

By Lemma 1.8, the norm η satisfies $|x|_\infty \leqslant \eta(x)$, where $|x|_\infty$ is with respect to the double dual basis $\{e_i\}$. Now we define the norm ρ_ε by

$$\rho_\varepsilon(x) = \sup(\eta(x), (1 + \varepsilon)\nu_\varepsilon(x)). \tag{5.4}$$

Since

$$\begin{aligned} \eta(x) &\leqslant \sup(\eta(x), (1 + \varepsilon)\nu_\varepsilon(x)) \\ &\leqslant \sup(\eta(x), (1 + \varepsilon)^2|x|_\infty) \\ &\leqslant \sup(\eta(x), (1 + \varepsilon)^2\eta(x)) \\ &= (1 + \varepsilon)^2\eta(x), \end{aligned}$$

we obtain $\eta(x)/\rho_\varepsilon(x) \leqslant 1$ and $\rho_\varepsilon(x)/\eta(x) \leqslant (1 + \varepsilon)^2$ for all x, or, using the semi-metric introduced in Definition 1.9,

$$d(\eta, \rho_\varepsilon) \leqslant 2 \log(1 + \varepsilon) \leqslant 2\varepsilon. \tag{5.5}$$

If we now show that ρ_ε allows only multiples of the identity as Hermitian elements, the density of such norms in $N(n)$, and consequently in $N_1(n)$, follows from (5.5).

Since $\eta(e_i) = \nu_\varepsilon(e_i) = 1$ for each basis vector e_i, the norm $(1 + \varepsilon)\nu_\varepsilon$ in (5.4) is active at each basis vector (cf. Definition 1.4). Then from Lemmas 1.3 and 1.5 it follows that there is a Euclidean neighborhood U_i of each basis vector e_i so that if $x \in U_i$ and $y||x$ with respect to ν_ε (or equivalently $(1+\varepsilon)\nu_\varepsilon$) then $y||x$ with respect to μ_ε. But for any pair of indices $1 \leqslant i, j \leqslant n$, $i \neq j$, the neighborhood U_i contains vectors $x_\pm = z_1 e_i \pm z_2 e_j$ where z_1, z_2 satisfy (4.8). As vector operations are continuous, z_1 can be allowed to vary in an open subset $U' \subset U$, where U was described in connection with (4.8). But with $l_1 \in L$ as described before (4.8), $l_1||x_+$ with respect to ν_ε. Likewise $l_2||x_-$ where l_2 differs from l_1 only in having ε replaced by $-\varepsilon$. Then the equations (4.42) through (4.48) are valid if h is a ρ_ε-Hermitian matrix, forcing h to be a real multiple of the identity.

The remainder of this section is aimed at showing the openness of the set of norms admitting only trivial Hermitians.

Remarks on Convexity 5.6. If y is a linear functional and c is a real number the set

$$A = \{x \in \mathbb{C}^n | \mathrm{Re}\langle y, x\rangle = c\}$$

is an *affine hyperplane*. If $\mathbb{C}^n$ is regarded as a $2n$-dimensional linear space over the real numbers, then A has codimension one. If K is a convex set in $\mathbb{C}^n$ one says that A is a *supporting hyperplane* for K at x_0 if

(1) $$\mathrm{Re}\langle y, x_0\rangle = c$$

and either

(2) $$x \in K \Rightarrow \mathrm{Re}\langle y, x\rangle \leqslant c$$

or

(2′) $$x \in K \Rightarrow \mathrm{Re}\langle y, x\rangle \geqslant c$$

are satisfied. Suppose K is a convex set which is *balanced*; that is, has the property that $x \in K$ implies $e^{i\theta}x \in K$ for all real θ. Then if K is supported by a hyperplane A at x_0 (in the sense of (1) and (2) above) the condition

$\mathrm{Re}\langle y, x_0\rangle = c$ implies $\langle y, x_0\rangle = c$. Otherwise one could obtain $\mathrm{Re}\langle y, e^{i\theta}x_0\rangle > c$ for a suitable θ.

If ρ is a norm on $\mathbb{C}^n$ we let

$$K_\rho = \{x \in \mathbb{C}^n | \rho(x) \leqslant 1\}.$$

Clearly K_ρ is balanced and from the foregoing discussion one easily derives the known result that given x with $\mu(x) = 1$, one has $y||x$ with respect to ρ if and only if $\{\tilde{x}|\mathrm{Re}\langle y, \tilde{x}\rangle = 1\}$ is a supporting hyperplane for K_ρ at x.

LEMMA 5.7. *Let ρ_0 be a norm and suppose $y_0||x_0$ with respect to ρ_0. Then, given $\varepsilon > 0$, there is a $\delta > 0$ such that if ρ is a norm and $d(\rho, \mu_0) < \delta$, there are vectors x, y satisfying $\chi(x - x_0) < \varepsilon$, $\chi(y - y_0) < \varepsilon$, and $y||x$ with respect to ρ.*

Recall that χ is the Euclidean norm.

Proof. We use the canonical basis $\{e_i\}$ for $\mathbb{C}^n$ so that with $x = \sum \alpha_i e_i$ and $y = \sum \beta_i e_i$, $\langle y, x\rangle = \sum \alpha_i \bar{\beta}_i$ and $\langle x, x\rangle = \chi^2(x)$.

Suppose that $y_0||x_0$ with respect to ρ_0 and that $\rho_0(x_0) = \rho_0{}^D(y_0) = 1$. We let

$$B_\omega = \{x \in \mathbb{C}^n | \chi(x - x_0 - y_0) \leqslant (1 + \omega)\chi(y_0)\},$$

where $0 \leqslant \omega < \omega_1 < 1$ and ω_1 is chosen so that the ball B_{ω_1} does not contain $x = 0$. Since y_0 is dual to x_0, we have $\mathrm{Re}\langle y_0, x\rangle \leqslant 1$ for each x in the unit ball K_{ρ_0}. If $x \in K_{\rho_0} \cap B_\omega$ and we write $x = x_0 + e$, then $\mathrm{Re}\langle y_0, e\rangle \leqslant 0$. Since $\chi(x - x_0 - y_0) \leqslant (1 + \omega)\chi(y_0)$, we have

$$\begin{aligned}(1 + \omega)^2\chi^2(y_0) &\geqslant \chi^2(x - x_0 - y_0)\\ &= \langle e - y_0, e - y_0\rangle\\ &= \langle e, e\rangle - 2\,\mathrm{Re}\langle y_0, e\rangle + \langle y_0, y_0\rangle\\ &\geqslant \chi^2(e) + \chi^2(y_0),\end{aligned}$$

yielding

$$\chi^2(e) \leqslant [(1 + \omega)^2 - 1]\chi^2(y_0)$$

or

$$\chi(e) \leqslant (3\omega)^{1/2}\chi(y_0). \tag{5.8}$$

Since $0 \notin B_{\omega_1}$, an intermediate value theorem allows one to choose $0 < \xi < 1$ so that, letting $x_1 = (1 - \xi)x_0$, the equality $\chi(x_1 - x_0 - y_0) = (1 + \omega_1)\chi(y_0)$ holds and hence, $x_1 \in B_{\omega_1}$. Now, with $0 < \delta < \xi/2$ we let ρ be any norm satisfying $d(\rho, \rho_0) < \delta$ with ρ normalized so that $\sup \rho_0(y)/\rho(y) = 1$. Recalling the definition of $d(\rho, \rho_0)$, we see that $\rho(x) \leqslant e^{\delta}\rho_0(x)$ for any x and consequently

$$\begin{aligned} \rho(x_1) &= \rho((1 - \xi)x_0) \\ &\leqslant e^{\delta}\rho_0((1 - \xi)x_0) \\ &\leqslant e^{\xi/2}(1 - \xi) \\ &\leqslant 1. \end{aligned}$$

We now have $x_1 \in K_\rho \cap B_{\omega_1}$ and we have normalized ρ so that $K_\rho \subseteq K_{\rho_0}$. As a consequence of inequality (5.8) with $\omega = 0$, one sees that the Euclidean distance from $x_0 + y_0$ to the set K_{ρ_0} is precisely the distance from $x_0 + y_0$ to x_0; that is, $\chi(y_0)$. Since $K_\rho \subseteq K_{\rho_0}$, the distance from $x_0 + y_0$ to K_ρ must be $(1 + \omega_2)\chi(y_0)$ for some $0 \leqslant \omega_2 \leqslant \omega_1$ and it is a standard result that the minimum distance is achieved at a unique point which we call x_2. It is known (cf. [15, p. 98]) that there is a hyperplane which is supporting for both B_{ω_2} and K_ρ at their common point x_2. Since the ball B_{ω_2} has a unique supporting hyperplane at x_2:

$$\{\tilde{x} | \mathrm{Re}\langle y_1, \tilde{x}\rangle = c\}$$

with $y_1 = x_0 + y_0 - x_2$ and $c = \mathrm{Re}\langle y_1, x_2\rangle$, it must be a supporting hyperplane for K_ρ. Denoting $c^{-1}y_1$ by y_2 we have $\rho(x_2) = 1$, $\langle y_2, x_2\rangle = \mathrm{Re}\langle y_2, x_2\rangle = 1$ and $\rho^D(y_2) = 1$, so y_2 and x_2 are dual with respect to ρ. It remains to be seen how far they are from y_0 and x_0, respectively.

Recall that δ depended upon ξ, and ξ, in turn, upon ω_1. Since x_2 is in $B_{\omega_2} \cap K_{\rho_0}$, $\chi(x_2 - x_0) < (3\omega)^{1/2}\chi(y_0)$, using (5.8). In terms of $x_2 - x_0$ we have

$$\begin{aligned} \chi(y_2 - y_0) &= \chi\left(\frac{1}{c}y_1 - y_0\right) \\ &= \chi\left(\frac{1}{c}((x_0 + y_0) - x_2) - y_0\right) \\ &= \chi\left(\frac{1}{\mathrm{Re}\langle y_0 - (x_2 - x_0), (x_2 - x_0) + x_0\rangle}(y_0 - (x_2 - x_0)) - y_0\right). \end{aligned}$$

Since $\operatorname{Re}\langle y_0, x_0\rangle = 1$, one easily sees that by choosing $\omega_1 > 0$ sufficiently small one can make $\chi(x_2 - x_0) < \varepsilon$ and $\chi(y_2 - y_0) < \varepsilon$. The value δ arising from the choice of ω_1, and the dual vectors $x = x_2$ and $y = y_2$ then serve for the conclusion of the lemma.

DEFINITION 5.9. *Let (X, d) be a metric space and let $\mathscr{S}$ be the collection of subspaces of a normed linear space V. We say that a map m from X to $\mathscr{S}$ is* upper semi-continuous *if the two conditions*

(1) $x_k \in X$ $(k = 1, 2, \ldots)$ *and x_k converges to x,*

(2) $v_k \in V$, $v_k \in m(x_k)$, *and v_k converges to v,*

imply that $v \in m(x)$.

In the following we identify the space $\mathbb{C}^{nn}$ of $n \times n$ complex matrices with the Euclidean space $\mathbb{C}^{n^2}$ and let $\mathscr{S}$ be the collection of subspaces of $\mathbb{C}^{nn}$.

THEOREM 5.10. *The map H on N_1 taking a norm ν to $H(\nu) \subset \mathbb{C}^{nn}$ is upper semi-continuous.*

Proof. Suppose ρ_k $(k = 1, 2, \ldots)$ is a sequence of norms converging to ρ, that h_k is Hermitian with respect to ρ_k, and that h_k converges to $h \in \mathbb{C}^{nn}$. We must show that $h \in H(\rho)$. We have a continuous map of $\mathbb{C}^n \times \mathbb{C}^{nn} \times \mathbb{C}^n$ (with the product topology) into $\mathbb{C}$ defined by taking (y, g, x) to $\langle y, gx\rangle$. Given $h \in \mathbb{C}^{nn}$, from above, and $y_0 \| x_0$ with respect to ρ, let $\langle y_0, hx_0\rangle = z$. If, for some dual pair (y_0, x_0), z is not real, then there exists an $\xi > 0$ so that the disc $D_\xi = \{w \in \mathbb{C} \mid |w - z| < \xi\}$ contains no real number. By the continuity of the expression $\langle y, gx\rangle$, there is an $\varepsilon > 0$ so that $\chi(y - y_0) < \varepsilon$, $\chi(x - x_0) < \varepsilon$, and $\chi(g - h) < \varepsilon$ imply $\langle y, gx\rangle \in D_\xi$. But using the hypotheses of the theorem together with the previous lemma, we can find an integer k and vectors y_k, x_k so that $\chi(h_k - h) < \varepsilon$, $\chi(x_k - x_0) < \varepsilon$, $\chi(y_k - y_0) < \varepsilon$, and $y_k \| x_k$ with respect to ρ_k. Then $\langle y_k, h_k x_k\rangle \in D_\xi$ and is also real, a contradiction. Hence z is real for each dual pair (y_0, x_0) and h is in $H(\rho)$.

THEOREM 5.11. *The set*

$$C = \{\rho \in N_1 \mid \exists h \in H(\rho),\ h \neq \alpha I,\ \alpha \in \mathbb{R}\}$$

is closed in N_1.

Proof. Suppose that $\rho_m \in C$ $(m = 1, 2, \ldots)$ and that $d(\rho_m, \rho) \to 0$ as $m \to \infty$ for some $\rho \in N_1$. By hypothesis, each space $H(\rho_m)$ contains a

Hermitian element h_m which is not a real multiple of I. With $(\cdot, \cdot)$ denoting the $\mathbb{C}^{n^2}$ inner product we may assume that $(h_m, I) = 0$ and $(h_m, h_m) = \chi^2(h_m) = 1$. Let

$$T = \{g \in \mathbb{C}^{nn} | \chi(g - \alpha I) < \tfrac{1}{2}, \text{ for some } \alpha \in \mathbb{R}\}$$

and

$$B = \{g \in \mathbb{C}^{nn} | \chi(g) \leqslant 2\}.$$

The set $B - T$ contains h_m for each integer m and, as $B - T$ is compact, we may assume without loss of generality that the matrices h_m converge as $m \to \infty$ to some $h \in B - T$. Using Theorem 5.10 we see that $h \in H(\rho)$, and as $h \notin T$ it follows that $\rho \in C$.

THEOREM 5.12. *The set of norms*

$$\{\rho \in N_1 | h \in H(\rho) \Rightarrow h = \alpha I, \alpha \in \mathbb{R}\}$$

is a dense, open subset of N_1.

Proof. The result is immediate from Theorem 5.1 and Theorem 5.11.

APPENDIX

We use the notation introduced in Section 2 and in addition let $\mathbb{R}^{nm}$ and $\mathbb{Z}^{nm}$ denote the spaces of matrices having real and integral entries, respectively, and having n rows and m columns. If $L \in \mathbb{R}^{np}$ and $M \in \mathbb{R}^{nm}$, then $[L, M]$ will denote the element of $\mathbb{R}^{n,m+p}$ obtained by situating L to the left of M. Similarly, if $L \in \mathbb{R}^{pm}$ and $M \in \mathbb{R}^{nm}$ then we denote by $\begin{bmatrix} L \\ M \end{bmatrix}$ the element of $\mathbb{R}^{n+p,m}$ obtained by situating L above M.

The four theorems stated below are all concerned with simultaneous diophantine approximation. The first three can be found in the references cited, but it is the fourth version we require and, while the result may be known to workers in number theory, we could find no reference for it. We thought it would be worthwhile to give all four theorems and show how each can be obtained starting with a restatement of the Perron result. (Perron states the result in terms of the rank and rational rank of the matrix M.) By $\mathbb{R}^n$ we denote both the space of row and column n-tuples with real elements; the context indicates which is intended.

THEOREM I (Perron [11, Theorem 63, p. 153, 1st edition, or Theorem 64, p. 159, 4th edition]). *The following two statements are equivalent for a given* $\alpha \in \mathbb{R}^n$ *and* $M \in \mathbb{R}^{nm}$:

A1. *For each* $\varepsilon > 0$ *there is a* $q \in \mathbb{Z}^m$ *such that* $|Mq - \alpha|_\infty < \varepsilon$.

B1. *For any* $\beta \in \mathbb{R}^n$,

$$(1) \qquad \beta M = 0 \in \mathbb{R}^m \Rightarrow \beta\alpha = 0 \in \mathbb{R},$$

and

$$(2) \qquad \beta M \in \mathbb{Z}^m \Rightarrow \beta\alpha \in \mathbb{Z}.$$

THEOREM II (Cassels [5, p. 53], Koksma [9, p. 83]). *Given* $\alpha \in \mathbb{R}^n$ *and* $M \in \mathbb{R}^{nm}$ *the following are equivalent*:

A2. *For each* $\varepsilon > 0$ *there is a* $q \in \mathbb{Z}^m$ *such that* $||M_q - \alpha|| < \varepsilon$.

B2. *For any* $\beta \in \mathbb{Z}^n$, $\beta M \in \mathbb{Z}^m \Rightarrow \beta\alpha \in \mathbb{Z}$.

Proof. (Assuming Theorem I).

A2 $\Rightarrow$ B2. Suppose $\beta \in \mathbb{Z}^n$ and $\beta M \in \mathbb{Z}^m$. Statement A2 says that for any given $\varepsilon > 0$ there is a $q \in \mathbb{Z}^m$, a vector $z \in \mathbb{Z}^n$, and a vector $\mu \in \mathbb{R}^n$ with $|\mu|_\infty < \varepsilon$ such that

$$Mq = \alpha + \mu + z.$$

Then

$$\beta Mq = \beta\alpha + \beta\mu + \beta z.$$

Since the terms βMq and βz are integers and since $|\beta\mu|_\infty < |\beta|_\infty n\varepsilon$, we can make $|\beta\mu|_\infty$ arbitrarily small by choosing $\varepsilon > 0$ small. Thus the number $\beta\alpha$, which is independent of ε, must be an integer.

B2 $\Rightarrow$ A2. Now suppose that B2 is satisfied and let $[I, M] = K$ be in $\mathbb{R}^{n,m+n}$. Clearly $\beta K = 0$ only if $\beta = 0$, so B1, Part (1) is satisfied. If $\beta \in \mathbb{R}^n$ and $\beta K \in \mathbb{Z}^{m+n}$, then $\beta \in \mathbb{Z}^n$ and $\beta M \in \mathbb{Z}^m$ which, by B2, implies $\beta\alpha \in \mathbb{Z}$. We have verified B1 and hence A1 holds. That is, there is a $q \in \mathbb{Z}^{m+n}$ such that $|Kq - \alpha|_\infty < \varepsilon$. Otherwise stated, $Mq = \alpha + \mu - q$ where $|\mu|_\infty < \varepsilon$. But then $||Mq - \alpha|| < \varepsilon$, showing that A2 is satisfied.

THEOREM III (Koksma [9, p. 83]). *Given* $\alpha \in \mathbb{R}^n$ *and* $M \in \mathbb{R}^{nm}$, *the following are equivalent*:

A3. *For each* $\varepsilon > 0$ *there is an* $x \in \mathbb{R}^m$ *such that* $||Mx - \alpha|| < \varepsilon$.

B3. *For any* $\beta \in \mathbb{Z}^n$, $\beta M = 0 \in \mathbb{Z}^m \Rightarrow \beta\alpha \in \mathbb{Z}$.

Proof. (Using Theorem II).

A3 $\Rightarrow$ B3. The proof is similar to that above, except that in this case βMx is the integer 0.

B3 $\Rightarrow$ A3. Suppose that B3 holds and that M has the special form

$$M = \begin{bmatrix} I \\ L \end{bmatrix}$$

where I is $m \times m$ and L is in $\mathbb{R}^{n-m,m}$ (note that in this case $n \geqslant m$). Write α in the form $\begin{bmatrix} \alpha^1 \\ \alpha^2 \end{bmatrix}$ where $\alpha^1 \in \mathbb{R}^m$ and $\alpha^2 \in \mathbb{R}^{n-m}$. Let γ be in $\mathbb{Z}^{n-m}$ and suppose $\gamma L \in \mathbb{Z}^m$. Let $\beta = [-\gamma L, \gamma]$ be the n-tuple having its first m entries equal to those of $-\gamma L$ and the last $n - m$ equal to those of γ. By assumption, $\beta \in \mathbb{Z}^n$ and $\beta M = -\gamma LI + \gamma L = 0 \in \mathbb{Z}^m$. Since we are assuming B3, we must have $\beta\alpha = -\gamma L\alpha^1 + \gamma\alpha^2 \in \mathbb{Z}$. Now, assuming that $\gamma \in \mathbb{Z}^{n-m}$ and that $\gamma L \in \mathbb{Z}^m$ we have shown that $\gamma(\alpha^2 - L\alpha^1) \in \mathbb{Z}$. Thus B2 holds for the matrix L and the vector $\alpha^2 - L\alpha^1 \in \mathbb{R}^{n-m}$. Hence A2 holds for the same pair. That is, for each $\varepsilon > 0$ there is a $q \in \mathbb{Z}^m$ such that $||Lq - (\alpha^2 - L\alpha^1)|| < \varepsilon$ or letting $x = q + \alpha^1$, $||Lx - \alpha^2|| < \varepsilon$. But $||Ix - \alpha^1|| = ||q + \alpha^1 - \alpha^1|| = 0$. Hence $||Mx - \alpha|| < \varepsilon$. This completes the proof in the case that

$$M = \begin{bmatrix} I \\ L \end{bmatrix}.$$

Suppose that M has rank m (and so $n \geqslant m$). Then by reordering rows, if necessary one can assume that

$$M = \begin{bmatrix} M_1 \\ M_2 \end{bmatrix}$$

where M_1 is a nonsingular $m \times m$ matrix and M_2 is $(n - m) \times m$. Write M as $\begin{bmatrix} I \\ L \end{bmatrix} M_1$ where $L = M_2 M_1^{-1}$. We are supposing that $\beta \in \mathbb{Z}^n$ and $\beta M = 0 \in \mathbb{Z}^m$ together imply $\beta\alpha = 0$. If $\beta \in \mathbb{Z}^n$ and

$$\beta \begin{bmatrix} I \\ L \end{bmatrix} = 0,$$

then $\beta M = 0$ and hence $\beta\alpha = 0$. According to what was proved in the special case above, given $\varepsilon > 0$, there is a $y \in \mathbb{R}^m$ such that

$$\left\| \begin{bmatrix} I \\ L \end{bmatrix} y - \alpha \right\| < \varepsilon.$$

Letting $x = M_1^{-1}y$ one has $\|Mx - \alpha\| < \varepsilon$. This completes the proof in the case that rank $M = m$.

Finally suppose M has rank $k < m$ (this includes the case $n < m$). By a permutation of rows and columns, if necessary, we can assume M is partitioned as

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}$$

where M_{11} is a $k \times k$ nonsingular matrix. We can then write

$$M = \begin{bmatrix} M_{11} & 0 \\ M_{21} & 0 \end{bmatrix} \cdot \tilde{M}$$

where $\tilde{M} \in \mathbb{R}^{mm}$ is nonsingular. Suppose that $\beta \in \mathbb{Z}^n$ and

$$\beta \begin{bmatrix} M_{11} \\ M_{21} \end{bmatrix} = 0.$$

Then $\beta M = 0 \in \mathbb{Z}^m$ and, as we are assuming B3, $\beta\alpha \in \mathbb{Z}$. Then from the previously established case, there is a $w^1 \in \mathbb{R}^k$ such that

$$\left\| \begin{bmatrix} M_{11} \\ M_{21} \end{bmatrix} w^1 - \alpha \right\| < \varepsilon.$$

We let w be the vector in $\mathbb{R}^m$ having its first k entries those of w^1 and the remaining ones zero. Then

$$\left\| \begin{bmatrix} M_{11} & 0 \\ M_{21} & 0 \end{bmatrix} w - \alpha \right\| < \varepsilon$$

and letting $x = \tilde{M}^{-1}w$ we have $\|Mx - \alpha\| < \varepsilon$.

THEOREM IV. *Given $\alpha \in \mathbb{R}^n$ and $M \in \mathbb{R}^{nm}$, the following are equivalent:*

A4. *For each $\varepsilon > 0$ and each real t, there is an $x \in \mathbb{R}^m$ such that $\|Mx - t\alpha\| < \varepsilon$.*

B4. *For any $\beta \in \mathbb{Q}^n$, $\beta M = 0$ in $\mathbb{Z}^m \Rightarrow \beta\alpha = 0$.*

Proof. (Using Theorem III).

A4 $\Rightarrow$ B4. Suppose A4 holds, and for a fixed $\beta \in \mathbb{Q}^m$, $\beta M = 0$. Then for some integer p, $p\beta \in \mathbb{Z}^m$ and $p\beta M = 0$. Choose a non-zero t so that

$|p\beta t\alpha| < \frac{1}{2}$. Then given ε such that $0 < 2pn|\beta|_\infty\varepsilon < 1$, there is an x such that $Mx = t\alpha + z + \mu$ where $z \in \mathbb{Z}^n$ and $|\mu|_\infty < \varepsilon$. Hence $0 = p\beta Mx = p\beta t\alpha + p\beta z + p\beta\mu$, and so $|p\beta z| \leqslant |p\beta t\alpha| + |p\beta\mu| \leqslant \frac{1}{2} + pn|\beta|_\infty\varepsilon < 1$. Since $p\beta z \in \mathbb{Z}$, we obtain $p\beta z = 0$. Thus $|p\beta t\alpha| = |p\beta\mu| \leqslant pn|\beta|_\infty\varepsilon$. Since ε can be chosen arbitrarily small, it follows that $\beta\alpha = 0$.

B4 $\Rightarrow$ A4. Assuming B4, we see that $\beta \in \mathbb{Z}^m$ and $\beta M = 0 \in \mathbb{Z}^m$ together imply that $\beta t\alpha = 0 \in \mathbb{Z}$; for every $t \in \mathbb{R}$. Hence B3 holds with α replaced by $t\alpha$, for $t \in \mathbb{R}$. The implication (B3 $\Rightarrow$ A3) now yields A4.

We wish to thank B. D. Saunders for his comments on a draft of this paper.

REFERENCES

1 F. L. Bauer, On the field of values subordinate to a norm, *Numer. Math.* 4(1962), 103–113.
2 G. Birkhoff, Extensions of Jentzsch's theorem, *Trans. Amer. Math. Soc.* 85(1957), 219–227.
3 H. F. Bohnenblust and S. Karlin, Geometrical properties of the unit sphere of Banach algebras, *Ann. of Math.* 62(1955), 217–229.
4 F. F. Bonsall and J. Duncan, Numerical ranges of operators and of elements of normed algebras, *London Math. Soc., Lecture Note Series 2* (1971).
5 J. W. S. Cassels, *An Introduction to Diophantine Approximation*, Cambridge Tract 45, Cambridge University Press (1965).
6 M. J. Crabb, Some results on the numerical range of an operator, *London Math. Soc.* 2, Series 2 (1970), 741–745.
7 M. J. Crabb, J. Duncan, and C. M. McGregor, Mapping theorems and the numerical radius, *Proc. London Math. Soc.* (3) 25(1972), 486–502.
8 N. Dunford and J. T. Schwartz, *Linear Operators*, Part I, Interscience, New York (1958).
9 J. F. Koksma, Diophantische Approximationen, *Ergebnisse Math. Grenzgebiete* 4(1936), 4.
10 N. Nirschl and H. Schneider, The Bauer fields of values, *Numer. Math.* 6(1964), 355–365.
11 O. Perron, *Irrationalzahlen*, 1st Ed., W. de Gruyter, Berlin (1921); 4th Ed. (1960).
12 E. C. Posner and H. Schneider, Hyperplanes and prime rings, *Arch. Math.* 11(1960), 322–326.
13 H. Schneider, Bound norms, *Linear Algebra and Appl.* 3(1970), 11–22.
14 H. Schneider and R. E. L. Turner, Matrices hermitian for an absolute norm, *Linear and Multilinear Algebra* 1(1973), 9–31.
15 J. Stoer and C. Witzgall, Convexity and optimization in finite dimensions I, *Die Grundlehren Math. Wissenschaften* 163, Springer Verlag, New York, Berlin, Heidelberg (1970).

Received March 1, 1973