The Journal of Fourier Analysis and Applications

Volume 8, Issue 5, 2002

Schmidt's Conjecture on Normality for Dimension Two

Gavin Brown, William Moran, and Andrew D. Pollington

Communicated by Andrew M. Odlyzko

ABSTRACT. The aim of this article is to prove the two-dimensional case of Wolfgang Schmidt's conjecture for normality with respect to matrices. Specifically we show that if S and T are two by two almost integer ergodic matrices, then normality with respect to S implies normality with respect to T if and only if $S^p = T^q$ for some positive integers p and q.

1. Introduction

We prove here Wolfgang Schmidt's Conjecture, [10], concerning normality with respect to matrices, for the case when the matrices are 2×2. Before stating the conjecture, we recall the one-dimensional situation. A real number is *normal* with respect to a given (integer) base r if each digit block in the decimal expansion in base r occurs with the expected frequency; that is, a digit block of length n recurs with density r^{-n}. This can be expressed in terms of uniform distribution: x is normal in base r if and only if the sequence of fractional parts $(\{r^n x\})$ is uniformly distributed in the unit interval. Borel [2] showed that almost all real numbers are normal in every base.

In a similar vein, we define, following Schmidt, normality for members of $\mathbf{R}^n$ with respect to an $n \times n$ matrix S: $\mathbf{x} \in \mathbf{R}^n$ is *S-normal* if the sequence $(S^n\mathbf{x})$ when considered modulo 1 in each coordinate is uniformly distributed over the n-torus $\mathbf{T}^n$. It is relatively easy to see that, at this level of generality, Borel's theorem is no longer valid: there are many non-zero (and even non-singular) matrices for which no element of $\mathbf{R}^2$ is normal. We take the opportunity to introduce the notation $B(S)$ to denote the set of members of $\mathbf{R}^n$ which are S-normal. When almost every member of $\mathbf{R}^2$ is normal for a matrix S (that is, when the complement of $B(S)$ has measure zero), we again follow Schmidt in calling the matrix *ergodic*.

We are interested in describing when $B(S)$ and $B(T)$ coincide for two ergodic matrices S and T. It is relatively straightforward (cf. [10]) to see that, if $S^p = T^q$, for integers p

Math Subject Classifications. 11K16.
Keywords and Phrases. normal numbers, matrix normality, Riesz products, uniform distribution.

ISSN 1069-5869

and q then $B(S) = B(T)$. Schmidt's conjecture concerns the converse of this result. He restricts attention to rational non-singular matrices all of whose eigenvalues are algebraic integers which are not roots of unity. These are automatically ergodic (cf. [10]). Indeed among rational matrices with algebraic integer eigenvalues these are exactly the ergodic matrices. We follow Schmidt in calling them *almost integer ergodic*. Schmidt's conjecture can now be stated precisely.

Conjecture.

Let S and T be almost integer ergodic $n \times n$ matrices. Then, either there are positive integers p, q such that $S^p = T^q$, or there are uncountably many points in $\mathbf{R}^n$ which are normal in base S but not normal in base T (and conversely).

Schmidt, [9], established his conjecture in one dimension and, in [10], he proved it in all dimensions under the additional hypotheses that every eigenvalue of T has modulus strictly greater than one and that the two matrices commute. Brown and Moran [3] removed the hypothesis on the size of the eigenvalues, and so obtained the result for any pair of commuting almost integer matrices. Anne Bertrand [1] has proved Schmidt's conjecture for some special cases of non-commuting pairs of matrices. She is able to prove the conjecture if one of the matrices has an eigenvalue which is a PV-number and the other matrix has a dominant real eigenvalue and no eigenvalue of this second matrix is a rational power of this PV-number. In this article we describe a method which proves Schmidt's conjecture in full generality for the case of 2 by 2 matrices. Our method seems to generalize to higher dimensions with some considerable complications. We hope to give a full proof of the conjecture by again applying suitable Riesz products in a subsequent article.

Theorem.

Let S and T be almost integer ergodic 2×2 matrices. Either there are positive integers p, q such that $S^p = T^q$ or there are uncountably many points in $\mathbf{R}^n$ which are normal in base S but not normal in base T (and conversely).

The innovation which allowed Brown and Moran to remove the eigenvalue restriction was the use of Riesz product measures rather than the Cantor-like measures used by Schmidt. The results of this article also use the Riesz product technique. In addition, we introduce here some extensions which we have utilized in tackling similar problems for non-integer bases [4, 7]. We should add that invoking the results of [5] (and in some special cases [1]) would allow simplifications in a few places. We have refrained from doing that, however, since that articles appeals to Baker's theorem on linear forms in logarithms in an essential way. The exposition here is entirely elementary.

2. Preliminaries

In this section, we describe some of the basic machinery used in the proof. First we give some facts about almost integer ergodic matrices. Let S be such a 2×2 matrix. Then, by a result of Schmidt [10], for some positive integer d, dS^r is an integer matrix for all positive integers r. We shall call d the *denominator* of S. Thus, if we choose $\mathbf{t}$ to be of the form $d.\mathbf{w}$ where $\mathbf{w}$ is in $\mathbf{Z}^2$, then the sequence $\mathbf{t}S^r$ is in $\mathbf{Z}^2$. Note that we think of matrices as operating on $\mathbf{R}^2$ and on the quotient $\mathbf{T}^2$ by multiplication on the left. These matrices operate on the dual groups, and in particular $\mathbf{Z}^2$ by multiplication on the right.

We shall say that S is *dissociate* if, for each $\mathbf{s} \neq 0$ in the dual group $\widehat{\mathbf{T}^2} = \mathbf{Z}^2$ of $\mathbf{T}^2$, the sequence $\mathbf{s}S^n$ is a dissociate sequence in the sense of Hewitt and Zuckerman (see [5]); that is,

$$\sum_{r=1}^{N} \varepsilon_r \mathbf{s} S^r \neq 0 \tag{2.1}$$

for all choices of $\varepsilon_r \in \{0, \pm 1, \pm 2\}$, not all equal to 0, and all N. Fix a member $\mathbf{t}$ of $\mathbf{Z}^n$. The effect of dissociateness of $\mathbf{t}T^n$ is that all expressions of the form

$$\mathbf{r} = \sum_{j=1}^{N} \varepsilon_j \mathbf{t} S^j \tag{2.2}$$

where $\varepsilon_j = \pm 1$ or 0 are distinct members of $\mathbf{R}^2$.

At this point, we note the following lemma from [3].

Lemma 1.

Let T be an almost integer ergodic matrix. Then there is some positive integer k such that $(\mathbf{t}T^{rk})$ is dissociate for all $\mathbf{t} \neq 0$ in $\mathbf{Q}^n$.

In Section 3 we shall define a probability measure to discriminate between S and T normality, for almost integer ergodic matrices S and T. Specifically, we define μ such that almost every $\mathbf{x} \in \mathbf{T}^2$ is normal in base S and non-normal in base T. The non-normality will be forced by our construction. We need to be able to say when $\mathbf{x}$ is normal with respect to S for almost all $\mathbf{x}$ with respect to the measure μ. The next result of this section is the key ingredient in this. It is effectively the lemma of Davenport, Erdős and LeVeque (cf. [8]).

Lemma 2.

Let μ be a probability measure on $\mathbf{T}^2$. If, for all $\mathbf{s} \neq 0$ in $\mathbf{Z}^2$,

$$\sum_{N=1}^{\infty} \frac{1}{N^3} \sum_{k=1}^{N} \sum_{j=1}^{k} \left| \hat{\mu}\left(\mathbf{s}\left(S^k - S^j\right)\right)\right| < \infty\,, \tag{2.3}$$

then almost every (with respect to ρ) $\mathbf{x} \in \mathbf{R}^2$ is normal in base S.

We will require later the following technical number theoretic lemma.

Lemma 3.

Let α be a positive irrational number, and let (p_k/q_k) $(\gcd(p_k, q_x) = 1)$ be its sequence of partial quotients. Write

$$A = \bigcup_k (q_k, 2q_k) \cap (q_k, q_{k+1})\,. \tag{2.4}$$

Fix positive constants K and C. Then for R sufficiently large, there are fewer than $2^{R/2}$ pairs (u, v) of integers such that

$$\begin{cases} \alpha = \frac{u}{v} + \delta\,, \\ u \leq r \leq K 2^R\,, \quad r \in A\,, \\ |\delta| < e^{-cR^2}\,. \end{cases} \tag{2.5}$$

Proof. Assume, on the contrary, that there are infinitely many such integers R, and at least $2^{R/2}$ pairs (u, v) satisfying (2.5). It follows that

$$|\delta| \leq e^{-C'(\log r)^2} < \frac{1}{r^3} < \frac{1}{2v^2} \tag{2.6}$$

for sufficiently large u. By Lagrange's Theorem u/v must be a convergent of α. Consequently, there is some integer d so that $u = dp_k$ and $v = dq_k$ where

$$\frac{p_k}{q_k} = [a_0; a_1, \ldots, a_k] \ . \tag{2.7}$$

Moreover,

$$\frac{1}{r^3} > \left|\alpha - \frac{p_k}{q_k}\right| > \frac{1}{q_k(q_k + q_{k+1})} > \frac{1}{2uq_{k+1}} > \frac{1}{2rq_{k+1}} \tag{2.8}$$

(see [6], Theorem 9.9), so that $r < q_{k+1}$. Since $r \in A$, it is less than $2q_k$ and $u < 2q_k$. It follows that $v \leq K'q_k$ for some constant K'. Now we see that there are at most K' times as many pairs u, v as there are denominators q_k from the convergents of α in the interval of integers $[1, 2^R]$. Since the q_k's increase exponentially, this is at most $C''R$ for some constant C'', and so contradicts the statement there are $2^{R/2}$ pairs u, v. The proof of the lemma is complete. □

3. The Riesz Product Equation

In this section, we construct Riesz product measures appropriate for our purposes. Given any two almost integer ergodic matrices S and T which do not satisfy an equation like $S^p = T^q$, we shall find a probability measure μ such that $\mu(B(S)\backslash B(T)) = 0$. We do this by the Riesz product construction for probability measures on the torus $\mathbf{T}^2$.

Let T be an almost integer ergodic matrix. Replacing T by T^r if necessary we may and do assume that $(\mathbf{t}T^n)$ is a dissociate sequence in $\mathbf{Z}^2$ for all $\mathbf{t} \neq 0$. We shall make several different choices of Riesz products at various points of the proof. Each will be defined in terms of a subset A of $\mathbf{N}$ of positive lower density and an element $\mathbf{t}$ of $\mathbf{Z}^2$ which is divisible (that is, each of its coordinates is divisible) by the denominator $d(T)$ of the matrix T. Consider the trigonometric polynomial

$$P_N(\mathbf{x}) = \prod_{\substack{1 \leq j \leq N \\ j \in A}} \left(1 + \cos 2\pi \mathbf{t}T^j\mathbf{x}\right) \ . \tag{3.1}$$

This is non-negative and its integral over $\mathbf{T}^2$ is equal to 1. It is straightforward to see that, as a result of the dissociateness of T, the sequence of probability measures $P_N.m$ converges in the weak$*$ sense where m is Lebesgue measure on $\mathbf{T}^2$. The limit measure which we write as

$$\mu = \prod_{j \in A} \left(1 + \cos 2\pi \mathbf{t}T^j\mathbf{x}\right) \ .m \tag{3.2}$$

is called a *Riesz product.* In fact, it is a special case of a somewhat more general construction, [5]. In our circumstances μ is always singular to Lebesgue measure, and when it is regarded

as a measure on the 2-torus its Fourier coefficients are given by the formula

$$\hat{\mu}(\mathbf{r}) = \begin{cases} \left(\frac{1}{2}\right)^{l(\mathbf{r})} & \text{if } \mathbf{r} = \sum_{\substack{1\le j\le N \\ j\in A}} \varepsilon_j \mathbf{t}T^j \text{ and } l(\mathbf{r}) = \sum_{\substack{1\le j\le N \\ j\in A}} |\varepsilon_j| \\ 0 & \text{otherwise.} \end{cases} . \tag{3.3}$$

This formula, which follows from the dissociateness of the sequence, and the fact that μ is a continuous probability measure, will be all the properties we need of Riesz products.

The next result is a slight strengthening of a proposition which may also be found in [3]. It is a straightforward consequence of the Weyl criterion for uniform distribution.

Proposition 1.

Let T be dissociate and almost integer ergodic, let A be a subset of $\mathbf{N}$ *of positive lower density, and let* $\mathbf{t}$ *be a non-zero element of* $\mathbf{Z}^2$ *divisible by d where d is the denominator of T. Let μ be the Riesz product measure defined according to* (3.2). *Then, for μ-almost every* $\mathbf{x}$ *in* $\mathbf{T}^2$, $\mathbf{x}$ *is non-normal with respect to the base T.*

Let S be another 2×2 almost integer ergodic matrix. We denote the eigenvalues of S and T by $\sigma_1,\ \sigma_2$ and $\tau_1,\ \tau_2$, respectively, where we use the convention that, for any 2×2 matrices, the first eigenvalue is of larger or equal absolute value. At this point we make no assumptions as to the diagonalizability of either of the two matrices.

Let $d(T)$ be the denominator of T. We replace T by an appropriate power of itself, so that $(\mathbf{t}T^r)$ is a dissociate sequence for any integer vector $\mathbf{t} \neq 0$. We choose $\mathbf{t}$ to be divisible by $d(T)$ and A to be some subset of $\mathbf{N}$ which has positive lower density. At a later stage we shall be more specific about the nature of $\mathbf{t}$ and of A.

Now we construct a Riesz product μ as in (3.2). By Proposition 1, it follows that, for almost all $\mathbf{x} \in \mathbf{T}^2$ with respect to μ, $\mathbf{x}$ is not normal in base T. We make the assumption that normality in base T implies normality in base S, that is that $B(T) \subset B(S)$, and then conclude that, for some p and q, $S^p = T^q$. We recall that it follows from Lemma 6 of Schmidt [10] that the normal set of the matrix S is unchanged by multiplying each vector in the set by a non-singular rational matrix. Thus in the application of Lemma 2 one can assume that in the case of almost integer ergodic matrices that the vector $\mathbf{s}$ is a multiple of the denominator of S. This is enough to prove our main theorem under the assumption, by Lemma 2, we have

$$\sum_{N=1}^{\infty} \frac{1}{N^3} \sum_{n=1}^{N} \sum_{m=1}^{n} \hat{\mu}\left(\mathbf{s}\left(S^n - S^m\right)\right) = \infty \tag{3.4}$$

for some integer vector, $\mathbf{s} \neq 0$. This being the case, and because of the form of the Fourier coefficients of Riesz products given in Equation (3.3), there are infinitely many pairs (m, n) such that

$$\mathbf{s}\left(S^n - S^m\right) = \mathbf{t} \sum_j \varepsilon_j T^j \tag{3.5}$$

holds, where the sum is a finite sum over j's in the set A. We refer to this equation as the *Riesz Product Equation.*

To obtain more precise estimates on the density of such pairs, we now follow the argument in [7]. We write

$$\hat{\mu}\left(\mathbf{t} \sum_j \varepsilon_j T^j\right) = 2^{-r(n,m)} \tag{3.6}$$

where $r(n,m)$ is the number of non-zero ε_j's and $r(n,m) = \infty$ if no equality of the form (3.5) exists. Thus, if we let $t(n) = n/(\log n)^2$ then, under the assumption that the sum in (3.4) is infinite,

$$\sum_N \frac{1}{N^3} \sum_{0\le n\le N} \sum_{0\le m<n-t(n)} \frac{1}{2^{r(n,m)}} = \infty \,. \tag{3.7}$$

We will use this fact to derive the following technical result. We use the notation I_R for the dyadic interval $[2^R, 2^{R+1})$ of integers.

Proposition 2.

Assume that $B(T) \subset B(S)$. Then, for infinitely many positive integers R, there is a matrix W_R, which is a polynomial in T^{-1}, such that, for at least $2^{R-3(\log R)\sqrt{R}}$ numbers $n \in I_R$

$$\mathbf{s}\left(S^n - S^m\right) = \mathbf{t}T^{r(n)}(W_R + B_n) \,, \tag{3.8}$$

where $m < n - t(n)$, $r(n)$ is some integer in A, and

$$B_n = \sum_{j>M+2^{\sqrt{R}}} \varepsilon'_j T^{-j} \,, \tag{3.9}$$

where M is the degree of the polynomial W_R, where $\varepsilon'_j = 0, \pm 1$ for all j.

Proof. It follows from (3.7) that

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \sum_{0\le m<n-t(n)} \frac{1}{2^{r(n,m)}} = \infty$$

and hence that

$$\sum_{n\in I_R} \sum_{0\le m<n-t(n)} \frac{1}{2^{r(n,m)}} \ge \frac{2^{2R}}{R^2} \,,$$

for infinitely many R. For infinitely many R's, then

$$\sum_{0\le m<n-t(n)} \frac{1}{2^{r(n,m)}} \ge \frac{2^{R+1}}{R^2} \,,$$

for at least $2^R/R^2$ n's in I_R. Moreover, for each of these n's, there are at least $2^R/R^2$ m's satisfying

$$0 \le m \le n - t(n) \text{ and } r(n,m) \le 2\log 2R \,. \tag{3.10}$$

Fix a pair (n,m) satisfying these conditions, and suppose that (3.10) holds. Then, somewhere in the sequence

$$\varepsilon_{r(n,m)}, \varepsilon_{r(n,m)-1}, \ldots, \varepsilon_{r(n,m)-s(R)} \,,$$

where $2R$, there is a block of $2^{\sqrt{R}}$ zeros. Let us write $H(n,m)$ for the smallest i such that $\varepsilon_{r(n,m)-i}$ is in a block of $2^{\sqrt{R}}$ zeros. Since $r(n,m) - H(n,m)$ is less than $s(R)$, the number of possible choices of

$$\varepsilon_{r(n,m)}, \varepsilon_{r(n,m)-1}, \ldots, \varepsilon_{r(n,m)-H(n,m)}$$

does not exceed $\tau(R) = (2s(R))^{\log R}$. It follows that, for some

$$W_R = \sum_{k=0}^{M} \varepsilon'_k T^{-k} ,$$

where $\varepsilon'_k = 0, \pm 1$, there are at least $\Omega(2^R/\tau(R)R^2)$ n's in I_R and some $m = m(n)$ for which

$$\mathbf{s}\left(S^n - S^m\right) = \mathbf{t}T^{r(n)}(W_R + B_n) ,$$

where, because of the gap in the ε's,

$$B_n = \sum_{j > M + e^{\sqrt{R}}} \varepsilon'_j T^{-j} .$$

This completes the proof. □

4. The Diagonalizable Case: Equality of Eigenvalues

Our aim in this section is to prove the following result.

Proposition 3.

If S and T are diagonalizable and are such that $B(T) \subset B(S)$, then S and T have the same eigenvalues.

Since S and T are diagonalizable, we can find invertible matrices U and V and diagonal matrices Δ and Γ such that

$$S = U^{-1}\Delta U \quad \text{and} \quad T = V^{-1}\Gamma V . \tag{4.1}$$

Let $\mathbf{t}' = \mathbf{t}V^{-1}$ and $\mathbf{s}' = \mathbf{s}U^{-1}$, so that Equation (3.5) becomes

$$\mathbf{s}'\left(\Delta^n - \Delta^m\right) U V^{-1} = \mathbf{t}' \sum_{j=1}^{r(n)} \varepsilon_j \Gamma^j , \tag{4.2}$$

and Equation (3.8) becomes

$$\mathbf{s}'\left(\Delta^n - \Delta^m\right) U V^{-1} = \mathbf{t}'\Gamma^{r(n)}(W_R + B_n) . \tag{4.3}$$

Now substitute

$$U V^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \mathbf{s}' = (s_1, s_2), \quad \mathbf{t}' = (t_1, t_2) \tag{4.4}$$

in (4.2) and (4.3) to obtain

$$\begin{aligned} as_1\left(\sigma_1^n - \sigma_1^m\right) + cs_2\left(\sigma_2^n - \sigma_2^m\right) &= t_1 \sum_{j=1}^{r(n)} \varepsilon_j \tau_1^j \\ bs_1\left(\sigma_1^n - \sigma_1^m\right) + ds_2\left(\sigma_2^n - \sigma_2^m\right) &= t_2 \sum_{j=1}^{r(n)} \varepsilon_j \tau_2^j , \end{aligned} \tag{4.5}$$

and

$$\begin{aligned} as_1\left(\sigma_1^n-\sigma_1^m\right)+cs_2\left(\sigma_2^n-\sigma_2^m\right) &= t_1\tau_1^{r(n)}(w_R(\tau_1)+\beta_n(\tau_1)) \\ bs_1\left(\sigma_1^n-\sigma_1^m\right)+ds_2\left(\sigma_2^n-\sigma_2^m\right) &= t_2\tau_2^{r(n)}(w_R(\tau_2)+\beta_n(\tau_2))\ , \end{aligned} \tag{4.6}$$

where

$$w_R(\tau)=\sum_{i=0}^{M_R}\varepsilon_i\tau^{-i} \quad \text{and} \quad \beta_n(\tau)=\sum_{j=M_R+2\sqrt{R}}\varepsilon_i\tau^{-j}\ . \tag{4.7}$$

Note that the ε_i's in the expression for $w_R(\tau)$ depend only on R, whereas those in $\beta_n(\tau)$ depend on n. Moreover, if $|\tau|>3$,

$$|\beta_n(\tau)|\le Ce^{-2^{\sqrt{R}}}\ , \tag{4.8}$$

where $C=C(\tau)$.

We first show that n and $r(n)$ are comparable in size. To do this we need to make sure that when an eigenvalue exceeds 1, it in fact exceeds 3. We do this by replacing both S and T by an appropriate power. Such a substitution does not affect the statement of the theorem, though it does change the Riesz product. In any case, we may assume that S and T have been adjusted once for all in such a way that if $|\tau_2|>1$ or $|\sigma_2|>1$, then, in fact, they are larger than 3.

Lemma 4.

If both S and T are diagonalizable and $\mathbf{t}$ is not an eigenvector of T, then, for some positive real number α,

$$n\alpha-C\le r(n)\le n\alpha+C\ . \tag{4.9}$$

Proof. First we deal with the case when $|\tau_2|>1$. Note that, in this case, both $|\beta_n(\tau_1)|$ and $|\beta_n(\tau_2)|$ do not exceed $Ce^{-2^{\sqrt{R}}}$.

We solve (4.6) for $s_1(\sigma_1^n-\sigma_1^m)$, which we can do since UV^{-1} is invertible. If $s_1\neq 0$ we immediately obtain $n\alpha<r(n)+C$, where $\alpha=\log|\sigma_1|/\log|\tau_1|$. The reverse inequality is an immediate consequence of (4.6) in this case. If $s_1=0$, set $\alpha=\log|\sigma_2|/\log|\tau_1|$ and the full inequality follows immediately.

Now we deal with the case when $|\tau_2|\le 1$. Note that it follows that $|\tau_1|>1$ as we have argued above. Moreover, the eigenvalue condition implies that $t_1.t_2\neq 0$. In this case we turn to Equations (4.5). First we see quickly that $r(n)\le n\alpha+C$. Then we solve for $s_1(\sigma_1^n-\sigma_1^m)$ in terms of the right sides of those equations. A simple estimate gives the inequality $n\alpha-C\le r(n)$, where $\alpha=\log|\sigma_1|/\log|\tau_1|$, unless $s_1=0$. If $s_1=0$, we look again at Equations (4.5) and obtain the inequality with σ_2 in place of σ_1. □

We now continue the proof of Proposition 3. We choose $\mathbf{t}$ not to be an eigenvector of T. Using the fact that $m<n-t(n)$ in (4.6), we can incorporate the m terms in the "error" term, thus:

$$\begin{aligned} as_1\sigma_1^n+cs_2\sigma_2^n &= t_1\tau_1^{r(n)}\left(w_R(\tau_1)+\beta_n^{(1)}\right) \\ bs_1\sigma_1^n+ds_2\sigma_2^n &= t_2\tau_2^{r(n)}\left(w_R(\tau_2)+\beta_n^{(2)}\right)\ , \end{aligned} \tag{4.10}$$

where the $\beta_n^{(1)} < C\exp(-2^{\sqrt{R}})$ and, if $\tau_2 > 1$, then $\beta_n^{(2)} < C\exp(-2^{\sqrt{R}})$ [cf. (4.8)].

Suppose, for the moment, that $s_1 \neq 0$. Recall that there are $2^{R-4\sqrt{R}}$ such n's in the dyadic interval I_R. It follows by a simple counting argument that, for infinitely many R, there is some k and $4C2^{R/2}$ pairs (n, n') of elements of J_R such that $n - n' = k$. By inequality (4.9), as (n, n') varies over all such pairs , $r(n) - r(n')$ can take at most $4C$ different values. Accordingly, for $2^{R/2}$ pairs (n, n'), we have

$$n - n' = k \quad \text{and} \quad r(n) - r\left(n'\right) = l \tag{4.11}$$

say.

We pick two such pairs $(n, n+k)$ and $(n', n'+k)$ and, if $as_1 \neq 0$, we write down the equations in (4.10) for n and n'. This gives us the following four equations:

$$\begin{aligned}
as_1\sigma_1^n + cs_2\sigma_2^n &= t_1\tau_1^{r(n)}\left(w_R(\tau_1) + \beta_n^{(1)}\right) \\
as_1\sigma_1^{n'} + cs_2\sigma_2^{n'} &= t_1\tau_1^{r(n')}\left(w_R(\tau_1) + \beta_{n'}^{(1)}\right) . \\
as_1\sigma_1^{n+k} + cs_2\sigma_2^{n+k} &= t_1\tau_1^{r(n)+l}\left(w_R(\tau_1) + \beta_{n+k}^{(1)}\right) \\
as_1\sigma_1^{n'+k} + cs_2\sigma_2^{n'+k} &= t_1\tau_1^{r(n')+l}\left(w_R(\tau_1) + \beta_{n'+k}^{()}\right) .
\end{aligned} \tag{4.12}$$

We multiply the first two equations by σ_2^k and subtract from the last two equations to obtain

$$\begin{aligned}
as_1\sigma_1^n\left(\sigma_1^k - \sigma_2^k\right) &= t_1\tau_1^{r(n)}\left(\tau_1^l - \sigma_2^k\right)(w_R(\tau_1) + \delta_n) \\
as_1\sigma_1^{n'}\left(\sigma_1^k - \sigma_2^k\right) &= t_1\tau_1^{r(n')}\left(\tau_1^l - \sigma_2^k\right)(w_R(\tau_1) + \delta_{n'}) ,
\end{aligned} \tag{4.13}$$

where $|\delta_n| \leq Ce^{-2^{\sqrt{R}}}$ and $|\delta_n'| \leq Ce^{-e^{\sqrt{R}}}$. If either of the equations

$$\sigma_2^k = \sigma_1^k$$

or

$$\sigma_2^k = \tau_1^k$$

is true, then so is the other. Failing that, we may divide the first of the Equations (4.13) by the other and, after taking logs obtain,

$$\left(n - n'\right)\log\sigma_1 = \left(r(n) - r\left(n'\right)\right)\log\tau_1 + \nu_{n,n'} + 2M\pi i , \tag{4.14}$$

where $|\nu_{n,n'}| \leq Ce^{-2^{\sqrt{R}}}$ for some integer M. Taking real parts, we have

$$\left(n - n'\right)\log|\sigma_1| = \left(r(n) - r\left(n'\right)\right)\log|\tau_1| + \Re\left(\nu_{n,n'}\right) . \tag{4.15}$$

Note that this is true for $C2^{R/2}$ values of n and n'. At this point we shall argue as in [7], using Lemma 3, to obtain that $\alpha = \log|\sigma_1|/\log|\tau_1|$ is rational. To this end we need to make a choice of A which depends on the continued fraction expansion of α.

We shall assume that $\alpha = \log|\sigma_1|/\log|\tau_1|$ is irrational to obtain a contradiction. We let

$$\frac{p_r}{q_r} = [a_0; a_1, a_2, \ldots, a_r] \tag{4.16}$$

denote the rth partial quotient of α. Now we define

$$A = \bigcup_r (q_r, 2q_r) \cap (q_r, q_{r+1}) \,, \tag{4.17}$$

where (a, b) denotes the interval of integers d such that $a < d < b$. Observe that the sequence (q_r) increases at least exponentially, and that the upper density of A is positive, in fact, at least $1/2$. This is sufficient to guarantee the almost everywhere (μ) non-normality of $\mathbf{x}$ with respect to T. Now we use Lemma 3 to deduce that α is rational.

Thus we can and do replace S and T by appropriate powers so that $|\sigma_1| = |\tau_1|$.

In (4.15) we now see that $n - n' = r(n) - r(n')$. So that, in the imaginary part of (4.14), we obtain

$$\left(n - n'\right)(\arg\sigma_1 - \arg\tau_1) = 2\pi M + \Im\left(v_{n,n'}\right) \,.$$

Writing $\gamma = (\arg\sigma_1 - \arg\tau_1)/2\pi$ we have

$$\left|\gamma - \frac{M}{n - n'}\right| < C \exp\left(2^{-\sqrt{R}}\right) \tag{4.18}$$

for at least $C' 2^{R/2}$ pairs M and $n - n'$ with n and n' in $A \cap I_R$. Lemma 3 now applies to show that $\arg\sigma_1 - \arg\tau_1$ is a rational multiple of 2π. We may replace S and T by appropriate powers so that $\sigma_1 = \tau_1$. It follows that, if σ_1 is irrational then $\sigma_2 = \tau_2$. If σ_1 is rational it is integral as are σ_2, τ_1 and τ_2. Now we choose $\mathbf{t}$ so that $t_1 = 0$ and use the second equation in (4.10) to show that either σ_1 or σ_2 equals τ_2. Size considerations yield that $\sigma_2 = \tau_2$.

The remaining case when $as_1 \neq 0$ entails

$$\sigma_1^k = \sigma_2^k = \tau_1^l \,. \tag{4.19}$$

After taking powers of S and T we may assume that $\sigma_1 = \sigma_2 = \tau_1$. This implies that they are integers, as is τ_2. Now we use the second equation in (4.10) and repeat the above argument to show that some power of τ_2 is a power of σ_1. Size considerations again show that $\tau_2 = \tau_1$. This completes the proof Proposition 3 when $as_1 \neq 0$.

If $as_1 = 0$ but $bs_1 \neq 0$ then we may do the entire argument using the second equation of (4.10). If both as_1 and bs_1 are zero then $s_1 = 0$.

In this case we use the first equation and repeat the arguments to obtain, after taking powers, that $\sigma_2 = \tau_1$. In the irrational case again this implies $\sigma_1 = \tau_2$ and indeed that $|\sigma_1| = |\sigma_2| = |\tau_1| = |\tau_2|$. In the rational case we argue as above using a choice of $\mathbf{t}$ with $t_2 = 0$ to obtain $\sigma_2 = \tau_1$, and again $|\sigma_1| = |\sigma_2| = |\tau_1| = |\tau_2|$.

This completes the proof of Proposition 3.

5. The Diagonalizable Case: Proof of the Theorem

We assume now that S and T are diagonalizable and have the same eigenvalues. Our aim is to show that $S = T$. This will complete the proof of the theorem in the diagonalizable case.

At this point we again split the argument into two cases:

1. $|\sigma_1| > |\sigma_2|$ and by replacing each of S and T by a power of themselves, if necessary, $|\sigma_1| > 3|\sigma_2|$;

2. $|\sigma_1| = |\sigma_2|$.

Each of these case will in turn be split into integral and non-integral eigenvalue cases.

First we consider case (1). We refer back to Equations (4.5),

$$\begin{aligned} as_1\left(\sigma_1^n - \sigma_1^m\right) + cs_2\left(\sigma_2^n - \sigma_2^m\right) &= t_1 \sum_{j=1}^{r(n)} \varepsilon_j \sigma_1^j \ , \\ bs_1\left(\sigma_1^n - \sigma_1^m\right) + ds_2\left(\sigma_2^n - \sigma_2^m\right) &= t_2 \sum_{j=1}^{r(n)} \varepsilon_j \sigma_2^j \ , \end{aligned} \tag{5.1}$$

where first $\mathbf{t}$ has been chosen so that neither coordinate is zero. Recall that m is much smaller than n. Then $as_1 = t_1$ and so $b = 0$. In the non-integer eigenvalue case we may conjugate (5.1) and apply the same argument to obtain $c = 0$. For the integer eigenvalue case we again choose $\mathbf{t}$ to be an eigenvector for the smaller eigenvalue, that is, $t_1 = 0$. The Equations (5.1) then force $s_1 = 0$ and so $c = 0$.

In either case UV^{-1} is diagonal. Hence

$$ST^{-1} = U^{-1}\Delta U V^{-1}\Gamma^{-1}V = U^{-1}\Delta U V^{-1}\Delta^{-1}V = I \ . \tag{5.2}$$

This completes the proof in the case when $|\sigma_1|$ and $|\sigma_2|$ are unequal.

In case (2), when the absolute values are equal, σ_1 and σ_2 are complex conjugates, or, without loss of generality, are equal integers. First we deal with the latter. In this case both S and T are multiples, by σ_1, of the identity matrix and so are identical.

For the complex case, we assume that $\rho = \sigma_2/\sigma_1$ is not a root of unity, otherwise we could take a power and make $\sigma_1 = \sigma_2$ and be back in the former case. We reinitialize the process with a Riesz product where A is now defined in terms of the partial quotients in the continued fraction expansion of the argument of σ_2/σ_1. Equation (4.10) becomes

$$\begin{aligned} (as_1 - t_1 w_R(\sigma_1))\sigma_1^n + cs_2\bar{\sigma_1}^n &= t_1\sigma_1^n\beta_n^{(1)} \\ \left(bs_1\sigma_1^n + (ds_2 - t_2 w_R(\sigma_2)\right)\bar{\sigma_1}^n &= t_2\bar{\sigma_1}^n\beta_n^{(2)} \ . \end{aligned} \tag{5.3}$$

Note that $as_1 = t_1 w_R(\sigma_1)$ if and only if $cs_2 = 0$ and $ds_2 = t_2 w_R(\bar{\sigma_1})$ if and only if $bs_1 = 0$. Assume, in order to achieve a contradiction, that one of these fails and choose the corresponding equation. Without loss of generality we assume that $as_1 \neq t_1 w_R(\sigma_1)$ and use the first of the equations. Otherwise we would use the second.

Rewrite the first equation as

$$(as_1 - t_1 w_R(\sigma_1)) + cs_2\rho^n = t_1\beta_n^{(1)} \ , \tag{5.4}$$

and again as

$$cs_2\rho^n = (t_1 w_R(\sigma_1) - as_1)\left(1 + \beta_n'\right) \tag{5.5}$$

where

$$\beta_n' = \frac{\beta_n^{(1)}}{t_1 w_r(\sigma_1) - as_1} \ .$$

We take two such equations, for n and n' and divide the former by the latter to obtain

$$\rho^{n-n'} = 1 + \gamma_{n,n'} \ , \tag{5.6}$$

where again $|\gamma_{n,n'}| \leq e^{-2^{\sqrt{R}}}$. Using Lemma 3 again, we find that ρ is a root of unity, giving the required contradiction. It follows that

$$as_1 = t_1, \quad bs_1 = 0, \quad cs_2 = 0, \quad \text{and} \quad ds_2 = t_2 \, . \tag{5.7}$$

Now we argue as in the previous case to show that $S = T$.

6. The Non-Diagonalizable Case

Finally in this section we deal with the case when one of S and T is not diagonalizable. First we suppose that T is not diagonalizable but that S is. Then the eigenvalues of T are integers and its (one-dimensional) eigenspace contains integer vectors which are divisible by the denominator. Initially, we choose one of these for $\mathbf{t}$. Note that we may replace Γ in Equation (4.1), by

$$\Gamma = \begin{pmatrix} \tau & 1 \\ 0 & \tau \end{pmatrix} \, . \tag{6.1}$$

Then the Equations (4.5) become

$$\begin{aligned} as_1 \left(\sigma_1^n - \sigma_1^m\right) + cs_2 \left(\sigma_2^n - \sigma_2^m\right) &= 0 \\ bs_1 \left(\sigma_1^n - \sigma_1^m\right) + ds_2 \left(\sigma_2^n - \sigma_2^m\right) &= t_2 \sum_j \varepsilon_j \tau^j \, . \end{aligned} \tag{6.2}$$

Suppose first that $|\sigma_1| > |\sigma_2|$. This forces as_1 and cs_2 to equal zero. Since both a and c cannot be zero, either s_1 or s_2 equals zero and one of a and c equals zero. If s_1 equals zero, then arguing as in the diagonalizable case, we may assume $\sigma_2 = \tau$ and, if not, then $\sigma_1 = \tau$.

Now we may choose a vector $\mathbf{t}$ so that

$$\mathbf{t}' \Gamma^n = \left(\tau^n, n\tau^{n-1}\right) \, . \tag{6.3}$$

Then we have the following equations:

$$\begin{aligned} as_1 \left(\sigma_1^n - \sigma_1^m\right) + cs_2 \left(\sigma_2^n - \sigma_2^m\right) &= \sum_j \varepsilon_j \tau^j \\ bs_1 \left(\sigma_1^n - \sigma_1^m\right) + ds_2 \left(\sigma_2^n - \sigma_2^m\right) &= \sum_j \varepsilon_j j \tau^{j-1} \, . \end{aligned} \tag{6.4}$$

It is easily seen from size considerations that these are incompatible with $\sigma_j = \tau$ and so we arrive at a contradiction.

Now suppose that S is diagonalizable with eigenvalues both equal to σ. Then S is a multiple of the identity. The choice of $\mathbf{t}$ as an eigenvector of T gives

$$\begin{aligned} (as_1 + cs_2) \left(\sigma^n - \sigma^m\right) &= 0 \\ (bs_1 + ds_2) \left(\sigma^n - \sigma^m\right) &= \sum_j \varepsilon_j \tau^{j-1} \, . \end{aligned} \tag{6.5}$$

These give $(as_1 + cs_2) = 0$ and using the methods of the diagonalizable case we see that $\sigma = \tau$. The remainder of the argument is as that following Equation (6.4) and yields a contradiction.

It follows that S too cannot be diagonalizable and so we may assume that Δ is of the form

$$\Delta = \begin{pmatrix} \sigma & 1 \\ 0 & \sigma \end{pmatrix} . \tag{6.6}$$

This gives equations

$$\begin{aligned} (as_1 + cs_2)\left(\sigma^n - \sigma^m\right) + cs_1\left(n\sigma^{n-1} - m\sigma^{m-1}\right) &= t_1 \sum_j \varepsilon_j \tau^j \\ (bs_1 + ds_2)\left(\sigma^n - \sigma^m\right) + ds_1\left(n\sigma^{n-1} - m\sigma^{m-1}\right) &= t_1 \sum_j \varepsilon_j j\tau^j \\ &\quad + t_2 \sum_j \varepsilon_j \tau^j . \end{aligned} \tag{6.7}$$

First we choose $t_1 = 0$ and use the diagonalizable arguments on the second equation to obtain $\sigma = \tau$. Order of magnitude arguments now show that $ds_1 = 0$. If $s_1 = 0$ then so is cs_2 which forces $c = 0$ and then $as_1 = t_1$, for every choice of t_1.

Now we choose $\mathbf{t}$ so that $t_2 = 0$. Then $d = 0$ gives a contradiction, by order of magnitude estimates, on the second equation. This yields $ds_1 = t_1$ and so $a = d$ but then UV^{-1} commutes with Δ and this gives $S = T$.

Finally we deal with the case when T is diagonalizable but S is not. We refer back to Equation (3.8) and let U, V be as in Equation (4.1) where now

$$\Delta = \begin{pmatrix} \sigma & 1 \\ 0 & \sigma \end{pmatrix} \quad \text{and} \quad \Gamma = \begin{pmatrix} \tau_1 & 0 \\ 0 & \tau_2 \end{pmatrix} . \tag{6.8}$$

The resulting equations now become

$$\begin{aligned} (c(s_1 n + s_2) + as_1)\sigma^n - (c(s_1 m + s_2) + as_1)\sigma^m \\ = t_1 \tau_1^{r(n)}(w_R(\tau_1) + \beta_n(\tau_1)) \\ (d(s_1 n + s_2) + bs_1)\sigma^n - (d(s_1 m + s_2) + as_1)\sigma^m \\ = t_2 \tau_2^{r(n)}(w_R(\tau_2) + \beta_n(\tau_2)) . \end{aligned} \tag{6.9}$$

The inequality (4.9) is replaced by

$$\alpha n - C \le r(n) \le \alpha n + \beta \log n$$

for $n \in I_R$. Using a counting argument which is a minor refinement of that given in the proof of Proposition 3, we obtain $2^{R/2}$ pairs (n, n') in I_R such that $n - n' = k$ and $r(n) - r(n') = l$, for some k and l. For such pairs, the first equations are of the form,

$$\begin{aligned} (An + B)\sigma^n &= t_1 \tau_1^{r(n)}\left(w_R(\tau_1) + \beta'_n(\tau_1)\right) \\ (A(n+k) + B)\sigma^{n+k} &= t_1 \tau_1^{r(n)+l}\left(w_R(\tau_1) + \beta'_n(\tau_1)\right) \end{aligned} \tag{6.10}$$

after absorbing the terms involving σ^m into the error term β'_n. We multiply the first equation by σ^k and subtract to obtain

$$Ak\sigma^{n+k} = t_1 \tau_1^{r(n)} w_R(\tau_1)\left(\tau_1^l - \sigma^k\right) + \tau_1^{r(n)} \gamma_{n,n'}$$

where $|\gamma_{n,n'}| < C\exp(-2^{\sqrt{R}})$. This equation for these values of n may be used, as in the diagonalizable case, to show that some power of σ is a power of τ_1. Replacing the matrices S and T by suitable powers of themselves we may assume $\tau_1 = \sigma$. This means that τ_1 is an integer and so we can choose $\mathbf{t}$ so that $t_1 = 0$. This forces some power of τ_2 to be a power of σ and so by size considerations $\tau_2 = \tau_1 = \sigma$. Moreover, using (6.9) we can show that

$$n - C \le r(n) \le n + C$$

for some constant $C > 0$. Returning to Equations (6.9) we see that $cs_1 = ds_1 = 0$. It follows that $s_1 = 0$. Now we have the equations

$$\begin{aligned} cs_2\left(\sigma^n - \sigma^m\right) &= t_1\sigma^{r(n)}\left(w_R(\sigma) + \beta_n(\sigma)\right) \\ ds_2\left(\sigma^n - \sigma^m\right) &= t_2\sigma^{r(n)}(w_R(\sigma) + \beta_n(\sigma)) \ . \end{aligned} \tag{6.11}$$

This forces

$$\frac{c}{d} = \frac{t_1}{t_2} \tag{6.12}$$

for any $\mathbf{t}$, and so a contradiction.

Finally we deal with the case where neither S nor T is diagonalizable. In this case Equation (3.8) becomes

$$\begin{aligned} &(as_1 + cs_2)\left(\sigma^n - \sigma^m\right) + cs_1\left(n\sigma^{n-1} - m\sigma^{m-1}\right) \\ &\qquad = t_1\tau_1^{r(n)}(w_R(\tau_1) + \beta_n(\tau_1)) \\ &(bs_1 + ds_2)\left(\sigma^n - \sigma^m\right) + ds_1\left(n\sigma^{n-1} - m\sigma^{m-1}\right) \\ &\qquad = t_1\tau(n)\tau_1^{r(n)}\left(w_R'(\tau_1) + \beta_n'(\tau_1)\right) + t_2\tau_2^{r(n)}(w_r(\tau_2) + \beta_n(\tau_2)) \ , \end{aligned} \tag{6.13}$$

where w_R and β_n are as before and

$$w_R'(\tau) = \sum_{j=0}^{M} \varepsilon_j \frac{r(n) - j}{r(n)}\tau^{-j} \text{ and } \beta_n'(\tau) = \sum_{j=M+2^{\sqrt{R}}} \varepsilon_j \frac{r(n) - j}{r(n)}\tau^{-j} \ . \tag{6.14}$$

First we choose $\mathbf{t}$ so that $t_1 = 0$. This forces first $cs_1 = 0$ and then $as_1 + cs_2 = 0$. Together these yield $c = 0$. We then use familiar arguments with the first of the equations in (6.13) to obtain $\sigma = \tau$ and

$$n - C \le r(n) \le n + C \ , \tag{6.15}$$

for the n's and $r(n)$'s in (6.13), at least after replacing S and T by appropriate powers. Without loss of generality, we may assume $r(n) = n$.

Now we choose $\mathbf{t}$ so that $t_2 = 0$. Dividing the first of these two equations by the second and taking n large, we find that

$$\frac{a}{d} = 1 \ . \tag{6.16}$$

Thus

$$UV^{-1} = \begin{pmatrix} a & b \\ 0 & a \end{pmatrix} \ , \tag{6.17}$$

which commutes with $\Delta = \Gamma$, so that $S = T$. This completes the proof of the theorem.

References

[1] Bertrand, A. (1997). Ensembles normaux relatifs à des matrices non-commutantes, *J. of Number Theory,* **63**(1), 180–190.

[2] Borel, E. (1909). Les probabilités dénombrables et leurs applications arithmétiques, *Rend. Circ. Mat. Palermo,* **27**, 247–271.

[3] Brown, G. and Moran, W. (1993). Schmidt's normality conjecture for commuting matrices, *Inventiones Mathematicae,* **111**, 449–463.

[4] Moran, W., Brown, G., and Pollington, A.D. (1997). Normality with respect to powers of a base, *Duke Mathematics J.,* **88**(2), 247–265.

[5] Graham, C.C. and McGeehee, O.C. (1979). *Essays in Commutative Harmonic Analysis,* vol. 238 of *Grundelehren der Mathematischen Wissenschafen,* Springer-Verlag, New York.

[6] LeVeque, W.J. (1977). *Fundamentals of Number Theory,* Addison-Wesley.

[7] Moran, W. and Pollington, A.D. (1997). The discrimination theorem for normality to non-integer bases, *Israel J. of Math.,* **100**, 339–347.

[8] Rauzy, G. (1976). *Propriétés statistiques de suites arithmétiques,* Presses Universitaires de France.

[9] Schmidt, W.M. (1960). On normal numbers, *Pacific J. of Math.,* **10**, 661–672.

[10] Schmidt, W.M. (1964). Normalität bezüglich Matrizen, *J. Reine Angew. Math.,* **214/215**, 227–260.

Received August 31, 1999

Revision received February 12, 2002

Vice-Chancellor, University of Sydney, NSW 2006, Australia
e-mail: Vice-Chancellor@vcc.usyd.edu.au

Department of Electrical and Electronic Engineering
University of Melbourne, Vic 3010, Australia
e-mail: b.moran@ee.mu.oz.au

Department of Mathematics, Brigham Young University, UT 84602, USA
e-mail: andy@math.byu.edu