Werk

Titel: Schmidt's conjecture on normality for commuting matrices.

Autor: Brown, Gavin; Moran, William

Jahr: 1993

PURL: https://resolver.sub.uni-goettingen.de/purl?356556735_0111|log35

Kontakt/Contact

Digizeitschriften e.V.
SUB Göttingen
Platz der Göttinger Sieben 1
37073 Göttingen

info@digizeitschriften.de

Invent. math. 111, 449–463 (1993)

Inventiones mathematicae
© Springer-Verlag 1993

Schmidt's conjecture on normality for commuting matrices

Gavin Brown[1] **and William Moran**[2]

[1] Department of Mathematics, University of Adelaide, S.A., Australia 5001
[2] School of Information Science & Technology, Flinders University of South Australia, S.A., Australia 5001

Oblatum 14-I-1992 & 23-VI-1992

1 Introduction

Our purpose here is to confirm for commuting matrices a long standing conjecture of Wolfgang Schmidt, [11], concerning normality. That overworked word is here used in the sense of Borel – a number is *normal* with respect to a given base if each digit block in the expansion occurs with the expected frequency. To motivate the higher dimensional analogue it is perhaps best to consider the dynamical system generated on the n-dimensional torus by an integer matrix S. A vector $\mathbf{x}$ is S-normal if the trajectory beginning from $\mathbf{x}$ is uniformly distributed over $\mathbf{T}^n$ so that, in the limit, time averages coincide with space averages. It is natural to call a system for which the typical trajectory has this property, *ergodic*, and it is natural to call ergodic systems, generated by S and T respectively, *normally equivalent* if S-normality and T-normality coincide. It is rather straightforward to see that, if $S^p = T^q$, for positive integers p and q then the systems generated by S and T are normally equivalent. A special case of our main result is that the following converse is true: if two normally equivalent ergodic dynamical systems are generated by commuting matrices S and T, then some power of S is a power of T.

Schmidt [9], gave a complete solution of the problem of normality with respect to different bases in one dimension: given two integers $t, s \geqq 2$, either

(1) if $s^p = t^q$ for some positive integers p and q then every number normal in base s is normal in base t and *vice versa*.

(2) if s and t are multiplicatively independent then there are uncountably many real numbers x normal in base s and not normal in base t (and *vice versa*).

It is the second statement whose proof demands effort and (as Cassels did for a special case in [5]) Schmidt constructs a probability measure μ for which almost all numbers have biased expansions in base t but with respect to which almost all numbers are s-normal. The result has been refined and its proof simplified in [1, 2, 3, 7, 10] but that aspect of the basic technique remains unaltered. In fact the essential difference which characterizes our approach is that μ is drawn from the class of Riesz product measures rather than from the class of Bernoulli convolutions. While this difference may at first sight appear relatively slight, the added power achieved by this change is what makes the proof of the conjecture possible.

Formally let us say that the point $\mathbf{x}$ of $\mathbf{R}^n$ is *normal* with respect to the $n \times n$ matrix T if the sequence $(T^r\mathbf{x})$ is uniformly distributed modulo one, that is, if for each rectangular box $C = (a_1, b_1) \times (a_2, b_2) \times \cdots \times (a_n, b_n)$ in $(0, 1)^n$

$$\frac{1}{N}|\{k \leqq N : T^k(\mathbf{x}) \in C \bmod \mathbf{Z}^n\}| \to \lambda(C)$$

as $N \to \infty$, where λ is Lebesgue measure on $\mathbf{R}^n$. When $n = 1$ this coincides with Borel's notion of normality and his famous result is, of course, that with respect to any base almost every real number is normal. In higher dimensions it is evident that not all matrix bases share the analogous property (consider, for example, the matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.)$$

We call an $n \times n$ matrix T *ergodic* if almost every $\mathbf{x} \in \mathbf{R}^n$ is normal in base T. In fact, following Schmidt [11], we consider *almost integer* matrices – $n \times n$ invertible matrices with rational entries and all of whose eigenvalues are algebraic integers. Matrices of that class which are also ergodic will be called *almost integer ergodic*. It is shown in [11] that an almost integer matrix is almost integer ergodic if and only if neither 0 nor a root of unity occurs as an eigenvalue. Now we are in a position to state our main theorem. Its proof will occupy the remainder of the paper.

Theorem. *Let S and T be commuting almost integer ergodic 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).*

In [11], Schmidt obtained the result under the additional hypothesis that every eigenvalue of T (respectively S) has modulus strictly greater than one. This appears to be an essential feature of the analytic construction of the measure μ to which we earlier referred. Our use of Riesz products, which we originally developed in this context with Charles Pearce in [1], frees us from such an *a priori* constraint and makes it easier to separate the necessary structure theory for matrices from the analytic tools. In this way we build on the linear algebra developed by Schmidt but we are also able considerably to simplify many of the proofs. As partial penalty we find ourselves replacing some elementary (but complicated) number theoretical estimates by an appeal to Baker's estimates for lower bounds of linear forms in logarithms, but it is by no means clear that this essential to the Riesz product method. Some sense of the power of this method can be gained by specializing the techniques of this paper to obtain a very short proof of the one-dimensional case (there is considerable clarification and simplification of the technicalities of [1]).

It should be noted that the full conjecture in [11] concerns the case where S and T do not commute. Even in the two dimensional case the question appears significantly more difficult. We have, however, obtained substantial partial results and intend to report on these soon.

2 Riesz products in $\mathbf{T}^n$

For the moment let T be an arbitrary integer valued $n \times n$ matrix which we think of as a linear operator on $\mathbf{R}^n$, and on the n-torus $\mathbf{T}^n (= \mathbf{R}^n/\mathbf{Z}^n)$. We shall need to

consider the dual action of T on $\mathbf{Z}^n$. Rather than resorting to a different notation for this action we prefer to think of T multiplying the 'row vectors' from $\mathbf{Z}^n$ on the right. We shall say that T is *dissociate* if for each $\alpha \neq 0$ in $\widehat{\mathbf{T}^n} = \mathbf{Z}^n$ the sequence (αT^n) is a dissociate sequence in the sense of Hewitt and Zuckermann (see [6]); that is,

$$\sum_{r=1}^{N} \varepsilon_r \alpha T^r \neq 0$$

for all choices of $\varepsilon_r \in \{0, \pm 1, \pm 2\}$, not all equal to 0, and all N. Fix a member α of $\mathbf{Z}^n$. Let us agree to call another member β of $\mathbf{Z}^n$ an α-*word* if it can be written in the form

$$\beta = \sum_{r=1}^{N} \varepsilon_r \alpha T^r, \tag{1}$$

where $\varepsilon_r = \pm 1$ or 0. Now consider the trigonometric polynomial

$$P_N(\mathbf{x}) = \prod_{r=1}^{N} (1 + \cos \alpha T^r \mathbf{x}). \tag{2}$$

This is non-negative and its integral over $\mathbf{T}^n$ is equal to 1. It is straightforward to see that as a result of the dissociateness of T, the sequence of probability measures $P_N \cdot m_1$ converges in the weak-$*$ sense where m_1 is Lebesgue measure on $\mathbf{T}^n$. The limit measure which we write as

$$\mu = \prod_{r=1}^{\infty} (1 + \cos \alpha T^r \mathbf{x}) \cdot m_1 \tag{3}$$

is called a *Riesz product*. More general Riesz product measures can be defined in this context but these are sufficient for our purposes. In these circumstances μ is always singular to Lebesgue measure, and when it is regarded as a measure on the n-torus its Fourier coefficients are given by the formula

$$\hat{\mu}(\beta) = \begin{cases} 0 & \text{unless } \beta \text{ is an } \alpha\text{-word} \\ (\frac{1}{2})^{\sum_{r=1}^{N} |\varepsilon_r|} & \text{if } \beta = \sum_{r=1}^{N} \varepsilon_r \alpha T^n. \end{cases} \tag{4}$$

It is this formula which is the salient feature of the Riesz product. From now on all we shall use about this measure is that it is a continuous probability measure on T^n which satisfies (4).

We shall need to perform the Riesz product construction with T an almost integer ergodic matrix. There are two obstacles to this programme. The first is that the matrix in question is not necessarily integral. This is quickly taken care of by a result of Schmidt [11] to the effect that an almost integer ergodic matrix has the property that, for some positive integer d, dT^r is integral for all positive integers r. We shall call d the *denominator* of T. Thus if we choose α to be of the form $d \cdot \gamma$ where γ is in $\mathbf{Z}^n$, then the sequence (αT^r) is in $\mathbf{Z}^n$. The second obstacle is the problem of dissociateness of T. The following lemma guarantees that the sequences is dissociate provided T is replaced by a suitably high power of itself.

Lemma 1 *Let T be an almost integer ergodic matrix. Then there is some positive integer k such that (αT^{rk}) is dissociate for all $\alpha \neq 0$ in $\mathbf{Q}^n$.*

Proof. First choose k so that for every eigenvalue λ of T with absolute value not equal to one,

$$\tfrac{1}{4} > |\lambda|^k \quad \text{or} \quad |\lambda|^k > 4.$$

We now replace T by T^k, so that it is enough to prove the result for the new T with $k = 1$. Fix a non-zero α in $\mathbf{Q}^n$. We decompose T over the field $\mathbf{Q}$ into a direct sum of matrices whose characteristic polynomials are of the form $q(x)^e$ where q is a monic irreducible over $\mathbf{Q}$. Evidently at least one component of α in this decomposition of $\mathbf{Q}^n$ will be non-zero. Now we replace T by one of its component blocks corresponding to a non-zero component of α. Thus we can and do assume that the characteristic polynomial of T is a power of an irreducible.

Recall that the eigenvalues of T are algebraic integers which are neither roots of unity nor 0, so, by Dirichlet's theorem and the earlier choice of k, one at least of them, say λ_1, has absolute value greater than 4 or smaller than $\frac{1}{4}$. We now consider T as a linear operator on $\mathbf{C}^n$, considered as row vectors, and decompose this space into the subspaces $V_\lambda = \ker(\lambda - T)^n$ as λ ranges over the eigenvalues of T. We claim that the component of α in each of these subspaces is non-zero. To see this note that the decomposition can in fact be done over the splitting field of q and the automorphism group of this field will permute the components of α with respect to this decomposition while leaving α unchanged since it is rational. Next we replace T and α by their components on the subspace V_{λ_1}. Choose s to be the largest integer such that $\beta = \alpha(\lambda_1 - T)^s \neq 0$.

Now suppose that for some N and $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_N$.

$$\sum_{r=1}^{N} \varepsilon_r \alpha T^r = 0.$$

Then

$$\sum_{r=1}^{N} \varepsilon_r \alpha(\lambda_1 - T)^s T^r = 0,$$

that is

$$\sum_{r=1}^{N} \varepsilon_r \lambda_1^r \beta = 0.$$

In view of the restriction on the absolute value of λ_1 this situation can only happen if $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_N$ are all zero. □

Since x is normal in base T if and only if it is normal in base T^k there is no loss in generality in assuming that T itself is dissociate. We shall do this from now on. The trick of replacing a matrix by a suitable power of itself will be used at several points in the paper. The proof of the theorem falls into two parts. The first, relatively straightforward, part is the proof that, for almost every choice of $\mathbf{x}$ with respect to μ, $\mathbf{x}$ is non-normal with respect to T. This follows.

Lemma 2 *Let T be dissociate and almost integer ergodic and let α be a non-zero element of $\mathbf{Z}^n$ divisible by d where d is the denominator of T. Let μ be the Riesz product measure defined according to* (3). *Then, for μ-almost every $\mathbf{x}$ in $\mathbf{T}^n$, $\mathbf{x}$ is non-normal with respect to the base T.*

Proof. The proof in this case is essentially the same as in the one-dimensional case (see [1]). We give it for completeness. Consider the sequence $(\exp 2\pi i\alpha T^r \mathbf{x})$ as a sequence of random variables on $\mathbf{T}^n$ with probability measure μ. Because of (4) this sequence is uncorrelated and each random variable has expectation equal to $\frac{1}{2}$, so by the Strong Law of Large Numbers (cf. [4, p. 97]).

$$\frac{1}{N}\sum_{r=1}^{N} \exp 2\pi i\alpha T^r \mathbf{x} \to \frac{1}{2}$$

almost everywhere with respect to μ. The result now follows by the Weyl Criterion for uniform distribution. □

3 Normality: Preliminary results

In view of what we have done so far it is enough to prove that almost every $\mathbf{x}$ with respect to μ is normal in base S where S is another almost integer ergodic matrix which is multiplicatively inequivalent to T. This will be done over the next three sections. Before tackling the problem we shall need to establish some preliminary results. The first is a lemma which is essentially due to Davenport, Erdös and Leveque. The neat formulation appears in Rauzy [8].

Lemma 3 *Let $\beta \in \mathbf{Z}^n$ be non-zero and let ρ be a probability measure on $\mathbf{T}^n$. Define*

$$H_N(\beta) = \frac{1}{N^2}\sum_{k=1}^{N}\sum_{j=1}^{k-1} |\hat{\rho}(\beta S^k - \beta S^j)|.$$

If $\sum_{N=1}^{\infty} H_N(\beta)/N < \infty$ then

$$\frac{1}{N}\sum_{r=1}^{N} \exp(2\pi i\beta S^r \mathbf{x}) \to 0$$

almost everywhere with respect to ρ.

At this point we replace ρ by the Riesz product measure μ as in (3) and use the special nature of the Fourier coefficients of such measures to obtain the following result. In fact we merely count the number of non-zero coefficients to obtain an estimate for H_N.

Lemma 4 *Let $t(N) \geqq 0$, and $B_N(\beta)$ be the set of all $k \leqq N$ such that for some j with $0 \leqq j \leqq k - t(N)$ and some $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{M(k,j)}$ $(\varepsilon_m \in \{0, \pm 1\})$*

$$\beta S^k - \beta S^j = \sum_{m=1}^{M(k,j)} \varepsilon_m \alpha T^m.$$

If $\sum_{N=1}^{\infty} N^{-2}(t(N) + |B_N(\beta)|) < \infty$, then

$$\frac{1}{N}\sum_{r=1}^{N} \exp(2\pi i\beta S^r \mathbf{x}) \to 0$$

almost everywhere with respect to μ.

Proof. It is enough to note that

$$\sum_{N=1}^{\infty} \frac{1}{N^2} \sum_{k=1}^{N} \sum_{j=1}^{k-1} |\hat{\rho}(\beta S^k - \beta S^j)| \leqq \sum_{N=1}^{\infty} \frac{1}{N^3} \left(\sum_{k \in B_N} k + \sum_{k=1}^{N} t(N) \right)$$

$$\leqq \sum_{N=1}^{\infty} N^{-2}(t(N) + |B_N|) < \infty . \quad \square$$

Corollary 1 *If* $\sum_{N=1}^{\infty} N^{-2}(t(N) + |B_N(\beta)|) < \infty$ *for all* $\beta \neq 0$ *then almost every* $\mathbf{x}$ *with respect to* μ *is normal in base* S.

Evidently then we need to count the solutions of equations of the form

$$\beta S^k - \beta S^j = \sum_{m=1}^{M(k,j)} \varepsilon_m \alpha T^m. \tag{5}$$

As part of the process of relating S and T we shall apply the one-dimensional techniques we developed in [1] to their eigenvalues. Thus we convert Eq. (5) into an equation involving the eigenvalues of the matrices. The next few lemmas, which include some of the methods of [1], will then allow us to obtain useful information from this. The proofs of Lemmas 5 and 6 we give here, however, are simpler than the corresponding proofs in [1], and, in particular, avoid the use of the concept of *discrepancy*.

Lemma 5 *Suppose that* $\min(|\lambda|, |\rho|) \geqq C, j \leqq k - t(N)$ *and*

$$\lambda^k - \lambda^j = K \sum_{i=1}^{M(k,j)} \varepsilon_i \rho^i \tag{6}$$

where K *is constant,* $M(k,j)$ *is an integer and* $\varepsilon_i \in \{0, \pm 1\}$ *with* $\varepsilon_{M(k,j)} = 1$. *Then* $\lambda^k / \rho^{M(k,j)}$ *belongs to a union of* $3^{t(N)}$ *discs each of radius* $O(C^{-t(N)})$, *lying in some fixed disc.*

Proof. For Eq. (6) to be satisfied

$$\frac{\lambda^k}{\rho^{M(k,j)}} = K\left(1 - \frac{\lambda^j}{\lambda^k}\right)^1 \left(1 + \sum_{j=1}^{M(k,j)-1} \varepsilon_i \rho^{-1}\right),$$

where $\varepsilon_i \in \{0, \pm 1\}$ and the second term in the product on the right hand side is $1 + O(C^{-t(N)})$. The third term in the product can be rewritten as

$$\left(1 + \sum_{i=1}^{t(N)} \varepsilon_i \rho^{-i}\right)\left\{1 + \left(1 + \sum_{i=1}^{t(N)} \varepsilon_i \rho^{-i}\right)^{-1} \sum_{j=t(N)+1}^{M(k,j)-1} \varepsilon_i \rho^{-1}\right\}$$

$$= \left(1 + \sum_{i=1}^{t(N)} \varepsilon_i \rho^{-i}\right)(1 + \rho^{-t(N)} W),$$

where $|W| \leqq S(S-1) \leqq 3/2, S = \sum_{j=1}^{\infty} |\rho|^{-i}$. $\square$

To prevent the appearance of circularity in our choices of constants it is necessary to suppose for the remainder of this section that the quantities $\lambda, \lambda', \rho, \rho'$, etc. are eigenvalues of absolute value greater than 1 of the fixed matrices S, T of the

type under consideration. In particular this means that they are algebraic integers and that there is some constant B (for Baker) such that, if $\lambda^n \neq \rho^m$, then

$$|n \log \lambda - m \log \rho| > B^{-\log N} \tag{7}$$

whenever $\max(|n|, |m|) \leqq N$ and whenever λ is an eigenvalue of S and ρ is an eigenvalue of T. Observe that the truth of this last statement is not altered if we replace S, T by S^r, T^r for some arbitrary positive integer power r. Accordingly we can do assume, without loss of generality, that

$$\min(|\lambda|, |\rho|) \geqq C, \tag{8}$$

where C is a constant which we may choose depending on B. (It turns out that a convenient choice is B^2.) We define B_N to be the number of $k \leqq N$ for which there exists $j \leqq k - t(N)$ and ε_i's such that (6) holds.

Lemma 6 *If $\lambda^p \neq \rho^q$ for all positive integers p, q; if $t(N) = D \log N$ with $D \log 3 < 1$; and if $C > B^{1/D}$, then*

$$\sum_{N=1}^{\infty} \frac{1}{N^2}(t(N) + |B_N|) < \infty$$

for all $\beta \neq \mathbf{0}$.

Proof. Evidently $\sum N^{-2} t(N)$ converges, so our only concern is to estimate $|B_N|$. From Lemma 5 we know that $k \log \lambda - M(k, j) \log \rho$ belongs to a union of $3^{t(N)}$ discs of radius $O(C^{-t(N)})$. If any disc contains two such numbers then we have an estimate

$$|(k - k') \log \lambda - (M(k. j) - M(k', j') \log \rho| = O(C^{-t(N)}). \tag{9}$$

This is impossible because we know from (7) that the right hand side of (9) must be at least $B^{-\log N}$, while we arranged that $C^D > B$. This shows that

$$|B_N| \leqq 3^{t(N)} = 3^{D \log N} = N^{D \log 3}.$$

The rest is clear. □

The following lemma corresponds to the final statement of Lemma 5, given that our other choices have been made. We isolate it for convenience.

Lemma 7 *Under the conditions of* Lemma 6, *there are infinitely many equations of the form* (5) *in which $M(k, j) - k$ is constant.*

Lemma 8 *Let λ, ρ and λ', ρ' each satisfy an equation of the form* (2) *for the same infinite set of (k, j)'s and suppose they have the same $M(k, j)$'s. Assume further that $\lambda^p = \rho^q$ and $\lambda'^{p'} = \rho'^{q'}$. Then*

$$\frac{p}{q} = \frac{p'}{q'}.$$

Proof. Let the two sets of equations be of the form:

$$\lambda^k - \lambda^{j(k)} = K \sum_{i=1}^{M(k. j)} \varepsilon_i \rho^i$$

and

$$\lambda'^{k} - \lambda'^{j'(k)} = K' \sum_{i=1}^{M(k,j)} \varepsilon_i' \rho'^{i}$$

We take logarithms and use the fact that $j, j' < N - t(N)$ to obtain

$$k\frac{\log \lambda}{\log \rho} = M(k,j) + o(k)$$

and

$$k\frac{\log \lambda'}{\log \rho'} = M(k,j) + o(k)$$

for those (k, j)'s in question. The result is now clear. □

Consideration of equations of the form (5) will ultimately yield the theorem. Recall that $B_N(\beta)$ is the set of all $k \leqq N$ such that for some j with $0 \leqq j \leqq k - t(N)$ and some $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{M(k,j)}$ $(\varepsilon_m \in \{0, \pm 1\})$

$$\beta S^k - \beta S^j = \sum_{m=1}^{M(k,j)} \varepsilon_m \alpha T^m,$$

and note that we have now fixed the choice of $t(N)$ to be $D \log N$, and S, T have the property that the absolute values of eigenvalues which are greater than 1 are in fact greater than C. We shall say that S is *weakly algebraically dependent* on T if for each $\alpha \in \mathbf{Q}^n$ there exists $\beta \in \mathbf{Q}^n$ with $\sum N^{-2} B_N(\beta) = \infty$ and if this statement remains true when S and T are replaced by powers of themselves. Observe that we may free ourselves from the constraint that α be divisible by the denominator by exploiting the linearity of the Eq. (5). In view of the corollary to Lemma 4 it is enough to prove the following proposition.

Proposition 1 *If S is weakly algebraically dependent on T then S and T are multiplicatively equivalent.*

4 The irreducible case

In this section we prove the theorem via Proposition 1 for the case when the algebra $\mathscr{A}(S, T)$ is irreducible. In other words we assume that $\mathbf{Q}^n$ has no invariant subspace for the algebra $\mathscr{A}(S, T)$. By breaking up the theorem in this way we are following the strategy of Schmidt's proof of the restricted result. Our proof of the irreducible case is, however, much easier than his. It will be convenient to replace (5) by a matrix equation and this is possible if the algebra $\mathscr{A}(S, T)$ is indecomposable, that is, if $\mathbf{Q}^n$ cannot be written as a non-trivial direct sum of $\mathscr{A}(S, T)$-invariant subspaces. The key property of indecomposable algebras of matrices we shall need here is the existence of a cyclic vector.

Lemma 9 *Let $\mathscr{A}(S, T)$ be indecomposable and suppose that S is weakly algebraically dependent on T. Then there exist an $n \times n$ matrix U over $\mathbf{Q}$ such that if E_N is the set of all k for which*

$$U(S^k - S^j) = \sum_{m=1}^{M(k,j)} \varepsilon_m T^m \tag{10}$$

for some $0 \leqq j \leqq k - t(N)$ *and* $\varepsilon_m \in \{0, \pm 1\}$ *then* $\sum N^{-2} E_N = \infty$. *Moreover* U *belongs to* $\mathscr{A}(S, T)$.

Proof. We know that for each $\beta \neq 0$ in $\mathbf{Q}^n$ there exists some $\alpha \in \mathbf{Q}^n$ such that

$$\alpha(S^k - S^j) = \sum_{m=1}^{M(k,j)} \varepsilon_m \beta T^m, \tag{11}$$

for $B_N(\beta)$ k's less than or equal to N, some j $(0 \leqq j \leqq k - t(N))$ and some $\varepsilon_1, \varepsilon_2, \ldots \varepsilon_{M(k,j)}$ $(\varepsilon_m \in \{0, \pm 1\})$ and that $\sum N^{-2} B_N(\beta) = \infty$. Choose a cyclic vector β_0 for $\mathscr{A}(S, T)$ and let H be the subspace of $\mathbf{Q}^n$ of those β for which (11) remains true for the same ε's and (k, j)'s as for β_0. Multiply both sides of (11) by arbitrary A in $\mathscr{A}(S, T)$ and use commutativity of the latter to observe that $A(\beta_0)$ belongs to H. Thus H is invariant under the action of $\mathscr{A}(S, T)$ and, because β_0 is cyclic, is the whole of $\mathbf{Q}^n$. For any particular choice of $k, j, \varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{M(k,j)}$ we may define U by

$$U = (S^k - S^j)^{-1} \sum_{i=1}^{M(k,j)} \varepsilon_i T^i$$

and then we find

$$\alpha = \beta U.$$

Evidently U belongs to the algebra generated by S and T. □

The next lemma is effectively in Schmidt's proof [11]. It allows us to reduce the irreducible case to the one-dimensional case.

Lemma 10 *Let* $\mathscr{A} \neq 0$ *be an irreducible commutative subalgebra of* $M(\mathbf{Q}^n)$. *Then there is a field isomorphism* $\phi : \mathscr{A} \to F$ *where* F *is a finite extension of* $\mathbf{Q}$. *For any* $A : \mathscr{A}$, $\phi(A)$ *is an eigenvalue of* A *and for a fixed* $A_0 \in \mathscr{A}$ *and eigenvalue* λ_0 *of* A_0, ϕ *may be chosen so that* $\phi(A_0) = \lambda_0$.

Proof. We observe that $\mathscr{A}$ is a field, since if $A \in \mathscr{A}$ and $\ker A \neq 0$ then by commutativity $\ker A$ is a proper invariant subspace of $\mathbf{Q}^n$. Thus $\mathscr{A}$ is isomorphic to a finite extension F of $\mathbf{Q}^n$. Moreover, if $\phi(A) = \lambda$ then, by the Cayley–Hamilton theorem, λ is an eigenvalue of A. Since the minimal polynomial m_{A_0} of any $A_0 \neq 0$ in $\mathscr{A}$ is irreducible the Galois group acts transitively on the roots of m_{A_0}. In consequence we may choose ϕ so that $\phi(A_0) = \lambda_0$. □

Now we are in a position to complete the proof of the irreducible case.

Proposition 2 *Let* S, T *generate an irreducible commutative subalgebra* $\mathscr{A}$ *of* $M(\mathbf{Q}^n)$. *If* S *is weakly algebraically dependent on* T *then* S *and* T *are multiplicatively equivalent.*

Proof. We begin by replacing S and T by sufficiently high powers so that eigenvalues of each of absolute value greater than 1 have absolute value greater than C (see (8)). We apply Lemma 10 and choose ϕ such that $|\phi(S)| > C$. This is possible because at least one eigenvalue of S lies outside the unit circle – a consequence of the fact that, for some integer d, dS^r is integral for all positive integers r. (In the particular case where the determinant of S has unit modulus, not all eigenvalues can lie on the unit circle in view of Dirichlet's theorem.) It follows from weak

algebraic dependence that $|\phi(T)| > 1$. Now we invoke Lemmas 6 and 8 with $\lambda = \phi(S)$ and $\rho = \phi(T)$ to deduce that $\lambda^p = \rho^q$ for some positive integers p and q. Since ϕ is an isomorphism $S^p = T^q$. □

5 The indecomposable case

Now we consider the case when the algebra $\mathscr{A}(S, T)$ generated by S and T is indecomposable. Here, as before, we follow Schmidt in the development of the algebraic aspects of the proof. We choose $\mathscr{A}$-invariant subspaces of $\mathbf{Q}^n$,

$$\mathbf{Q}^n = B_r \supset B_{r-1} \supset \ldots \supset B_1 \supset B_0 = \{0\} \tag{12}$$

so that the sequence has maximal length, k, and the length does not increase if either S or T is replaced by a power of itself. We may ensure this by replacing S and T by powers if necessary. The algebra $\mathscr{A}$ acts on each of the quotient spaces $C_i = B_i/B_{i-1}$ in an irreducible way. Call the resulting algebra of matrices $\mathscr{A}_i$ and write A_i for the linear operator acting on C_i induced by $A \in \mathscr{A}$. The corresponding operator on $\mathbf{Q}^n/B_i$ is denoted by ${}_iA$. Our aim here is to prove the following proposition which is the extension of Proposition 2 to the indecomposable case.

Proposition 3 *If the algebra $\mathscr{A}$ generated by S and T is indecomposable and S is weakly algebraically dependent on T then S and T are multiplicatively equivalent.*

Our proof of this requires a number of lemmas, and will be by induction on the length r of the sequence (12). Evidently the case $r = 1$ is covered by Proposition 2. It is also clear that if S is weakly algebraically dependent on T, then S_i and T_i and ${}_iS$ and ${}_iT$ are similarly related. Since the algebra $\mathscr{A}_i$ generated by S_i and T_i is irreducible we may use Proposition 2 to conclude that $S_i^{pi} = T_i^{qi}$ for some p_i and q_i. By Lemma 8, $p_i/q_i = p_j/q_j$, for all i, j. Let p and q be the largest of $p_1, p_2, \ldots, p_r$ and $q_1, q_2, \ldots, q_r$ respectively. We replace S and T by S^p and T^q respectively, so that now $S_i = T_i$ for all i. This means of course that

$$\mathscr{A}_i = \mathscr{A}(S_i, T_i) = \mathscr{A}(T_i)$$

so that the minimal polynomial m_i of T_i is irreducible. Since $\mathscr{A}(S, T)$ is indecomposable each $A \in \mathscr{A}$ has a minimal polynomial m_A which is a power of an irreducible; in particular, $m_T = m^d$ where m is irreducible. Now it is clear that each m_i must equal m and of course $d = r$, so $rs = n$ where $s = \dim B_i/B_{i+1}$ for all i.

In fact we shall need to work with T^{-1} and S^{-1} rather than T and S. Write $\tilde{m}$ for the common minimal polynomial of all the T_i^{-1} and S_i^{-1}.

Our aim, of course, is to prove that $S = T$ in these circumstances. Since we shall induct on the length of the sequence of $\mathscr{A}(S, T)$-invariant subspaces it will be convenient to isolate the result we need as follows.

Proposition 4 *Let S be weakly algebraically dependent on T and suppose that $\mathscr{A}(S, T)$ is indecomposable with a maximal sequence of invariant subspaces*

$$\mathbf{Q}^n = B_r \supset B_{r-1} \supset \cdots \supset B_1 \supset B_0 = \{0\}. \tag{13}$$

Suppose further that $S_i = T_i$ for each i. Then $S = T$.

Our inductive assumption allows us to assume the result for sequences (13) of length $r-1$, so that we may assume that ${}_{r-1}S = {}_{r-1}T$ and that the maps induced by S and T on the invariant subspace B_{r-1} are equal.

Now we need a structural result of Schmidt [11, Lemma 28]. Schmidt obtains this result for S and T. It will be convenient for us to have it for their inverses. The extra assumptions we are able to impose allow us to obtain a much simpler proof and a sharper conclusion than Schmidt.

Lemma 11 *The space* $\mathbf{Q}^n$ *can be decomposed as a direct sum*

$$\mathbf{Q}^n = W_r \oplus W_{r-1} \oplus \ldots \oplus W_1$$

where $W_i \oplus B_{i-1} = B_i$. *Furthermore a basis of each of the spaces* W_i *may be chosen so that with respect to the resulting basis* $\alpha_0, \alpha_1, \ldots, \alpha_{n-1}$ *for* $\mathbf{Q}^n$, T^{-1} *has the matrix representation*

$$T^{-1} = \begin{pmatrix} J & P & 0 & & \cdots & 0 \\ 0 & J & P & 0 & \cdots & 0 \\ 0 & 0 & J & P & \cdots & 0 \\ \vdots & & \ddots & \ddots & & \vdots \\ 0 & 0 & \cdots & 0 & J & P \\ 0 & 0 & \cdots & 0 & 0 & J \end{pmatrix},$$

$$P = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & \cdots & 0 & 0 \end{pmatrix}$$

and J *is the companion matrix of* $\tilde{m}$, *that is*

$$J = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{s-1} \end{pmatrix},$$

where $\tilde{m}(x) = c(a_0 + a_1 x + \cdots + a_{s-1} x^{s-1} + x^s)$. *Furthermore* S^{-1} *has the representation*

$$S^{-1} = \begin{pmatrix} J & P & 0 & & \cdots & A \\ 0 & J & P & 0 & \cdots & 0 \\ 0 & 0 & J & P & \cdots & 0 \\ \vdots & & \ddots & \ddots & & \vdots \\ 0 & 0 & \cdots & 0 & J & P \\ 0 & 0 & \cdots & 0 & 0 & J \end{pmatrix},$$

if $r > 2$ *and*

$$S^{-1} = \begin{pmatrix} J & A \\ 0 & J \end{pmatrix},$$

if $r = 2$. *The matrix* A ($A - P$ *if* $r = 2$) *belongs to the algebra* $\mathscr{A}(J)$ *generated by* J.

Proof. First we make the observation that $\alpha \in B_i$ is a cyclic vector for $\mathscr{A}|_{B_i}$ if and only if $\alpha \notin B_{i-1}$. Necessity is obvious, and, for sufficiency, let $W(\alpha)$ be the subspace generated by $\alpha, \alpha T^{-1}, \alpha T^{-2}, \ldots, \alpha T^{-(s-1)}$. This is s-dimensional in the quotient by B_{i-1} since the quotient algebra $\mathscr{A}(J)$ is irreducible. Now $\alpha \tilde{m}(T^{-1})$ is in B_{i-1} and not in B_{i-2}, otherwise the subspace of such vectors would be an $\mathscr{A}(S, T)$-invariant subspace of B_i containing B_{i-1}. Thus we can express any vector as an $\mathscr{A}(S, T)$-image of α.

Now pick a cyclic vector α_0 for $\mathscr{A}(S, T)$. Of necessity this will lie outside B_{r-1}. Form $W(\alpha_0)$ and observe that $W(\alpha) \cap B_{r-1} = (0)$ since if not there is some polynomial p of degree less than s such that $\alpha_0 p(T^{-1}) \tilde{m}(T^{-1})^{r-1} = 0$ which contradicts what we know about the minimal polynomial of T^{-1}. By a dimension argument $\mathbf{Q}^n = W(\alpha_0) \oplus B_{r-1}$.

Let $\alpha_1 = \alpha_0 \tilde{m}(T^{-1})$. This clearly belongs to B_{r-1} and is a cyclic vector by the earlier observation, since it cannot be zero. Now we repeat the argument with α_1 in place of α_0 and find a subspace W_{r-1} of B_{r-1} which complements B_{r-2}. Continuing in this way we obtain the direct sum decomposition. The basis of $\mathbf{Q}^n$ is obtained by putting together the bases of W_i in the orders given.

Now the various properties of the matrix for T^{-1} may be easily seen. For S^{-1} we merely use the fact (from the induction hypothesis) that it agrees with T^{-1} on B_{r-1} and $\mathbf{Q}^n/B_1$. The properties of A result from the fact that it commutes with J and $\mathscr{A}(J)$ is irreducible. □

Now we write $D = ST^{-1}$, so that with respect to the basis given above D has the form

$$D = \begin{pmatrix} I & 0 & 0 & & \cdots & D_1 \\ 0 & I & 0 & 0 & \cdots & 0 \\ 0 & 0 & I & 0 & \cdots & 0 \\ \vdots & & \ddots & \ddots & & \vdots \\ 0 & 0 & \cdots & 0 & I & 0 \\ 0 & 0 & \cdots & 0 & 0 & I \end{pmatrix},$$

where $D_1 = AJ^{-1}$ if $r > 2$ and $D_1 = (A - P)J^{-1}$ if $r = 2$. Note that D_1 commutes with J and so belongs to $\mathscr{A}(J)$, since J is irreducible.

At this point we rewrite the basic sequence of matrix Eq. (10) as

$$U(D^k - T^{j-k} D^j) = \sum_{m=1}^{M(k,j)} \varepsilon_m T^{m-k}.$$

In view of Lemma 7 we may, on replacing U by UT^t for some integer t and relabelling the ε_i's, rewrite this equation as

$$U(D^k - T^{j-k} D^j) = \sum_{m=1}^{k} \varepsilon_m T^{-m}. \tag{14}$$

Since by dissociateness the right hand side has trivial kernel, U is invertible, and

$$U^{-1} = \begin{pmatrix} U_1 & U_2 & U_3 & & \cdots & U_r \\ 0 & U_1 & U_2 & U_3 & \cdots & U_{r-1} \\ 0 & 0 & U_1 & U_2 & \cdots & U_{r-2} \\ \vdots & & \ddots & \ddots & & \vdots \\ 0 & 0 & \cdots & 0 & U_1 & U_2 \\ 0 & 0 & \cdots & 0 & 0 & U_1 \end{pmatrix},$$

because U belongs to $\mathscr{A}(S, T)$.

We now regard the matrices as members of $M_n(\mathbf{C})$ and choose λ to be an eigenvalue of T of largest absolute value. The subspace V_λ of $\mathbf{C}^n$ of β such that $\beta(\lambda I - T)^m = 0$ for some m is clearly $\mathscr{A}(S, T)$ invariant. It is possible to write down a basis of this subspace by choosing an eigenvector corresponding to λ for J^{-1} in the complexification of W_r and then taking the vector with the same coordinates in the complexifications of each of the other W_i's. To see this note that these vectors belong to the subspace V_λ and are linearly independent. The dimension of this subspace is just the multiplicity of λ in the minimal polynomial for T and so equals r. We write A_λ for the restriction of a member A of $\mathscr{A}(S, T)$ to this subspace.

Lemma 12 $\| T_\lambda^{-n} \| = O((2/\lambda)^n)$.

Proof. From what we know of the minimal polynomial of T^{-1}, we see that the Jordon canonical form for T_λ^{-1} has λ's on the diagonal and 1's in the superdiagonal places. Thus with respect to the basis giving this form

$$T_\lambda^{-n} = \begin{pmatrix} \lambda^{-n} & n\lambda^{-(n-1)} & n(n-1)\lambda^{-(n-2)} & \cdots & & \binom{n}{r-1}\lambda^{-(n-r+1)} \\ 0 & \lambda^{-n} & n\lambda^{-(n-1)} & \cdots & & \binom{n}{r-2}\lambda^{-(n-r+2)} \\ 0 & 0 & \lambda^{-n} & \cdots & & \binom{n}{r-3}\lambda^{-(n-r+3)} \\ \vdots & & \ddots & & \vdots & \\ 0 & 0 & \cdots & 0 & \lambda^{-n} & n\lambda^{-(n-1)} \\ 0 & 0 & \cdots & 0 & 0 & \lambda^{-n} \end{pmatrix}.$$

The result is now obvious. □

Returning to the basis for V_λ described earlier, we observe that

$$D_\lambda = \begin{pmatrix} 1 & d_2 & d_3 & d_4 & & \cdots & d_{r-1} \\ 0 & 1 & 0 & 0 & & \cdots & 0 \\ 0 & 0 & 1 & 0 & & \cdots & 0 \\ 0 & 0 & 0 & 1 & & \cdots & 0 \\ \vdots & & & \ddots & \ddots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 1 & 0 \\ 0 & 0 & & \cdots & 0 & 0 & 1 \end{pmatrix}.$$

The next lemma is the final step in the proof of the theorem for $\mathscr{A}(S, T)$ indecomposable.

Lemma 13 *$d_i = 0$ for $i = 2, 3, \ldots, r-1$.*

Proof. When restricted to the subspace V_λ the Eq. (14) becomes

$$(D_\lambda^k - T_\lambda^{j-k} D_\lambda^j) = U_\lambda^{-1} \sum_{m=1}^{k} \varepsilon_m T_\lambda^{-m}.$$

In view of Lemma 12 the right hand side of this equation remains bounded and the second term on the left hand side tends to 0 as k tends to infinity; the latter follows because j is restricted to be less than $k - t(N)$. The (1, 2)-term in D_λ^k on the other hand is kd_1 and so $d_2 = 0$. Assume that we have proved $d_2, d_3, \ldots, d_l = 0$, then the $(1, l+1)$-term in D_λ^k is kd_{l+1} and so $d_{l+1} = 0$. By induction the result is proved. □

The proof of the last lemma is an induction within an induction so it is as well to summarize the plot thus far. For the case where $\mathscr{A}(S, T)$ is indecomposable, currently under discussion, the induction is on r, the maximal length of the nest of invariant subspaces B_i. The case $r = 1$ is the irreducible case covered by Sect. 4. For the inductive step, we assumed $S_i = T_i$ for each i, that the quotient maps on the space $\mathbf{Q}^n/B_1$ are equal and that the restrictions of S and T to B_{r-1} coincide. The result of Lemma 13 shows that the $r \times r$ matrix D_λ equals the identity, and hence that the restrictions of S and T to V_λ coincide. Because we can transfer via the Galois group to other eigenvalues, this shows, in fact, that $S = T$.

6 The general case

Now we consider a general pair of almost integer ergodic $n \times n$ matrices S and T which satisfy the property that for each $\alpha \in \mathbf{Q}^n$ there exists $\beta \in \mathbf{Q}^n$ such that $\sum N^{-2} B_N(\beta) = \infty$. We may assume that the eigenvalues of S and T, which are in absolute value greater than 1, in fact have absolute value greater than $\max(4, C)$. The space $\mathbf{Q}^n$ may be decomposed into a direct sum of indecomposable $\mathscr{A}(S, T)$-invariant subspaces and correspondingly we may write $S = S(1) \oplus S(2) \oplus S(3) \oplus \cdots \oplus S(l)$ and $T = T(1) \oplus T(2) \oplus T(3) \oplus \cdots \oplus T(l)$ so that $\mathscr{A}(S(i), T(i))$ is indecomposable.

Evidently if α is chosen from one of the invariant subspaces then β must also come from the same space. Thus we may apply the result for the indecomposable case to each summand in turn to obtain $S(i)^{p_i} = T(i)^{q_i}$ for some integers p_i, q_i. Now choose some α which is non-zero in every summand, and fix an eigenvalue λ_i of largest absolute value for each $S(i)$. In view of what we have already proved this eigenvalue occurs in an irreducible component of $\mathscr{A}(S(i), T(i)))$, and there is a corresponding eigenvalue ρ_i of $T(i)$ which is picked out by the field isomorphism in Lemma 10. Applying the field isomorphism in each component, we obtain

$$\lambda_i^k - \lambda_i^j = c_i \sum_{m=1}^{M(k,j)} \varepsilon_m \alpha \rho_i^m, \tag{15}$$

where now the k's and j's are independent of i. By Lemma 8 this yields that p_i/q_i is independent of i. Hence we can find p and q such that $S(i)^p = T(i)^q$ for all i and therefore $S^p = T^q$. This completes the proof of the theorem.

References

1. Brown, G., Moran, W., Pearce, C.E.M.: Riesz products and normal numbers. J. Lond. Math. Soc. **32**, 12–18 (1985)
2. Brown, G., Moran, W., Pearce, C.E.M.: A decomposition theorem for numbers in which the summands have prescribed normality properties. J. Number Theory **24**, 259–271 (1986)
3. Brown, G., Moran, W., Pearce, C.E.M.: Riesz products, Hausdorff dimension and normal numbers. Math. Proc. Camb. Philos. Soc. **101**, 529–540 (1987)
4. Chung, K.L.: A Course in Probability Theory. New York Chicago San Francisco: Harcourt, Brace and World 1968
5. Cassels, J.W.S. On a problem of Steinhaus about normal numbers. Colloq. Math. **7** 95–101 (1959),
6. Hewitt E. and Zuckerman, H.S. Singular measures with absolutely continuous convolution squares. Proc. Camb. Philos. Soc. **62**, 399–420 (1966); Corrigendum, ibid **63**, 367–368 (1967)
7. A.D. Pollington, The Hausdorff dimension of a set of normal numbers. Pac. J. Math. **95** (1981), 193–204.
8. Rauzy, G. Propriétés statistiques de suites arithmétiques. Paris: Presses Universitaires de France 1976
9. Schmidt, W.M. On normal numbers. Pac. J. Math. **10**, 661–672 (1960).
10. Schmidt, W.M. Über die Normalität von Zahlen zu verschiedenen Basen. Acta Arithmetica **7**, 299–301 (1962).
11. Schmidt, W.M. Normalität bezüglich Matrizen. J. Reine Angew. Math. **214**(5) 227–260 (1964)