Pacific Journal of Mathematics

Volume 190 No. 2 October 1999

PACIFIC JOURNAL OF MATHEMATICS

http://www.pjmath.org

Founded in 1951 by

E. F. Beckenbach (1906–1982) F. Wolf (1904–1989)

See inside back cover or www.pjmath.org for submission instructions.

Regular subscription rate for 2006: $425.00 a year (10 issues). Special rate: $212.50 a year to individual members of supporting institutions. Subscriptions, requests for back issues from the last three years and changes of subscribers address should be sent to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163, U.S.A. Prior back issues are obtainable from Periodicals Service Company, 11 Main Street, Germantown, NY 12526-5635. The Pacific Journal of Mathematics is indexed by Mathematical Reviews, Zentralblatt MATH, PASCAL CNRS Index, Referativnyi Zhurnal, Current Mathematical Publications and the Science Citation Index.

The Pacific Journal of Mathematics (ISSN 0030-8730) at the University of California, c/o Department of Mathematics, 969 Evans Hall, Berkeley, CA 94720-3840 is published monthly except July and August. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices. POSTMASTER: send address changes to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163.

PUBLISHED BY PACIFIC JOURNAL OF MATHEMATICS
at the University of California, Berkeley 94720-3840
A NON-PROFIT CORPORATION
Typeset in LaTeX

PACIFIC JOURNAL OF MATHEMATICS
Vol. 190, No. 2, 1999

ON KRONECKER PRODUCTS OF COMPLEX REPRESENTATIONS OF THE SYMMETRIC AND ALTERNATING GROUPS

C. Bessenrodt and A. Kleshchev

In this paper we study the homogeneous tensor products of simple modules over symmetric and alternating groups.

1. Introduction.

Kronecker or inner tensor products of representations of symmetric groups (and many other groups) have been studied for a long time. But even for the symmetric groups no reasonable formula for decomposing Kronecker products of two irreducible complex representations into irreducible components is available (cf. [**7**, **5**]). An equivalent problem is to decompose the inner product of the corresponding Schur functions into a linear combination of Schur functions.

In recent years, a number of partial results have been obtained. For example, the products of characters labelled by hook partitions or by two-row partitions [**3**, **8**] have been computed, and special constituents, in particular of tensor squares, have been considered [**10**, **11**, **12**]. For general products, Dvir [**2**] and Clausen-Meier [**1**] determined the largest part and the maximal number of parts in a constituent of a product (this result is crucial in this paper).

In general, Kronecker products of irreducible representations have very many irreducible constituents (see e.g. [**4**, 2.9]). In this paper, we first consider the simple question: 'when is the Kronecker product of two irreducible S_n-characters again irreducible?' We prove that in fact such a product is *always* reducible, and even inhomogeneous, except for the obvious exception where one of the characters is of degree 1. Then we turn to the same question for the representations of the alternating group A_n. Here one can easily construct examples of non-trivial irreducible tensor products (actually, we observed this first using calculations with the MAPLE packages SF (by Stembridge) and ACE (by Veigneau et al.)). It turns out that the problem for A_n reduces to the classification of certain products of S_n-characters with 2 constituents. So we classify in general the Kronecker products of S_n-characters with 2 constituents, and even more generally, with two homogeneous components. We also obtain some partial results for products with

4 homogeneous components and conjecture a complete classification of the pairs (L_1, L_2) of irreducible complex S_n-representations such that $L_1 \otimes L_2$ has at most 4 homogeneous components.

2. Preliminaries.

We denote by $\mathbb{N}$ the set $\{1, 2, \dots\}$ of the natural numbers.

If G and H are two groups, L is a $\mathbb{C}G$-module and M is a $\mathbb{C}H$-module we write $L \boxtimes M$ for the *outer* tensor product of L and M (which is a module over $G \times H$). If N is another $\mathbb{C}G$-module we write $L \otimes N$ for the *inner* tensor (or *Kronecker*) product of L and N (which is a G-module).

A $\mathbb{C}G$-module is called *homogeneous* if it is isomorphic to a direct sum of copies of one simple module. Every $\mathbb{C}G$-module can be (uniquely) decomposed into a direct sum of its *homogeneous components*. Similarly we speak of the homogeneous characters and the homogeneous components of the characters.

We use the notions and notation of the representation theory of S_n and A_n and refer the reader to [4] for the most basic ones. In particular, we write $\lambda = (\lambda_1, \dots, \lambda_k) \vdash n$ if λ is a *partition* of n; in this case we also write $|\lambda|$ for n. We often gather together equal parts of a partition and write, for example, $(5^2, 3^3)$ for $(5, 5, 3, 3, 3)$. The partition *conjugate* to λ is denoted by λ'. If $\lambda = \lambda'$ we say that λ is *symmetric*. We do not distinguish between a partition λ and its *Young diagram* $\lambda = \{(i, j) \in \mathbb{N} \times \mathbb{N} \mid j \leq \lambda_i\}$. Elements $(i, j) \in \mathbb{N} \times \mathbb{N}$ are called nodes. If $\lambda = (\lambda_1, \lambda_2, \dots)$ and $\mu = (\mu_1, \mu_2, \dots)$ are two partitions we write $\lambda \cap \mu$ for the partition $(\min(\lambda_1, \mu_1), \min(\lambda_2, \mu_2), \dots)$ whose Young diagram is just the intersection of those for λ and μ. A node $(i, \lambda_i) \in \lambda$ is called *removable* (for λ) if $\lambda_i > \lambda_{i+1}$. A node $(i, \lambda_i + 1)$ is called *addable* (for λ) if $i = 1$ or $i > 1$ and $\lambda_i < \lambda_{i-1}$. We denote by

$$\lambda_A = \lambda \setminus \{A\} = (\lambda_1, \dots, \lambda_{i-1}, \lambda_i - 1, \lambda_{i+1}, \dots)$$

a partition of $n - 1$ obtained by removing a removable node $A = (i, \lambda_i)$ from λ. Similarly

$$\lambda^B = \lambda \cup \{B\} = (\lambda_1, \dots, \lambda_{i-1}, \lambda_i + 1, \lambda_{i+1}, \dots)$$

is a partition of $n + 1$ obtained by adding an addable node $B = (i, \lambda_i + 1)$ to λ.

We denote by

$$h_{ij} = h_{ij}^{\lambda} = \lambda_i - j + \lambda'_j - i + 1$$

the (i, j)-*hook length*. If a partition λ has r nodes on the main diagonal and there are α_i (resp., β_i) nodes to the right of (resp., below) the node (i, i) then we may write λ in the *Frobenius notation* (cf. [6]):

$$F(\lambda) = \begin{pmatrix} \alpha_1 & \cdots & \alpha_r \\ \beta_1 & \cdots & \beta_r \end{pmatrix}.$$

If $H_\lambda \cong S_{\lambda_1} \times S_{\lambda_2} \times \cdots < S_n$ is a Young subgroup we write M^λ for the permutation module $\mathbb{C}S_n \otimes_{\mathbb{C}H_\lambda} 1_{H_\lambda}$. The *Specht module* S^λ is explicitly defined as a submodule of M^λ (cf. [4]). The set $\{S^\lambda \mid \lambda \vdash n\}$ is a complete set of irreducible $\mathbb{C}S_n$-modules (up to isomorphism). We write $[\lambda]$ (or $[\lambda_1, \lambda_2, \dots]$) for the character of S^λ. Thus, $\{[\lambda] \mid \lambda \vdash n\}$ is a complete set of the irreducible characters of S_n. It is well known that S^λ is self-dual. Another fact (to be used without comment) is that $S^{(1^n)}$ is the 1-dimensional sign representation and $S^\lambda \otimes S^{(1^n)} \cong S^{\lambda'}$. The standard inner product on the class functions on a group (symmetric or alternating, depending on the context) is denoted by $\langle \cdot, \cdot \rangle$. If χ and ψ are two class functions we write $\chi \cdot \psi$ for the function $[g \mapsto \chi(g)\psi(g)]$. The character of $S^\lambda \otimes S^\mu$ is $[\lambda] \cdot [\mu]$. For $\lambda, \mu, \nu \vdash n$ we define the numbers $d(\mu, \nu; \lambda)$ via

$$[\mu] \cdot [\nu] = \sum_\lambda d(\mu, \nu; \lambda)[\lambda].$$

If $\alpha = (\alpha_1, \alpha_2, \dots)$ and $\beta = (\beta_1, \beta_2, \dots)$ are two partitions then we write $\beta \subseteq \alpha$ if $\beta_i \le \alpha_i$ for all i. In this case we also consider the skew partition α/β. We do not distinguish between α/β and its Young diagram, which is the set of nodes $\alpha \setminus \beta$.

If α/β is a skew Young diagram and $A = (i, j)$ is some node we say A is *connected* with α/β if at least one of the nodes $(i \pm 1, j), (i, j \pm 1)$ belongs to α/β. Otherwise A is *disconnected* from α/β.

If $\beta \vdash m, \gamma \vdash n, \alpha \vdash m+n$ we write $c^\alpha_{\beta\gamma}$ for the corresponding *Littlewood-Richardson coefficient*, which may be defined as the multiplicity of S^α in the induced module

$$S^\beta \hat{\otimes} S^\gamma := (S^\beta \boxtimes S^\gamma) \uparrow^{S_{m+n}}_{S_m \times S_n} .$$

The character of this module will be denoted $[\beta]\hat{\otimes}[\gamma]$. The *Littlewood-Richardson rule* [4, 6] gives a combinatorial description of the coefficients $c^\alpha_{\beta\gamma}$ and will be repeatedly used in this paper. It says that $c^\alpha_{\beta\gamma}$ is the number of semistandard tableaux of skew shape α/β and content γ, which give a lattice permutation when the entries are read from right to left along the rows starting from the top row.

Let α and β be two partitions. Then the *skew character* $[\alpha/\beta]$ is defined to be the sum

$$[\alpha/\beta] = \sum_\gamma c^\alpha_{\beta\gamma}[\gamma].$$

Note that $[\alpha/\beta] = 0$ unless $\beta \subseteq \alpha$.

The following four results will be used repeatedly.

Theorem 2.1 ([2, 1.6], [1, 1.1])**.** *Let μ, ν be partitions of n. Then*

$$\max\{\lambda_1 \mid d(\mu, \nu; \lambda) \neq 0 \textit{ for some } \lambda = (\lambda_1, \lambda_2, \dots)\} = |\mu \cap \nu|$$

and

$$\max\{m \mid d(\mu,\nu;\lambda) \neq 0 \ \textit{ for some } \ \lambda = (\lambda_1 \geq \cdots \geq \lambda_m > 0)\} = |\mu \cap \nu'|.$$

Since the skew characters can in principle be decomposed into the irreducible characters, the following theorem provides a recursive formula for the coefficients $d(\mu,\nu;\lambda)$.

Theorem 2.2 ([2, 2.3])**.** *Let μ, ν and $\lambda = (\lambda_1, \lambda_2, \dots)$ be partitions of n, and set $\hat{\lambda} = (\lambda_2, \lambda_3, \dots)$. Define*

$$Y(\lambda) = \{\eta \mid \eta \vdash n, \eta_i \geq \lambda_{i+1} \geq \eta_{i+1} \ \textit{ for all } \ i \geq 1\}.$$

Then

$$d(\mu,\nu;\lambda) = \sum_{\substack{\alpha \vdash \lambda_1 \\ \alpha \subseteq \mu \cap \nu}} \langle [\mu/\alpha] \cdot [\nu/\alpha], [\hat{\lambda}] \rangle - \sum_{\substack{\eta \in Y(\lambda) \\ \eta \neq \lambda \\ \eta_1 \leq |\mu \cap \nu|}} d(\mu,\nu;\eta).$$

Corollary 2.3 ([2, 2.4], [1, 2.1(d)])**.** *Let μ, ν and $\lambda = (\lambda_1, \lambda_2, \dots)$ be partitions of n, and set $\hat{\lambda} = (\lambda_2, \lambda_3, \dots)$, $\gamma = \mu \cap \nu$. Assume that $\lambda_1 = |\mu \cap \nu|$. Then*

$$d(\mu,\nu;\lambda) = \langle [\mu/\gamma] \cdot [\nu/\gamma], [\hat{\lambda}] \rangle.$$

Corollary 2.4 ([2, 2.4′])**.** *Let μ and ν be partitions of n, and $m = |\mu \cap \nu'|$. Let λ be a partition of n with m non-zero parts. Define $\bar{\lambda} = (\lambda_1 - 1, \lambda_2 - 1, \dots, \lambda_m - 1)$. Then*

$$d(\mu,\nu;\lambda) = \langle [\mu/(\mu \cap \nu')] \cdot [\nu/(\mu' \cap \nu)], [\bar{\lambda}] \rangle.$$

3. Homogenous S_n-products.

Lemma 3.1. *Let α, β, a, b be positive integers. Then*

(1)
$$\min(\alpha+\beta+1, a+b+1) < \min(\alpha,a) + \min(\beta,b) + \min(\alpha,b) + \min(\beta,a).$$

Proof. We may assume that $\alpha \leq \beta$, $a \leq b$ and $\alpha + \beta \leq a + b$. So the left hand side in (1) is $\alpha + \beta + 1$.

If $\beta \leq b$, then the right hand side of (1) equals

$$\min(\alpha,a) + \beta + \alpha + \min(\beta,a)$$

which is greater than $\alpha + \beta + 1$ since all numbers in this expression are positive integers.

If $b < \beta$, then the right hand side of (1) is

$$\min(\alpha,a) + b + \min(\alpha,b) + a \geq \min(\alpha,a) + \min(\alpha,b) + \alpha + \beta > \alpha + \beta + 1,$$

as claimed. □

Lemma 3.2. *Let μ, ν be partitions of n, both different from (n) and (1^n). Then*

$$\min(h_{11}^{\mu}, h_{11}^{\nu}) < |\mu \cap \nu| + |\mu \cap \nu'| - 2.$$

Proof. We write μ and ν in the Frobenius notation:

$$F(\mu) = \begin{pmatrix} \alpha_1 & \cdots & \alpha_r \\ \beta_1 & \cdots & \beta_r \end{pmatrix}, \qquad F(\nu) = \begin{pmatrix} a_1 & \cdots & a_s \\ b_1 & \cdots & b_s \end{pmatrix}.$$

We may assume that $r \leq s$. Then

$$h_{11}^{\mu} = \alpha_1 + \beta_1 + 1, \qquad h_{11}^{\nu} = a_1 + b_1 + 1,$$

$$|\mu \cap \nu| = r + \sum_{i=1}^{r} (\min(\alpha_i, a_i) + \min(\beta_i, b_i)),$$

$$|\mu \cap \nu'| = r + \sum_{i=1}^{r} (\min(\alpha_i, b_i) + \min(\beta_i, a_i)).$$

Since $r \geq 1$, it suffices to prove that

$$\min(\alpha_1 + \beta_1 + 1, a_1 + b_1 + 1) < \min(\alpha_1, a_1) + \min(\beta_1, b_1) + \min(\alpha_1, b_1) + \min(\beta_1, a_1).$$

But this follows from Lemma 3.1 since our assumption on the partitions ensures that $\alpha_1, \beta_1, a_1, b_1 > 0$. □

Theorem 3.3. *Let μ, ν be partitions of n, both different from (n) and (1^n). If $[\lambda]$ is a constituent of $[\mu] \cdot [\nu]$, then $h_{11}^{\lambda} < |\mu \cap \nu| + |\mu \cap \nu'| - 1$.*

Proof. Put $\ell = |\mu \cap \nu| + |\mu \cap \nu'| - 1$. Take π to be an ℓ-cycle in S_n. By Lemma 3.2, either μ or ν does not have a hook of length ℓ. Hence, by the Murnaghan-Nakayama Rule [4, 2.4.7], either $[\mu](\pi) = 0$ or $[\nu](\pi) = 0$. So

$$([\mu] \cdot [\nu])(\pi) = 0. \tag{2}$$

By Theorem 2.1, any constituent $[\lambda]$ of $[\mu] \cdot [\nu]$ satisfies

$$\lambda_1 \leq |\mu \cap \nu| \;\text{ and }\; \lambda_1' \leq |\mu \cap \nu'|$$

where $\lambda = (\lambda_1, \dots)$, $\lambda' = (\lambda_1', \dots)$. So the maximal possible hook length in λ is ℓ. Moreover, λ contains a hook of length ℓ if and only if $\lambda_1 = |\mu \cap \nu|$ and $\lambda_1' = |\mu \cap \nu'|$, in which case this is the $(1,1)$-hook whose leg length is $|\mu \cap \nu'| - 1$. In this case, using the Murnaghan-Nakayama Rule again, we get

$$[\lambda](\pi) = (-1)^{|\mu \cap \nu'| - 1} [\lambda \setminus H_{11}](1) \neq 0$$

where $\lambda \setminus H_{11}$ is the partition obtained from λ by removing the $(1,1)$-hook H_{11}. Hence for every constituent $[\lambda]$ of $[\mu] \cdot [\nu]$ containing an ℓ-hook we get a contribution on π of the same sign, and so no cancellation can occur. But this contradicts Equation (2). □

Theorem 3.4. *Let μ, ν be partitions of n, both different from (n) and (1^n). Then $[\mu] \cdot [\nu]$ is not homogenous.*

Proof. By Theorem 2.1, $[\mu]\cdot[\nu]$ has a constituent $[\lambda]$ with $\lambda_1 = |\mu\cap\nu|$ and a constituent $[\kappa]$ with $\kappa_1' = |\mu\cap\nu'|$. If $\lambda=\kappa$, then $h_{11}^\lambda = |\mu\cap\nu| + |\mu\cap\nu'| - 1$, which is impossible by Theorem 3.3. $\square$

Corollary 3.5. *A product $[\mu]\cdot[\nu]$ is irreducible if and only if at least one of the two characters $[\mu]$, $[\nu]$ is of degree* 1.

4. Kronecker products of S_n-representations with few components.

The main result of this section is a description of the products of S_n-representations with two homogeneous components. First we need to know the product of any character with the character $[n-1,1]$:

Lemma 4.1. *Let $n\geq 3$ and μ be a partition of n. Then*

$$[\mu]\cdot[n-1,1] = \sum_A\sum_B \left[(\mu_A)^B\right] - [\mu]$$

where the first sum is over all removable nodes A for μ, and the second sum runs over all addable nodes B for μ_A.

Proof. This follows from the isomorphisms $M^{(n-1,1)} \cong S^{(n-1,1)} \oplus S^{(n)}$ and $S^\mu \otimes M^{(n-1,1)} \cong (S^\mu\downarrow_{S_{n-1}})\uparrow^{S_n}$. $\square$

Corollary 4.2. *Let $n\geq 3$ and μ be a partition of n. Then:*

(i) *$[\mu]\cdot[n-1,1]$ has exactly one homogeneous component if and only if μ is (n) or (1^n).*

(ii) *$[\mu]\cdot[n-1,1]$ has exactly two homogeneous components if and only if μ is a rectangle (a^b) for some $a,b>1$. In this case we have*

$$[a^b]\cdot[n-1,1] = [a+1, a^{b-2}, a-1] + [a^{b-1}, a-1, 1].$$

(iii) *$[\mu]\cdot[n-1,1]$ has exactly three homogeneous components if and only if $n=3$ and $\mu=(2,1)$. In this case we have*

$$[2,1]\cdot[2,1] = [3]+[2,1]+[1^3].$$

(iv) *$[\mu]\cdot[n-1,1]$ has exactly four homogeneous components if and only if one of the following happens:*
  (a) *$n\geq 4$ and $\mu=(n-1,1)$ or $(2,1^{n-2})$;*
  (b) *$\mu=(k+1,k)$ or $(2^k,1)$ for $k\geq 2$.*
  We then have:

$$\begin{aligned}
[n-1,1]\cdot[n-1,1] &= [n]+[n-1,1]+[n-2,2]+[n-2,1^2],\\
[k+1,k]\cdot[2k,1] &= [k+2,k-1]+[k+1,k]\\
&\quad +[k+1,k-1,1]+[k^2,1],
\end{aligned}$$

and the remaining products are obtained by conjugation.

Proof. The "if" parts and the decompositions of the products follow from Lemma 4.1.

We now prove the "only if" directions. We are going to use Lemma 4.1 again. First, observe that $[\mu]$ appears as a constituent in the product $[\mu] \otimes [n-1,1]$ unless μ is a rectangle. Also note that $(\mu_A)^B = (\mu_{A'})^{B'}$ for two different pairs (A, B), (A', B') if and only if $A = B$ and $A' = B'$, in which case $(\mu_A)^B = (\mu_{A'})^{B'} = \mu$.

A partition with r removable nodes has exactly $r+1$ addable nodes. So if μ has at least 2 removable nodes, say A_1 and A_2, then μ_{A_1} and μ_{A_2} both have at least 2 addable nodes, which gives 4 composition factors in the product with the only common constituent $[\mu]$. This proves the "only if" part of (i) and (ii). If μ has at least 3 removable nodes, then a similar argument shows that $[\mu] \cdot [n-1,1]$ has at least 5 non-isomorphic constituents. So we may assume that μ has exactly two removable nodes: A_1 and A_2. For $[\mu] \cdot [n-1,1]$ to have exactly 3 components, both μ_{A_1} and μ_{A_2} should have only one removable node. This is only possible if $n = 3$ and $\mu = (2,1)$. Finally, for $[\mu] \cdot [n-1,1]$ to have exactly 4 components, one of μ_{A_1} and μ_{A_2} should have only one removable node and the other one should have two. This occurs exactly if μ or μ' is $(n-1,1)$, $n \geq 4$, or $(k+1,k)$, $k \geq 2$. ☐

Lemma 4.3. *Let λ be a partition of n. Then the square $[\lambda]^2$ has at most 4 homogeneous components if and only if one of the following holds:*

- (i) $\lambda = (n)$ *or* (1^n), *when* $[\lambda]^2 = [n]$;
- (ii) $n \geq 4$, $\lambda = (n-1,1)$ *or* $(2,1^{n-2})$, *when* $[\lambda]^2 = [n] + [n-1,1] + [n-2,2] + [n-2,1^2]$;
- (iii) $n = 3$, $\lambda = (2,1)$, *when* $[\lambda]^2 = [3] + [2,1] + [1^3]$;
- (iv) $n = 4$, $\lambda = (2^2)$, *when* $[\lambda]^2 = [4] + [2^2] + [1^4]$;
- (v) $n = 6$, $\lambda = (3^2)$ *or* (2^3), *when* $[\lambda]^2 = [6] + [4,2] + [3,1^3] + [2^3]$.

Proof. The "if" part follows from Corollary 4.2 and [**4**, Tables I.I].

In the other direction, let $[\lambda]^2$ have at most 4 homogeneous components. We may assume that λ is not one of $(n), (1^n), (n-1,1), (2,1^{n-2})$, and that $n > 8$ since for $n \leq 8$ the results hold by [**4**, Tables I.I].

Clearly $[\lambda]^2$ always contains $[n]$. Furthermore, by [**10**, Lemmas 1-3] and [**12**, 4.3] or by [**11**, 6.3], $[\lambda]^2$ contains $[n-2,2]$, and unless λ is a rectangle, it also contains $[n-1,1]$, $[n-2,1^2]$ and $[n-3,3]$. So we only have to deal with the case where $\lambda = (a^b)$ is a rectangle. We already know that $[\lambda]^2$ has the constituents $[n]$ and $[n-2,2]$. If $b > 2$, then $[\lambda]^2$ also has the constituent $[n-3,3]$ by [**10**, Lemma 3] or [**11**, 6.3]. If $n > 12$, then also $[n-4,4]$ occurs, see [**10**, Lemma 4]. Furthermore, by [**11**, 6.3], $[n-3,1^3]$ appears as a constituent. Hence we can restrict ourselves to the cases $\lambda = (k,k)$ or $\lambda = (4^3)$.

Suppose $\lambda = (k,k)$ $(k \geq 5)$. By Corollary 2.4, the components $[\mu]$ of $[k,k]^2$ with $\mu'_1 = 4 = |\lambda \cap \lambda'|$ are of the form $(\rho_1+1, \rho_2+1, \rho_3+1, \rho_4+1)$, where $[(\rho_1, \rho_2, \rho_3, \rho_4)]$, is a constituent of $[k-2, k-2]^2$. By what we have already proved, there are at least 3 such constituents. Thus $[k, k]^2$ has at least 5 components.

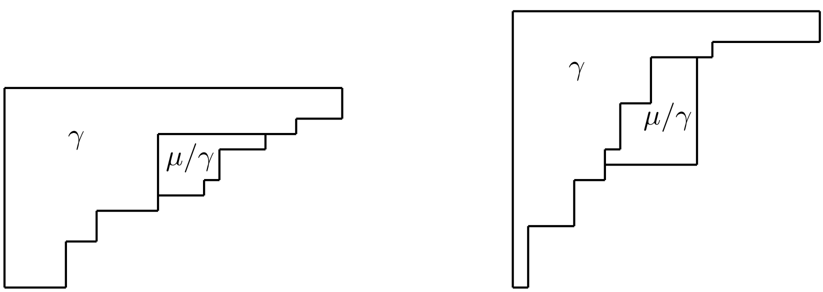
Now, let $\lambda = (4^3)$. We already know that $[\lambda]^2$ contains $[12]$, $[10, 2]$, $[9, 3]$ and $[9, 1^3]$. But it also contains some $[\mu]$ with $\mu_1' = 9 = |\lambda \cap \lambda'|$, thanks to Theorem 2.1. Alternatively, one may calculate $[4^3]^2$ on a computer and find 52 (!) homogeneous components. □

Lemma 4.4. *Let μ, γ be partitions, $\gamma \subset \mu$. Set $I = \{i \mid \gamma_i < \mu_i\}$. Then the following assertions are equivalent:*

(i) *$[\mu/\gamma]$ is homogeneous;*
(ii) *$[\mu/\gamma]$ is irreducible;*
(iii) *$I = \{j, j+1, \dots, k\}$ for some $j \le k$, and one of the following holds:*
  (a) *$\gamma_j = \gamma_{j+1} = \dots = \gamma_k$;*
  (b) *$\mu_j = \mu_{j+1} = \dots = \mu_k$.*
  Moreover, in this case $[\mu/\gamma] = [\alpha]$, where α is the partition with the parts $\mu_i - \gamma_i$, $i \in I$, sorted in the weakly decreasing order.

Proof. This follows from the Littlewood-Richardson Rule. □

Remark. The situations described in (iii)(a) and (iii)(b) above correspond respectively to the pictures



Lemma 4.5. *In the notation of Lemma 4.4 (and under the same assumptions), let A be a removable node of γ.*

(1) *If A is disconnected from μ/γ then*

$$[\mu/\gamma_A] = \sum_B [\alpha^B]$$

where B runs over the addable nodes of α.

(2) *Let A be connected with μ/γ.*
In the case (iii)(a) we have

$$[\mu/\gamma_A] = \sum_{B \neq B_0} [\alpha^B]$$

where B runs over the addable nodes of α, except for the bottom addable node B_0.

In the case (iii)(b) we have

$$[\mu/\gamma_A] = \left[\alpha^B\right]$$

where B is an addable node of α.

Proof. Again, this follows by the Littlewood-Richardson Rule. □

The following two lemmas will be used in the proof of the main theorem of this section.

Lemma 4.6. *Let $\mu \neq \nu$ be partitions of n, both different from (n), (1^n), $(n-1,1)$ and $(2,1^{n-2})$. Put $\gamma = \mu \cap \nu$, $m = |\gamma|$. Assume that ν/γ is a row and that $[\mu/\gamma]$ is an irreducible character $[\alpha_1, \alpha_2, \dots]$. Then $[m, \alpha_1, \alpha_2, \dots]$ appears in $[\mu] \cdot [\nu]$. Moreover if an S_{n-m+1}-character $[\theta_1, \theta_2, \dots]$ appears in*

$$\sum_{A \text{ removable for } \gamma} [\mu/\gamma_A] \cdot [\nu/\gamma_A] - \sum_{B \text{ addable for } \alpha} \left[\alpha^B\right] \tag{3}$$

with a positive coefficient then $[m-1, \theta_1, \theta_2, \dots]$ appears in $[\mu] \cdot [\nu]$.

Proof. We have $[\nu/\gamma] = [n-m]$. So Theorem 2.1 and Corollary 2.3 yield:

$$\langle [\mu] \cdot [\nu], [m, \alpha_1, \alpha_2, \dots] \rangle = 1, \tag{4}$$

and

$$\text{if } \lambda \neq (m, \alpha_1, \alpha_2, \dots) \text{ and } \langle [\mu] \cdot [\nu], [\lambda] \rangle \neq 0 \text{ then } \lambda_1 < m. \tag{5}$$

If λ is a partition of n with $\lambda_1 = m-1$, then in the notation of Theorem 2.2, we may write

$$\begin{aligned}&\{\eta \in Y(\lambda) \mid \eta \neq \lambda, \eta_1 \leq m\} \\ &\qquad = \{(m, \lambda_2, \dots, \lambda_{i-1}, \lambda_i - 1, \lambda_{i+1}, \dots) \mid i \geq 1,\ \lambda_i > \lambda_{i+1}\}.\end{aligned}$$

So (4) and (5) imply $\sum_{\substack{\eta \in Y(\lambda) \\ \eta \neq \lambda \\ \eta_1 \leq m}} d(\mu, \nu; \eta) = \varepsilon$, where

$$\varepsilon = \begin{cases} 1 & \text{if } \hat{\lambda} = \alpha^B \text{ for some addable node } B \text{ of } \alpha \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

Now, by Theorem 2.2, for a partition λ of n with $\lambda_1 = m-1$ we have

$$\langle [\mu] \cdot [\nu], [\lambda] \rangle = \sum_A \langle [\mu/\gamma_A] \cdot [\nu/\gamma_A], [\hat{\lambda}] \rangle - \varepsilon \tag{7}$$

where the sum is over all removable nodes A of γ.

Let $[\theta]$ be a constituent of $[\mu/\gamma_A] \cdot [\nu/\gamma_A]$. Then $[\theta]$ is a constituent of $[\beta] \cdot [\delta]$ with $[\beta]$ a constituent of $[\mu/\gamma_A]$ and $[\delta]$ a constituent of $[\nu/\gamma_A]$. It follows from the definition of skew characters that $\beta \subseteq \mu$, $\delta \subseteq \nu$. Hence $\beta \cap \delta \subseteq \mu \cap \nu = \gamma$. In view of Theorem 2.1, this implies

$$\theta_1 \leq |\beta \cap \delta| \leq |\mu \cap \nu| = m\,.$$

If $\theta_1 = m$, then $\beta \cap \delta = \gamma$, therefore $\beta \supseteq \gamma$ and $\delta \supseteq \gamma$. However, ν/γ_A is a union of a row and a node, so either $\delta = (n-m+1)$ or $\delta = (n-m, 1)$. If $\delta = (n-m+1)$, then $\mu \cap \nu \subseteq \delta$ implies $\mu \cap \nu = (m)$. But then either μ or ν is (n), which contradicts the assumptions of the lemma. If $\delta = (n-m, 1)$, then we conclude similarly that $\mu \cap \nu = (m-1, 1)$. Since neither μ nor ν is equal to $(n-1, 1)$ or its conjugate and μ/γ should be connected by Lemma 4.4, then the only possibilities are: $\nu = (m-1, n-m+1)$, $\mu = (m-1, 1^{n-m+1})$ or $\nu = (n-2, 1^2)$, $\mu = (n-2, 2)$ (in the latter case $n-m = 1$). In both cases $m-1 \geq n-m+1$, so $\theta_1 \leq m-1$ since θ is a partition of $n-m+1$. This contradiction shows that we may assume that $\theta_1 \leq m-1$ for any $[\theta]$ appearing in $[\mu/\gamma_A] \cdot [\nu/\gamma_A]$.

This, together with (7), shows that any S_{n-m+1}-character $[\theta_1, \theta_2, \ldots]$ appearing in (3) gives rise to the character $[m-1, \theta_1, \theta_2, \ldots]$ appearing in $[\mu] \cdot [\nu]$. □

Lemma 4.7. *Let $\mu \neq \nu$ be partitions of n, both different from $(n), (1^n), (n-1, 1)$, and $(2, 1^{n-2})$. Put $\gamma = \mu \cap \nu$. Assume that ν/γ is a row, $[\mu/\gamma]$ is irreducible, and $[\mu] \cdot [\nu]$ has 2 homogeneous components.*

If there exists a removable node A_0 of γ, disconnected from ν/γ, then the following condition holds:

(*) $[\mu/\gamma_{A_0}]$ is 1-dimensional, μ/γ is connected with all removable nodes of γ, ν/γ is connected with all removable nodes of γ except A_0.

Proof. Let A_0 be a removable node of γ disconnected from ν/γ, and put $m = |\gamma|$. Since $\mu \neq \nu$, we have $n-m > 0$. Let α be the partition of $n-m$ defined by $[\mu/\gamma] = [\alpha]$. Note that $[\nu/\gamma] = [n-m]$.

By Lemma 4.6, it suffices to show that the expression (3) contains at least two distinct irreducible characters unless the conditions (*) hold.

Since A_0 is disconnected from ν/γ, we have by Lemma 4.5(1):

$$[\nu/\gamma_{A_0}] = [n-m+1] + [n-m, 1]. \tag{8}$$

In view of Lemmas 4.4 and 4.5, we have three cases to consider: (a) When A_0 is disconnected from μ/γ; (b) when A_0 is connected with μ/γ and we are in the case (iii)(a) of Lemma 4.4; (c) when A_0 is connected with μ/γ and we are in the case (iii)(b) of Lemma 4.4 (the cases (b) and (c) overlap when μ/γ is a rectangle).

(a) In this case A_0 is disconnected from μ/γ. Then, by Lemma 4.5(1), we get

$$[\mu/\gamma_{A_0}] = \sum_B \left[\alpha^B\right]$$

where the sum runs over all addable nodes B of α. So (3) contains

$$\begin{aligned}&([n-m+1]+[n-m,1])\cdot\left(\sum_B\left[\alpha^B\right]\right)-\sum_B[\alpha^B]\\&=[n-m,1]\cdot\left(\sum_B\left[\alpha^B\right]\right).\end{aligned}$$

If there is a non-linear character among the $[\alpha^B]$, we are done by Corollary 4.2(i). Otherwise $\alpha=(1)$, but even in this case the expression above contains two different characters: $[2]$ and $[1^2]$. This completes the case (a).

In particular, we now may assume that every removable node A of γ disconnected from ν/γ is connected with μ/γ.

Note that $[\mu/\gamma_{A_0}]$ contains $[\alpha^{B_1}]$ for some addable node B_1, see Lemma 4.5. So, in view of (8) and Lemma 4.1, $[\mu/\gamma_{A_0}]\cdot[\nu/\gamma_{A_0}]$ contains $\sum_B[\alpha^B]$. Hence any removable node $A_1\neq A_0$ of γ yields a positive contribution of $[\mu/\gamma_{A_1}]\cdot[\nu/\gamma_{A_1}]$ to the expression (3). If A_1 is disconnected from ν/γ then $[\nu/\gamma_{A_1}]=[n-m,1]+[n-m+1]$, and the product $[\mu/\gamma_{A_1}]\cdot[\nu/\gamma_{A_1}]$ is not homogeneous. If A_1 is connected with ν/γ but disconnected from μ/γ then, by Lemma 4.5, $[\mu/\gamma_{A_1}]$ is not irreducible and $[\nu/\gamma_{A_1}]$ is $[n-m,1]$ or $[n-m+1]$. So the product $[\mu/\gamma_{A_1}]\cdot[\nu/\gamma_{A_1}]$ is not homogeneous again, thanks to Lemmas 4.1 and 4.5. Thus we may always assume that:

(**) μ/γ is connected with all removable nodes of γ, and ν/γ is connected with all removable nodes of γ different from A_0.

(b) In this case Lemma 4.5 yields

$$[\mu/\gamma_{A_0}]=\sum_{B\neq B_0}\left[\alpha^B\right]$$

where the sum runs over all addable nodes B of α except for the bottom one B_0. Consider the constituent $[\mu/\gamma_{A_0}]\cdot[\nu/\gamma_{A_0}]-\sum_B[\alpha^B]$ of (3). By (8), it is equal to

$$\begin{aligned}&([n-m+1]+[n-m,1])\cdot\left(\sum_{B\neq B_0}\left[\alpha^B\right]\right)-\sum_B\left[\alpha^B\right]\\&=[n-m,1]\cdot\left(\sum_{B\neq B_0}\left[\alpha^B\right]\right)-\left[\alpha^{B_0}\right].\end{aligned}\tag{9}$$

Since $\alpha\neq\emptyset$, it has at least 2 addable nodes. Let B_1 be an addable node of α, different from B_0, and let r be the number of removable nodes of α^{B_1}.

Then, using Lemma 4.1, we can rewrite (9) as follows:

$$[n-m,1]\cdot\left(\left[\alpha^{B_1}\right]+\sum_{B\neq B_0,B_1}\left[\alpha^{B}\right]\right)-\left[\alpha^{B_0}\right]$$

$$=(r-1)\left[\alpha^{B_1}\right]+\left[\alpha^{B_0}\right]+\sum_{B\neq B_0,B_1}\left[\alpha^{B}\right]+\sum_{C,D}\left[\left(\alpha^{B_1}\right)_C{}^{D}\right]$$

$$+[n-m,1]\cdot\left(\sum_{B\neq B_0,B_1}\left[\alpha^{B}\right]\right)-\left[\alpha^{B_0}\right]$$

$$=(r-1)\left[\alpha^{B_1}\right]+([n-m+1]+[n-m,1])\cdot\left(\sum_{B\neq B_0,B_1}\left[\alpha^{B}\right]\right)$$

$$+\sum_{C,D}\left[\left(\alpha^{B_1}\right)_C{}^{D}\right]$$

where the sum $\sum_{C,D}$ is over all removable nodes C of α^{B_1}, different from B_1, and over all addable nodes D of $(\alpha^{B_1})_C$, different from C.

If α is not a rectangle, then $\sum_{B\neq B_0,B_1}$ is non-empty, so our expression involves at least two different irreducible characters. Let α be a rectangle. If $[\alpha]$ is not of degree 1, then α^{B_1} is not a rectangle, so $r>1$, and thus our expression involves $[\alpha^{B_1}]$. Moreover, α^{B_1} has a removable node $C\neq B_1$, so for an addable node $D\neq C$ of $(\alpha^{B_1})_C$ we get the contribution $[(\alpha^{B_1})_C{}^{D}]\neq[\alpha^{B_1}]$. Finally, let $[\alpha]=[\mu/\gamma]$ be of degree 1. If $[\alpha^{B_1}]$ is not of degree 1, then it is $[2,1^{(n-m-1)}]$. So for $n-m\geq 2$, we have $r=2$, and so $[2,1^{n-2}]$ and $[3,1^{n-3}]$ appear in our expression. However, if $n-m=1$, then $[\mu/\gamma_{A_0}]$ is of degree 1. So, in view of (**), all the conditions in (*) hold.

(c) In this case by Lemma 4.5 we have

$$[\mu/\gamma_{A_0}]=\left[\alpha^{B_1}\right]$$

for some addable node B_1 of α. Then the constituent $[\mu/\gamma_{A_0}]\cdot[\nu/\gamma_{A_0}]-\sum_B[\alpha^B]$ of (3) is

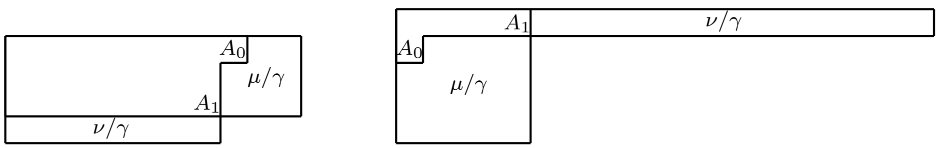
$$([n-m+1]+[n-m,1])\cdot\left[\alpha^{B_1}\right]-\sum_{B}\left[\alpha^{B}\right]\tag{10}$$

$$=\sum_{C,D}\left[\left(\alpha^{B_1}\right)_C{}^{D}\right]-\sum_{B}\left[\alpha^{B}\right]$$

$$=\sum_{B}\left[\alpha^{B}\right]+\sum_{C,D;\ C\neq B_1}\left[\left(\alpha^{B_1}\right)_C{}^{D}\right]-\sum_{B}\left[\alpha^{B}\right]$$

$$= \sum_{C,D;\ C\neq B_1} \left[\left(\alpha^{B_1}\right)_C{}^D \right].$$

In the last sum C runs through the removable nodes of α^{B_1}, different from B_1, and D runs through the addable nodes of $(\alpha^{B_1})_C$. So (10) has at least two different irreducible constituents, unless it is empty. Hence we may assume that α^{B_1} is a rectangle. If $[\alpha]$ is of degree 1 then $[\alpha^{B_1}] = [\mu/\gamma_{A_0}]$ is also of degree 1, and, in view of (**), we are in the exceptional case (*). So we may assume that $\alpha = (a^{b-1}, a-1)$ for some $a > 1$, $b > 1$, and $\alpha^{B_1} = (a^b)$. This together with Lemma 4.4 implies that γ has a removable node A_1, different from A_0. We know that it must be connected with ν/γ and μ/γ, thanks to (**). If there was a third removable node of γ, A_2 say, then again by (**), both A_1 and A_2 would be connected with both ν/γ and μ/γ. But this is impossible since $\gamma = \mu\cap\nu$. So we may assume that γ has exactly two removable nodes. Now, by Lemma 4.5(2), we have $[\mu/\gamma_{A_1}] = [\alpha^{B_2}]$ with B_2 the top or the bottom, but not the middle, addable node of α, and $[\nu/\gamma_{A_1}]$ is either $[n-m,1]$ or $[n-m+1]$. The corresponding pictures are:



In the first case, $[\alpha^{B_2}]\cdot[n-m,1]$ contributes at least two constituents by Theorem 3.4. In the second case $\nu = (n-1,1)$. $\square$

Theorem 4.8. *Let μ, ν be partitions of n. Then $[\mu]\cdot[\nu]$ has exactly two homogenous components if and only if one of the partitions μ,ν is a rectangle (a^b) with $a, b > 1$, and the other is $(n-1,1)$ or $(2,1^{n-2})$. In these cases we have:*

$$[n-1,1]\cdot[a^b] = [a+1, a^{b-2}, a-1] + [a^{b-1}, a-1, 1],$$
$$[2,1^{n-2}]\cdot[a^b] = [b+1, b^{a-2}, b-1] + [b^{a-1}, b-1, 1].$$

Proof. The "if" part is proved in Corollary 4.2 (note that $S^{(2,1^{n-2})} \cong S^{(n-1,1)}\otimes \mathrm{sign}$). To prove the "only if" part, assume that

$$[\mu]\cdot[\nu] = x[\kappa] + y[\lambda] \quad \text{for some } x, y \in \mathbb{N}\,,$$

with $\kappa > \lambda$ in the lexicographic order. Clearly, $\mu,\nu \notin \{(n),(1^n)\}$. If μ or ν is $(n-1,1)$ or $(2,1^{n-2})$ the result follows from Corollary 4.2. Assume $\mu,\nu \notin \{(n-1,1),(2,1^{n-2})\}$. By Theorems 2.1 and 3.3, we have

$$\kappa_1 = |\mu\cap\nu|, \quad \lambda_1' = |\mu\cap\nu'| \quad \text{and} \quad \lambda_1 < |\mu\cap\nu| = \kappa_1.$$

By Lemma 4.3, $\mu \neq \nu$, and hence $\kappa_1 < n$. Put $\gamma = \mu \cap \nu$, $m = |\gamma|$. By Corollary 2.3, we must have

$$[\mu/\gamma] \cdot [\nu/\gamma] = x[\hat{\kappa}]$$

where $\hat{\kappa} = (\kappa_2, \kappa_3, \dots)$. So, in view of Theorem 3.4, one of the following happens:

(i) $x = 1$ and one of the characters $[\mu/\gamma]$, $[\nu/\gamma]$ is of degree 1, while the other is irreducible;

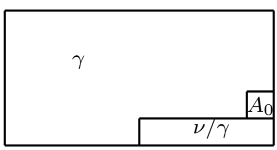
(ii) one of the characters $[\mu/\gamma]$, $[\nu/\gamma]$ is equal to $[\hat{\kappa}]$, the other is of the form $z[n-m]+w[1^{n-m}]$ with some $z, w \in \mathbb{N}$, and $\hat{\kappa} = \hat{\kappa}'$. By the Littlewood-Richardson rule, a skew character contains both $[n-m]$ and $[1^{n-m}]$ only if its diagram is a set of disconnected nodes. So we must have $n - m = 2$, since otherwise such a skew character has more than 2 constituents. But there is no symmetric partition of 2, i.e. $\hat{\kappa} \neq \hat{\kappa}'$. This contradiction allows us to assume that we are in the case (i).

Without loss of generality, suppose that $[\nu/\gamma]$ is of degree 1 and $[\mu/\gamma] = [\alpha]$ is irreducible. Then the shape of ν/γ is a row or a column. Passing, if necessary, from μ, ν to μ', ν', we may assume that ν/γ is a row. Now, by Lemma 4.7 we may assume that one of the following holds:

(a) ν/γ is connected with every removable node of γ.

(b) There exists a removable node A_0 of γ disconnected from ν/γ, $[\mu/\gamma]$ and $[\mu/\gamma_{A_0}]$ are of degree 1, μ/γ is connected with every removable node of γ, and ν/γ is connected with every removable node of γ different from A_0.

Case (a). In this case ν must be a rectangle, and γ must have a removable node A_0 such that $[\nu/\gamma_{A_0}] = (n-m, 1)$ for otherwise $\nu = (n)$.



Let us first assume that μ/γ is disconnected from A_0. Then, in view of Lemmas 4.5(1) and 4.1, the expression (3) contains

$$\begin{aligned}
(11) \qquad & [\nu/\gamma_{A_0}] \cdot [\mu/\gamma_{A_0}] - \sum_B \left[\alpha^B\right] \\
&= [n-m, 1] \cdot \left(\sum_B \left[\alpha^B\right]\right) - \sum_B \left[\alpha^B\right] \\
&= \sum_B \sum_C \sum_D \left[\left(\alpha^B\right)_C{}^D\right] - 2\sum_B \left[\alpha^B\right]
\end{aligned}$$

$$= \sum_B (r_B - 2)\left[\alpha^B\right] + \sum_B \sum_C \sum_{D \neq C} \left[\left(\alpha^B\right)_C{}^D\right]$$

where B runs over the addable nodes of α, C runs over the removable nodes of α^B (for the respective node B), D runs over the addable nodes of $(\alpha^B)_C$ and r_B denotes the number of removable nodes of α^B.

If α has at least 3 addable nodes, say B_0, B_1, B_2, then we have the following contribution to the expression above:

$$\begin{aligned}&(r_{B_0} - 2)\left[\alpha^{B_0}\right] + \left[\alpha^{B_1}\right] + \left[\alpha^{B_2}\right] + (r_{B_1} - 2)\left[\alpha^{B_1}\right] + \left[\alpha^{B_0}\right] + \left[\alpha^{B_2}\right]\\ &\quad +(r_{B_2} - 2)\left[\alpha^{B_2}\right] + \left[\alpha^{B_0}\right] + \left[\alpha^{B_1}\right]\\ &= r_{B_0}\left[\alpha^{B_0}\right] + r_{B_1}\left[\alpha^{B_1}\right] + r_{B_2}\left[\alpha^{B_2}\right].\end{aligned}$$

By Lemma 4.6, this yields 3 irreducible components in $[\mu] \cdot [\nu]$.

So α has exactly two addable nodes, say B_0, B_1, i.e. α is a rectangle. Then we have the following contribution to the expression (11):

$$(r_{B_0} - 2)\left[\alpha^{B_0}\right] + \left[\alpha^{B_1}\right] + (r_{B_1} - 2)\left[\alpha^{B_1}\right] + \left[\alpha^{B_0}\right].$$

If α is not a row or a column then both r_{B_0}, r_{B_1} are at least 2, and in view of Lemma 4.6, we get two irreducible constituents for $[\mu] \cdot [\nu]$, both different from $[\kappa]$. Let α be a row or a column. Assume that α is a row, the column case being similar. Then (11) equals $[n-m, 1] + [n-m-1, 2] + [n-m-1, 1^2]$ if $n - m > 2$, and $[2, 1] + [1^3]$ if $n - m = 2$. By Lemma 4.6, this yields at least two constituents in $[\mu] \cdot [\nu]$ different from $[\kappa]$. Finally, let $n - m = 1$. Then (11) equals 0. Note that γ must have a removable node $A_1 \neq A_0$, since otherwise $\nu = (1^n)$. If μ/γ is disconnected from A_1, then

$$[\mu/\gamma_{A_1}] \cdot [\nu/\gamma_{A_1}] = [2] + [1^2],$$

and we are done by Lemma 4.6. If μ/γ is connected with A_1, then $\mu = (2^k, 1^2), \nu = (2^{k+1})$ (and $k > 1$ since μ is not of the form $(2, 1^{n-2})$). Then the expression (3) equals $[1^2]$. So, by Lemma 4.6, $[n-1, 1]$ and $[n-2, 1^2]$ are constituents of $[\mu] \cdot [\nu]$. But $|\mu \cap \nu'| = 4$, so there also must be a constituent with 4 non-zero rows, thanks to Theorem 2.1.

This completes the consideration of the case where $[\mu/\gamma]$ is disconnected from A_0.

Let μ/γ be connected with A_0. Then, in view of Lemmas 4.4 and 4.5(2), we have

$$[\mu/\gamma_{A_0}] = \sum_{B \neq B_0} \left[\alpha^B\right]$$

where B_0 is the bottom addable node of α. Let B_1 be the top addable node of α. Then we get a contribution to (3) from the following expression:

$$[\nu/\gamma_{A_0}] \cdot [\mu/\gamma_{A_0}] - \sum_B \left[\alpha^B\right] \tag{12}$$

$$
\begin{aligned}
&= [n-m,1]\cdot \sum_{B\neq B_0} [\alpha^B] - \sum_B [\alpha^B] \\
&= \sum_{B\neq B_0}\sum_C\sum_D \left[\left(\alpha^B\right)_C{}^D \right] - \sum_{B\neq B_0} [\alpha^B] - \sum_B [\alpha^B] \\
&= \sum_{C\neq B_1}\sum_D \left[\left(\alpha^{B_1}\right)_C{}^D \right] + \sum_{B\neq B_0,B_1}\sum_C\sum_D \left[\left(\alpha^B\right)_C{}^D \right] - \sum_{B\neq B_0} [\alpha^B]
\end{aligned}
$$

where B runs through the addable nodes of α, C runs through the removable nodes of α^B (for the respective node B) and D runs through the addable nodes of $(\alpha^B)_C$.

If α has a third addable node, say B_2, then α^{B_1} is not a rectangle, and hence there exists a node $C_1 \neq B_1$ which is removable from α^{B_1}. This shows that the first sum in (12) contains $[\alpha^{B_1}]$. Moreover, the second sum in (12) contains $\sum_D[\alpha^D]$, and so both $[\alpha^{B_0}]$ and $[\alpha^{B_1}]$ are constituents of (12). Now we can apply Lemma 4.6.

If B_0 and B_1 are the only addable nodes of α, then α is a rectangle. Let C_1 be the corner node of α.

If α is not a row, then α^{B_1} also has the removable node C_1. In this case, (12) is

$$\sum_D \left[\left(\alpha^{B_1}\right)_{C_1}{}^D \right] - \left[\alpha^{B_1}\right]$$

which gives at least two contributions, except in the case where $\alpha = (1^2)$ when (12) equals $[3]$. If γ has a further removable node A_1, then this leads to a further contribution $[2,1]$ to (3). But if γ is a rectangle, then $\mu = (2^3)$ and $\nu = (3^2)$, and we can apply Lemma 4.3.

If α is a row then γ must have a removable node $A_1 \neq A_0$, since otherwise $\mu = (n)$. Note that (12) equals $-[n-m+1]$. Also $[\nu/\gamma_{A_1}]\cdot[\mu/\gamma_{A_1}] = [n-m+1] + [n-m,1]$. By Lemma 4.6, the product $[\mu]\cdot[\nu]$ contains $[m, n-m]$ and $[m-1, n-m, 1]$. Note that our assumptions yield $\mu = (k+n-m, k-n+m)$, $\nu = (k,k)$ with $k-n+m \geq 2$. But in this case $|\mu \cap \nu'| \geq 4$. So Theorem 2.1 implies that $[\mu]\cdot[\nu]$ has a constituent with 4 rows.

Case (b). Since $[\mu]$ is not of degree 1, the assumption $[\mu/\gamma]$ and $[\mu/\gamma_{A_0}]$ are of degree 1 implies that γ must have a removable node $A_1 \neq A_0$. By assumption, A_1 is connected with both μ/γ and ν/γ, and since $[\mu/\gamma]$ is of degree 1, A_1 and A_0 are the only removable nodes of γ.

Since $[\mu/\gamma_{A_0}]$ is 1-dimensional, we conclude from Lemmas 4.4 and 4.5 that $[\mu/\gamma_{A_1}] = [n-m,1]$ or the conjugate. So if $n-m > 1$ and $[\nu/\gamma_{A_1}] = [n-m,1]$ then (3) equals $[n-m,1]\cdot[n-m,1]$ or the conjugate. Now we apply Corollary 4.2 and Lemma 4.6. Otherwise $\mu = (k,k), \nu = (k+n-m, k-n+m)$ or $\mu = (2^k), \nu = (2^{k-1}, 1^2)$. But these cases have already been considered. □

Thus we have classified all pairs μ, ν such that $[\mu] \cdot [\nu]$ has at most 2 homogeneous components. The "if-parts" of the following conjecture are proved in Corollary 4.2 and Lemma 4.3.

Conjecture.

(i) $[\mu] \cdot [\nu]$ has 3 homogeneous components if and only if $n = 3$ and $\mu = \nu = (2,1)$ or $n = 4$ and $\mu = \nu = (2,2)$.

(ii) $[\mu]\cdot[\nu]$ has 4 homogeneous components if and only if one of the following happens:

  (a) $n \geq 4$ and $\mu, \nu \in \{(n-1,1), (2,1^{n-2})\}$;

  (b) $n = 2k+1$ for some $k \geq 2$, and one of μ, ν is in $\{(2k,1),(2,1^{2k-1})\}$ while the other one is in $\{(k+1,k),(2^k,1)\}$;

  (c) $n = 6$ and $\mu, \nu \in \{(2^3),(3^2)\}$.

The following theorem proves the conjecture in the special case when both μ and ν are symmetric.

Theorem 4.9. *Let μ and ν be symmetric partitions of n. Then $[\mu]\cdot[\nu]$ has at most 4 homogeneous components if and only if one of the following holds:*

(i) $n = 1$;

(ii) $n = 3$, $\mu = \nu = (2,1)$, *when* $[\mu]^2 = [3] + [2,1] + [1^3]$;

(iii) $n = 4$, $\mu = \nu = (2^2)$, *when* $[\mu]^2 = [4] + [2^2] + [1^4]$.

Proof. Let $\gamma = \mu \cap \nu$, $m = |\gamma|$. Then γ is a symmetric partition, and at least one of the skew diagrams μ/γ, ν/γ has no box on the main diagonal. Say it is μ/γ. Because of the symmetry, we can then write μ/γ as a disjoint union $\alpha \cup \alpha'$, where α and α' are some skew shapes which are conjugate to each other. In particular, $n-m$ is even. By [**6**, (5.7)],

$$[\mu/\gamma] = [\alpha]\hat{\otimes}[\alpha'].$$

If every constituent of $[\alpha]\hat{\otimes}[\alpha']$ belongs to $M = \{[n-m],[1^{n-m}],[n-m-1,1],[2,1^{n-m-2}]\}$ then by the Littlewood-Richardson Rule, every constituent of $[\alpha]$ and $[\alpha']$ would have to belong to $\{[(n-m)/2],[1^{(n-m)/2}],[(n-m)/2-1,1],[2,1^{(n-m)/2-2}]\}$. But even then, if $n-m \geq 6$, the Littlewood-Richardson Rule implies that there are components of $[\alpha]\hat{\otimes}[\alpha']$ not in M.

Assume first that $n-m \geq 6$. Then, by the Littlewood-Richardson Rule again, $[\nu/\gamma]$ contains a constituent not in M. Now Theorems 3.4 and 4.8 imply that $[\mu/\gamma]\cdot[\nu/\gamma]$ contains at least three different irreducible constituents, say $[\hat{\rho}_1]$, $[\hat{\rho}_2]$, $[\hat{\rho}_3]$. Then $[\mu]\cdot[\nu]$ contains the corresponding constituents $[\rho_1]$, $[\rho_2]$, $[\rho_3]$, thanks to Corollary 2.3. Since μ and ν are symmetric, $[\mu]\cdot[\nu]$ also contains the conjugate constituents $[\rho_1']$, $[\rho_2']$, $[\rho_3']$. Now, by Theorem 3.3 no constituent can have at the same time the maximal length and width among all the constituents. Hence $[\rho_i] \neq [\rho_j']$ for all i,j. Thus we have found 6 different irreducible constituents.

The case $n-m=0$ follows from Lemma 4.3. So we may now assume that $n-m=2$ or 4. Note that in the first case $n>7$ since for $n\le 7$ there is only one symmetric partition, and in the second case $n>8$, since the intersection of the two different symmetric partitions for $n=8$ is a partition of 6. Then by the Littlewood-Richardson Rule and Corollary 2.3, we know that $[\mu]\cdot[\nu]$ has the constituents $[n-2,2]$, $[n-2,1^2]$ and their conjugates if $n-m=2$, and it has the constituents $[n-4,3,1]$ and $[n-4,2,1^2]$ and their conjugates if $n-m=4$. By the remark above, n is sufficiently large in both cases so that the four constituents are all different.

Assume that these are all the constituents of $[\mu]\cdot[\nu]$. Consider the case $n-m=4$. We compute the character values on $(n-1)$-cycles and $(n-2)$-cycles. Since $|\gamma|=n-4$, we know that $\min(h^\mu_{11},h^\nu_{11})<n-2$. Hence on an $(n-1)$-cycle z_{n-1} and an $(n-2)$-cycle z_{n-2} in S_n we have by the Murnaghan-Nakayama rule:

$$[\mu](z_{n-1})\cdot[\nu](z_{n-1})=0=[\mu](z_{n-2})\cdot[\nu](z_{n-2}).$$

On the other hand, if n is even, then

$$[n-4,2,1^2](z_{n-1})=-1=[4,2,1^{n-6}](z_{n-1})$$

and

$$[n-4,3,1](z_{n-1})=0=[3,2^2,1^{n-7}](z_{n-1})$$

gives a contradiction. If n is odd, then similarly

$$[n-4,2,1^2](z_{n-2})=0=[4,2,1^{n-6}](z_{n-2})$$

and

$$[n-4,3,1](z_{n-2})=1=[3,2^2,1^{n-7}](z_{n-2})$$

gives a contradiction. The case $n-m=2$ is considered similarly using z_n and z_{n-1}. □

5. Homogeneous Kronecker products of A_n-representations.

We first recall the classification of the complex irreducible A_n-representations (cf. [**4**, 2.5]). If μ is a non-symmetric partition of n then the restrictions $S^\mu\downarrow_{A_n}$ and $S^{\mu'}\downarrow_{A_n}$ are irreducible and isomorphic to each other. We denote the corresponding irreducible A_n-module by T^μ or $T^{\mu'}$. Thus $T^\mu\cong T^{\mu'}$ for $\mu\neq\mu'$. On the other hand, if $\mu=\mu'$ then $S^\mu\downarrow_{A_n}$ splits into a direct sum of two non-isomorphic A_n-modules, say T^μ_+ and T^μ_-. Moreover, the modules T^μ_+ and T^μ_-, as μ runs over all symmetric partitions of n, together with the modules T^μ, as μ runs over a system of representatives of the pairs $\{\mu,\mu'\}$ for the non-symmetric partitions μ of n, form a complete system of the non-isomorphic irreducible A_n-modules. It is well known that $T^\mu_\pm$ is obtained from $T^\mu_\mp$ by twisting with an automorphism of A_n, which comes from a conjugation by an element $g \in S_n \setminus A_n$. The character of $T^\mu_{(\pm)}$ will be denoted by $\{\mu\}_{(\pm)}$.

Lemma 5.1. *Let μ, ν be non-symmetric partitions of n, both different from (n) and (1^n). Then $T^\mu \otimes T^\nu$ is homogeneous if and only if $S^\mu \otimes S^\nu \cong x\, S^\lambda \oplus y\, S^{\lambda'}$ for some $\lambda \neq \lambda'$, $x, y \in \mathbb{N}$.*

Proof. This follows from the definition of the modules T^μ and Theorem 3.4. □

Lemma 5.2. *Let μ, ν be partitions of n, both different from $(n), (1^n)$. Assume that $\mu \neq \mu'$, $\nu = \nu'$. Then $T^\mu \otimes T^\nu_\pm$ is homogeneous if and only if $S^\mu \otimes S^\nu \cong x\, S^\lambda \oplus y\, S^{\lambda'}$ for some $\lambda \neq \lambda'$, $x, y \in \mathbb{N}$.*

Proof. The "if-part" is clear.

If $T^\mu \otimes T^\nu_+ \cong x\, T^\lambda_\pm$ for some $\lambda = \lambda'$, then, conjugating by $g \in S_n \setminus A_n$, we get $T^\mu \otimes T^\nu_- \cong x\, T^\lambda_\mp$. So

$$T^\mu \otimes (T^\nu_+ \oplus T^\nu_-) \cong x(T^\lambda_+ \oplus T^\lambda_-).$$

The lift to S_n gives $S^\mu \otimes S^\nu \cong x\, S^\lambda$, which is impossible by Theorem 3.4.

If $T^\mu \otimes T^\nu_+ \cong x\, T^\lambda$ for some $\lambda \neq \lambda'$, then as above we have $T^\mu \otimes T^\nu_- \cong x\, T^\lambda$, so the lift to S_n gives $S^\mu \otimes S^\nu \cong y\, S^\lambda \oplus z\, S^{\lambda'}$ (with $y + z = x$). □

Lemma 5.3. *Let ν be a symmetric partition of n, and let ϕ, ψ be irreducible A_n-characters both different from $\{\nu\}_+$ and $\{\nu\}_-$. Then*

$$\langle \psi \cdot \{\nu\}_+, \phi \rangle \;=\; \langle \psi \cdot \{\nu\}_-, \phi \rangle.$$

Proof. By [4, 2.5.13], we have

$$\begin{aligned}
\langle \psi \cdot \{\nu\}_\pm, \phi \rangle &= \frac{1}{|A_n|} \sum_{g \in A_n} \psi(g)\{\nu\}_\pm(g)\overline{\phi(g)} \\
&= \frac{1}{|A_n|} \left(\sum_{g \in A_n \setminus (C_\nu^+ \cup C_\nu^-)} \psi(g)\{\nu\}_\pm(g)\overline{\phi(g)} \right. \\
&\quad + \sum_{g \in C_\nu^+} \psi(g) \frac{1}{2} \left(\varepsilon_\nu \pm \sqrt{\varepsilon_\nu \prod_i h^\nu_{ii}} \right) \overline{\phi(g)} \\
&\quad \left. + \sum_{g \in C_\nu^-} \psi(g) \frac{1}{2} \left(\varepsilon_\nu \mp \sqrt{\varepsilon_\nu \prod_i h^\nu_{ii}} \right) \overline{\phi(g)} \right)
\end{aligned}$$

where $\varepsilon_\nu = (-1)^{(n-k)/2}$ and $C^\pm_\nu$ denote the *two* conjugacy classes in A_n which consist of elements of cycle type $(h^\nu_{11}, \ldots, h^\nu_{kk})$. Since ψ, ϕ correspond to partitions different from ν, each of them takes the same value on C_ν^+ and C_ν^-, so the last expression is the same for $\{\nu\}_+$ and $\{\nu\}_-$. □

Lemma 5.4. *Let ν be a symmetric partition of n and let ψ be an irreducible A_n-character different from $\{\nu\}_+$ and $\{\nu\}_-$. Then*

$$\langle \psi \cdot \{\nu\}_+, \{\nu\}_+\rangle = \langle \psi \cdot \{\nu\}_-, \{\nu\}_-\rangle \qquad \text{and}$$
$$\langle \psi \cdot \{\nu\}_+, \{\nu\}_-\rangle = \langle \psi \cdot \{\nu\}_-, \{\nu\}_+\rangle.$$

Proof. We compute the scalar products using [**4**, 2.5.13] as in the previous proof, and use the facts that $\{\nu\}_+(g) = \{\nu\}_-(g)$ for any $g \in A_n \setminus (C_\nu^+ \cup C_\nu^-)$ and $\psi(g) = \psi(h)$ for any $g, h \in C_\nu^+ \cup C_\nu^-$. □

From the previous two results we deduce:

Proposition 5.5. *Let μ, ν be symmetric partitions of n, $\mu \neq \nu$. Then $\{\mu\}_+ \cdot \{\nu\}_+$ is homogeneous if and only if $\{\mu\}_+ \cdot \{\nu\}_-$ is homogeneous.*

Now we can classify the homogeneous Kronecker products of irreducible A_n-characters. Note that if $n > 4$ then the only 1-dimensional character is the trivial one. For $n = 3$ and 4 there are two more 1-dimensional characters in each case: $\{2,1\}_\pm$ and $\{2^2\}_\pm$.

Theorem 5.6. *Let ϕ, ψ be irreducible A_n-characters both of degrees greater than 1. Then $\phi \cdot \psi$ is homogeneous if and only if $n = a^2$ for some $a > 2$ and one of the characters is $\{n-1, 1\}$, while the other is $\{a^a\}_+$ or $\{a^a\}_-$. In the exceptional case:*

$$\{n-1, 1\} \cdot \{a^a\}_\pm = \{a+1, a^{a-2}, a-1\}.$$

Proof. The "if-part" follows from Corollary 4.2(ii).

Let ϕ and ψ correspond to partitions μ and ν, respectively. If μ and ν are both non-symmetric, then by Lemma 5.1 and Theorem 4.8 the tensor product $T^\mu \otimes T^\nu$ is not homogeneous. If one of the partitions μ, ν is symmetric and the other is not, use Lemma 5.2 and Theorem 4.8. So we may assume that μ and ν are both symmetric. If $\mu \neq \nu$, then by Lemmas 5.3, 5.4 and 5.5, if one of the four products $\{\mu\}_\pm \cdot \{\nu\}_\pm$ is homogeneous then the product $[\mu] \cdot [\nu]$ has at most two homogeneous components, contradicting Theorems 3.4 and 4.8. Indeed, consider for example the case where $\{\mu\}_- \cdot \{\nu\}_-$ is homogeneous. Since $\{\lambda\}_\pm$ is obtained from $\{\lambda\}_\mp$ by conjugating with an element $g \in S_n \setminus A_n$, we conclude that $\{\mu\}_+ \cdot \{\nu\}_+$ is also homogeneous. Moreover, if $\{\mu\}_- \cdot \{\nu\}_- = x\{\lambda\}$ then $\{\mu\}_+ \cdot \{\nu\}_+ = x\{\lambda\}$, and if $\{\mu\}_- \cdot \{\nu\}_- = x\{\kappa\}_\pm$ then $\{\mu\}_+ \cdot \{\nu\}_+ = x\{\kappa\}_\mp$. By Proposition 5.5, we also have that $\{\mu\}_\pm \cdot \{\nu\}_\mp$ are homogeneous. Moreover, Lemmas 5.3, 5.4 imply $\{\mu\}_\pm \cdot \{\nu\}_\mp = \{\lambda\}$ or $\{\kappa\}_{\pm \text{ or } \mp}$. Thus $[\mu] \cdot [\nu]$ is $x[\lambda] + y[\lambda']$ or $x[\kappa]$.

Now let $\mu = \nu$ be symmetric. We have to consider three cases: $\{\mu\}_\pm \cdot \{\mu\}_\pm$ and $\{\mu\}_+ \cdot \{\mu\}_-$. Using conjugation with $g \in S_n \setminus A_n$, we can eliminate one of them, and work only with $\{\mu\}_+ \cdot \{\mu\}_+$ and $\{\mu\}_+ \cdot \{\mu\}_-$. Let us consider the first case (the second one is similar). So let $\{\mu\}_+ \cdot \{\mu\}_+ = x\psi$ for some irreducible A_n-character ψ.

If the dual character $\{\mu\}_+^*$ is equal to $\{\mu\}_+$, then

$$\langle \{n\}, \{\mu\}_+ \cdot \{\mu\}_+ \rangle \ = \ \langle \{\mu\}_+, \{\mu\}_+ \rangle \ = 1,$$

so we deduce $\{\mu\}_+ \cdot \{\mu\}_+ = \{n\}$, which is impossible as $\{\mu\}$ is not of degree 1.

If $\{\mu\}_+^* = \{\mu\}_-$, then

$$\langle \{n\}, \{\mu\}_+ \cdot \{\mu\}_+ \rangle \ = \ \langle \{\mu\}_-, \{\mu\}_+ \rangle \ = 0$$

and

$$\begin{aligned} \langle \{n-1,1\}, \{\mu\}_+ \cdot \{\mu\}_+ \rangle \ &= \ \langle \{n\} + \{n-1,1\}, \{\mu\}_+ \cdot \{\mu\}_+ \rangle \\ &= \ \langle \{n-1\} \uparrow^{A_n}, \{\mu\}_+ \cdot \{\mu\}_+ \rangle \\ &= \ \langle \{\mu\}_- \downarrow_{A_{n-1}}, \{\mu\}_+ \downarrow_{A_{n-1}} \rangle. \end{aligned}$$

Consider the case where μ is not a square. Then, by the Branching Rule, both restrictions in the last expression contain some $\{\lambda\}$ where λ is a non-symmetric partition of $n-1$. So the scalar product above is non-zero, whence $\{\mu\}_+ \cdot \{\mu\}_+ = x\{n-1,1\}$. Take $z \in A_n$ of cycle type $(n-2,2)$, if n is even and of cycle type $(n-2,1,1)$, if n is odd. As μ is symmetric it does not have a hook of length $n-2$. Hence by [**4**, 2.5.13] and the Murnaghan-Nakayama Rule we have

$$\{\mu\}_+(z)\{\mu\}_+(z) = 0\,.$$

On the other hand, $x\{n-1,1\}(z) = \pm x \neq 0$, when n is odd or even, respectively. This is a contradiction.

It remains to deal with the case where $\{\mu\}_+^* = \{\mu\}_-$ and μ is a square. Consider

$$\langle \{n-2\} \uparrow^{A_n}, \{\mu\}_+ \cdot \{\mu\}_+ \rangle = \langle \{\mu\}_- \downarrow_{A_{n-2}}, \{\mu\}_+ \downarrow_{A_{n-2}} \rangle.$$

By the Branching Rule, the last scalar product is non-zero. But

$$\{n-2\} \uparrow^{A_n} = \{n\} + 2\{n-1,1\} + \{n-2,2\} + \{n-2,1^2\},$$

and the product $\{\mu\}_+ \cdot \{\mu\}_+$ can not be of the form $x\{n\}$ or $x\{n-1,1\}$ by the same arguments as before. So we may assume that

$$\{\mu\}_+ \cdot \{\mu\}_+ = x\{n-2,2\} \quad \text{or} \quad \{\mu\}_+ \cdot \{\mu\}_+ = x\{n-2,1^2\}.$$

In the first case, we evaluate both sides on an element of cycle type $(n-2,1^2)$ if n is odd, and on an element of cycle type $(n-1,1)$ if n is even. Then the left hand side gives zero whereas the right hand side is $\pm x$, giving a contradiction.

In the second case, we evaluate both sides on an element of cycle type (n) if n is odd, and on an element of cycle type $(n-3,1^3)$ if n is even. This gives zero on the left hand side and $\pm x$ on the right hand side. $\square$

Acknowledgement. We are grateful to the Mathematical Sciences Research Institute, University of Oregon and the Isaac Newton Institute where the research for this paper was done. The authors were also supported by the NSF (grants # DMS 9022140 and # DMS-9600124) and the Deutsche Forschungsgemeinschaft (grant Be 923/6-1).

Note added in proof.

After this paper had been accepted we learned of the paper of I. Zisser "Irreducible products of characters in A_n", Israel J. Math., **84** (1993), 147-151. The main result of the Zisser's paper is that A_n has a pair of non-linear characters, whose product is irreducible, if and only if n is a perfect square. Even though Zisser does not classify all such pairs (which is done in our paper), he does prove that one of the characters must correspond to the square diagram. Moreover, he also proves that the product of two non-linear S_n-characters is never irreducible, using his previous results on decomposing the squares of irreducible characters. However, we believe that the short direct proof of the more general fact that such a product is never homogeneous given in Section 3 of our paper (Theorem 3.4) might be useful. Generally, our approach allows us to consider more general questions concerning few homogeneous components rather than few irreducible components.

References

[1] M. Clausen and H. Meier, *Extreme irreduzible Konstituenten in Tensordarstellungen symmetrischer Gruppen*, Bayreuther Math. Schriften, **45** (1993), 1-17.

[2] Y. Dvir, *On the Kronecker product of S_n characters*, J. Algebra, **154** (1993), 125-140.

[3] A.M. Garsia and J. Remmel, *Shuffles of permutations and the Kronecker product*, Graphs and Combin., **1** (1985), 217-263.

[4] G. James and A. Kerber, *The representation theory of the symmetric group*, Addison-Wesley, London, 1981.

[5] D.E. Littlewood, *The Kronecker product of symmetric group representations*, J. London Math. Soc., **31** (1956), 89-93.

[6] I.G. Macdonald, *Symmetric functions and Hall polynomials*, 2nd edition, Oxford Univ. Press, Oxford, 1995.

[7] F.D. Murnaghan, *The analysis of the Kronecker product of irreducible representations of the symmetric group*, Amer. J. Math., **60** (1938), 761-784.

[8] J. Remmel, *A formula for the Kronecker products of Schur functions of hook shapes*, J. Algebra, **120** (1989), 100-118.

[9] J. Remmel and T. Whitehead, *On the Kronecker product of Schur functions of two row shapes*, Bull. Belg. Math. Soc., **1** (1994), 649-683.

[10] J. Saxl, *The complex characters of the symmetric groups that remain irreducible in subgroups*, J. Algebra, **111** (1987), 210-219.

[11] E. Vallejo, *On the Kronecker product of irreducible characters of the symmetric group*, preprint, 1997.

[12] I. Zisser, *The character covering numbers of the alternating groups*, J. Algebra, **153** (1992), 357-372.

Received December 15, 1997 and revised July 1, 1998.

Otto-von-Guericke-Universität Magdeburg
D-39016 Magdeburg
Germany
E-mail address: bessen@mathematik.uni-magdeburg.de

University of Oregon
Eugene, Oregon, 97403
E-mail address: klesh@darkwing.uoregon.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 190, No. 2, 1999

L^2 ESTIMATES ON CHORD-ARC CURVES

Joaquim Bruna and María J. González

We characterize those domains in the plane whose boundary is a chord arc curve in terms of some L^2 integrals, which are mainly a version of Green's theorem. As a consequence of this we obtain a "converse" to a theorem due to Laurentiev that states that for such domains harmonic measure and arc length are A_∞ equivalent.

Let Γ be a locally rectifiable Jordan curve in the plane that passes through ∞, and let Ω_+, Ω_- be the two domains bounded by Γ.

Given a function f defined on Γ, its Cauchy integral

$$Cf(z) = \int_\Gamma \frac{f(\zeta)}{\zeta - z}\, d\zeta, \quad z \notin \Gamma$$

defines an analytic function off Γ.

If C_+f, C_-f denote the restrictions of Cf to Ω_+ and Ω_-, and if f_+, f_- denote their boundary values, then

$$f_\pm(z) = \pm\frac{1}{2}f(z) + \frac{1}{2\pi i}\ \text{P.V.} \int_\Gamma \frac{f(\zeta)}{\zeta - z}\, d\zeta, \quad z \in \Gamma.$$

G. David has shown in [**D**] that the Cauchy integral is bounded in $L^2(\Gamma)$ if and only if Γ is regular, that is, there exists a constant C such that for all $z_0 \in \mathbb{C}$ and all $R > 0$, the arclength of $B(z_0, R) \cap \Gamma$ is at most CR, where $B(z_0, R)$ denotes the ball centered at z_0 and radius R.

Several proofs have been given of the boundedness of the Cauchy integral under stronger hypothesis on Γ. We shall concentrate on the first proof presented in [**C-J-S**] which is based on complex variables methods. They show the result for Lipschitz graphs, i.e.,

$$\Gamma = \{x + iA(x) : x \in \mathbb{R}\} \text{ with } A' \in L^\infty.$$

By following their argument very closely one can notice that the theorem is a consequence of the fact that for any F holomorphic in $\Omega_\pm$ that decays to zero at ∞, the following two integrals are equivalent:

$$\iint_{\Omega_\pm} |F'(z)|^2 \delta(z)\, dx\, dy \cong \int_\Gamma |F|^2\, ds$$

where $\delta(z) = \text{dist}(z, \Gamma)$.

It is a well known result, [**J-K**], that such an equivalence holds if Γ is a chord-arc curve (the length of the arc is comparable to the chord). The main purpose of this paper is to show that the chord-arc condition is also necessary for the equivalence to hold.

To avoid problems at ∞, we will assume that the curves Γ and the functions F are analytic at ∞. In particular $F(z) = O\left(\frac{1}{z}\right)$ at ∞. Note that if Γ is the real line and Ω is the upper half plane, the equivalence of the integrals is just Green's theorem applied to the functions $u(z) = |F(z)|^2$ and $v(z) = y$ in the domain $\Omega_R = \{z \in \mathbb{R}^+_2;\ |z| \le R\}$ [**G**, p. 236]. Since $F(z) = O\left(\frac{1}{z}\right)$ the terms involving the line integral on $\{z = \mathrm{R}e^{i\theta};\ 0 < \theta < \pi\}$ will tend to 0 as R tends to ∞.

Before stating the results we need to recall a few definitions:

A function $\varphi \in L^1_{\mathrm{loc}}(\mathbb{R})$ lies in $\mathrm{BMO}(\mathbb{R})$ if

$$\sup_I \frac{1}{|I|} \int_I |\varphi - \varphi_I|\, dt = \|\varphi\|_* < \infty$$

where $I \subset \mathbb{R}$ is any bounded interval and $\varphi_I = \frac{1}{|I|}\int_I \varphi\, dt$. The space $\mathrm{BMOA}(\mathbb{R})$ denotes the space of holomorphic functions in the upper half plane that are Poisson integrals of functions in $\mathrm{BMO}(\mathbb{R})$.

A positive measure μ defined on the upper half plane is called a *Carleson measure* if there is a constant $N(\mu)$ such that

$$\mu(Q) \le N(\mu) l(Q)$$

for all cubes

$$Q = \{x_0 < x < x_0 + l(Q),\ 0 < y < l(Q)\}.$$

There is a close connection between BMO functions and Carleson measures: A function $\varphi \in \mathrm{BMO}(\mathbb{R})$ if and only if $|\nabla\varphi(z)|^2 y\, dx\, dy$ is a Carleson measure where $\varphi(z)$ denotes the harmonic extension of φ. See [**G**, p. 240].

We are ready now to state the results:

Theorem 1. *Let Γ be a locally rectifiable Jordan curve analytic at ∞ and let Ω be a domain bounded by Γ.*

Denote by Φ the conformal mapping from $\mathbb{R}^+_2$ onto Ω with $\Phi(\infty) = \infty$. Then $\log \Phi' \in \mathrm{BMOA}(\mathbb{R})$ if and only if there is a constant c, depending only on the BMO constant, such that

$$\iint_\Omega |F'|^2 \delta(z)\, dx\, dy \le c \int_\Gamma |F|^2\, ds \tag{1}$$

for any F holomorphic in Ω with $F(z) = O\left(\frac{1}{z}\right)$ at ∞.

Note that the boundary values of Φ' are defined a.e. on $\mathbb{R}$ because of our assumptions on Γ.

Theorem 2. *Let Γ be a locally rectifiable Jordan curve bounding the domains Ω_+, Ω_-. Suppose there exists a constant c such that*

$$\int_\Gamma |F|^2\,ds \le c\iint_{\Omega_+} |F'|^2\delta(z)\,dx\,dy \tag{2}$$

and

$$\int_\Gamma |G|^2\,ds \le c\iint_{\Omega_-} |G'|^2\delta(z)\,dx\,dy$$

for any holomorphic function $F(G)$ on $\Omega_+(\Omega_-)$ vanishing at ∞. Then Γ is a chord-arc curve.

As we mentioned before its converse is also true. Also note that if (2) holds then (1) holds, that is because if Ω is bounded by a chord-arc curve, $\log\Phi' \in \mathrm{BMOA}(\mathbb{R})$.

It will become clear from the proof of the theorem that (2) can be replaced by

$$\int_\Gamma |\varphi|^2\,ds \cong \iint_{\mathbb{C}} |(C\varphi)'|^2\delta(z)\,dx\,dy$$

where $\varphi = \chi_I$ for any arc $I \subset \Gamma$.

It is also interesting to see what happens if we consider functions of the form $F(z) = \frac{1}{|z-w|}$, $w \notin \Gamma$. Then the result is the following:

Theorem 3. *Let Γ be a locally rectifiable curve, then Γ is regular if and only if there exist constants c_1, c_2 such that*

$$c_1\frac{1}{\delta(w)} \le \int_\Gamma \frac{|dz|}{|z-w|^2} \le c_2\frac{1}{\delta(w)}, \quad \textit{for all } w \notin \Gamma. \tag{3}$$

The proofs of these theorems are contained in Section 1. Further remarks and corollaries will be given in Section 2. Finally we would like to thank, M. Melnikov for suggesting some questions and for many helpful conversations, and the referee for his comments which improved the presentation of this paper.

1. Proofs of the Theorems.

Proof of Theorem 1. First note that by changing variables and by using Koebe's distortion theorem, (1) is equivalent to

$$\iint_{\mathbb{R}^+_2} |f'|^2|\Phi'|y\,dx\,dy \le c\int_{\mathbb{R}} |f|^2|\Phi'|\,dx \tag{4}$$

where f is a holomorphic function on $\mathbb{R}^+_2$ with $f(z) = O\left(\frac{1}{z}\right)$ at ∞.

Consider now $g = f(\Phi')^{1/2}$. Then applying Green's Theorem as in the remark of the introduction, we get

$$\int_{\mathbb{R}} |g|^2\,dx = 4\iint_{\mathbb{R}^+_2} |g'|^2 y\,dx\,dy.$$

Since $f'(\Phi')^{1/2} = g' - \frac{1}{2} g \frac{\Phi''}{\Phi'}$

$$\begin{aligned}\iint_{\mathbb{R}_2^+} |f'|^2 |\Phi'| y\, dx\, dy &= \iint_{\mathbb{R}_2^+} \left| g' - \frac{1}{2} g \frac{\Phi''}{\Phi'} \right|^2 y\, dx\, dy \\ &\leq 2 \iint_{\mathbb{R}_2^+} \left(|g'|^2 + \frac{1}{4} |g|^2 \left| \frac{\Phi''}{\Phi'} \right|^2 \right) y\, dx\, dy \\ &\leq \frac{1}{2} \left(\int_{\mathbb{R}} |g|^2\, dx + \iint_{\mathbb{R}_2^+} |g|^2 \frac{|\Phi''|^2}{|\Phi'|^2} y\, dx\, dy \right).\end{aligned}$$

By the remark at the end of the introduction if $\log \Phi' \in \mathrm{BMOA}(\mathbb{R})$, then $\frac{|\Phi''|^2}{|\Phi'|} y$ is a Carleson measure and (4) holds.

On the other hand, if (4) holds then

$$\iint_{\mathbb{R}_2^+} |g|^2 \left| \frac{\Phi''}{\Phi'} \right|^2 y\, dx\, dy = 4 \iint_{\mathbb{R}_2^+} |g' - (f')^2 \Phi'|^2 y\, dx\, dy \leq 4c \int_{\mathbb{R}} |g|^2\, dx$$

which is equivalent to $\frac{|\Phi''|^2}{|\Phi'|^2} y$ being a Carleson measure ([**G**, p. 33]). □

Proof of Theorem 2. Let I be an arc on Γ with length $l(I)$ and endpoints α, β.

Set $f = \chi_I$ and consider the functions $C_\pm f(z)$ defined in the introduction. Since $f = f_+ - f_-$, (2) implies

$$\begin{aligned}\int_\Gamma |f|^2\, ds &\leq c \left(\iint_{\Omega_+} |(C_+ f)'|^2 \delta(z)\, dx\, dy + \iint_{\Omega_-} |(C_- f)'|^2 \delta(z)\, dx\, dy \right) \\ &= c \iint_{\mathbb{C} \setminus \Gamma} |(Cf)'|^2 \delta(z)\, dx\, dy\end{aligned}$$

that is

$$l(I) \leq C \iint_{\mathbb{C} \setminus \Gamma} \delta(z) \left| \int_I \frac{d\zeta}{(\zeta - z)^2} \right| dx\, dy.$$

Let $\zeta(s)$, $s \in [a, b]$ be a parameterization of I by arclength, then

$$\int_I \frac{d\zeta}{(\zeta - z)^2} = \int_a^b \frac{\zeta'(s)}{(\zeta(s) - z)^2}\, ds = \frac{1}{\zeta(a) - z} - \frac{1}{\zeta(b) - z} = \frac{\beta - \alpha}{(\alpha - z)(\beta - z)}.$$

Therefore

$$l(I) \leq |\beta - \alpha|^2 \iint_{\mathbb{C} \setminus \Gamma} \frac{\delta(z)}{|z - \alpha|^2 |z - \beta|^2}\, dx\, dy.$$

It only remains to estimate the last integral. To do so we split it into three integrals. Let B_1 be the ball centered at α with radius $\frac{|\beta - \alpha|}{2}$ and let B_2 be the corresponding one centered at β. Then

$$\iint_{B_1} \frac{\delta(z)}{|z-\alpha|^2|z-\beta|^2}\,dx\,dy \le \frac{4}{|\beta-\alpha|^2}\iint_{B_1} \frac{dx\,dy}{|z-\alpha|} = \frac{c}{|\beta-\alpha|}.$$

By a similar argument one can show that the same estimate holds on B_2 and outside $B_1 \cup B_2$. Therefore

$$l(I) \le c|\beta-\alpha|.$$

□

Proof of Theorem 3. Suppose first that Γ is a regular curve. Fix a point $w \notin \Gamma$, choose $z_0 \in \Gamma$ such that $\delta(w) = |w-z_0|$ and consider the ball B centered at z_0 with radius $2\delta(w)$. So:

$$\int_\Gamma \frac{|dz|}{|z-w|^2} = \int_{\Gamma\cap B} \frac{|dz|}{|z-w|^2} + \int_{\Gamma\setminus B} \frac{|dz|}{|z-w|^2}.$$

If $z \in \Gamma \cap B$, then $|z-w| \cong \delta(w)$. Also, since Γ is regular $l(\Gamma\cap B) \cong \delta(w)$. Therefore, trivially

$$\int_{\Gamma\cap B} \frac{|dz|}{|z-w|^2} \cong \frac{1}{\delta(w)}.$$

On the other hand

$$\int_{\Gamma\setminus B} \frac{|dz|}{|z-w|^2} = \sum_{k=1}^{\infty} \int_{A_k} \frac{|dz|}{|z-w|^2}$$

where $A_k = \{z \in \Gamma : 2^k\delta(w) \le |z-z_0| \le 2^{k+1}\delta(w)\}$.

If $z \in A_k$, $|z-w| \cong 2^k\delta(w)$. Since $l(A_k) \cong 2^k\delta(w)$ we get

$$\int_{\Gamma\setminus B} \frac{|dz|}{|z-w|^2} \cong \frac{1}{\delta(w)}$$

which proves the first part of the theorem.

Suppose now that (3) holds. Choose any $r > 0$ and any point $z_0 \in \mathbb{C}$ and consider the ball B centered at z_0 of radius r. Let A be the annulus $A = \{2r < |z-z_0| < 3r\}$ and let $w \in A$ be a point with the property that

$$\delta(w) = \sup_{z\in A} \delta(z).$$

We claim that there is a constant c depending only on c_1, c_2 such that $\delta(w) \ge cr$. Assuming the claim let us finish the proof of the theorem:

$$\frac{l(\Gamma\cap B)}{r^2} \le \int_{\Gamma\cap B} \frac{|dz|}{|z-w|^2} \le \int_\Gamma \frac{|dz|}{|z-w|^2} \cong \frac{1}{\delta(w)} \le \frac{c}{r}.$$

Therefore $l(\Gamma\cap B)\le cr$, i.e. Γ is regular. To prove the claim consider a grid on A of size $\delta(w)$. Then, because of the choice of w, any square of the grid contains points of Γ. So, letting $N\cong r/\delta(w)$, we have

$$\frac{c_2}{\delta(w)}\ge\int_\Gamma\frac{|dz|}{|z-w|^2}\ge\sum_{k=1}^N\int_{\substack{\{z\in\Gamma:\\ k\delta(w)<|z-w|<(k+1)\delta(w)\}}}\frac{|dz|}{|z-w|^2}$$

$$\ge c\sum_{k=1}^N\frac{1}{k\delta(w)}\cong\frac{c}{\delta(w)}\log r/\delta(w).$$

Therefore $r/\delta(w)\le c$ which proves the claim. $\square$

Note that the same result holds if we replace $\frac{1}{|z-w|^2}$ by $\frac{1}{|z-w|^\alpha}$ for any $\alpha>1$. Then instead of (3) we get

$$\int_\Gamma\frac{|dz|}{|z-w|^\alpha}\cong(\delta(w))^{-\alpha+1}.$$

The proof is the same.

2. Further remarks.

Let $w(x)>0$ be locally integrable on $\mathbb{R}$.

Set $w(E)=\int_E w(x)\,dx$, and let $|E|$ denote the Lebesgue measure of E. We say that w is an A_∞ weight if for every $\varepsilon>0$, there is a $\delta>0$ such that if I is any interval and $E\subseteq I$, then

$$\frac{|E|}{|I|}<\delta\Rightarrow\frac{w(E)}{w(I)}<\varepsilon.$$

If ω is an A_∞ weight, then $\log w\in\mathrm{BMO}$. For a proof of this fact and some related ones see [**S**].

As before, given an unbounded simply connected domain Ω other than the plane itself, Φ will denote the conformal mapping from $\mathbb{R}^+_2$ onto Ω fixing ∞.

There is a theorem due to Laurentiev which states that if Ω is a domain bounded by a chord-arc curve, then arc-length and harmonic measure on $\partial\Omega$ are A_∞-equivalent. That is, $|\Phi'|$ is an A_∞-weight.

A version of a converse is given in [**J-K**]. Before stating it we need some more definitions.

A Jordan curve Γ that passes through ∞ is called a quasi-circle if it satisfies the three-point condition, that is there is a constant c such that for any three points $z_1,z_2\in\Gamma$ and z_3 on the arc joining z_1 and z_2, $|z_1-z_3|\le c|z_1-z_2|$. Obviously a chord-arc curve is a quasicircle.

A domain is called a Smirnov domain if $\log|\Phi'(z)|$ is represented by its Poisson integral. In particular domains bounded by regular curves are Smirnov [**Z**].

Theorem 4 ([**J-K**])**.** *Suppose Ω is a Smirnov domain, $\partial\Omega$ is a quasicircle and harmonic measure is A_∞-equivalent to arc length. Then $\partial\Omega$ is a chord-arc curve.*

Note that its converse is also true.

Using Theorem 2 we give another "converse" to Laurentiev's theorem which is very similar to [**J-K**].

Corollary 1. *Suppose that the two sides of a curve Γ are Smirnov domains and that on each domain harmonic measure is A_∞-equivalent to arc-length. Then Γ is a chord-arc curve.*

As before note that its converse is also true.

Proof. Let Ω be one of the sides of Γ and let $\Phi : \mathbb{R}^+_2 \to \Omega$ be its conformal mapping, $\Phi(\infty) = \infty$. We are assuming that $|\Phi'|$ is an A_∞ weight, therefore [**G-W**],

$$\int_{\mathbb{R}} |F(x)|^2 |\Phi'(x)|\, dx \le c \int_{\mathbb{R}} \left(\iint_{\Gamma_x} |F'(z)|^2\, dA(z) \right) |\Phi'(x)|\, dx$$

where $\Gamma(x)$ is a cone centered at x:

$$\Gamma_x = \{(s, y) : |x - s| < ay\} \text{ for some } a \text{ fixed}$$

and F is a holomorphic function on $\mathbb{R}^+_2$ vanishing at ∞ as before. The constant c depends only on the opening of the cone and the A_∞-constant.

By Fubini's theorem the integral on right-hand side is equivalent to

$$\iint_{\mathbb{R}^+_2} |F'(z)|^2 \sigma(I_z)\, dx\, dy$$

where $\sigma(I_z) = \int_{I_z} |\Phi'(t)|\, dt$ and I_z is the interval on $\mathbb{R}$ centered at x and length $2ay$.

Since $\log\Phi' \in \mathrm{BMO}$,

$$\left| P_y * \log|\Phi'| - \frac{1}{|I_z|} \int_{I_z} \log|\Phi'| \right| \le c$$

with c depending on the BMO-constant of $\log|\Phi'|$ [**G**].

On the other hand, Ω being a Smirnov domain implies that $P_y * \log|\Phi'| = \log|\Phi'(z)|$ and $|\Phi'| \in A_\infty$ is equivalent to saying that

$$\exp\left(\frac{1}{|I|} \int_I \log|\Phi'|\, dt \right) \cong \frac{1}{|I|} \int_I |\Phi'|\, dt$$

for any interval $I \subset \mathbb{R}$.

So,

$$|\Phi'(z)| \cong \frac{1}{|I_z|} \int_{I_z} |\Phi'(t)|\, dt = \sigma(I_z)/2ay.$$

Hence, there is a constant c such that

$$\int_{\mathbb{R}} |F(x)|^2 |\Phi'(x)|\, dx \leq c \iint_{\mathbb{R}_2^+} |F'(z)|^2 |\Phi'(z)| y\, dx\, dy.$$

Since this inequality holds on both sides of Γ, by changing variables and using Koebe's distortion theorem we get the hypothesis of Theorem 2. Therefore Γ is chord-arc. $\square$

Next corollary involves quasiconformal mappings. The result we need to use is the quasiconformal analogue of Koebe's distortion theorem [**A-G**]: Suppose that Ω and Ω' are domains in $\mathbb{R}^2$ and that $\rho : \Omega \to \Omega'$ is K-quasiconformal with Jacobian J_ρ. For each $z \in \Omega$, define

$$a_\rho(z) = \frac{1}{|B_z|} \iint_{B_z} (J_\rho(\zeta))^{1/2}\, d\zeta\, d\bar{\zeta}$$

where B_z is the disk of center z and radius $\delta(z)$. Then

$$\delta(\rho(z)) \cong a_\rho(z)\delta(z).$$

Using this fact and a change of variables in (2) we get the following:

Corollary 2. *Let Γ be a locally rectifiable quasicircle analytic at ∞ bounding the domain Ω, and let ρ be a quasiconformal mapping that sends $\mathbb{R}_2^+$ onto Ω. Then Γ is a chord-arc curve if and only if*

$$\int_{\mathbb{R}} |F(x)|^2 J_\rho^{1/2}(x)\, dx \cong \iint_{\mathbb{C}} JF(z) a_\rho(z) y\, dx\, dy$$

for any quasiregular mapping F satisfying $\bar{\partial} F = \mu \partial F$ where μ is the dilatation of ρ and $F(z) = O\left(\frac{1}{z}\right)$ at ∞.

References

[A-G] K. Astala and F.W. Gehring, *Quasiconformal analogues of theorems of Koebe and Hardy-Littlewood*, Michigan Math. J., **32** (1985), 99-107.

[C-J-S] R.R. Coifman, P.W. Jones and S. Semmes, *Two elementary proofs of the L^2 boundedness of Cauchy integrals on Lipschitz curves*, J. Amer. Math. Soc., **2** (1989), 553-564.

[D] G. David, *Opérateurs intégraux singuliers sur certaines courbes du plan complex*, Ann. Sci. École Norm. Sup., **17** (1984), 157-189.

[G] J. B. Garnett, *Bounded Analytic Functions*, Academic Press, 1981.

[G-W] R. Gundy and R. Wheeden, *Weighted integral inequalities for the nontangential maximal functions, Lusin area integral and Welsh-Paley series*, Studia Math., **49** (1974), 107-124.

[J-K] D. Jerison and C. Kenig, *Hardy spaces, A_∞, and singular integrals in chord-arc domains*, Math. Scand., **50** (1982), 221-247.

[S] E. M. Stein, *Harmonic Analysis*, Princeton University Press, Princeton, 1993.

[Z] M. Zinsmeister, *Représentation conforme et courbes presque Lipschitziennes*, Ann. Inst. Fourier (Grenoble), **34** (1984), 29-44.

Received November 1, 1996 and revised August 5, 1998. This research was supported by DGICYT Grant PB95-0956-C02-02 and by CIRIT grant 1996SGR00026.

Universitat Autònoma de Barcelona
08193 Bellaterra Barcelona
Spain
E-mail address: bruna@mat.uab.es

Universitat Autònoma de Barcelona
08193 Bellaterra Barcelona
Spain
E-mail address: majose.gonzalez@uca.es

PACIFIC JOURNAL OF MATHEMATICS
Vol. 190, No. 2, 1999

ENTROPY OF CUNTZ'S CANONICAL ENDOMORPHISM

Marie Choda

Let $\{S_i\}_{i=1}^n$ be generators of the Cuntz algebra $\mathcal{O}_n$ and let Φ be the *-endomorphism of $\mathcal{O}_n$ defined by $\Phi(x) = \sum_{i=1}^n S_i x S_i^*$. Then both of Connes–Narnhofer–Thirring's entropy $h_\phi(\Phi)$ and Voiculescu's topological entropy $ht(\Phi)$ are $\log n$, where ϕ is the unique $\log n$-KMS state of $\mathcal{O}_n$. Also Longo's canonical endomorphism for $N \subset M$ have the same entropy $\log n$, where the inclusion $N \subset M$ comes from $\mathcal{O}_n$.

1. Introduction.

Connes-Stϕrmer entropy $H(\cdot)$ extended the entropy invariant of Kolmogorov-Sinai to trace preserving automorphisms of finite von Neumann algebras ([**CS**]). Replacing a finite trace to an invariant state ϕ, Connes-Narnhofer-Thirring entropy $h_\phi(\cdot)$ is defined for automorphisms of C^*-algebras as a generalization of $H(\cdot)$ ([**CNT**]). These entropies depend on an invariant state under a given automorphism.

The first typical interesting example to compute the entropy is the Bernoulli shift β_n on the infinite product space of n-point sets.

In the context of operator algebras (von Neumann algebras or C^*-algebras), the non-commutative Bernoulli shift α_n takes the place of the the Bernoulli shift β_n. It is the shift automorphism on the infinite tensor product $A = \bigotimes_{i=-\infty}^{\infty} A_i$ (where A_i is the $n \times n$-matrix algebra) and $H(\alpha_n) = \log n = h_\tau(\alpha_n)$ ([**CS**], [**CNT**]), where τ is the unique tracial state of A.

Ler γ be an aperiodic automorphism of an algebra B. Then there exists an implimenting unitary operator u for γ in the crossed product $M = B \rtimes_\gamma \mathbb{Z}$. The inner automorphism $Ad_u, (Ad_u(x) = uxu^*)$ of M is an extension of γ to M. In general, the entropy of γ is less than the entropy of Ad_u. Stϕrmer [S] asked if the equality between the entropies of γ and Ad_u holds.

Voiculescu [**V**] defined topological entropy $ht(\cdot)$ for automorphisms of nuclear C^*-algebras (cf. [**Hu**], [**T**]), which does not depend on any state but is based on approximations. As an application, he showed that his topological entropy satisfies the equality for the Bernoulli shift β_n, so that Connes-Narnhofer-Thirring entropy does too.

In this paper, we show the equality for both of the automorphism α_n and the unital *-endomorphism of the type of the non-commutative Bernoulli shift.

In §3, we denote only the fact that

$$H(Ad_u) = h_\tau(Ad_u) = ht(Ad_u) = \log n,$$

where τ is the unique tracial state of the reduced crossed product $A \rtimes_{\alpha_n} \mathbb{Z}$. These are proved by similar method as in §4 and §5.

The definition of Connes-Stϕrmer entropy is available to trace preserving *-endomorphisms on finite von Neumann algebras. Similarly, we can apply the definition of Connes-Narnhofer-Thirring entropy to unital and state preserving *-endomorphisms of C^*-algebras, and also Voiculescu's topological entropy to unital *-endomorphisms of nuclear C^*-algebras. We apply here, in particular, to the unital *-endomorphism which is an extension of the *-endomorphism coming from the non-commutative Bernoulli shift α_n as follows.

If we restrict our algebra A to the *half side* infinite C^*-tensor product (or von Neumann tensor product) $B = \bigotimes_{i=0}^{\infty} A_i$ of matrix algebras, then the restriction of α_n to B defines a unit preserving *-endomorphism σ_n of B, which is canonical in the sense of [**Ch2**, **Ch3**]. Then we have the extension algebra $\langle B, \sigma_n \rangle$ of B by σ_n ([**Ch2**, **Ch3**]). In the case of C^* algebras, $\langle B, \sigma_n \rangle$ is the crossed product $B \rtimes_\rho \mathbb{N}$ of B by the corner endomorphism ρ in [**R**, **I2**], which is defined by σ_n using the canonical property of σ_n. Further, the canonical extension $\hat{\sigma}_n$ (in the sense of [**Ch2**, **Ch3**]) of σ_n to $\langle B, \sigma_n \rangle$ is obtained. The *-endomorphism $\hat{\sigma}_n$ of $\langle B, \sigma_n \rangle$ is defined by a modification of the automorphism Ad_u of $A \rtimes_{\alpha_n} \mathbb{Z}$ and has the property like the canonical extension in the sense of [**I1**, **HS**]. In the case of C^*-algebras, the extension algebra $\langle B, \sigma_n \rangle$ is the Cuntz algebra $\mathcal{O}_n$ and $\hat{\sigma}_n$ is nothing but Cuntz's canonical inner endomorphism Φ of $\mathcal{O}_n$ defined by $\Phi(x) = \sum_{i=1}^{n} S_i x S_i^*$, $(x \in \mathcal{O}_n)$ for generators $\{S_1, \ldots, S_n\}$ of $\mathcal{O}_n$. In the case of von Neumann algebras, $\langle B, \sigma_n \rangle$ is the unique injective type $\mathrm{III}_{1/n}$ factor and $\hat{\sigma}_n$ is Longo's canonical endomorphism for the subfactor of $\langle B, \sigma_n \rangle$, which appears naturally in the construction of the extension algebra $\langle B, \sigma_n \rangle$ by the canonical *-endomorphism σ_n ([**Ch3**]).

In §4, we show that

$$ht(\Phi) = \log n = ht(\sigma_n).$$

Applying to Connes-Narnhofer-Thirring's entropy $h_\phi(\cdot)$ relative to the unique $\log n$-KMS state ϕ of $\mathcal{O}_n$, we have

$$h_\phi(\Phi) = \log n = h_\phi(\sigma_n).$$

This relation implies the same relation for Longo's canonical endomorphism. Thus the canonical extension of the non-commutative Bernoulli shift has the same entropy with the original one in the case of *-endomorphisms too.

The author thanks F. Hiai for his interest in this work and encouragement during the preparation.

2. Preliminaries.

2.1. Let H_0 be a Hilbert space of dimension $n < \infty$. Put $H_i = H_0,\ i \in \mathbb{Z}$. For two integers i and j with $i < j$, we put

$$H_{[i,j]} = H_i \otimes H_{i+1} \otimes \cdots \otimes H_j.$$

Let $\{\delta(i) : i = 1, ..., n\}$ be an orthonormal basis of H_0. The emmbedding $H_{[i,j]} \hookrightarrow H_{[i-1,j+1]}$ is given by $\xi \in H_{[i,j]} \to \delta(1) \otimes \xi \otimes \delta(1) \in H_{[i-1,j+1]}$. We denote by $\mathcal{H}_i$ the inductive limit of $\{H_{[i,i+j]} : j = 0, 1, ...\}$ and by $\mathcal{H}$ the inductive limit of the incleasing sequence $\{\mathcal{H}_i : i = 0, -1, ...\}$.

Given $k, l \in \mathbb{Z}\ \ k < l$, let

$$W^n_{[k,l]} = \{\mu = (\mu_k, \ldots, \mu_l) : \mu_i \in \{1, \ldots, n\},\ (k \le i \le l)\}.$$

Let $\mu \in W^n_{[k,l]}$ and $\nu \in W^n_{[l+1,m]}$. We put

$$\mu \cdot \nu = (\mu_k, \ldots, \mu_l, \nu_{l+1}, \ldots, \nu_m).$$

Further, let

$$W^n_0 = \{0\}, \quad W^n_{[0,\infty]} = \cup_{k=0}^{\infty} W^n_{[0,k]} \quad \text{and} \quad W^n_\infty = \cup_{k=0}^{\infty} W^n_{[-k,k]}.$$

The shift $\alpha : i \in \mathbb{Z} \to i+1$ induces the mapping on W^n_∞, which we denote by the same notation α.

For $\mu \in W^n_{[k,l]}$, we put

$$\delta(\mu) = \delta(\mu_k) \otimes \cdots \otimes \delta(\mu_l) \in H_{[k,l]}.$$

Then $\{\delta(\mu) : \mu \in W^n_{[k,l]}\}$ is an orthonormal basis in $H_{[k,l]}$.

Let $A_0 = B(H_0)$ and $\{e(i,j) : i, j = 1, ..., n\}$ be the matrix unit of A_0 with respect to the orthonormal basis $\{\delta(i) : i = 1, ..., n\}$. We denote the trace (1/n)Tr of A_0 by τ_0. Put $A_i = A_0, (i \in \mathbb{Z})$ and $\tau_i = \tau_0$. For two integers $i < j$, let

$$A_{[i,j]} = A_i \otimes A_{i+1} \otimes \cdots \otimes A_j.$$

For $\mu, \nu \in W^n_{[k,l]}$, we put

$$e(\mu, \nu) = e(\mu_k, \nu_k) \otimes \cdots \otimes e(\mu_l, \nu_l) \in A_{[k,l]}.$$

Then $\{e(\mu, \nu) : \mu, \nu \in W^n_{[k,l]}\}$ is a matrix units of $A_{[k,l]}$.

2.2. We apply the entropy of Connes-Narnhofer-Thirring and Voiculescu's topological entropy to both of automorphisms and unital *-endomorphisms on C^*-algebras. To fix notations, we recall the definition of the topological entropy. Let B be a nuclear C^*-algebra with unity. Let $CAP(B)$ be triples (ρ, η, C), where C is a finite dimensional C^*-algebra, and $\rho : B \to C$ and $\eta : C \to B$ are unital completely positive maps. Let Ω be the set of finite subsets of B. For an $\omega \in \Omega$, put

$$rcp(\omega; \delta) = \inf\{\text{rank } C : (\rho, \eta, C) \in CAP(B), \|\eta \cdot \rho(a) - a\| < \delta, a \in B\},$$

where rank C means the dimension of a maximal abelian self-adjoint subalgebra of C. For a unital *-endomorphism β of B, put

$$ht(\beta,\omega\ ;\delta) = \overline{\lim}_{\mathrm{N}\to\infty}\frac{1}{\mathrm{N}}\log rcp\left(\omega\cup\beta(\omega)\cup\cdots\cup\beta^{\mathrm{N}-1}(\omega);\delta\right)$$

and

$$ht(\beta,\omega) = \sup_{\delta>0} ht(\beta,\omega;\delta).$$

Then the topological entropy $ht(\beta)$ of β is defined by

$$ht(\beta) = \sup_{\omega\in\Omega} ht(\beta,\omega).$$

Assume that there exists an increasing sequence $(\omega_j)_{j\in\mathbb{N}}$ of finite subsets of B such that the linear span of $\cup_{j\in\mathbb{N}}\ \omega_j$ is dense in B. Even in the case of *-endomorphisms which are not automorphisms, by the obvious analogoues of [**V**, Proposition 4.3], $ht(\cdot)$ is obtained as the following form which we use later:

$$ht(\beta) = \sup_{j\in\mathbb{N}} ht(\beta,\omega_j).$$

Let ϕ be a state of B with $\phi\cdot\beta=\phi$. The essential relation between $ht(\beta)$ and Connes-Narnhofer-Thirring entropy $h_\phi(\beta)$ is by [**V**, Proposition 4.6]

$$h_\phi(\beta)\le ht(\beta).$$

3. Entropy of Ad_u for non-commutative Bernoulli shift.

In this section, we only state results without proof. We remark that these are proved by similar methods as in §4 and §5.

3.1. Let $n(2\le n<\infty)$ be an integer. Let $A_i,\ \tau_i(i\in\mathbb{Z})$ be as in §2.1 and let A be the infinite C^*-tensor product $A=\bigotimes_{i\in\mathbb{Z}} A_i$. We denote the unique tracial state of A by τ. The non-commutative Bernoulli shift α_n is the automorphism of the C^*-algebra A induced by the shift $\alpha : i(\in\mathbb{Z})\to i+1$. Let u be the implimenting unitary in the reduced C^*-crossed product $A\rtimes_{\alpha_n}\mathbb{Z}$ for α_n. Let E be the conditional expectation of $A\rtimes_{\alpha_n}\mathbb{Z}$ onto A with $E(u^j)=0,(j\ne 0)$. Then $\tau\cdot E$ is a tracial state of $A\rtimes_{\alpha_n}\mathbb{Z}$ which is invariant under Ad_u. We denote by the same notation α_n the extension of α_n to the hyperfinite II_1 factor $\bigotimes_{i\in\mathbb{Z}}(A_i,\tau_i)\rtimes_{\alpha_n}\mathbb{Z}$.

Theorem 3.2. *Under the above notations,*

$$ht(\alpha_n) = ht(Ad_u) = h_{\tau\cdot E}(Ad_u) = h_\tau(\alpha_n) = \log n = H(\alpha_n) = H(Ad_u).$$

4. Entropy of Cuntz's canonical endomorphism.

In this section, we apply the definition of Connes-Narnhofer-Thirring entropy and Voiculescu's topological entropy for unital *-endomorphisms of nuclear C^*-algebras. All facts for automorphisms, which we need here, work for unital *-endomorphisms by the analogues of definitions and proofs in [**CNT**] and [**V**].

Let n $(2 \leq n < \infty)$ be an integer. Given n isometries $\{S_i\}$ on a Hilbert space such that $\sum_{i=1}^n S_i S_i^* = 1$, Cuntz defined the Cuntz algebra $\mathcal{O}_n$ as the C^*-algebra generated by $\{S_i\}_i$ ([**Cu1**]). So called Cuntz's canonical endomorphism Φ of $\mathcal{O}_n$ is defined by

$$\Phi(x) = \sum_{i=1}^n S_i x S_i^*, \quad (x \in \mathcal{O}_n).$$

The $\mathcal{O}_n$ has exactly one $\log n$-KMS state ϕ ([**OP**]). In this section we compute Voiculescu's topological entropy of Φ and Connes-Narnhofer-Thiring's entropy $h_\phi(\Phi)$. Applying to the factor generated by $\pi_\phi(\mathcal{O}_n)$, we get the entropy of Longo's canonical endomorphism.

4.1. To compute the entropy of Φ, we recall some of the representation for the Cuntz algebra $\mathcal{O}_n$ as a crossed product in [**Cu1**], (cf., [**Ch2**, **I2**, **P**, **R**]). Let $A_i, \tau_i, (i \in \mathbb{Z})$ and $e(i,j), (i,j \in \mathbb{N})$ be the same as in §2.1. For a $j \in \mathbb{Z}$, $\mathcal{A}_j$ is given as the infinite tensor product:

$$\mathcal{A}_j = \bigotimes_{i=j}^{\infty} A_i.$$

Define embeddings

$$\mathcal{A}_j \hookrightarrow \mathcal{A}_{j-1} \hookrightarrow \mathcal{A}_{j-2} \hookrightarrow \cdots$$

by $x \in \mathcal{A}_j \to e_{j-1}(1,1) \otimes x \in \mathcal{A}_{j-1}$, where $e_{j-1}(i,l)$ is a copy of $e(i,l)$ in A_{j-1}. The inductive limit of this sequence is denoted by $\mathcal{A}$. Since the embedding $\mathcal{A}_j \hookrightarrow \mathcal{A}_{j-1}$ and the embedding $\mathcal{H}_i \hookrightarrow \mathcal{H}_{i-1}$ in §2.1 are compatible, we can consider $\mathcal{A}$ acting faithfully on $\mathcal{H}$.

The automorphism σ of $\mathcal{A}$ is induced by the shift $\alpha : i(\in \mathbb{Z}) \to i+1$.

Then the crossed product $\mathcal{A} \rtimes_\sigma \mathbb{Z}$ acts faithfully on the Hilbert space

$$K = \sum_{i \in \mathbb{Z}} \bigoplus u^i \mathcal{H},$$

where u is the implimenting unitary in $\mathcal{A} \rtimes_\sigma \mathbb{Z}$ for the automorphism σ of $\mathcal{A}$. Let p be the unit of $\mathcal{A}_0 \subset \mathcal{A} \rtimes_\sigma \mathbb{Z}$ and put

$$w = up.$$

We remark $u^j p = w^j$.

Then Cuntz algebra $\mathcal{O}_n$ is reresented as $p(\mathcal{A} \rtimes_\sigma \mathbb{Z})p$, which is the C^* subalgebra $C^*(\mathcal{A}_0, w)$ of $(\mathcal{A} \rtimes_\sigma \mathbb{Z})$ generated by $\{\mathcal{A}_0,\ w\}$. There exists a conditional expectation E of $C^*(\mathcal{A}_0, w)$ onto $\mathcal{A}_0$ with $E(w^j) = 0$ for all $j = 1, 2, \ldots$. The unique tracial state τ of $\mathcal{A}_0$ is extended to the state ϕ of $C^*(\mathcal{A}_0, w)$ by $\phi = \tau \cdot E$. Then ϕ is the unique $\log n$-KMS state of $C^*(\mathcal{A}_0, w)$ ([**OP**]).

4.2. Since

$$\sigma^j(p)(\mathcal{H}) = \mathcal{H}_j, \quad j \in \mathbb{Z},$$

the algebra $p(\mathcal{A} \rtimes_\sigma \mathbb{Z})p$ is acting faithfully on

$$pK = \sum_{i \in \mathbb{Z}} \bigoplus u^i \mathcal{H}_{-i}.$$

The restriction $\sigma|_{\mathcal{A}_0}$ of σ to $\mathcal{A}_0$ is the one sided non commutative Bernoulli shift. Cuntz's canonical inner endomorphism Φ of $\mathcal{O}_n$ is nothing but the extension of $\sigma|_{\mathcal{A}_0}$ to the Cuntz algebra $C^*(\mathcal{A}_0, w)$ which maps

$$a \to \sigma(a),\ (a \in \mathcal{A}_0), \quad \text{and} \quad w \to vw,$$

where

$$v = \sum_{j=1}^{n} e((j,1),(1,j)) \in A_{[0,1]},$$

([**Cu2**], cf. [**Ch2**]).

4.3. Let $k, m \in \mathbb{N}$. We define

$$K(k,m) = \sum_{l=-k}^{k} \bigoplus u^l H_{[-l,-l+m]}$$

and we denote the orthogonal projection of K onto $K(k,m)$ by $Q(k,m)$. The set $\{u^j\delta(\mu) : -k \le j \le k,\ \mu \in W^n_{[-j,-j+m]}\}$ is an orthonomal basis of $K(k,m)$. We denote by $E((j,\mu),(l,\nu))$ the partial isometry in $B(K(k,m))$ such that

$$E((j,\mu),(l,\nu)) : u^l\delta(\nu) \to u^j\delta(\mu), \quad \left(\mu \in W^n_{[-j,-j+m]},\ \nu \in W^n_{[-l,-l+m]}\right).$$

Then the set

$$\mathcal{E}(k,m) = \left\{E((j,\mu),(l,\nu)) : -k \le j,l \le k,\ \mu \in W^n_{[-j,-j+m]},\ \nu \in W^n_{[-l,-l+m]}\right\}$$

is a matrix units of $B(K(k,m))$.

4.4. Let $k, m \in \mathbb{N}$. We define the completely positive unital linear map

$$\varphi_{k,m} : p(\mathcal{A} \rtimes_\sigma \mathbb{Z})p \to B(K(k,m))$$

by

$$\varphi_{k,m}(x) = Q(k,m)xQ(k,m)|_{K(k,m)}, \quad x \in p(\mathcal{A} \rtimes_\sigma \mathbb{Z})p.$$

We remark that if $e(\mu,\nu)w^j \neq 0$ for ν in $W^n_{[0,b]}$, $(b \geq j)$, then

$$\nu = (1, \dots, 1, \nu_j, \dots, \nu_b) \quad \text{and} \quad \delta(\nu) \in H_{[j,b]}.$$

For two integers a and b with $a < b$, we let

$$\omega_{a,b} = \left\{ e(\mu,\nu)w^j : 0 \leq j \leq a \quad \text{and} \quad \mu, \nu \in W^n_{[0,b]} \right\}.$$

Let $e(\mu,\nu)w^j \in \omega_{a,b}$ for $a, b \in \mathbb{N}$, $(a < b)$ and $e(\mu,\nu)w^j \neq 0$. Since $\sigma^{-l}(p)\delta(\mu) = \delta(\mu)$ for $u^l\delta(\mu) \in K(k,m)$, we have that if $k \geq a$ and $m \geq b$ then

$$\varphi_{k,m}(e(\mu,\nu)w^j) = \sum_{l=-k}^{k-j} E((j+l,\ \alpha^{-(j+l)}(\mu)\cdot\beta_l),\ (l,\ \alpha^{-(j+l)}(\nu)\cdot\gamma_l)),$$

where

$$\beta_l = (1, \dots, 1) \in W^n_{[-(j+l)+b+1, -(j+l)+m]},$$
$$\gamma_l = (1, \dots, 1) \in W^n_{[-l+b+1, -l+m]}.$$

We remark that

$$\delta(\alpha^{-(j+l)}(\nu)) \in H_{[-l,-l+b+1]},$$

so that $E((j+l,\ \alpha^{-(j+l)}(\mu)\cdot\beta_l),\ (l,\ \alpha^{-(j+l)}(\nu)\cdot\gamma_l)) \in \mathcal{E}(k,m)$.

4.5. We define the linear map

$$\psi_{k,m} : B(K(k,m)) \to p(\mathcal{A} \rtimes_\sigma \mathbb{Z})p$$

by

$$\psi_{k,m}(E((j,\mu),(l,\nu))) = \frac{1}{2k+1} pu^j e(\mu,\nu)u^{*l}p,$$

for $E((j,\mu),(l,\nu)) \in \mathcal{E}(k,m)$.

Let $T_j, (j \in \mathbb{Z})$ be the unitary operator on K defined by

$$T_j(u^i\delta(\mu)) = u^{i+j}\delta(\alpha^{-j}(\mu)), \quad i \in \mathbb{Z},\ \mu \in W^n_\infty.$$

Then we have

$$w - \lim_{r \to \omega} \sum_{i=-r}^{r} T_i E((j,\mu),(l,\nu)) T_i^* = u^j e(\mu,\nu)u^{*l}$$

for any $E((j,\mu),(l,\nu)) \in B(K(k,m))$. Here ω is a nontrivial ultrafilter on $\mathbb{N}$. Hence $\psi_{k,m}$ is a unital completely positive map from $B(K(k,m))$ to $p(\mathcal{A} \rtimes_\sigma \mathbb{Z})p$. Since $u^j p = w^j$, we have

$$\psi_{k,m} \cdot \varphi_{k,m}(e(\mu,\nu)w^j) = \frac{2k-j+1}{2k+1} e(\mu,\nu)w^j,$$

for all $e(\mu,\nu)w^j \in \omega_{a,b}$, $a \le k$ and $b \le m$.

Theorem 4.6. *Let Φ be Cuntz's canonical inner endomorphism of $\mathcal{O}_n$. Then*

$$ht(\Phi) = \log n.$$

Proof. Let $e(\mu,\nu)w^j \in \omega_{a,b}$. Then we have, for $a \le k$ and $b \le m$, by §4.5

$$\|\psi_{k,m} \cdot \varphi_{k,m}(e(\mu,\nu)w^j) - e(\mu,\nu)w^j\| = \frac{j}{2k+1}\|e(\mu,\nu)w^j\| \le \frac{a}{2k+1}$$

and we have for an $i \in \mathbb{N}$

$$\begin{aligned}\Phi^i(e(\mu,\nu)w^j) &= \sigma^i(e(\mu,\nu)) \sum_{s=1}^{n} e(\beta_s,\gamma_s)w^j \\ &= \sum_{s=1}^{n} e(\bar\beta_s \cdot \alpha^i(\mu), \gamma_s \cdot \nu_j)w^j.\end{aligned}$$

Here

$$\begin{aligned}\beta_s &= (1,\dots,1,\underset{i-1}{s},1,\dots,1) \in W^n_{[0,j+i-1]}, \\ \gamma_s &= (1,\dots,1,s) \in W^n_{[0,j+i-1]}\end{aligned}$$

and

$$\bar\beta_s = (1,\dots,1,\underset{i-1}{s}) \in W^n_{[0,i-1]}, \quad \nu_j = (\nu_{j+i},\dots,\nu_b) \in W^n_{[j+i,b+i]}.$$

Hence for $k \ge a$ and $m \ge b+i$ we have

$$\|\psi_{k,m} \cdot \varphi_{k,m}(\Phi^i(e(\mu,\nu)w^j)) - \Phi^i(e(\mu,\nu)w^j)\| \le \frac{an}{2k+1}.$$

Therefore, we have for $\mathrm{N} \in \mathbb{N}$

$$\begin{aligned} rcp\left(\bigcup_{i=0}^{\mathrm{N}} \Phi^i\left(\omega_{a,b} \cup (\omega_{a,b})^* : \frac{an}{2k+1}\right)\right) &\le \operatorname{rank} B(K(k,\mathrm{N}+b+1)) \\ &= (2k+1)n^{\mathrm{N}+b+1},\end{aligned}$$

where $(\omega_{a,b})^* = \{x^*; x \in \omega_{a,b}\}$. This implies that for all integers a,b with $a < b$,

$$ht\left(\Phi, \omega_{a,b} \cup (\omega_{a,b})^*; \frac{an}{2k+1}\right) \le \overline{\lim}_{\mathrm{N}\to\infty} \frac{1}{\mathrm{N}} \log\left((2k+1)n^{\mathrm{N}+b+1}\right) = \log n.$$

Increasing k, we have $ht(\Phi, \omega_{a,b} \cup (\omega_{a,b})^*) \leq \log n$, for all $a, b \in \mathbb{N}$ with $a < b$. Put $\omega_a = \omega_{a,2a} \cup (\omega_{a,2a})^*$, for $a \in \mathbb{N}$. Then the set $\{\omega_a : a \in \mathbb{N}\}$ is an increasing sequence of finite subsets of $p(\mathcal{A} \rtimes_\sigma \mathbb{Z})p$ and the linear span of $\bigcup\{\omega_a : a \in \mathbb{N}\}$ is dense in $p(\mathcal{A} \rtimes_\sigma \mathbb{Z})p$. Hence

$$ht(\Phi) = \sup_{a \in \mathbb{N}} ht(\Phi, \omega_a) \leq \log n.$$

On the other hand, the restriction $\Phi|\mathcal{A}_0$ of Φ to $\mathcal{A}_0$ is $\sigma|_{\mathcal{A}_0} = \alpha_n|_{\mathcal{A}_0}$ and $h_\tau(\alpha_n|_{\mathcal{A}_0}) = h_\tau(\alpha_n) = \log n$. Since there exists a conditional expectation of $\mathcal{O}_n$ onto $\mathcal{A}_0$ and $\tau \cdot \alpha_n|_{\mathcal{A}_0} = \tau$,

$$\log n = h_\tau(\alpha_n|_{\mathcal{A}_0}) \leq ht(\Phi|_{\mathcal{A}_0}) \leq ht(\Phi) \leq \log n$$

by the version for unital *-endomorphisms of [**V**, Proposition 4.4]. Therefore, $ht(\Phi) = \log n$. □

Corollary 4.7. *Let ϕ be the unique $\log n$-KMS state of $\mathcal{O}_n$. Then*

$$h_\phi(\Phi) = \log n.$$

Proof. Let τ be the unique tracial state of $\mathcal{A}_0$ and E be the conditional expectation of $p(\mathcal{A} \rtimes_\sigma \mathbb{Z})p$ onto $\mathcal{A}_0$, then $\phi = \tau \cdot E$. Hence $\phi \cdot \Phi = \phi$. This relation implies, by the endomorphism version of [**V**, Proppsition 4.6],

$$\log n = h_\tau(\sigma|\mathcal{A}_0) \leq h_\phi(\Phi) \leq ht(\Phi) = \log n.$$

Therefore $h_\phi(\Phi) = \log n$. □

5. Entropy of Longo's canonical endomorphism.

In this section we apply the result in §4 to Longo's canonical endomorphism. We use the same notations as in §4.

5.1. Let τ_i be the tracial state of A_i, for $i \in \mathbb{N}$ and let

$$\tilde{A} = \bigotimes_{i=0,}^{\infty} (A_i, \tau_i).$$

The $\tilde{A}$ has the canonical trace $\bigotimes_{i=0,}^{\infty} \tau_i$, which we denote by τ. The shift $\sigma|\mathcal{A}_0$ is extended to the *-endomorphism γ of the hyperfinite II_1 factor $\tilde{A}$. The γ is canonical in the sense of [**Ch3**]. Hence we have the extension algebra $\tilde{M} = \langle \tilde{A}, \sigma \rangle$, which is the injective type $\mathrm{III}_{1/n}$ factor generated by $\tilde{A}$ and an isometry W. Then γ is extended to the canonical *-endomorphism Γ of $\tilde{M}$ and

$$\Gamma(a) = \gamma(a), a \in \tilde{A}, \quad \text{and} \quad \Gamma(W) = \pi_\phi(vW).$$

The Γ is Longo's canonical endomorphism for the inclusion $\tilde{N} \subset \tilde{M}$ [**Ch3**, Theorem 6.10]. Here the subfactor $\tilde{N}$ is obtained naturally in the step of constructing $\tilde{M}$. The factor $\tilde{M}$ is the von Neumann algebra generated by $\pi_\phi(\langle \mathcal{A}_0, \sigma|_{\mathcal{A}_0}\rangle)$ and the C^*-algebra $\langle \mathcal{A}_0, \sigma|_{\mathcal{A}_0}\rangle$ is $\mathcal{O}_n$. Hence Γ is the extension of Φ to $\tilde{M}$. Since Φ is ϕ-preserving, as an application of 4.7 Corollary, we have the following by [**CNT**, Theorem VII.2]:

Corollary 5.2. *Let M be the von Neumann algera generated by $\pi_\phi(\mathcal{O}_n)$ and let Γ be the extension of Cuntz's canonical endomorphism Φ of $\mathcal{O}_n$ to M. Then Γ is Longo's canonical endomorphism and*

$$h_\phi(\Gamma) = \log n.$$

References

[Ch1] M. Choda, *Entropy for *-endomorphisms and relative entropy for subalgebras*, J. Operator Theory, **25** (1991), 125-140.

[Ch2] ———, *Canonical *-endomorphisms and simple C^*-algebras*, J. Operator Theory, **33** (1995), 235-248.

[Ch3] ———, *Extension algebras of* II_1 *factors via *-endomorphisms*, Int. J. Math., **5** (1994), 635-655.

[CNT] A. Connes, H. Narnhofer and W. Thirring, *Dynamical entropy of C^* algebras and von Neumann algebras*, Commun. Math. Phys., **112** (1987), 691-719.

[CS] A. Connes and E. Størmer, *Entropy of* II_1 *von Neumann algebras*, Acta Math., **134** (1975), 289-306.

[Cu1] J. Cuntz, *Simple C^*-algebras generated by isometries*, Commun. Math. Phys., **57** (1977), 173-185.

[Cu2] ———, *Regular actions of Hopf algebras on the C^*-algebra generated by a Hilbert space*, in 'Operator algebras, mathematical physics, and low dimensional topology' (Istanbul, 1991), 87-100, Res, Notes Math., **5**, Wellesley, MA, (1993).

[HS] U. Haagerup and E. Størmer, *Equivalence of normal states on von Neumann algebras and flow of weights*, Adv. Math., **83** (1990), 180-262.

[Hu] T. Hudetz, *Topological entropy for appropriately approximated C^*-algebras*, J. Math. Phys., **35** (1994), 4303-4333.

[I1] M. Izumi, *Canonical extension on factors, in Subfactors*, H. Araki et al. (eds.), World Scientific, Singapore, (1994), 129-138.

[I2] ———, *Subalgebras of infinite C^*-algebras with finite indices*, II, Cuntz-Krieger algebras, preprint.

[L1] R. Longo, *Index of subfactors and statistics of quantum fields*, I, Commun. Math. Phys., **126** (1989), 145-155.

[L2] ———, *Duality for Hopf algebras and for subfactors*, Commun. Math. Phys., **159** (1994), 133-150.

[LR] R. Longo and J.E. Roberts, *A theory of dimension*, preprint.

[OP] D. Olesen and G.K. Pedersen, *Some C^*-dynamical systems with a single KMS state*, Math. Scand., **42** (1978), 111-118.

[P] W. Paschke, *The crossed product of a C^*-algebra by an endomorphism*, Proc. Amer. Math. Soc., **80** (1980), 113-118.

[R] M. Rørdam, *Classification of certain infinite simple C^*-algebras*, J. Funct. Anal., **131** (1995), 415-458.

[S] E. Størmer, *Entropy in operator algebras*, preprint in University of Oslo, (1994).

[T] K. Thomsen, *Topological entropy for endomorphisms of local C^*-algebras*, Comm. Math. Phys., **164** (1994), 181-193.

[V] D. Voiculescu, *Dynamical approximation entropies and topological entropy in operator algebras*, Comm. Math. Phys., **170** (1995), 249-281.

Received July 1, 1997 and revised October 2, 1998.

Osaka Kyoiku University
Kashiwara 582-8582
Japan
E-mail address: marie@cc.osaka-kyoiku.ac.jp

PACIFIC JOURNAL OF MATHEMATICS
Vol. 190, No. 2, 1999

GENERALIZED SOLENOIDS AND C*-ALGEBRAS

Valentin Deaconu

We present the continuous graph approach for some generalizations of the Cuntz-Krieger algebras. These algebras are simple, nuclear, and purely infinite, with rich K-theory. They are tied with the dynamics of a shift on an infinite path space. Interesting examples occur when the vertex spaces are unions of tori, and the shift is not necessarily expansive. We also show how the algebra of a continuous graph could be thought as a Pimsner algebra.

Introduction.

Recent papers are dealing with different generalizations of the Cuntz-Krieger algebras $\mathcal{O}_A$ (see [**Pi**], [**P1**], [**D2**], [**AR**], etc). The exact relationship between these approaches remains to be explored, but certainly there are overlaps. In [**Pi**], the author considers a Hilbert bimodule H over a C*-algebra, and creation operators on a corresponding Fock space. These operators generate the Toeplitz algebra $\mathcal{T}_H$ and, taking a quotient of this, one obtains the algebra $\mathcal{O}_H$. If the Hilbert bimodule is projective and finitely generated over an abelian, finite dimensional C*-algebra, then one recovers the algebras $\mathcal{O}_A$.

In [**P1**], the starting point is a Smale space (a compact metric space endowed with an expansive homeomorphism with canonical coordinates), on which one defines the stable and unstable equivalence relations. The associated C*-algebras have natural shift automorphisms, and the crossed products are the so called Ruelle algebras. These are strongly Morita equivalent to particular Cuntz-Krieger algebras if the Smale space is a topological Markov shift.

Our point of view is to start with a continuous oriented graph (or diagram) E, to consider the space of one-sided infinite paths (obtained by concatenation of edges in E), and to associate a groupoid (à la Renault) using the unilateral shift on this path space. The C*-algebra of this groupoid plays the role of a continuous version of the Cuntz-Krieger algebras, since these could be obtained by the same construction from a finite graph defined by a 0-1 matrix. In many cases, this groupoid algebra is simple, purely infinite, with computable K-theory. This approach offers more freedom for constructing easy, concrete examples, with prescribed K-theory. It should be mentioned that C*-algebras associated with discrete graphs were studied in [**KPRR**], [**KPR**], [**KP**]. See also the survey [**K2**].

The continuous graph approach is very similar to the point of view of polymorphisms or correspondences, introduced earlier in a measure theoretical context by Vershik and Arzumanian (see [**AR**] for a precise definition and references).

Even though our groupoid algebras could be obtained also by using the Pimsner approach, with a right choice of the Hilbert bimodule, we feel that the present point of view has certain advantages, beeing tied with the dynamics of a shift. For example, even in a case where this shift is not expansive, so the space of two-sided infinite paths has no obvious Smale space structure, we will prove that the corresponding algebra is simple and purely infinite.

In the particular case when the vertex space is a disjoint union of tori, we call the corresponding space of paths a *generalized solenoid*, and we obtain results similar to those of Brenken (see [**B**]). It is interesting to notice how these fairly complicated dynamical systems appear in a natural way from embeddings of toral algebras.

Acknowledgements. Thanks are due to several people who helped me while this paper was growing, especially Paul Muhly, Alex Kumjian, Jean Renault, Jack Spielberg, Berndt Brenken, Ian Putnam.

1. Continuous graphs and dynamical systems.

Definition 1.1. By a *continuous graph* we mean a closed subset

$$E \subset V \times \{1, 2, ..., m\} \times V,$$

where V is a compact metric space. The elements of V are called *vertices*, and the elements of E are called *edges*. The set $\{1, 2, ..., m\}$ is used to label different edges between the same pair of vertices. The graph is *oriented* when for each edge $e = (v, k, w)$ we specify the origin $o(e) = v$ and the terminus $t(e) = w$.

In this paper we consider dynamical systems $(X_+, \sigma_+), (X, \sigma)$ built from a continuous oriented graph E. The space X_+ is the space of one-sided infinite paths,

$$X_+ = \{(x_i, k_i)_{i=0}^{\infty} \mid (x_i, k_i, x_{i+1}) \in E, i \geq 0\},$$

and $\sigma_+ : X_+ \to X_+$ is the unilateral shift,

$$\sigma_+(x_i, k_i)_p = (x_{p+1}, k_{p+1}).$$

The space X is the space of two-sided infinite paths, and σ is the bilateral shift. The dynamical system (X_+, σ_+) unifies in a natural way the notion of a continuous map $T : V \to V$, a (finitely-generated) semigroup or group of continuous transformations $\mathcal{S} : V \to V$ and the (unilateral) Markov shifts (when V is a finite set and E is defined by a 0-1 matrix)(see [**F**]). For example, if $T : V \to V$ is a continuous map, we can take $E = \Gamma(T)$, the graph of T (in this case $m = 1$, and we omit it). Then X_+ is homeomorphic to V, and σ_+ is conjugated to T.

Proposition 1.2. *The dynamical system (X, σ) could be obtained from (X_+, σ_+) by the usual inverse limit process by which one associates a homeomorphism to a continuous onto map.*

Proof. Indeed, let

$$\tilde{X} = \left\{ (\xi_n) \in \prod_1^\infty X_+ \mid \sigma_+(\xi_{n+1}) = \xi_n \right\}.$$

We have $\pi : \tilde{X} \to X_+, \pi(\xi_1\xi_2...) = \xi_1$, and $\tilde{\sigma}_+ : \tilde{X} \to \tilde{X}, \tilde{\sigma}_+(\xi_1\xi_2...) = (\sigma_+(\xi_1)\xi_1\xi_2...)$, such that $\sigma_+\pi = \pi\tilde{\sigma}_+$. Since

$$X_+ = \left\{ (e_n) \in \prod_1^\infty E \mid t(e_n) = o(e_{n+1}) \right\},$$

$\tilde{X} \subset \prod_1^\infty \prod_1^\infty E$ could be identified with X, the space of two-sided infinite paths, and $\tilde{\sigma}_+$ with the bilateral shift σ. □

Definition 1.3. For each continuous oriented graph E we define its dual (or opposite) graph $\hat{E}$ by

$$\hat{E} = \{(x, k, y) \mid (y, k, x) \in E\}.$$

This way we get dynamical systems $(\hat{X}_+, \hat{\sigma}_+), (\hat{X}, \hat{\sigma})$, where $\hat{X}_+$, $\hat{X}$ are constructed from $\hat{E}$, and $\hat{\sigma}_+$, $\hat{\sigma}$ are the unilateral and bilateral shift, respectively. Of course, the systems (X, σ) and $(\hat{X}, \hat{\sigma}^{-1})$ are conjugated. But (X_+, σ_+) and $(\hat{X}_+, \hat{\sigma}_+)$ could be very different.

Example 1.4. Take $V = \mathbf{T}$, the unit circle, and E the graph of the map $z \mapsto z^2$,

$$E = \{(z, z^2) \mid z \in \mathbf{T}\}.$$

Then $X_+ = \mathbf{T}, \sigma_+(z) = z^2$, and $\hat{X}_+$ is a solenoid,

$$\hat{X}_+ = \{(z_1, z_2, ...,) \mid z_n \in \mathbf{T}, z_{n+1}^2 = z_n, n \geq 1\},$$

$$\hat{\sigma}_+(z_1, z_2, ...) = (z_2, z_3, ...).$$

Note that if V has a group structure and $E \subset V \times V$ is a subgroup, then X_+ and X have also natural group structures, with componentwise multiplication.

2. The C*-algebra of a continuous graph.

In the case the two projections $o, t : E \to V, o(x,k,y) = x$ and $t(x,k,y) = y$ are onto local homeomorphisms, we can associate to the graph E a C*-algebra $C^*(E)$, using the Renault groupoid of the dynamical system (X_+, σ_+). The space X_+ is endowed with a metric defining the product topology. If δ denotes the metric on V, then one can take

$$d((x_i, k_i), (x'_i, k'_i)) = \sum_{i \geq 0} \frac{\delta(x_i, x'_i) + \mid k_i - k'_i \mid}{2^i}$$

as a metric on X_+. Similarly, we obtain a metric on X.

The unilateral shift σ_+ is a local homeomorphism, and we consider the following locally compact r-discrete groupoid:

$$\begin{aligned} \Gamma &= \Gamma(X_+, \sigma_+) \\ &= \{(x, n, y) \in X_+ \times \mathbf{Z} \times X_+ \mid \exists k, l \geq 0, n = k - l, \sigma_+^k(x) = \sigma_+^l(y)\}. \end{aligned}$$

The range map, the source map, and the operations are given as follows:

$$r(x,n,y) = x, \quad s(x,n,y) = y,$$
$$(x,n,y)(y,p,z) = (x, n+p, z), \quad (x,n,y)^{-1} = (y, -n, x).$$

The unit space of Γ is X_+, if we identify $(x, 0, x)$ with x. A basis of open sets for Γ is given by

$$Z(U, V, k, l) = \{(x, k-l, (\sigma_+^l \mid_V)^{-1} \circ \sigma_+^k(x)), x \in U\},$$

where U and V are open subsets of X_+, and k, l are such that $\sigma_+^k \mid_U$ and $\sigma_+^l \mid_V$ are homeomorphisms with the same open range.

Definition 2.1. Given a continuous oriented graph E with the maps o, t onto local homeomorphisms, we define its C*-algebra $C^*(E)$ to be $C^*(\Gamma)$, the C*-algebra of the Renault groupoid associated with the dynamical system (X_+, σ_+).

To understand the structure of $C^*(E)$, consider the homomorphism $c : \Gamma \to \mathbf{Z}$, $c(x,n,y) = n$, and let's denote by R^∞ the subgroupoid $c^{-1}(0)$. If we denote by B the C^*-algebra of the equivalence relation R^∞, the local homeomorphism σ_+ induces a *-endomorphism α of B by the formula

$$\alpha(f)(x,y) = \frac{1}{\sqrt{p(\sigma_+(x))p(\sigma_+(y))}} f(\sigma_+(x), \sigma_+(y)), f \in C_c(R^\infty),$$

where for $x \in X_+$, $p(x)$ is the number of paths z such that $\sigma_+(z) = x$. Moreover, assuming that σ_+ is not one-to-one, α is induced by a non unitary isometry v, in the sense that $\alpha(f) = vfv^*$, where

$$v(x,n,y) = \begin{cases} (p(\sigma_+(x)))^{-1/2}, & \text{if } n = 1 \text{ and } y = \sigma_+(x) \\ 0, & \text{otherwise.} \end{cases}$$

Indeed, $v^*v = 1$, and

$$vv^*(x,n,y) = \begin{cases} p(\sigma_+(x))^{-1} & \text{if } \sigma_+(x) = \sigma_+(y) \\ 0, & \text{otherwise.} \end{cases}$$

Thus, α is a proper corner endomorphism of B, and $C^*(E)$ is isomorphic to the crossed product $B \times_\alpha \mathbf{N}$ (see [**R1**]).

In order to compute the K-theory of $C^*(E)$, we can use the exact sequence

$$\begin{array}{ccccc} K_0(C^*(R^\infty)) & \xrightarrow{\text{id}-\alpha_0} & K_0(C^*(R^\infty)) & \xrightarrow{i_0} & K_0(C^*(E)) \\ {\scriptstyle\partial_1}\uparrow & & & & \downarrow{\scriptstyle\partial_0} \\ K_1(C^*(E)) & \xleftarrow{i_1} & K_1(C^*(R^\infty)) & \xleftarrow{\text{id}-\alpha_1} & K_1(C^*(R^\infty)) \end{array}$$

where $i : C^*(R^\infty) \to C^*(E)$ is the inclusion map.

If on E we consider the equivalence relation R defined by t: two edges (x,k,y) and (x',k',y') are equivalent iff $y = y'$, then the C*-algebra $C^*(R)$ is a continuous trace algebra with spectrum V, and there is a canonical embedding

$$\Phi : C(V) \to C^*(R),$$

$$\Phi(f)((x,k,y),(x',k',y)) = \begin{cases} f(x), & \text{if } x = x' \text{ and } k = k' \\ 0, & \text{otherwise.} \end{cases}$$

Using the same method as in the Main Result of [**D2**], we get:

Theorem 2.2. *If Φ_0 and Φ_1 are the maps induced on K-theory by the embedding $\Phi : C(V) \to C^*(R)$, and if the K-theory groups $K^0(V)$ and $K^1(V)$ are free and finitely generated, then*

$$K_0(C^*(E)) = \ker(\text{id} - \Phi_1) \oplus K^0(V)/(\text{id} - \Phi_0)K^0(V),$$
$$K_1(C^*(E)) = \ker(\text{id} - \Phi_0) \oplus K^1(V)/(\text{id} - \Phi_1)K^1(V).$$

Using this theorem, we can get interesting examples of simple purely infinite C*-algebras with prescribed K-theory groups. In particular, in the next example, we construct C*-algebras A_n with $K_0(A_n) = 0$ and $K_1(A_n) = \mathbf{Z}_n$.

Example 2.3. Let $V = V_1 \cup V_2$, where $V_i,\ i = 1,2$ are copies of the unit circle, and

$$E = \{(v,w) \in V_1 \times V_1 \mid v = w^2\} \cup \{(v,w) \in V_1 \times V_2 \mid v = w\} \cup$$
$$\{(v,w) \in V_2 \times V_1 \mid v = w\} \cup$$
$$\{(v,k,w) \in V_2 \times \{1,2,...,n+2\} \times V_2 \mid w = v^n\}.$$

Then

$$\Phi : C(V_1) \oplus C(V_2) \longrightarrow C(V_1) \otimes \mathbf{M}_2 \oplus C(V_2) \otimes \mathbf{M}_{n(n+2)+1},$$

$$\Phi(f\oplus g)=\begin{pmatrix}\sigma_2 f & 0\\ 0 & \sigma_1 g\end{pmatrix}\oplus\begin{pmatrix}\sigma_1 f & 0 & 0 & 0\\ 0 & \hat{\sigma}_n g & 0 & 0\\ 0 & 0 & \ddots & 0\\ 0 & 0 & 0 & \hat{\sigma}_n g\end{pmatrix}.$$

Here $\sigma_k f(z)=f(z^k)$, and $\hat{\sigma}_k$ is the k-times around embedding (the homomorphism compatible with the covering $z\to z^k$). There are $n+2$ copies of $\hat{\sigma}_n g$ in the definition of Φ. Note that

$$\Phi_0=\begin{pmatrix}1 & 1\\ 1 & n(n+2)\end{pmatrix},\ \ \Phi_1=\begin{pmatrix}2 & 1\\ 1 & n+2\end{pmatrix}.$$

It follows that

$$\ker(\mathrm{id}-\Phi_0)=0,\quad \mathbf{Z}^2/(\mathrm{id}-\Phi_0)\mathbf{Z}^2=0,$$
$$\ker(\mathrm{id}-\Phi_1)=0,\quad \mathbf{Z}^2/(\mathrm{id}-\Phi_1)\mathbf{Z}^2=\mathbf{Z}_n,$$

therefore the corresponding C*-algebra $C^*(E)$ has $K_0=0,\ K_1=\mathbf{Z}_n$. One can check that every orbit with respect to the equivalence relation R^∞ is dense, therefore $C^*(R^\infty)$ and $C^*(E)$ are simple. The latter algebra is purely infinite because it appears as a crossed product of an inductive limit of circle algebras by an endomorphism that does not preserve any trace (see Theorem 2.1 in [**R2**]).

Definition 2.4. Recall that $\sigma_+ : X_+\to X_+$ is (positive) expansive if there is a constant $c>0$ such that $x\neq y$ implies $d(\sigma_+^n(x),\sigma_+^n(y))\geq c$ for some integer $n\geq 0$. An element $x\in X_+$ is *eventually periodic* if there are two integers $p\neq q$ with $\sigma_+^p(x)=\sigma_+^q(x)$.

In [**De**], Proposition 4.2, it is proved that if σ_+ is expansive and the eventually periodic points form a dense set with empty interior, then $C^*(\Gamma)$, and therefore $C^*(E)$, is nuclear, purely infinite, and belongs to the bootstrap class $\mathcal{N}$.

Note that in the above hypotheses, the groupoid $\Gamma=\Gamma(X_+,\sigma_+)$ is essentially free, i.e. the set of points in the unit space with trivial isotropy is dense.

We will see in the last section that even for non-expansive σ_+, the C*-algebra $C^*(E)$ could be purely infinite. Of course, it can not be finite as long as the endomorphism α is induced by a non unitary isometry v. If σ_+ is minimal (i.e. each orbit with respect to the equivalence relation R^∞ is dense), then this C*-algebra is also simple.

Remark 2.5. When σ_+ is expansive, there are other C*-algebras associated with the continuous graph E. According to [**AR**], in this case, the space X of two-sided infinite paths has a Smale space structure, and one may consider the stable equivalence relation:

$$R_s=\{(x,y)\in X\times X\ \mid d(\sigma^n(x),\sigma^n(y))\to 0 \text{ as } n\to+\infty\}.$$

Then $C^*(R_s)$ is strongly Morita equivalent to $C^*(R^\infty)$, and its Ruelle algebra $C^*(R_s)\times\mathbf{Z}$ is strongly Morita equivalent to $C^*(E)$ (see [**AR**], Theorem 4.5).

Another C*-algebra which could be associated with the continuous graph E is the crossed product $C(X)\times_\sigma\mathbf{Z}$.

3. The connection with the Pimsner algebras $\mathcal{O}_H$.

In this paragraph, we recall the Pimsner construction from [**Pi**], and we show how the C*-algebra of a continuous graph could be thought as $\mathcal{O}_H$, for a particular Hilbert bimodule H. To a pair (H,A), where H is a (right) Hilbert module over a C*-algebra A, and A acts to the left on H via a $*$-homomorphism $\varphi : A\to L(H)$, Pimsner constructs a C*-algebra $\mathcal{O}_H$, which generalizes both the crossed products by $\mathbf{Z}$ and the Cuntz-Krieger algebras. The algebra $\mathcal{O}_H$ is a quotient of the generalized Toeplitz algebra $\mathcal{T}_H$, generated by the creation operators $T_\xi,\ \xi\in H$ on the Fock space $\mathcal{H}_+=\bigoplus_{n=0}^{\infty}H^{\otimes n}$. Here $H^{\otimes 0}=A$, and for $n\geq 1$, $H^{\otimes n}$ denotes the n-th tensor power of H, balanced via the map φ. By definition, $T_\xi a=\xi a$, for $a\in A$, and $T_\xi(\xi_1\otimes ...\otimes\xi_n)=\xi\otimes\xi_1\otimes ...\otimes\xi_n$, for $\xi_1\otimes ...\otimes\xi_n\in H^{\otimes n}$.

To give another description of the algebra $\mathcal{O}_H$, Pimsner considers a new pair $(H_\infty,\mathcal{F}_H)$, where $\mathcal{F}_H$ is the C*-algebra generated by all the compact operators $K(H^{\otimes n}),\ n\geq 0$ in $\varinjlim L(H^{\otimes n})$, and $H_\infty=H\otimes\mathcal{F}_H$. The advantage is that H_∞ becomes an $\mathcal{F}_H-\mathcal{F}_H$ bimodule, such that the adjoint H_∞^* is also an $\mathcal{F}_H-\mathcal{F}_H$ bimodule. The C*-algebra $\mathcal{O}_H$ is represented on the two-sided Fock space

$$\mathcal{H}_\infty=\bigoplus_{n\in\mathbf{Z}}H_\infty^{\otimes n},$$

where for $n<0, H_\infty^{\otimes n}$ means $(H_\infty^*)^{\otimes -n}$. In fact, it is isomorphic to the C*-algebra generated by the multiplication operators $M_\xi\in L(\mathcal{H}_\infty)$, where $\xi\in H_\infty$, and $M_\xi\omega=\xi\otimes\omega$.

Given a continuous graph E such that the origin and terminus maps $E\to V$ are onto local homeomorphisms, let $A=C(V)$, and let $H=C(E)$ (as a vector space), with the structure of Hilbert A-module given by

$$(\xi f)(e)=\xi(e)f(t(e)),\ \xi\in H, f\in A, e\in E,$$

$$\langle\xi,\eta\rangle(v):=\sum_{t(e)=v}\overline{\xi(e)}\eta(e), v\in V,\ \xi,\eta\in H.$$

In other words, the inner product is given by $\langle \xi, \eta \rangle = P(\bar{\xi}\eta)$, where $P : C(E) \to C(V)$ is the conditional expectation

$$(P\xi)(v) = \sum_{t(e)=v} \xi(e).$$

The left module structure is given by

$$\varphi : A \to L(H),\ (\varphi(f)\xi)(e) = f(o(e))\xi(e)\ f \in A, \xi \in H.$$

Note that indeed $\varphi(f)$ is in $L(H)$, having the adjoint $\varphi(\bar{f}), f \in A$.

To prove that $\mathcal{O}_H$ with this choice of A, H and φ is isomorphic to $C^*(E)$, let's identify the C*-algebra $\mathcal{F}_H$ in this case.

Note that $H \otimes_\varphi H$ is a quotient of $C(E) \otimes C(E)$, where we identify $\xi f \otimes \eta$ with $\xi \otimes \varphi(f)\eta$ for any $\xi, \eta \in H$ and any $f \in A$. Therefore $H \otimes_\varphi H$ could be identified as a vector space with the continuous functions on the set

$$\{(e_1, e_2) \in E \times E \mid\ t(e_1) = o(e_2)\}.$$

This set will be denoted by X_2, and is precisely the set of paths of length 2. In a similar way, $H^{\otimes n}$ is identified (as a vector space) with $C(X_n)$, where X_n is the set of paths of length n. The Hilbert A-module structure on $H^{\otimes n}$ for $n \geq 2$ is given by

$$(\xi f)(x) = \xi(x) f(t_n(x)), x \in X_n$$

where $t_n : X_n \to V, t_n(e_1 e_2 ... e_n) = t(e_n)$, and by

$$\langle \xi, \eta \rangle_n = P_n(\bar{\xi}\eta).$$

Here P_n is the conditional expectation

$$P_n : C(X_n) \to C(V), P_n(\xi)(v) = \sum_{t_n(x)=v} \xi(x).$$

Proposition 3.1. *The C^*-algebra $K(H)$ is isomorphic with $C^*(R)$, where*

$$R = \{(e_1, e_2) \in E \times E \mid\ t(e_1) = t(e_2)\}$$

is the equivalence relation associated with the map t. The map $\varphi : A \to L(H)$ could be identified with the embedding $\Phi : C(V) \to C^*(R)$, defined before Theorem 2.2. Moreover, $K(H^{\otimes n}) \simeq C^*(R_n)$, where

$$R_n = \{(x, y) \in X_n \times X_n \mid\ t_n(x) = t_n(y)\}$$

is the equivalence relation associated with t_n.

Proof. Taking into account the fact that o and t are local homeomorphisms, we have $L(H) = K(H)$, since H is algebraically finitely generated.

Now $K(H) = H \otimes H^*$, the tensor product balanced over A, where H^* is the adjoint of H. Since $\xi f \otimes \eta^* = \xi \otimes f\eta^*$, it follows that, as a set, $K(H) = C(R)$. The multiplication of compact operators turns out to be the convolution product on $C(R)$, therefore, as C*-algebras, $K(H) = C^*(R)$. $\square$

Corollary 3.2. *We have $\mathcal{F}_H = \varinjlim C^*(R_n)$. Therefore, $\mathcal{F}_H$ is isomorphic to the algebra $C^*(R^\infty)$.*

Proof. Note that for $n \geq 1$, the inclusion $\phi_n : \mathbf{C}^*(R_n) \to \mathbf{C}^*(R_{n+1})$,

$$(\phi_n)(f)(x_1...x_{n+1}, y_1...y_{n+1}) = \begin{cases} f(x_1...x_n, y_1...y_n) & \text{if } x_{n+1} = y_{n+1} \\ 0, & \text{otherwise} \end{cases}$$

is just the map $K(H^{\otimes n}) \to K(H^{\otimes n+1}),\ \ T \mapsto T \otimes I$. Here R_1=R. □

In order to establish an isomorphism between $\mathbf{C}^*(\Gamma)$ and $\mathcal{O}_H$, we show that they appear as the $\mathbf{C}^*$-algebras associated to isomorphic Fell bundles over the group $\mathbf{Z}$. This point of view was suggested by Abadie, Eilers and Exel in [**AEE**]. The definition of a Fell bundle and of the associated C^*-algebra is taken from [**K1**].

To the pair $(H_\infty, \mathcal{F}_H)$, we can associate the Fell bundle $\mathcal{B}$, where $\mathcal{B}_n := H_\infty^{\otimes n}, n \in \mathbf{Z}$. The multiplication is given by the tensor product, identifying $H_\infty^* \otimes H_\infty$ with $\mathcal{F}_H$ and $H_\infty \otimes H_\infty^*$ with the ideal $\mathcal{F}_H^1$ of $\mathcal{F}_H$, generated by $K(H^{\otimes n})$ with $n \geq 1$. But $\mathcal{F}_H^1$ is equal to $\mathcal{F}_H$ in our case. The involution is obvious. Then

$$L^2(\mathcal{B}) = \mathcal{H}_\infty = \bigoplus_{n \in \mathbf{Z}} H_\infty^{\otimes n}.$$

Since $\mathcal{H}_\infty$ is generated by $\mathcal{F}_H$ and H_∞, it follows that the C^*-algebra generated by the operators M_ξ is isomorphic to $C^*(\mathcal{B})$. Hence, $\mathcal{O}_H \simeq C^*(\mathcal{B})$.

For the groupoid Γ and $l \in \mathbf{Z}$, take

$$\Gamma_l := \{(x, k, y) \in \Gamma \mid k = l\} = \{(x, y) \in X \times X \mid x_n = y_{n+l} \text{ for large } n\},$$

and $\mathcal{D}_l = \overline{C_c(\Gamma_{-l})}$ (closure in $C^*(\Gamma)$). This way, we obtain a $\mathbf{Z}$-grading on $C^*(\Gamma)$, and it is easy to see that this C*-algebra could be recovered as $C^*(\mathcal{D})$. But

$$\mathcal{D}_0 = C^*(R^\infty) \simeq \mathcal{F}_H = \mathcal{B}_0,$$

and

$$\mathcal{D}_1 = \overline{C_c(\Gamma_{-1})} \simeq H \otimes_A \mathcal{F}_H = H_\infty = \mathcal{B}_1.$$

We get

Proposition 3.3. *With the above choice of A, H and φ, the C^*-algebras $C^*(E)$ and $\mathcal{O}_H$ are isomorphic.*

4. Generalized solenoids.

A solenoid is a compact connected abelian group of finite dimension. For example, if $\mathbf{T}$ is the unit circle,

$$\mathbf{T}(m) = \{z \in \mathbf{T}^{\mathbf{Z}} \mid \; z_k^m = z_{k+1}, \; k \in \mathbf{Z}\}$$

is such a group, for any integer $m > 1$. The bilateral shift σ on $\mathbf{T}(m)$, $\sigma(z)_p = z_{p+1}$ is a homeomorphism, and in many respects it is an analogue of the Bernoulli shift. In [**B**], Brenken considered the dynamical system (G_0, σ) for G_0 the connected component of the identity of the group

$$G = \{z \in (\mathbf{T}^d)^{\mathbf{Z}} \;\mid\; Fz_k = Mz_{k+1}, \; k \in \mathbf{Z}\},$$

where M, F are surjective endomorphisms of the d-torus, given by matrices $M, F \in \mathbf{M}_d(\mathbf{Z})$ with nonzero determinant. (Note that the case $d = 1, M = 1, F = m$ corresponds to the above example $\mathbf{T}(m)$.) The space G_0 has a natural local product structure, being a principal bundle over $\mathbf{T}^d$ with fiber the Cantor set. Moreover, it has a Smale space structure, and the author identifies the C^*-algebras associated with the stable and unstable equivalence relations.

Definition 4.1. By a *generalized solenoid* we mean the space X of two-sided infinite paths with edges in the graph E described bellow. Let $V = \mathbf{T}_1^d \sqcup ... \sqcup \mathbf{T}_N^d$ be the disjoint union of N copies of the d-dimensional torus $\mathbf{T}^d$, and let $L = (l(i,j))_{i,j}$ be an $N \times N$ matrix with positive integer entries (the "incidence" matrix of the graph). We require that in the matrix L each row and each column has at least a nonzero entry. For each pair (i,j) with $l(i,j) \geq 1$, consider a family of closed, connected subgroups $G_1^{ij}, G_2^{ij}, ..., G_{l(i,j)}^{ij}$ of $\mathbf{T}_i^d \times \mathbf{T}_j^d$, not necessarily distinct, such that all the projections on $\mathbf{T}_i^d$ and $\mathbf{T}_j^d$ are surjective. For the pairs (i,j) with $l(i,j) = 0$, this family is empty by definition. We take E to be the disjoint union of all the groups G_k^{ij}, $1 \leq i,j \leq N, 1 \leq k \leq l(i,j)$, with obvious origin and terminus maps.

It is known (see [**KS**]) that there are families of $d \times d$ nonsingular matrices with integer entries,

$$\mathcal{A}_{ij} = \{A_1^{ij}, ..., A_{l(i,j)}^{ij}\}, \;\; \mathcal{B}_{ij} = \{B_1^{ij}, ..., B_{l(i,j)}^{ij}\}.$$

such that

$$G_l^{ij} = \{(z,w) \in \mathbf{T}_i^d \times \mathbf{T}_j^d \mid A_l^{ij} z = B_l^{ij} w\}.$$

The matrices A_k^{ij}, B_k^{ij} are not necessarily distinct. Note that a generalized solenoid X has no longer a group structure, and the dynamical system (X, σ) is an analogue of the Matkov shift.

Example 4.2. Let $d = 2, N = 2$,

$$A_1^{11} = \begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix}, B_1^{11} = \begin{pmatrix} -1 & 0 \\ 2 & 3 \end{pmatrix}, A_1^{12} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, B_1^{12} = \begin{pmatrix} 3 & 0 \\ 1 & 1 \end{pmatrix},$$

$$A_2^{12} = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}, B_2^{12} = \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix},$$

$$A_1^{22} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, B_1^{22} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}, A_2^{22} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, B_2^{22} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The space of edges is

$$\{((z,w),1,(t,u)) \in \mathbf{T}_1^2 \times \{1\} \times \mathbf{T}_1^2 \mid \ z^3 w = t^{-1},\ zw = t^2 u^3\} \cup$$

$$\{((z,w),1,(t,u)) \in \mathbf{T}_1^2 \times \{1\} \times \mathbf{T}_2^2 \mid \ z = t^3, w^2 = tu\} \cup$$

$$\{((z,w),2,(t,u)) \in \mathbf{T}_1^2 \times \{2\} \times \mathbf{T}_2^2 \mid \ z^2 w = t^{-1}u, w = t\} \cup$$

$$\{((z,w),1,(t,u)) \in \mathbf{T}_2^2 \times \{1\} \times \mathbf{T}_2^2 \mid \ z = t^2, w = u\} \cup$$

$$\{((z,w),2,(t,u)) \in \mathbf{T}_2^2 \times \{2\} \times \mathbf{T}_2^2 \mid \ z = t, w = u\}.$$

The corresponding embedding $C(V) \to C^*(R)$ of toral algebras is

$$\Phi : C(\mathbf{T}^2) \oplus C(\mathbf{T}^2) \to C(\mathbf{T}^2) \otimes \mathbf{M}_3 \oplus C(\mathbf{T}^2) \otimes \mathbf{M}_7,$$

$$\Phi(f \oplus g) = \Phi_{11}(f) \oplus \begin{pmatrix} \Phi_{12}(f) & 0 \\ 0 & \Phi_{22}(g) \end{pmatrix},$$

where

$$\Phi_{11}(f) = \hat{\sigma}_{B_1^{11}} \circ \sigma_{A_1^{11}}(f), \Phi_{12}(f) = \begin{pmatrix} \hat{\sigma}_{B_1^{12}} \circ \sigma_{A_1^{12}}(f) & 0 \\ 0 & \hat{\sigma}_{B_2^{12}} \circ \sigma_{A_2^{12}}(f) \end{pmatrix},$$

and

$$\Phi_{22}(g) = \begin{pmatrix} \hat{\sigma}_{B_1^{22}} \circ \sigma_{A_1^{22}}(g) & 0 \\ 0 & \hat{\sigma}_{B_2^{22}} \circ \sigma_{A_2^{22}}(g) \end{pmatrix}.$$

Here $\sigma_A : C(\mathbf{T}^2) \to C(\mathbf{T}^2)$ denotes the $*$-homomorphism induced by the covering A, defined by $(\sigma_A f)(z) = f(Az)$, and $\hat{\sigma}_A : C(\mathbf{T}^2) \to C(\mathbf{T}^2) \otimes \mathbf{M}_{|\det A|}$ is the homomorphism compatible with A, in the sense that $\hat{\sigma}_A \circ \sigma_A(f) = f \otimes I_{|\det A|}$.

Remark 4.3. Given a generalized solenoid X, let's denote by K the space of two-sided infinite paths in the discrete graph with N vertices, where from the vertex i to the vertex j there are $l(i,j)$ edges. On the Cantor set K consider the Markov shift τ. Note that there is a natural map $\rho : X \to K$, $\rho((x_n, k_n)_{n\in\mathbf{Z}}) = (k_n)_{n\in\mathbf{Z}}$. Moreover, $\rho\sigma = \tau\rho$. Therefore, (X,σ) is in fact an extension of a Markov shift (K, τ). In the above example, the incidence matrix L is

$$L = \begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix}.$$

Note that the fiber of ρ has a group structure, therefore X could be seen as a group bundle over the Cantor set K. The groups are in fact solenoids if they are connected.

The space X is also fibered over $V = \mathbf{T}_1^d \sqcup \ldots \sqcup \mathbf{T}_N^d$ by the map $\pi : X \to V, \pi((x_n, k_n)_{n\in\mathbf{Z}}) = x_0$. The fibers of π are totally disconnected, since the set $\{x_n \in V \mid \ \pi((x_n, k_n)_{n\in\mathbf{Z}}) = x_0\}$ is finite for each fixed $x_0 \in V$ and $n \in \mathbf{Z}$.

The following example arose in a discussion with Jack Spielberg.

Example 4.4. Let $V = \mathbf{T}$, the unit circle, and

$$E = \{(z, 1, z^2) \mid z \in V\} \cup \{(z^3, 2, z) \mid z \in V\}.$$

Then

$$X = \{(z_n, k_n) \in (V \times \{1, 2\})^{\mathbf{Z}} \mid k_n = 1 \Rightarrow \\ z_{n+1} = z_n^2,\ k_n = 2 \Rightarrow\ z_{n+1}^3 = z_n\}.$$

We will show that $\sigma : X \to X,\ \sigma(z_n, k_n)_p = (z_{p+1}, k_{p+1})$ is not expansive, therefore the space (X, σ) has not a Smale space structure.

It suffices to show that for any $\varepsilon > 0$, we can find two distinct sequences (z_n, k_n) and (w_n, k_n) such that $\delta(z_n, w_n) \geq \varepsilon$ for all $n \in \mathbf{Z}$. Fix $z_0, w_0 \in V$. The idea is that, taking in a certain order squares, cubes, square roots and cubic roots, the corresponding vertices remain close together. We can choose two sequences of integers $(a_n)_{n\geq 1}, (b_n)_{n\geq 1}$ such that

$$\lim_{n\to\infty} \frac{2^{a_1+\ldots+a_n}}{3^{b_1+\ldots+b_n}} = 1.$$

Consider the symmetric sequence $(k_n)_{n\in\mathbf{Z}}$ described as

$$\ldots \underbrace{2\ldots2}_{b_2}\underbrace{1\ldots1}_{a_2}\underbrace{2\ldots2}_{b_1}\underbrace{1\ldots1}_{a_1}\underbrace{\bar{1}\ldots1}_{a_1}\underbrace{2\ldots2}_{b_1}\underbrace{1\ldots1}_{a_2}\underbrace{2\ldots2}_{b_2}\ldots,$$

where the bar indicates k_0. Given $\varepsilon > 0$, we can choose z_0 and w_0 sufficiently close together (but distinct), and z_n and w_n in a consistent way (when we take square or cubic roots) such that $\delta(z_n, w_n) \geq \varepsilon$. It follows that

$$d(\sigma^p(z_n, k_n), \sigma^p(w_n, k_n)) \geq \varepsilon\ \ \forall p \in \mathbf{Z},$$

and the shift is not expansive.

Nevertheless, the orbits with respect to R^∞ are dense in X_+, and there is no shift invariant trace, therefore the C*-algebra $C^*(E)$ is simple and purely infinite.

Note that in this example, the dynamical system (X_+, σ_+) is an extension of the Bernoulli shift $(\{1,2\}^{\mathbf{N}}, \tau)$. The fibers of the map $\rho : X_+ \to \{1,2\}^{\mathbf{N}}$ are circles over the sequences which contain only a finite numbers of 2's, and solenoids over the sequences containing infinitely many 2's.

It is interesting to notice that $C^*(E)$ and $C^*(\hat{E})$ are both simple, purely infinite, with K-theory

$$K_0(C^*(E)) = K_1(C^*(\hat{E})) = \mathbf{Z}_2, \;\; K_1(C^*(E)) = K_0(C^*(\hat{E})) = \mathbf{Z}_3.$$

In [**P1**] it is proved that the Ruelle algebra associated to the graph of the map $z \mapsto z^p$ on the unit circle is isomorphic to the one obtained from the dual graph. Whether this is true for more general graphs is an open question.

References

[AEE] B. Abadie, S. Eilers and R. Exel, *Morita equivalence for crossed products by Hilbert* C*-*bimodules*, Trans. AMS, **350** (1998), 3043-3054.

[AR] V. Arzumanian and J. Renault, *Examples of pseudogroups and their C*-algebras*, in 'Operator Algebras and Quantum Field Theory', Editors: S. Doplicher, R. Longo, J.E. Roberts and L. Zsido, International Press, 1997.

[B] B. Brenken, *The local product structure of expansive automorphisms of solenoids and their associated C*-algebras*, Canad. J. Math., **48**(**4**) (1996), 692-709.

[D1] V. Deaconu, *Groupoids associated with endomorphisms*, Trans. AMS, **347** (1995), 1779-1786.

[D2] ———, *Generalized Cuntz-Krieger algebras*, Proc. AMS, **124** (1996), 3427-3435.

[De] C. Anantharaman-Delaroche, *Purely infinite C*-algebras arising from dynamical systems*, Bull. Soc. Math. France, **125**(**2**) (1997), 199-225.

[F] S. Friedland, *Entropy of graphs, semigroups and groups*, in 'Ergodic theory of $\mathbf{Z}^d$-actions', 319-343, Cambridge, 1996.

[KPS] J. Kaminker, I. Putnam and J. Spielberg, *Operator algebras and hyperbolic dynamics*, Operator Algebras and Quantum Field Theory, Editors: S. Doplicher, R. Longo, J.E. Roberts and L. Zsido, International Press 1997.

[KS] B. Kitchens and K. Schmidt, *Automorphisms of compact groups*, Ergod. Th & Dynam. Sys., **9** (1989), 691-735.

[K1] A. Kumjian, *Fell bundles over groupoids*, Proc. AMS, **126**(**4**) (1998), 1115-1125.

[K2] ———, *Notes on C*-algebras of graph*, Operator algebras and operator theory (Shanghai, 1997), 189-200; Contemp. Math., **228**, AMS, Providence, RI, 1998.

[KP] A. Kumjian and D. Pask, *C*-algebras of directed graphs and group actions*, to appear in Ergodic Theory Dyn. Sys.

[KPR] A. Kumjian, D. Pask and I. Raeburn, *Cuntz-Krieger algebras of directed graphs*, Pacific J. Math., **184**(**1**) (1998), 161-174.

[KPRR] A. Kumjian, D. Pask, I. Raeburn and J. Renault, *Graphs, groupoids and Cuntz-Krieger algebras*, J. Funct. Anal., **144** (1997), 505-541.

[Pi] M.V. Pimsner, *A class of C*-algebras generalizing both Cuntz-Krieger algebras and crossed products by* **Z**, in 'Free Probability Theory', edited by D.V. Voiculescu, Fields Institute Communications, Vol. 12, 1996, 189-212.

[P1] I. Putnam, *C*-algebras from Smale spaces*, Can. J. Math., **48** (1996), 175-195.

[P2] ———, *Hyperbolic Systems and Generalized Cuntz-Krieger Algebras*, Lecture Notes, August 1996.

[R1] M. Rørdam, *Classification of Certain Infinite Simple C*-algebras*, J. Funct. Anal., **131**(**2**) (1995), 415-458.

[R2] ———, *Classification of Certain Infinite Simple C*-algebras*, III, Operator Algebras and Their Applications, Fields Institute Communications, No. 13, 1997, 257-282.

Received February 4, 1998. This research was partially supported by NSF grant DMS-9706982.

University of Nevada
Reno, NV 89557
E-mail address: vdeaconu@math.unr.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 190, No. 2, 1999

NONHYPERBOLIC DEHN FILLINGS ON HYPERBOLIC 3-MANIFOLDS

Mario Eudave-Muñoz and Ying-Qing Wu

In this paper we will give three infinite families of examples of nonhyperbolic Dehn fillings on hyperbolic manifolds. A manifold in the first family admits two Dehn fillings of distance two apart, one of which is toroidal and annular, and the other is reducible and ∂-reducible. A manifold in the second family has boundary consisting of two tori, and admits two reducible Dehn fillings. A manifold in the third family admits a toroidal filling and a reducible filling with distance 3 apart. These examples establish the virtual bounds for distances between certain types of nonhyperbolic Dehn fillings.

1. Introduction.

Given a slope r on a torus boundary component T_0 of a 3-manifold M, the *Dehn filling* of M along the slope r, denoted by $M(r)$, is the manifold obtained by gluing a solid torus V to M along ∂V and T_0 so that r bounds a meridian disk on V. A manifold is *simple* if it is irreducible, ∂-irreducible, atoroidal, and anannular. Thus a simple manifold is either hyperbolic, or a small Seifert fiber space, or it would be a counter example to the Geometrization Conjecture. In particular, if $M(r)$ has nonempty toroidal boundary, then it is simple if and only if it is hyperbolic [**Th**]. A Dehn filling $M(r)$ is of type S (resp. D, T, A) if $M(r)$ contains an essential S^2 (resp. D^2, T^2, A^2), so it is reducible (resp. ∂-reducible, toroidal, annular). The bound $\Delta(X, Y)$ is the least nonnegative number n such that if M is a hyperbolic manifold which admits two Dehn fillings $M(r_1), M(r_2)$ of type X, Y, respectively, then $\Delta(r_1, r_2) \leq n$. The bounds $\Delta(X, Y)$ have been established, via the work of many people, for all the 10 possible choices of (X, Y); see [**GW2**] for more details.

In some cases, the upper bound of $\Delta(X, Y)$ is reached only by a few manifolds. For example, it was shown in [**GW1**] that if $M(r_1)$ is annular and $M(r_2)$ is toroidal, then $\Delta(r_1, r_2) \leq 3$ unless M is one of three special manifolds, for which $\Delta(r_1, r_2)$ is 4 or 5; moreover, there are infinitely many manifolds which admit two such Dehn fillings with $\Delta(r_1, r_2) = 3$. Thus $\Delta(A, T) = 5$, but the "virtual bound" to be defined below is 3. Similarly for $\Delta(T, T)$, see [**Go**]. The main results of this paper are the following.

Theorem 0.1. *There are infinitely many hyperbolic manifolds M which admit two nonhyperbolic Dehn fillings $M(r_1)$ and $M(r_2)$, such that $M(r_1)$ is toroidal and annular, $M(r_2)$ is reducible and ∂-reducible, and $\Delta(r_1, r_2) = 2$.*

Theorem 0.2. *There are infinitely many hyperbolic manifolds M with two torus boundary components, each of which admits two reducible Dehn fillings $M(r_1), M(r_2)$, with $\Delta(r_1, r_2) = 1$.*

Theorem 0.3. *There are infinitely many hyperbolic manifolds M which admit two nonhyperbolic Dehn fillings $M(r_1)$ and $M(r_2)$, such that $M(r_1)$ is reducible, $M(r_2)$ is toroidal, and $\Delta(r_1, r_2) = 3$.*

These theorems follow immediately from Theorems 2.6, 3.6 and 4.2 below. Very few examples as in the theorems were known before. The only known example satisfying the conditions in Theorem 0.1 was found by Hayashi and Motegi [**HM**], and the only known example as in Theorem 0.2 was the one given by Gordon and Litherland [**GLi**].

Similar to $\Delta(X, Y)$, we define the *virtual bound* $\Delta_v(X, Y)$ of distances between type X and type Y Dehn fillings to be the maximal integer n such that there are infinitely many hyperbolic manifolds M which admit two Dehn fillings $M(r_1), M(r_2)$ of type X, Y respectively, with $\Delta(r_1, r_2) = n$. If no such infinite family exist, define $\Delta_v(X, Y) = 0$. Thus $\Delta_v(X, Y) \leq \Delta(X, Y)$. The above theorems and some known results determine the virtual bounds of distances between certain types of nonhyperbolic Dehn fillings. The following is a table of $\Delta_v(X, Y)$.

X \ Y	T	A	S	D
D	2	2	0	1
S	3	2	1	
A	3	3--5		
T	5			

Table 1.1. Virtual bound $\Delta_v(X, Y)$.

As we can see, except for $\Delta_v(A, A)$, all the other $\Delta(X, Y)$ have been completely determined. In the table, $\Delta_v(T, T)$ is determined by Gordon [**Go**], $\Delta_v(T, A)$ by Gordon and Wu [**GW1**]. The upper bounds of the other entries in Table 1.1 are the same as that in [**GW2**], and the lower bounds of them are determined by Theorem 0.1 for $\Delta_v(D, T)$, $\Delta_v(D, A)$, and $\Delta_v(S, A)$; by Theorem 0.3 for $\Delta_v(S, T)$; by Gabai [**Ga**] and Berge [**Be**] for $\Delta_v(D, D)$; by Gordon and Wu [**GW1**] for $\Delta_v(A, A)$; and by Gordon and Litherland [**GLi**] for $\Delta_v(S, S)$. Theorem 0.2 gives a stronger result about type S-S fillings, namely the manifolds can be chosen to have an extra torus boundary components. Also, it provides infinitely many examples of two essential planar surfaces in 3-manifolds with distinct boundary slopes, one of which has unbounded number of boundary components.

We would like to thank Cameron Gordon and John Luecke for some interesting discussion on this topic.

2. Toroidal/annular fillings and reducible/∂-reducible fillings.

In this section we prove Theorem 2.6, which shows that there are infinitely many hyperbolic manifolds which admit two Dehn fillings of distance two apart, one of which is toroidal and annular, and the other is reducible and ∂-reducible. Let $Y = S^2 \times I$. Consider the tangles ξ_p in Y as shown in Figure 2.1, where a rectangle labeled by an integer n denotes a rational tangle of slope $1/n$; in other words, it contains two vertical strings with n left hand half twists.

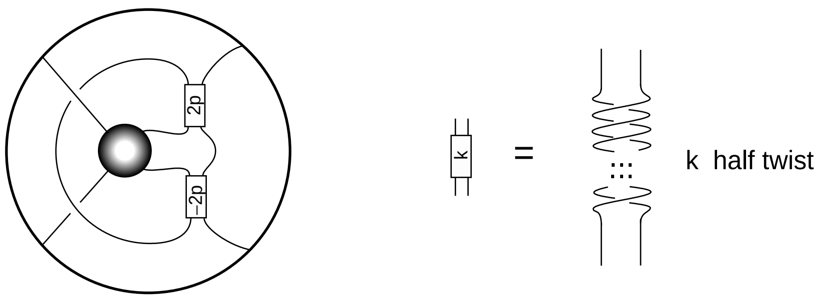


Figure 2.1.

Let $\xi_p(r)$ be the tangle obtained by filling the inside sphere S_0 of Y with a rational tangle of slope r. The tangles $\xi_p(r)$ are drawn in Figure 2.2(a)-(d) for $r = \infty, 0, -1, -1/2$, respectively. From the pictures we have the following lemma. We use $\xi(r,s)$ to denote a Montesinos tangle consisting of two rational tangles associated to the rational numbers r and s respectively. See [**Wu2**] or [**Mo1**, **Co**] for more details about Montesinos tangles and algebraic tangles.

Lemma 2.1. (1) $\xi_p(\infty)$ *is the connected sum of a trivial tangle and a Hopf link.*

(2) $\xi_p(0)$ *is the Montesinos tangle* $\xi[\frac{1}{2p-1}, \frac{-1}{2p+1}]$.

(3) $\xi_p(-1)$ *is the Montesinos tangle* $\xi[\frac{1}{2p+1}, \frac{-1}{2p-1}]$.

(4) $\xi_p(\frac{-1}{2})$ *is an algebraic tangle obtained by summing a Montesinos tangle* $\xi[\frac{1}{2p}, \frac{-1}{2p}]$ *with a rational tangle* $\xi[\frac{1}{2}]$. *It is not a Montesinos tangle.*

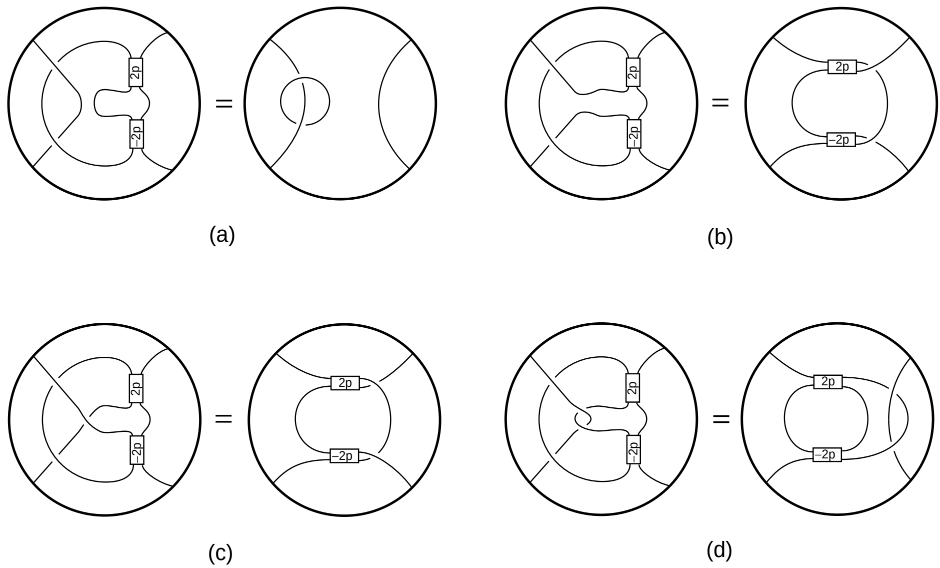


Figure 2.2.

Let M_p be the double branched covering of Y with branch set the tangle ξ_p. Then M_p is a compact orientable 3-manifold with boundary consisting of two tori T_0 and T_1, where T_0 is the lift of the inside sphere S_0. The ∞ and 0 slopes on S_0 lift to a meridian-longitude pair on T_0, with respect to which the Dehn filling manifold $M_p(r)$ is the double covering of the 3-ball branched along the tangle $\xi_p(r)$. See [**Mo2**] for more details. Denote by $Q(r,s)$ the double branched cover of a Montesinos tangle $\xi[\frac{1}{r}, \frac{1}{s}]$. Note that when $|r|, |s| > 1$, $Q(r,s)$ is a Seifert fiber space with orbifold $D(r,s)$, which by definition is a disk with two cone points of angle $2\pi/|r|$ and $2\pi/|s|$. Denote by $C(r,s)$ the cable space of type (r,s), that is, the exterior of a knot K in a solid torus V which is parallel to a curve on ∂V representing $rl + sm$ in $H_1(\partial V)$, where (m,l) is a meridian-longitude pair of ∂V. The above facts and Lemma 2.1 lead to the following lemma.

Lemma 2.2. *Suppose $p \geq 2$. The manifolds M_p have the following properties.*

(1) $M_p(\infty)$ *is the connected sum of a solid torus and the projective space* RP^3*;*
(2) $M_p(0) = Q(2p-1, -2p-1)$*;*
(3) $M_p(-1) = Q(2p+1, -2p+1)$*;*
(4) $M_p(-1/2)$ *is a non Seifert fibered graph manifold containing a unique essential torus T, cutting it into a cable space $C(2,1)$ and a Seifert fiber space $Q(2p,-2p)$.*

Proof. (1) follows from the fact that the double branched cover of the Hopf link is RP^3, and connected sum of links and tangles downstairs corresponds to connected sum of manifolds upstairs. (2) and (3) follow from the definition of $Q(r, s)$.

To prove (4), notice that the Conway sphere in $\xi_p(-1/2)$ cutting off the tangle $\xi(2p, -2p)$ lifts to an essential torus T upstairs, which cuts $M_p(-1/2)$ into $Q(2p, -2p)$ and $C(2, 1)$. Since $\xi_p(-1/2)$ is not a Montesinos tangle, the fibers on the two sides of T do not match. Seifert fibration on $C(2, 1)$ is unique, and since $p \geq 2$, the Seifert fibration on $Q(2p, -2p)$ is also unique [**Ja**, Theorem IV.18]. Therefore, $M_p(-1/2)$ is not a Seifert fiber space, so T is the Jaco-Shalen-Johannson decomposing torus because each side of it is a Seifert fiber space. Since each of $C(2, 1)$ and $Q(2p, -2p)$ are atoroidal, T is the unique essential torus in $M_p(-1/2)$. □

Note that when $p = 1$, $M_p(0)$ and $M_p(-1)$ are solid tori. Also, $M_p(-1/2)$ is a Seifert fiber space with orbifold a Mobiüs band with a cone point of angle $\pi/2$, so the conclusion of (4) is not true for $p = 1$. Thus the argument below will fail in this case. Actually, one can see that ξ_1 contains a nontrivial Conway sphere, so the manifold M_1 is toroidal.

In the following, we will assume $M = M_p$ and $p \geq 2$, and show that M is hyperbolic. Since M has toroidal boundary, by [**Th**] we need only show that M is irreducible, ∂-irreducible, non Seifert fibered, and atoroidal.

Lemma 2.3. *If $p \geq 2$, then M is irreducible, ∂-irreducible, and non Seifert fibered.*

Proof. If M is reducible, let S be a reducing sphere. S is separating, otherwise it would be a reducing sphere in all $M(r)$, contradicting Lemma 2.2(2). Let W, W' be the two components of M cut along S, with W the one containing T_0. Let $\widehat{W}'$ be W' with S capped off by a 3-ball. Since $M(0)$ is the Seifert fiber space $Q(2p-1, -2p-1)$, which is irreducible, $W(0)$ must be a 3-ball, so $\widehat{W}' = M(0) = Q(2p-1, -2p-1)$. But then we have

$$M(\infty) = \widehat{W}' \# \widehat{W}(\infty) = Q(2p-1, -2p-1) \# \widehat{W}(\infty) \neq (S^1 \times D^2) \# RP^3,$$

which is a contradiction. Therefore M is irreducible.

If M is ∂-reducible, then after ∂-compression one of the T_i becomes a sphere separating the two components of ∂M, hence is a reducing sphere, contradicting the above conclusion.

If M is Seifert fibered, then $M(r)$ is Seifert fibered for all but at most one r, for which $M(r)$ is reducible. Since $M(-1/2)$ is irreducible and is not a Seifert fiber space, this is not possible. □

Lemma 2.4. *Suppose T is an essential separating torus in an irreducible 3-manifold M, and suppose it is compressible in $M(r_1), M(r_2)$ with $\Delta(r_1, r_2) \geq 2$, where r_i are slopes on $T_0 \subset \partial M$. Then T and T_0 bound a cable space in M, with cabling slope r_0 satisfying $\Delta(r_0, r_i) = 1$, $i = 1, 2$.*

Proof. Cut M along T and let X be the component containing T_0. Then T is compressible in $X(r_i)$ and $\Delta(r_1, r_2) \geq 2$, so by [**Wu1**, Theorem 1] there is an essential annulus A in X with one boundary on T and the other on T_0, with slope r_0, say. Since T is essential in M, it is not parallel to T_0, so by [**CGLS**, Theorem 2.4.3] T is compressible in $X(r)$ only if $\Delta(r_0, r) \leq 1$. We must have $\Delta(r_0, r_i) = 1$, because if $r_0 = r_1$ then we would have $\Delta(r_0, r_2) = \Delta(r_1, r_2) = 2$, a contradiction. Now the manifold $X(r_i)$ is homeomorphic to the manifold Y obtained by cutting X along A, so the torus component of ∂Y corresponding to T under the homeomorphism is compressible in Y. Since M is irreducible, this implies that Y is a solid torus. It follows that X is a cable space with cabling slope r_0. □

Lemma 2.5. *M is atoroidal.*

Proof. Assuming the contrary, let T be an essential torus in M. Then T must be separating, otherwise $M(r)$ would contain a nonseparating torus or, if T becomes compressible in $M(r)$, a nonseparating sphere, for all r, which contradicts Lemma 2.2(1).

Let W, W' be the two components of M cut along T, with W the one containing T_0. Since M contains no nonseparating essential torus, by the Haken finiteness theorem (cf. [**Ja**, Page 49]), we may choose T to be outermost in the sense that W' contains no essential torus.

Claim. T is compressible in $M(-1/2)$.

Recall from Lemma 2.2(4) that $M(-1/2)$ has a unique essential torus T'. So if T is incompressible in $M(-1/2)$ then either it is boundary parallel or it is isotopic to T'. The first case is impossible, because then $M(-1/2) = W(-1/2) \cup W' = (T \times I) \cup W' \cong W'$, so T' would be an essential torus in W', contradicting the choice of T. Therefore T must be isotopic to T' in $M(-1/2)$. It follows that either $W' = C(2,1)$, or $W' = Q(2p, -2p)$.

Since $M(0)$ is atoroidal, either T is boundary parallel in $M(0)$ or it is compressible in $M(0)$. In the first case we would have $Q(2p-1, -2p-1) = M(0) = W' = C(2,1)$ or $Q(2p, -2p)$, which is absurd. In the second case let D be a compressing disk of T in $W(0)$, and let $\widehat{W}'$ be the manifold obtained by capping off the sphere boundary component of $W' \cup N(D)$ with a 3-ball. Then $\widehat{W}'$ is a summand of $M(0) = Q(2p-1, -2p-1)$, so either $\widehat{W}' = Q(2p-1, -2p-1)$ or $\widehat{W}' = S^3$. However, this is impossible whether $W' = C(2,1)$ or $W' = Q(2p, -2p)$ because $\widehat{W}'$ is obtained from W' by Dehn filling on T along certain slope, and it is easily seen that when $p \geq 2$ none of the Dehn fillings on such W' could produce $Q(2p-1, -2p-1)$ or S^3. This completes the proof of the claim.

Since $M(\infty)$ contains no incompressible torus, T is compressible in $M(\infty)$. By the claim above, T is also compressible in $M(-1/2)$. Since $\Delta(\infty, -1/2) = 2$, it follows from Lemma 2.4 that W is a cable space $C(p, q)$ with cabling slope r_0 satisfying $\Delta(r_0, \infty) = \Delta(r_0, -1/2) = 1$. Solving these equalities, we have $r_0 = 0$ or -1. Now we have $W(r_0) = L(p,q)\#(S^1 \times D^2)$, so $M(r_0)$ should have a lens space summand. On the other hand, we have shown that $r_0 = 0$ or -1, and in either case by Lemma 2.2 $M(r_0)$ is a prime manifold with torus boundary. This contradiction completes the proof that M is atoroidal. □

Theorem 2.6. *The manifolds M_p, $p \geq 2$, are mutually distinct hyperbolic manifolds, each admitting two nonhyperbolic Dehn fillings $M(r_1)$ and $M(r_2)$, such that $M(r_1)$ is toroidal and annular, $M(r_2)$ is reducible and ∂-reducible, and $\Delta(r_1, r_2) = 2$.*

Proof. Consider the manifold M_p which is the double cover of $Y = S^2 \times I$ branched along the tangle ξ_p in Figure 2.1. By Lemmas 2.3 and 2.5, M_p are hyperbolic for all $p \geq 2$. By Lemma 2.2, $M_p(\infty)$ is reducible and ∂-reducible, and $M_p(-1/2)$ is the union of $C(2,1)$ and $Q(2p, -2p)$ along a torus, hence is toroidal and annular because there is an essential annulus in $C(2,1)$ with both boundary components on the outside torus T_1. Since $\Delta(\infty, -1/2) = 2$, M_p satisfy all the conditions of the theorem. It remains to show that M_p and M_q are non homeomorphic when $p, q \geq 2$ and $p \neq q$.

Let T_0 (resp. T_0') be the torus of ∂M_p (resp. ∂M_q) on which the Dehn fillings are performed. Let (m, l) (resp. (m', l')) be the meridian-longitude pair on T (resp. T') chosen as in Lemma 2.2. Let $f : M_p \to M_q$ be a homeomorphism.

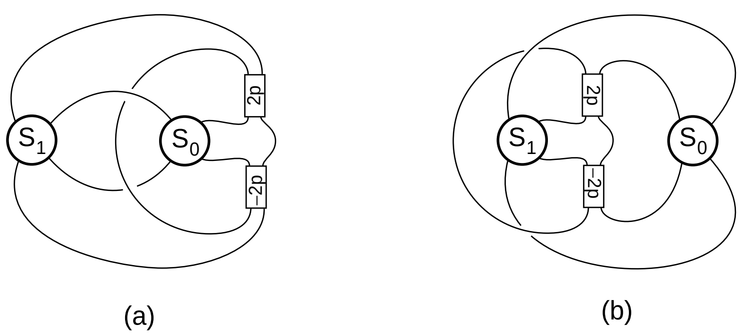


Figure 2.3.

There is a homeomorphism of Y interchanging the two sphere boundary components, and leaving ξ_p invariant, which induces a self homeomorphism of M_p interchanging the two boundary components. This can be seen by redrawing the tangle in Figure 2.1 as in Figure 2.3(a), where the sphere S_0 represents the inside sphere in Figure 2.1, and S_1 the outside sphere. After an isotopy the picture becomes that in Figure 2.3(b). (Note that the isotopy have changed the position of the endpoints of the tangle on the spheres, but that does not matter.) Now blow up the sphere S_0, we get the same picture as that in Figure 2.1, with S_0 and S_1 interchanged. Thus without loss of generality we may assume that f maps T_0 to T_0'.

Since $M_p(\infty)$ is ∂-reducible, by [**Sch**] $M_p(r)$ is irreducible for all $r \neq \infty$. Hence the reducing slope ∞ is unique, so f must send m to m'. Assume $f(l) = l' + km'$. Because of uniqueness of Seifert fibration, neither of $M_p(0)$ or $M_p(-1)$ is homeomorphic to $M_q(0)$ or $M_q(-1)$ when $p, q \geq 2$ and $p \neq q$. Hence $k \neq 0, \pm 1$. Now f sends the slope $-1/2$ to $(2k-1)/2$, so both $M_q(-1/2)$ and $M_q((2k-1)/2)$ are toroidal. We have $\Delta(-1/2, (2k-1)/2) = |4k| \geq 8$. On the other hand, by [**Go**], this happens only if M_q is the Figure 8 knot complement or the Whitehead sister link complement. Since M_q have two boundary components, this is impossible. □

3. Manifolds admitting two reducible Dehn fillings.

In this section we will show that there are infinitely many hyperbolic manifolds with two torus boundary components, each admitting two reducible Dehn fillings. Consider the tangles ξ_p in $Y = S^2 \times I$ as shown in Figure 3.1, where, as in Figure 2.1, a rectangle labeled by an integer n denotes a rational tangle of slope $1/n$.

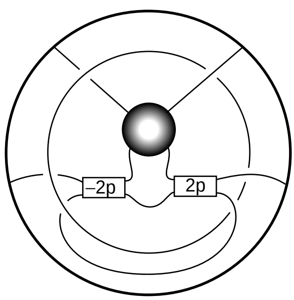


Figure 3.1.

As in Section 2, we denote by M_p the double branched cover of Y branched along ξ_p, and by $\xi_p(r)$ the tangle obtained by filling the inside sphere S_0 with a rational tangle of slope r. Then the Dehn filling manifold $M_p(r)$ is the double cover of Y branched along $\xi_p(r)$. The tangles $\xi_p(\infty)$ and $\xi_p(0)$ are drawn in Figure 3.2(a)–(b). We can see that $\xi_p(\infty)$ is the connected sum of $\xi(1/2, -1/2)$ and a Hopf link, while $\xi_p(0)$ is the connected sum of a Montesinos tangle $\xi(1/2p, -1/2p)$ and a Hopf link. Recall that $Q(r, s)$ denotes the Seifert fiber space which double branch covers the tangle $\xi(1/r, 1/s)$, and the double branched cover of a Hopf link is the projective space RP^3. Therefore we have the following lemma.

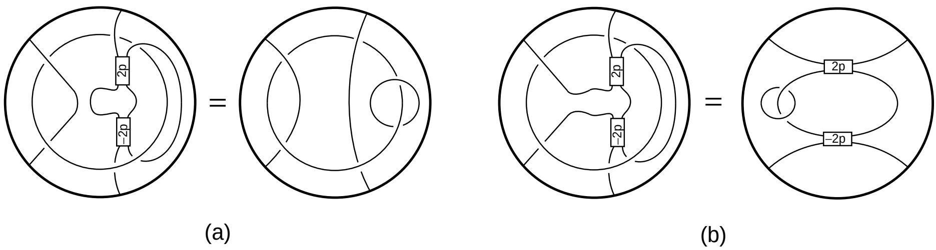


Figure 3.2.

Lemma 3.1. *The manifolds M_p, $p \neq 0$, have the following properties.*
(1) $M_p(\infty) = Q(2,-2)\#RP^3$;
(2) $M_p(0) = Q(2p,-2p)\#RP^3$.

Thus each M_p admits two reducible Dehn fillings. In what follows, we will assume $M = M_p$, and $p \geq 2$. We need to show that M is hyperbolic. Let T_0 be the component of ∂M on which the Dehn fillings are performed. Thus T_0 covers the inside sphere S_0 in Figure 3.1. Let T_1 be the component of ∂M covering the outside sphere S_1.

Lemma 3.2. *M is irreducible.*

Proof. Assuming the contrary, let S be a reducing sphere of M. Clearly S is separating, otherwise $M(0)$ would contain a nonseparating reducing sphere, contradicting Lemma 3.1. Let W, W' be the components of M cut along S, with W the one containing T_0. Denote by $\widehat{W}$ the manifold W with the sphere boundary capped off by a 3-ball. Similarly for $\widehat{W}'$. Then $\widehat{W}'$ is a summand of both $M(0)$ and $M(\infty)$, so by Lemma 3.1 we must have $\widehat{W}' = RP^3$. This also shows that the reducing sphere in M is unique up to isotopy, because if S and S' bound different punctured RP^3, then tubing them together would give a sphere which does not bound a punctured RP^3.

Let ρ be the involution of M which induces the branched covering. Since the reducing sphere S is unique up to isotopy, by the equivariant sphere theorem [**MSY**], it can be chosen to be invariant under the involution ρ, hence it double branch covers a sphere S' in the manifold Y downstairs, which must cut off a 3-ball B because one side of S is W', which does not contain the preimage of S_0 or S_1. Extending the involution $\rho|_S$ trivially over a 3-ball D, we get a double branched cover $\widehat{W}' \to S^3 = B \cup D'$, with branch set L the union of $\xi' = \xi_p \cap B$ and a trivial arc in the attached 3-ball D', which is the image of D under the branched covering map. Since $\widehat{W}' = RP^3 = L(2,1)$, the link L is the 2-bridge link associated to the number $1/2$, which is the Hopf link. Therefore, $\xi' = \xi_p \cap B$ is a tangle in B consisting of an unknotted arc and a trivial circle C around it.

We want to shown that no such pair (B, ξ') exists in (Y, ξ_p). Assuming the contrary, then (B, ξ') would remain the same after filling the sphere boundaries S_0, S_1 of Y with any rational tangles. The tangle ξ_p has two circle components C_1, C_2, where C_1 denotes the one on the left in Figure 3.1. The circle component C of ξ' must be one of the C_i. However, after filling both S_i with 0-tangle, C_2 has linking number $p \geq 2$ with one of the components of the resulting link, while after filling S_0 with 1-tangle and S_1 with ∞-tangle the circle C_1 has linking number 2 with one of the components of the resulting link, either case contradicting the fact that C bounds a disk in B intersecting the resulting link only once. □

Lemma 3.3. *M is ∂-irreducible, and is not a Seifert fiber space.*

Proof. Since ∂M consists of two tori, M being ∂-reducible would imply that it is reducible, which would contradict Lemma 3.2. If M is Seifert fibered (with two torus boundary components), then $M(r)$ would be reducible for at most one r, which would contradict Lemma 3.1. □

Lemma 3.4. *Let X be an irreducible and ∂-irreducible 3-manifold. If both $X(r_1)$ and $X(r_2)$ are reducible and ∂-reducible, then $r_1 = r_2$.*

Proof. Let T_0 be the Dehn filling component of ∂X. Assume $r_1 \neq r_2$. Since $X(r_1)$ is ∂-reducible and $X(r_2)$ is reducible, by Scharlemann's theorem [**Sch**, Theorem 6.1], r_2 is a cabling slope, so there is an essential annulus A_2 in X with boundary two copies of r_2 of opposite orientations. Similarly, we have an essential annulus A_1 in X with boundary consisting of two copies of r_1 of opposite orientations. Isotope A_1 to intersect A_2 essentially. Then $A_1 \cap A_2$ consists of essential arcs on A_i, running from one boundary component to the other. By the parity rule on [**CGLS**, Page 279], if an arc component of $A_1 \cap A_2$ connects two components of ∂A_1 which have opposite orientations on T_0, then it must connect two components of ∂A_2 with the same orientation on T_0. This is a contradiction because the two boundary components of each A_i have opposite orientations on T_0. □

Lemma 3.5. *M is atoroidal.*

Proof. Consider an essential torus T in M. Clearly T is separating, otherwise $M(0)$ would contain a nonseparating torus or sphere, which would contradict Lemma 3.1. Let W, W' be the two components of M cut along T, where W contains T_0. Note that T cannot be boundary parallel in $M(0)$ or $M(\infty)$, otherwise W', and hence M, would be reducible, which would contradict Lemma 3.2. Hence T is compressible in both $W(0)$ and $W(\infty)$ because by Lemma 3.1 they are atoroidal. After compression, T becomes a sphere in $W(0)$ and $W(\infty)$, so if W contained T_1, then both $W(0)$ and $W(\infty)$ would also be reducible, which is impossible by Lemma 3.4. Hence we conclude that any essential torus in M must separate the two boundary components of M.

Let $\rho : M \to M$ be the involution which induces the branch covering, and let X be the fixed point set of ρ. Then X covers the tangle ξ_p in the manifold Y downstairs. Since ξ_p contains four arcs running from S_0 to S_1, X has four arcs running from T_0 to T_1, hence each essential torus T intersect X at least four times.

By the equivariant torus theorem [**MS**, Theorem 8.6], there is a set of essential tori $\mathcal{T}$ in M such that $\rho(\mathcal{T}) = \mathcal{T}$. Let T be a component of $\mathcal{T}$. Since X intersects T in at least four points, we must have $\rho(T) = T$. Calculating the Euler number of T/ρ, we see that X cannot intersect T in more than four points. Hence T intersects X exactly four times, and $S = T/\rho$ is a sphere in Y which intersects each of the four arc components of ξ_p exactly once, and is disjoint from the circle components of ξ_p. Since the two circle components of ξ_p have linking number 1, they must lie on the same side of S.

Let Y_1, Y_2 be the two components of Y cut along S, with Y_1 the one disjoint from the circle components of ξ_p. Let W_1, W_2 be the components of M cut along T, with W_i covering Y_i. Consider the tangle ξ'_p consisting of the arc components of ξ_p. Let M' be the double cover of Y branched along ξ'_p, let T' be the torus in M' that covers S, and let W'_i be the part of M' that covers Y_i. It can be seen from Figure 3.1 that ξ'_p is isotopic to four straight arcs running from S_0 to S_1; hence $M' = T^2 \times I$. Since T' is a torus separating the two components of $\partial M'$, it is isotopic to a horizontal torus $T^2 \times x$, so each W'_i is also homeomorphic to $T^2 \times I$. Now we have $\xi_p \cap Y_1 = \xi'_p \cap Y_1$, therefore W_1, as the double cover of Y_1 branched along $\xi_p \cap Y_1$, is the same as W'_1, hence is a product $T^2 \times I$. But then T is boundary parallel, contradicting the assumption that T is an essential torus in M. □

Theorem 3.6. *The manifolds M_p, $p \geq 2$, are distinct hyperbolic manifolds, each admitting two reducible Dehn fillings $M(r_1), M(r_2)$ with $\Delta(r_1, r_2) = 1$.*

Proof. We have shown in Lemmas 3.1-3.5 that M_p are hyperbolic manifolds admitting two reducible Dehn fillings $M_p(0)$ and $M_p(\infty)$, so it remains to show that the manifolds are all different.

Suppose $f : M_p \to M_q$ is a homeomorphism, $p > q \geq 2$. As in the proof of Theorem 2.6, it is easy to see that there is a self homeomorphism of M_p interchanging the two boundary components, hence we may assume that f maps T_0 to T'_0, where T'_0 and T'_1 are the boundary tori of M_q, with T'_0 the one covering the inside sphere.

By [**GLu1**], M_i admits at most three reducible Dehn fillings, with mutual distance 1. Since $M_p(0) = Q(2p, -2p) \# RP^3$ is homeomorphic to neither $M_q(0)$ nor $M_q(\infty)$, f maps the slope 0 to another reducing slope of M_q, which must be ± 1 because it has distance 1 from 0 and ∞. Thus the only reducible Dehn filling of M_q homeomorphic to $M_p(\infty)$ is $M_q(\infty)$, so f sends the ∞ slope on T_0 to ∞ on T'_0. Similarly, it sends the ∞ slope on T_1 to ∞ on T'_1. Denote by $M_p(r, s)$ the manifold obtained by r filling on T_0 and s filling on T_1. Then we have $M_p(0, \infty) = M_q(\pm 1, \infty)$.

The manifold $M_k(r,s)$ is a double cover of $\xi_k(r,s)$, which is obtained from ξ_k by filling the inside sphere with a rational tangle of slope r and the outside sphere with one of slope s. One can check that $\xi_p(0,\infty)$ is the split union of a Hopf link and a trivial knot, while $\xi_q(\pm 1,\infty)$ is the connected sum of a Hopf link and a 2-bridge link associated to the rational number $\pm\frac{1}{4}$. Thus $M_p(0,\infty) = S^1 \times S^2 \# RP^3$, and $M_q(\pm 1,\infty) = L(4,\pm 1)\# RP^3$. Since these two manifolds are not homeomorphic, this is a contradiction. □

4. Reducible and toroidal fillings.

In this section we show that there are infinitely many hyperbolic manifolds which admit a reducible filling and a toroidal filling of distance 3 apart. Consider the tangles ξ_p ($p \geq 3$) in Y, as shown in Figure 4.1(a), where Y is the 3-ball obtained by deleting the interior of the 3-ball B in the figure from S^3. As before, let $\xi(r)$ be the union of (Y,ξ_p) with a rational tangle of slope r, and let $M_p(r)$ be the double branched cover of S^3 branched along $\xi_p(r)$.

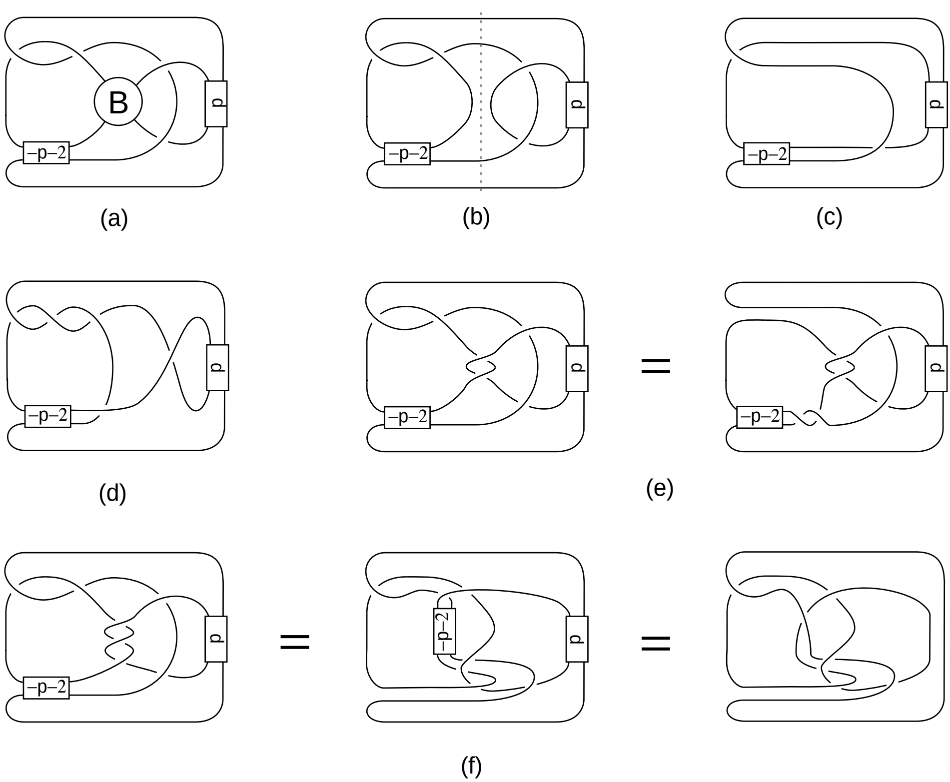


Figure 4.1.

Lemma 4.1. *The manifold M_p admits the following Dehn fillings.*

(1) $M_p(\infty)$ *is a non Seifert fibered, irreducible, toroidal manifold;*
(2) $M_p(0)$ *is a lens space* $L((p-1)(p+3)+1, p+3)$*;*
(3) $M_p(1)$ *and* $M_p(1/2)$ *are small Seifert fibered manifolds, but not lens spaces;*
(4) $M_p(1/3) = L(3,1)\#L(2,1)$.

Proof. The tangles $\xi(\infty), \xi(0), \xi(1), \xi(1/2), \xi(1/3)$ are shown in Figure 4.1(b)-(f), respectively. We can see that $\xi(\infty)$ is the union of $\xi[\frac{1}{2}, \frac{1}{-(p+2)}]$ and $\xi[\frac{1}{2}, \frac{1}{p}]$, and is not a Montesinos link; $\xi(0)$ is a 2-bridge link associated to the rational number $1/((p-1)+1/(p+3)) = (p+3)/((p+3)(p-1)+1)$; $\xi(1)$ and $\xi(1/2)$ are Montesinos links consisting of three rational tangles; and $\xi(1/3)$ is the connected sum of a trefoil knot and a Hopf link. The result now follows by taking the double cover of S^3 branched along the corresponding links. Note that $p \geq 3$ guarantees that the Seifert fibrations on the two sides of the essential torus in $M_p(\infty)$ are unique, which can be used to show that $M_p(\infty)$ is not a Seifert fiber space. See the proof of Lemma 2.2. □

Theorem 4.2. *The manifolds $M = M_p$, $p \geq 3$, are mutually distinct hyperbolic manifolds, each admitting two Dehn fillings $M(r_1)$ and $M(r_2)$, such that $M(r_1)$ is reducible, $M(r_2)$ is toroidal, and $\Delta(r_1, r_2) = 3$.*

Proof. Let $r_1 = 1/3$, and $r_2 = \infty$. Then $\Delta(r_1, r_2) = 3$, and by Lemma 4.1, $M(r_1)$ is reducible, $M(r_2)$ is toroidal. We need to show that M_p are hyperbolic and mutually distinct.

M is irreducible, otherwise a closed summand would survive after all Dehn fillings; but since $M(0)$ and $M(1)$ are non homeomorphic prime manifolds, this is impossible. M is not a Seifert fiber space because two Dehn fillings $M(\infty)$ and $M(1/3)$ are non Seifert fibered. These imply that M is ∂-irreducible. To prove M is hyperbolic, it remains to show that M is atoroidal.

If T is an essential torus in M, then it is compressible in $M(0)$, $M(1)$, $M(1/2)$ and $M(1/3)$. Since $M(0)$ is irreducible, T must be separating. Let W, W' be the components of M cut along T, with W the one containing T_0. Since $\Delta(1, 1/3) = 2$, by Lemma 2.4, W is a cable space $C(r, s)$, with cabling slope r_0 satisfying $\Delta(r_0, 1) = \Delta(r_0, 1/3) = 1$. Solving these equalities, we have $r_0 = 0$ or $1/2$; but since $M(r_0)$ contains a lens space $L(r, s)$, we must have $r_0 = 0$.

Let δ_0 and δ_1 be the slopes on T which bound disks in $W(0)$ and $W(1/3)$, respectively. Since 0 is the cabling slope, we have $\Delta(\delta_0, \delta_1) = |r| > 1$. Now $W(0)$ is the connected sum of a solid torus and $L(r, s)$, while $W(1/3)$ is a solid torus, so we have

$$M(0) = L(r,s)\#W'(\delta_0),$$
$$M(1/3) = W'(\delta_1).$$

Comparing the first equation with Lemma 4.1(2), we see that W' is the exterior of a knot in S^3 with δ_0 the meridional slope. But then since $\Delta(\delta_0, \delta_1) > 1$, by [**GLu2**] the manifold $M(1/3)$ would be irreducible, which would contradict Lemma 4.1(4). This completes the proof that M is atoroidal, and hence hyperbolic.

It remains to show that the manifolds M_p are mutually distinct. Assume there is a homeomorphism $f : M_p \cong M_q$, $p > q \geq 3$. Let (m, l) and (m', l') be the meridian-longitude pair of M_p and M_q, respectively. By [**CGLS**], [**GLu1**] and [**BZ**, Theorem 0.1], a hyperbolic manifold admits a total of at most three reducible or cyclic Dehn fillings, with mutual distance 1. Thus two of the four slopes $0, 1/3, f(0), f(1/3)$ on ∂M_q must be the same. But since $M_p(0)$ is not homeomorphic to $M_q(0)$ or $M_q(1/3)$, we must have $f(1/3) = 1/3$, and $f(0)$ is of distance 1 from 0 and $1/3$, so $f(0) = 1/2$ or $1/4$. The first is impossible because $M_q(1/2)$ is not a lens space. Hence $f(0) = 1/4$. Now $f(m) = f((m+3l) - 3l) = (m' + 3l') \pm 3(m' + 4l')$, and we have $\Delta(m', f(m)) \geq 9$. Since both m' and $f(m)$ are toroidal Dehn filling slopes on ∂M_q, this contradicts [**Go**]. □

References

[Be] J. Berge, *The knots in $D^2 \times S^1$ with nontrivial Dehn surgery yielding $D^2 \times S^1$*, Topology Appl., **38** (1991), 1-19.

[BZ] S. Boyer and X. Zhang, *The semi norm and Dehn filling*, preprint.

[Co] J. Conway, *An enumeration of knots and links, and some of their algebraic properties*, in 'Computational problems in abstract algebra', 329-358, New York and Oxford, Pergamon, 1970.

[CGLS] M. Culler, C. Gordon, J. Luecke and P. Shalen, *Dehn surgery on knots*, Annals Math., **125** (1987), 237-300.

[Ga] D. Gabai, *Surgery on knots in solid tori*, Topology, **28** (1989), 1-6.

[Go] C. Gordon, *Boundary slopes of punctured tori in 3-manifolds*, Trans. Amer. Math. Soc., **350** (1998), 1713-1790.

[GLi] C. Gordon and R. Litherland, *Incompressible planar surfaces in 3-manifolds*, Topology Appl., **18** (1984), 121-144.

[GLu1] C. Gordon and J. Luecke, *Reducible manifolds and Dehn surgery*, Topology, **35** (1996), 385-409.

[GLu2] ———, *Only integral Dehn surgeries can yield reducible manifolds*, Math. Proc. Camb. Phil. Soc., **102** (1987), 97-101.

[GW1] C. Gordon and Y.-Q. Wu, *Toroidal and annular Dehn fillings*, Proc. London Math. Soc., to appear.

[GW2] ———, *Annular and boundary reducing Dehn fillings*, preprint.

[HM] C. Hayashi and K. Motegi, *Dehn surgery on knots in solid tori creating essential annuli*, Trans. Amer. Math. Soc., **349** (1997), 4897-4930.

[Ja] W. Jaco, *Three manifold topology*, Regional Conference Series in Math., **43** (1980).

[MS] W. Meeks and P. Scott, *Finite group actions on* 3*-manifolds*, Invent. Math., **86** (1986), 287-346.

[MSY] W. Meeks, L. Simon and S.-T. Yau, *Embedded minimal surfaces, exotic spheres, and manifolds with positive Ricci curvature*, Annals Math., **116** (1982), 621-659.

[Mo1] J. Montesinos, *Una familia infinita de nudos representados no separables*, Revista Math. Hisp.-Amer., **33** (1973), 32-35.

[Mo2] ———, *Surgery on links and double branched covers of* S^3, Annals of Math. Studies, **84** (1975), 227-260.

[Sch] M. Scharlemann, *Producing reducible* 3*-manifolds by surgery on a knot*, Topology, **29** (1990), 481-500.

[Th] W. Thurston, *Three dimensional manifolds, Kleinian groups and hyperbolic geometry*, Bull. Amer. Math. Soc., **6** (1982), 357-381.

[Wu1] Y.-Q. Wu, *Incompressibility of surfaces in surgered* 3*-manifolds*, Topology, **31** (1992), 271-279.

[Wu2] ———, *The classification of nonsimple algebraic tangles*, Math. Ann., **304** (1996), 457-480.

Received September 15, 1997 and revised February 2, 1998. The second author's research at MSRI was supported in part by NSF grant #DMS 9022140.

UNAM
Ciudad Universitaria
04510 Mexico D.F.
Mexico
E-mail address: eudave@servidor.unam.mx

University of Iowa
Iowa City, IA 52242
E-mail address: wu@math.uiowa.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 190, No. 2, 1999

TEST WORDS, GENERIC ELEMENTS AND ALMOST PRIMITIVITY

B. Fine, G. Rosenberger, D. Spellman, and M. Stille

A test element in a group G is an element g with the property that if $f(g) = g$ for an endomorphism f of G to G then f must be an automorphism. A test element in a free group is called a test word. Nielsen gave the first example of a test word by showing that in the free group on x, y the commutator $[x, y]$ satisfies this property. T. Turner recently characterized test words as those elements of a free group contained in no proper retract. Since free factors are retracts, test words are therefore very strong forms of non-primitive elements. In this paper we give some new examples of test words and examine the relationship between test elements and several other concepts, in particular generic elements and almost-primitive elements (APE's). In particular we show that an almost primitive element which lies in a certain type of verbal subgroup must be a test word. Further using a theorem of Rosenberger on equations in free products we prove a result on APE's, generic elements and test words in certain free products of free groups. Finally we examine test elements in non-free groups and introduce the concept of the test rank of a group.

1. Introduction.

A **test element** in a group G is an element g with the property that if $f(g) = g$ for an endomorphism f of G to G then f must be an automorphism. A test element in a free group is called a **test word**. Nielsen [**N**] gave the first non-trivial example of a test word by showing that in the free group on x, y the commutator $[x, y]$ satisfies this property. T. Turner [**T**] recently characterized test words as those elements of a free group which do not lie in any proper retract. Using this characterization he was able to give several straightforward criteria to determine if a given element of a free group is a test word. Using these criteria, Comerford [**C**] proved that it is effectively decidable whether elements of free groups are test words. Since free factors are retracts, Turner's result implies that no test word can fall in a proper free factor. Therefore being a test word is a very strong form of non-primitivity.

In this paper we consider relationships between test words and two related concepts — **almost primitive elements** (APE's) and **generic elements**.

We give the formal definitions in the next section where we also prove that an almost primitive element of a free group which lies in a certain type of verbal subgroup must be a test word (Theorem 1). This is quite surprising given the strong non-primitivity of test words. In Section 3 we use a theorem of Rosenberger [**R1**] on equations in free products to prove a result on APE's, generic elements and test words in certain free products of free groups. In Section 4, using Nielsen transformations, we produce a set of generic elements in the free group of rank two. Using the theorem of Rosenberger mentioned above, these examples can be extended to finding generic elements in higher rank free groups. Finally in Section 5 we give some straightforward results on extensions of these concepts to arbitrary non-free groups. As pointed out by Turner the characterization of test elements in general is more subtle and difficult than in the free group case.

We note that a few of the results appear in the Diplomarbeit of N. Isermann [**I**] however the proofs given here are somewhat different.

2. Test Words, Almost Primitive Elements and Generic Elements.

A **test element** in a group G is an element g with the property that if $f(g) = g$ for an endomorphism f of G to G then f must be an automorphism. A test element in a free group is called a **test word**. Nielsen [**N**] gave the first non-trivial example of a test word by showing that in the free group on x, y the commutator $[x, y]$ satisfies this property. Other examples of test words have been given by Zieschang [**Z1**, **Z2**], Rosenberger [**R1**, **R2**, **R3**] Kalia and Rosenberger [**K-R**], Hill and Pride [**H-P**] and Durnev [**D**]. Gupta and Shpilrain [**G-S**] have studied the question as to whether the commutator $[x, y]$ is a test element in various quotients of the free group on x, y.

Recall that a subgroup H of a group G is a **retract** if there exists a homomorphism $f : G \to H$ which is the identity on H. Clearly in a free group F any free factor is a retract. However there do exist retracts in free groups which are not free factors. Recently T. Turner [**T**] characterized test words as those elements of a free group which do not lie in any proper retract. Using this characterization he was able to give several straightforward criteria to determine if a given element of a free group is a test word. Using these criteria, Comerford [**C**] proved that it is effectively decidable whether elements of free groups are test words. Since free factors are retracts, Turner's result implies that no test word can fall in a proper free factor. Therefore being a test word is a very strong form of non-primitivity. Shpilrain [**S1**, **S2**] defined the **rank** of an element w in a free group F as the smallest rank of a free factor containing w. Clearly in a free group of rank n a test word has maximal rank n. Shpilrain conjectured that the converse was also true but Turner gave an example showing this to be false. However

Turner also proved that Shpilrain's conjecture is true if only test words for monomorphisms are considered.

As a direct consequence of the characterization Turner obtains the following result [**T**, Example 5] which shows that there is a fairly extensive collection of test words in a free group of rank two.

Proposition 1 ([**T**]). *In a free group of rank two any non-trivial element of the commutator subgroup is a test word.*

Proof. Let F be a free group of rank two and suppose H is a proper retract. Then the rank of H must be one and hence H is abelian. Suppose $g \in F'$ the commutator subgroup of F. If $g \in H$ then there exists an endomorphism $f : F \to H$ which is the identity on H. Therefore $f(g) = g$. But $f(g) = 1$ if f is any homomorphism of F into an abelian group. Therefore $g = 1$. It follows that no non-trivial element of F' can lie in any proper retract and therefore by Turner's characterization must be a test word. □

An **almost primitive element** - (APE) - is an element of a free group F which is not primitive in F but which is primitive in any proper subgroup of F containing it. This can be extended to arbitrary groups in the following manner. An element $g \in G$ is **primitive** in G if g generates an infinite cyclic free factor of G, that is g has infinite order and $G = \langle g \rangle \star G_1$ for some $G_1 \subset G$. g is then an APE if it is not primitive in G but primitive in any proper subgroup containing it. Rosenberger [**R1**] proved that in the free group $F = F(x_i, y_i, z_j)$; $1 \leq i \leq m$, $1 \leq j \leq n$, of rank $2m + n$ the element

$$[x_1, y_1]...[x_m, y_m]z_1^{p_1} \dots z_n^{p_n}$$

where the p_i are not necessarily distinct primes, is an APE in F. Rosenberger [**R1**] proved, in a different setting that if A, B are arbitrary groups containing APE's a, b respectively, then the product ab is either primitive or an APE in the free product $A \star B$. This was reproved by Brunner, Burns and Oates-Williams [**B-B-O**] who also prove the more difficult result that if a and b are tame APE's in groups A, B respectively then their product normally is a tame APE in $A \star B$. An APE w in a group G ia a **tame APE** if whenever $w^\alpha \in H \subset G$ with $\alpha \geq 1$ minimal, then either w^α is primitive in H or the index $[G : H]$ is α. It follows easily that $[a_1, b_1] \dots [a_g, b_g]$, $g \geq 1$, is a tame APE in the free group on $a_1, b_1 \dots a_g, b_g$, (see [**R3**]). We note that Brunner, Burns and Oates-Williams give a more technical definition of a tame APE.

Let $\mathcal{U}$ be a variety defined by a set of laws $\mathcal{V}$. (We refer to the book of H. Neumann [**Ne**] for relevant terminology.) For a group G we let $\mathcal{V}(G)$ denote the verbal subgroup of G defined by $\mathcal{V}$. An element $g \in G$ is $\mathcal{U}$**-generic** in G if $g \in \mathcal{V}(G)$ and whenever H is a group, $f : H \to G$ a homomorphism and $w = f(u)$ for some $u \in \mathcal{V}(H)$ it follows that f is surjective. Equivalently $g \in G$ is $\mathcal{U}$-generic in G if $g \in \mathcal{V}(G) \subset G$ but $g \notin \mathcal{V}(K)$ for every proper subgroup K of G [**St**]. An element is **generic** if it is $\mathcal{U}$-generic for some variety $\mathcal{U}$. Let $\mathcal{U}_n$ be the variety defined by the set of laws $\mathcal{V}_n = \{[x,y], z^n\}$. For $n = 0$ we have $\mathcal{U}_n = \mathcal{A}$ the abelian variety. Stallings [**St**] and Dold [**Do**] have given sufficient conditions for an element of a free group to be $\mathcal{U}_n$-generic. Using this it can be shown that $x_1^n x_2^n \dots x_m^n$ is $\mathcal{U}_n$-generic in the free group on $x_1, \dots, x_m$ for all $n \geq 2$ and if m is even $[x_1, x_2], \dots [x_{m-1}, x_m]$ is $\mathcal{U}_n$-generic in the free group on $x_1, \dots, x_m$ for $n = 0$ and for all $n \geq 2$. These facts are also consequences of a result of Rosenberger [**R2**, **R3**].

Comerford [**C**] points out that if G is Hopfian, which is the case if G is free, then being generic implies being a test word. Thus for free groups we have

$$\text{generic} \longrightarrow \text{test word} .$$

Comerford also shows that there is no converse. In particular he shows that in a free group of rank 3 on x, y, z the word $w = x^2[y^2, z]$ is a test word but is not generic. We can also show that in general, generic does not imply APE. Suppose $F = F(x,y)$ is the free group of rank two on x, y and let $w = x^4 y^4$. Then w is $\mathcal{U}_4$-generic but w is not an APE since $w \in \langle x^2, y^2 \rangle$ and is not primitive in this subgroup while this subgroup is not all of F.

Further, in general it is not true that being an APE implies being a test word. Again let $F = F(x,y)$ and let $w = x^2 y x^{-1} y^{-1}$. Brunner, Burns and Oates-Williams, after a private communication with G. Rosenberger, show that w is an APE. However Turner shows that w is not a test word. Since generic elements are test words in a Hopfian group this example shows further that APE does not imply generic in general. This is really to be expected since test words are strongly non-primitive. However our first result shows that many APE's are indeed generic and therefore test words.

Recall that a variety $\mathcal{U}$ defined by the set of laws $\mathcal{V}$ is a non-trivial variety if it consists of more than just the trivial group. In this case $\mathcal{V}(F) \neq F$ for any free group F.

Theorem 1. *Let F be a free group and $\mathcal{B}$ a non-trivial variety defined by the set of laws $\mathcal{V}$. Let $w \in \mathcal{V}(F)$. If w is an APE then w is $\mathcal{B}$-generic. In particular w is a test word.*

Proof. Let $w \in \mathcal{V}(F)$ be an APE and let $\phi : H \to F$ be a homomorphism with $\phi(u) = w$ for some $u \in \mathcal{V}(H)$. As in the statement of the theorem, $\mathcal{V}$ is the set of laws defining the non-trivial variety $\mathcal{B}$. Let K be a proper subgroup of F. If $w \notin K$ then clearly $w \notin \mathcal{V}(K)$. If $w \in K$ then since w is an APE, w is primitive in K since K is a proper subgroup of F. Further since $\mathcal{B}$ is a non-trivial variety and K is free we have that $K \neq \mathcal{V}(K)$. It follows then from the primitivity of w in K that $w \notin \mathcal{V}(K)$. Therefore $w \in \mathcal{V}(F)$ and for any proper subgroup K of F we have $w \notin \mathcal{V}(K)$ and hence w is $\mathcal{B}$-generic. Since free groups are Hopfian, w must then be a test word. $\square$

In particular let $F(n)$ be the subgroup of the free group F generated by all commutators and n-th powers, $n \geq 2$ or $n = 0$, that is $F(n) = \mathcal{V}_n(F)$. Then:

Corollary 1. *Let $w \in F(n)$ with $n \geq 2$ or $n = 0$. If w is an APE then w is $\mathcal{U}_n$-generic and w is a test word.*

3. APE's in Certain Free Products of Free Groups.

In this section we give a result on APE's, generic elements and test words on certain free products of free groups. The result depends on the following theorem of Rosenberger [**R1**].

Theorem 2 ([**R1**])**.** *Let $G = H_1 \star \cdots H_n$, $n \geq 2$, be the free product of groups $H_1, \ldots, H_n$. Let $a_j \in H_j, a_j \neq 1$, and let p be the number of a_j which are proper powers in H_j, $(1 \leq j \leq n)$. Let $\{x_1, \ldots, x_m\} \subset G$, $m \geq 1$, and let H be the subgroup of G generated by $x_1, \ldots, x_m$. If $a = a_1 ... a_n \in H$ then one of the following cases holds:*

(1) *There is a Nielsen transformation from $\{x_1, \ldots, x_m\}$ to a system $\{y_1, \ldots, y_m\}$ with $y_1 = a_1 \cdots a_n$.*
(2) *It is $m \geq 2n - p$, and there is a Nielsen transformation from $\{x_1, \ldots, x_m\}$ to a system $\{y_1, \ldots, y_m\}$ with $y_i \in H_j$, $1 \leq j \leq n$, $1 \leq i \leq 2n - p$; and moreover a_j can be written as a word in those y_k, $1 \leq k \leq m$, which are contained in H_j,$1 \leq j \leq n$.*

Proof. Here we give a more detailed proof than in [**R1**]. This is done in order to explain more extensively the concept of semistable letters and the blockwise description of letters in a product of generators.

We regard G as the free product $G = H_1 \star \cdots \star H_n$ together with the length L and an order with respect to this factorization. We refer to the papers [**Z3**] and [**F-R-S**] for the terminology and properties related to the length L and Nielsen cancellation methods in such free products. Consider the sets of elements $\{x_1, \ldots, x_m\}$ and $\{a_1, \ldots, a_n\}$ as in the statement of the theorem. We may assume that $\{x_1, \ldots, x_m\}$ is Nielsen reduced. For this system we then have an equation

$$\prod_{k=1}^{q} x_{\nu_k}^{\epsilon_k} = a_1 \cdots a_n \tag{1}$$

where $\epsilon_k = \pm 1, \epsilon_k = \epsilon_{k+1}$ if $\nu_k = \nu_{k+1}$.

Among the equations as in (1) there is one for which q is minimal and let us assume that this is the case in Equation (1). Further we may also assume that each $x_i \neq 1$ and that each x_i occurs in (1). If some x_i occurs only once in (1) as either x_i or x_i^{-1} then case (1) of the theorem holds. That is $\{x_1, \ldots, x_m\}$ can be carried by a Nielsen transformation to a system

$\{y_1, \dots, y_m\}$ with $y_1 = a_1 \dots a_n$. If this is not the case we will show that there is no $\lambda \in \{1, \dots, m\}$ such that always

$$L(x_\nu^\epsilon x_\lambda x_\mu^\eta) > L(x_\nu) - L(x_\lambda) + L(x_\mu)$$

for $\nu, \mu \in \{1, \dots, m\}, \epsilon, \eta = \pm 1$ and $\nu \neq \lambda \neq \mu$ or $\nu = \lambda \neq \mu, \epsilon = 1$ or $\nu \neq \lambda = \mu, \eta = 1$ or $\nu = \lambda = \mu, \epsilon = \eta = 1$. It follows then that for $\lambda \in \{1, \dots, m\}$ there is always some $\nu, \mu \in \{1, \dots, m\}$ such that

$$L(x_\nu^\epsilon x_\lambda x_\mu^\eta) \leq L(x_\nu) - L(x_\lambda) + L(x_\mu)$$

for $\epsilon, \eta = \pm 1$ and $\nu \neq \lambda \neq \mu$ or $\nu = \lambda \neq \mu, \epsilon = 1$ or $\nu \neq \lambda = \mu, \eta = 1$ or $\nu = \lambda = \mu, \epsilon = \eta = 1$. This means that each x_λ is conjugate to some element of some H_s and hence necessarily either we return to case (1) or case (2) holds proving the theorem.

We may assume that each x_i either occurs twice in Equation (1) with the same exponent $\epsilon = \pm 1$ or occurs in (1) exactly once with exponent $+1$ and once with exponent -1. In either case we always have

$$L(x_{\nu_k}^{\epsilon_k} \cdots x_{\nu_h}^{\epsilon_h}) \geq L(x_{\nu_l}^{\epsilon_l})$$

for $1 \leq k \leq l \leq h \leq q$ and

$$L(x_{\nu_k}^{\epsilon_k} x_{\nu_{k+1}}^{\epsilon_{k+1}} x_{\nu_{k+2}}^{\epsilon_{k+2}}) \geq L(x_{\nu_k}) - L(x_{\nu_{k+1}}) + L(x_{\nu_{k+2}})$$

for $1 \leq k \leq q - 2$.

Assume that there is a $\lambda \in \{1, \dots, m\}$ such that always

$$L(x_\nu^\epsilon x_\lambda x_\mu^\eta) > L(x_\nu) - L(x_\lambda) + L(x_\mu)$$

for $\nu, \mu \in \{1, \dots, m\}, \epsilon, \eta = \pm 1$ and $\nu \neq \lambda \neq \mu$ or $\nu = \lambda \neq \mu, \epsilon = 1$ or $\nu \neq \lambda = \mu, \eta = 1$ or $\nu = \lambda = \mu, \epsilon = \eta = 1$. Suppose in particular that $\lambda = \nu_k$. We write $u_i = x_{\nu_i}^{\epsilon_i}$, $1 \leq i \leq q$. Let

$$u_i = l_{i_1} \cdots l_{i_{m_i}} k_i r_{i_{m_i}} \cdots r_{i_1}$$

be the symmetric normal form of u_i (see [**Z3**] and [**F-R-S**]). We call

$$l_{i_1}, \dots, l_{i_{m_i}}, k_i, r_{i_{m_i}}, \dots, r_{i_1}$$

the **places** of u_i. For brevity we write v_i for a place of u_i. In the following we write

$$z \equiv z_1 \cdots z_p$$

to stand for the equality of the elements= together with the fact that

$$L(z) = L(z_1) + \cdots + L(z_p).$$

Given the x_λ above and its places, there is an a_t, $1 \leq t \leq n$, such that one of the following holds:

(a) $u_k \equiv p_k v_k q_k$ and $a_t = v_k b_t$ where $v_k \in H_t \setminus \{1\}$ is a place of u_k and $b_t = 1$ or $1 \neq b_t \in H_t$ and $u_{k+1} \cdots u_{k+l} \equiv q_k^{-1} b_t q_k$ for some l with $1 \leq l \leq q - k$;

(b) $u_k \equiv p_k v_k q_k$ and $a_t = b_t v_k$ where $v_k \in H_t \setminus \{1\}$ is a place of u_k and $b_t = 1$ or $1 \neq b_t \in H_t$ and $u_{k-l} \cdots u_{k-1} \equiv p_k b_t p_k^{-1}$ for some l with $1 \leq l \leq k-1$;

(c) $u_k \equiv p_k v_k q_k$, $u_{k+l+1} \equiv q_k^{-1} v_{k+l+1} q_{k+l+1}$ and $a_t = v_k b_t v_{k+l+1}$ where $0 \leq l \leq q-k-1$, $v_k \in H_t \setminus \{1\}$ is a place of u_k,$v_{k+l+1} \in H_t \setminus \{1\}$ is a place of u_{k+l+1} and $b_t = 1$ or $1 \neq b_t \in H_t$ and $u_{k+1} \cdots u_{k+l} \equiv q_k^{-1} b_t q_k$ for some l with $1 \leq l \leq q-k-1$;

(d) $u_k \equiv p_k v_k q_k$, $u_{k-l-1} \equiv p_{k-l-1} v_{k-l-1} p_k^{-1}$ and $a_t = v_{k-l-1} b_t v_k$ where $0 \leq l \leq k-2$, $v_k \in H_t \setminus \{1\}$ is a place of u_k, $v_{k-l-1} \in H_t \setminus \{1\}$ is a place of u_{k-l-1} and $b_t = 1$ or $1 \neq b_t \in H_t$ and $u_{k-l} \cdots u_{k-1} \equiv p_k b_t p_k^{-1}$ for some l with $1 \leq l \leq k-2$.

Note that if $u_{k+1} \ldots u_{k+l} \equiv q_k^{-1} b_t q_k$ or $u_{k-l} \ldots u_{k-1} \equiv p_k b_t p_k^{-1}$, $b_t \neq 1$, respectively then each u_i occurring in this product is conjugate to an element of H_t.

To see all this assume that there is an a_t with $1 \leq t \leq n$ for whose formation some $u_i, 1 \leq i < k$, and some $u_j, k < j \leq q$, contribute. Then u_{k-1} cancels the whole leading half of u_k and u_{k+1} cancels the whole rear half of u_k and the kernel of u_k has a share in the formation of a_t. But then $L(u_{k-1} u_k u_{k+1}) \leq L(u_{k-1}) - L(u_k) + L(u_{k+1})$ giving a contradiction.

Now suppose we have the blockwise description of a_t above and we assume as before that $x_\lambda = x_{\nu_k}$ occurs twice in Equation (1). Then $\nu_k = \nu_h$ for some $h, 1 \leq h \leq q$, with $k \neq h$. Without loss of generality let $k < h$ and recall that all the a_t, $1 \leq t \leq n$, are all different from different factors. If $\epsilon_h = \epsilon_k$ then the leading half of $x_\lambda = x_{\nu_h} = x_{\nu_k}$ is inverse to the rear half of $x_\lambda = x_{\nu_k}$, that is x_{ν_k} is conjugate to an element of some factor H_s with $1 \leq s \leq n$. But then

$$L(x_{\nu_k}^3) \leq L(x_{\nu_k}) = L(x_{\nu_k}) - L(x_{\nu_k}) + L(x_{\nu_k})$$

which gives a contradiction.

Now let $\epsilon_h = -\epsilon_k$. Then we have the following situation:

$u_k \equiv p_k v_k q_k$, $u_h = u_{k+l+1} = u_k^{-1} \equiv q_k^{-1} v_k^{-1} p_k^{-1}$, $u_{k+1} \cdots u_{k+l} \equiv q_k^{-1} b_t q_k$ with $v_k, b_t \in H_t \setminus \{1\}$ for some t, $1 \leq t \leq n$ and some l, $1 \leq l \leq q-k-1$. We may choose v_k in such a way that $|L(p_k) - L(q_k)| \leq 1$. Assume that $p_k \neq 1$. Since the a_t are all different from different factors, p_k or p_k^{-1} must be cancelled completely in Equation (1) by a u_i with $i < k$ or $i >= h$ respectively, which is not conjugate to an element of some factor. For such an element u_i we always have (see [**F-R-S**])

$$L(u_l^\epsilon u_i u_j^\eta) > L(u_l) - L(u_i) + L(u_j)$$

for $l, j \in \{1, \ldots, q\}, \epsilon, \eta = \pm 1$ and $l \neq i \neq j$ or $l = i \neq j, \epsilon = 1$ or $l \neq i = j, \eta = 1$ or $l = i = j, \epsilon = \eta = 1$. We have $i < k$ or $i > h$ so an inductive argument gives a contradiction because of the blockwise description of the a_t.

Therefore we have $p_k = 1$. This gives $q_k \neq 1, L(q_k) = 1$ since u_k is not conjugate to an element of some factor. But then $u_{k+1} = q_k^{-1} d_t q_k$ for some $d_t \in H_t \backslash \{1\}$, $u := u_k u_{k+1} u_k^{-1} = v_k d_t v_k^{-1}$ and $L(u) = 1 < L(u_{k+1}) = 3$ which contradicts the fact that $\{x_1, \dots, x_m\}$ is Nielsen reduced. Hence $x_\lambda = x_{\nu_k}$ occurs only once in Equation (1) contradicting the assumption that each x_i occurs twice in Equation (1). Therefore there is no λ such that always

$$L(x_\nu^\epsilon x_\lambda x_\mu^\eta) > L(x_\nu) - L(x_\lambda) + L(x_\mu)$$

for $\nu, \mu \in \{1, \dots, m\}, \epsilon, \eta = \pm 1$ and $\nu \neq \lambda \neq \mu$ or $\nu = \lambda \neq \mu, \epsilon = 1$ or $\nu \neq \lambda = \mu, \eta = 1$ or $\nu = \lambda = \mu, \epsilon = \eta = 1$. As described before this statement completes the theorem. □

We note the following. Suppose $\{x_1, \dots, x_m\} \subset G = H_1 \star \cdots \star H_n$, $m \geq 1, n \geq 2$, is a Nielsen reduced system as above with $x_i \neq 1$ for all i. Let $\lambda \in \{1, \dots, m\}$ be such that always

$$L(x_\nu^\epsilon x_\lambda x_\mu^\eta) > L(x_\nu) - L(x_\lambda) + L(x_\mu)$$

for $\nu, \mu \in \{1, \dots, m\}, \epsilon, \eta = \pm 1$ and $\nu \neq \lambda \neq \mu$ or $\nu = \lambda \neq \mu, \epsilon = 1$ or $\nu \neq \lambda = \mu, \eta = 1$ or $\nu = \lambda = \mu, \epsilon = \eta = 1$. Let $w = a_{i_1} \cdots a_{i_r} \in G$, $r \geq 1$, be given in normal form and let $w \in H = \langle x_1, \dots, x_m \rangle$. Let

$$w = \prod_{k=1}^{q} x_{\nu_k}^{\epsilon_k}$$

$\epsilon_k = \pm 1$, $\epsilon_k = \epsilon_{k+1}$ if $\nu_k = \nu_{k+1}$ with q minimal. Assume that x_λ occurs in this equation, for instance suppose $x_\lambda^{\epsilon_k} = x_{\nu_k}^{\epsilon_k} =: u_k$. Then there is an a_{i_j} related to $u_k = x_\lambda^\epsilon$ which is described via the block relation to a place v_k of u_k as in the proof of the theorem. Such a v_k is called a **semistable letter** of u_k. The advantage of a semistable letter is that it can be influenced in such an equation as above, only from one side.

Using the theorem we obtain the following result on APE's in free products of free groups.

Theorem 3. *Let F be a finitely generated free group with basis B. Let $B_1, \dots, B_n$, $n \geq 2$, be pairwise disjoint, non-empty subsets of B and let F_j be the subgroup of F generated by B_j, $1 \leq j \leq n$. Let $a_j \in F_j$ with $a_j \neq 1$, $1 \leq j \leq n$ and let $a = a_1 \dots a_n$. Then:*

(1) *If each a_j is an APE in F_j then a is an APE in F.*
(2) *Let $\mathcal{U}$ a non-trivial variety defined by the set of laws $\mathcal{V}$.*
  (a) *Let $a_j \in \mathcal{V}(F_j)$. If each a_j is $\mathcal{U}$-generic in F_j then $a \in \mathcal{V}(F)$ and a is $\mathcal{U}$-generic in F.*
  (b) *Let $a \in \mathcal{V}(F)$. If a is $\mathcal{U}$-generic in F then each $a_j \in \mathcal{V}(F_j)$ and each a_j is $\mathcal{U}$-generic in F_j.*
(3) (a) *Let $a_j \in F_j^q F_j'$, $q = 0$ or $q = 2$, for each $j, 1 \leq j \leq n$. If each a_j is a test word in F_j then a is a test word in F.*

(b) *Let $a \in F^qF'$, $q = 0$ or $q = 2$. If a is a test word in F then each a_j is a test word in F_j.*

Proof. (1) Let $a_j \in F_j$ with $a_j \neq 1$, $1 \leq j \leq n$, and let $a = a_1 \cdots a_n$. Then a cannot be primitive in F because in that case at least one a_j has to be primitive in F_j contradicting that each a_j is an APE. Let K be a proper subgroup of F with $a \in K$. From Theorem 2, a is primitive in K or without loss of generality, we may assume that K has a finite basis X which is the disjoint union of n subsets X_j of F_j such that $a_j \in K_j \subset F_j$ for each $j, 1 \leq j \leq n$, where K_j is the subgroup generated by X_j. We consider this latter situation. If $K_j = F_j$ for each j then $K = \langle K_1, \ldots, K_n\rangle = \langle F_1, \ldots, F_n\rangle = F$ contradicting the fact that $K \neq F$. Hence K_j is a proper subgroup of F_j for at least one j. Suppose $K_1 \subset F_1, K_1 \neq F_1$. Then a_1 is primitive in K_1 since a_1 is an APE in F_1 and hence $a = a_1 \cdots a_n$ is primitive in K. This completes part (1).

(2)(a) Let each a_j be $\mathcal{U}$-generic in F_j. Since each $a_j \in \mathcal{V}(F_j)$ we have $a \in \mathcal{V}(F)$. Let $\phi : H \to F$ be a homomorphism with $\phi(u) = a$ for some $u \in \mathcal{V}(H)$. Without loss of generality assume H to be finitely generated. a cannot be primitive in $K = \phi(H)$ because $\mathcal{U}$ is non-trivial. Let A be a finite generating system for H. Then $X = \phi(A)$ is a finite generating system for K. We apply Theorem 2 and the fact that a Nielsen transformation from X to a system Y defines an epimorphism from K onto K. Hence without loss of generality we assume that X is the disjoint union of n subsets X_j of F_j such that $a_j \in K_j \subset F_j$ for each j with $1 \leq j \leq n$ where K_j is the subgroup generated by X_j. Let $H_j = \phi^{-1}(K_j)$ for each j. Then $\phi_j = \phi_{|H_j}$ defines a homomorphism $\phi_j : H_j \to F_j$ with $a_j = \phi_j(u_j), 1 \leq j \leq n$, for some $u_j \in \mathcal{V}(H_j)$. Since each a_j is $\mathcal{U}$-generic, ϕ_j is an epimorphism for each j. Therefore ϕ is an epimorphism completing part (2)(a).

(2)(b) Certainly each $a_j \in \mathcal{V}(F_j)$ if $a \in \mathcal{V}(F)$. For each j let $\phi_j : H_j \to F_j$ be a homomorphsim from some group H_j such that $\phi(u_j) = a_j$ for some $u_j \in \mathcal{V}(H_j)$. Let $H = H_1 \star \cdots \star H_n$ and let $\phi : H \to F$ be the induced homomorphism with $\phi_{|H_j} = \phi_j$. Then $\phi(u_1 \ldots u_n) = \phi(a_1 \cdots a_n) = \phi(a)$ and $u_1 \ldots u_n \in \mathcal{V}(H)$. Since a is $\mathcal{U}$-generic, ϕ is an epimorphism. Hence each ϕ_j is an epimorphism and therefore each a_j is $\mathcal{U}$-generic. This completes part (2)(b).

(3)(a) For a group G we let G^qG',$q = 0$ or $q \geq 2$, denote the subgroup of G generated by the q-th powers and the commutators in G.

Let each a_j be a testword in F_j and suppose each $a_j \in F_j^qF_j', q = 0$ or $q \geq 2$. Then $a \in F^qF', q = 0$ or $q \geq 2$. Let $\phi : F \to F$ be an endomorphism with $\phi(a) = a$. Let $K = \phi(F)$. Since $a \in F^qF'$ we also have $a \in K^qK', q = 0$ or $q \geq 2$. Hence a is not primitive in K and we must show that $K = F$.

Let $X = \phi(B)$. We show first that X is a basis for K. From Theorem 2, X can be carried by a Nielsen transformation, relative to $a = a_1 \cdots a_n$ into a free basis Y of K which contains a subset Z which is the disjoint union of n subsets Z_j of F_j such that $a_j \in K_j \subset F_j$ for each $j, 1 \leq j \leq n$, where K_j is the subgroup generated by Z_j. Since each a_j is a testword in F_j we must have $|Z_j| = |B_j|$ for each j and hence $|Z| = |B|$. This gives

$$|Z| \leq |Y| \leq |X| \leq |B| = |Z|$$

and hence $Y = Z$ and X is a basis of K. Now the Nielsen transformation from X to the above system Y defines an automorphsim α of K. Hence we may already assume that $X = Y = Z$ because of the free product decomposition $F = F_1 \star \cdots \star F_n$ and the description of a as $a = a_1 \cdots a_n$. Starting, with a permutation of B, if necessary, we obtain this way an endomorphism $\psi : F \to F$ such that $\psi_j = \psi_{|F_j}$ defines an endomorphism $\psi_j : F_j \to F_j$ with $\psi_j(a_j) = a_j$ for $1 \leq j \leq n$. Since each a_j is a testword in F_j we have that each ψ_j is an automorphism of F_j. Hence by combination, ψ is an automorphism of F. Therefore by construction, ϕ is also an automorphism of F and it follows that a is a testword in F.

(3)(b) Since $a \in F^q F', q = 0$ or $q \geq 2$, we have $a_j \in F_j^q F_j', q = 0$ or $q \geq 2$. For each j let $\phi_j : F_j \to F_j$ be an endomorphsim with $\phi_j(a_j) = a_j$. Then $\phi : F \to F$ with $\phi_{|F_j} = \phi_j$ defines an endomorphism $\phi : F \to F$ with $\phi(a) = \phi(a_1 \cdots a_n) = a_1 \cdots a_n = a$. Since a is a testword in F, ϕ is an automorphism of F. Hence each ϕ_j is an automorphism of F_j and therefore each a_j is a testword in F_j. This completes the theorem. □

Corollary 2. *Let $F = \langle x_1, y_1, \ldots, x_g, y_g; \rangle, g \geq 1$. Let $a_j = a_j(x_j, y_j) \neq 1$ for $j = 1, \ldots, g$ and let both x_j and y_j occur in the freely reduced expression of a_j. Let $|a_j|_{x_j}$ be the total exponent sum of x_j in a_j and let $|a_j|_{y_j}$ be the total exponent of y_j in a_j. Let F_j be the subgroup generated by x_j and y_j.*

1. *Let each a_j be not a proper power in F_j. Then a is a testword in F if and only if $\gcd(|a_j|_{x_j}, |a_j|_{y_j}) \neq 1$ for each j.*
2. *Let a be an element of the commutator subgroup of F and suppose a is a product $a = a_1 a_2 \cdots a_g$ where each a_j is a non-trivial element of the commutator subgroup in F_j. Then a is a testword.*

Proof. This follows directly from Theorem 3 and Example 4 in Turner's paper [**T**]. □

4. A Class of Generic Elements.

In this section we give a class of examples of generic elements.

Theorem 4. *Let F be a free group on a, b and let $X = \langle x_1, \ldots, x_k \rangle, k \geq 1$ be a finitely generated subgroup of F. Suppose that X contains the element $[a^n, b^m]$ for positive integers n, m. Then $\{x_1, \ldots, x_k\}$ can be carried by a Nielsen transformation into a free basis $\{y_1, \ldots, y_p\}$, $1 \leq p \leq k$, for X for which one of the following cases occurs.*

(1) $y_1 = [a^n, b^m]$ *is a primitive element of* X*;*
(2) $y_1 = a^\alpha, 1 \le \alpha \le n, \alpha | n$ *and*
$y_2 = b^\beta, 1 \le \beta \le m, \beta | m$*;*
(3) $y_1 = a^\alpha, 1 \le \alpha \le n, \alpha | n$ *and*
$y_2 = b^m a^\beta b^{-m}, 1 \le \beta \le n, \beta | n$*;*
(4) $y_1 = b^\alpha, 1 \le \alpha < m, \alpha | m$ *and*
$y_2 = a^n b^\beta a^{-n}, 1 \le \beta \le m, \beta | m$*;*
(5) $y_1 = a^\alpha, 1 \le \alpha \le n, \alpha | n$ *and*
$y_2 = b^m a^\beta, 1 \le \beta < \alpha$*;*
(6) $y_1 = b^\alpha, 1 \le \alpha \le m, \alpha | m$ *and*
$y_2 = a^n b^\beta, 1 \le \beta < \alpha$*;*
(7) $y_1 = a^n b^m$*,* $y_2 = a^\alpha$*,* $1 \le \alpha \le 2n, \alpha | 2n$ *and*
$y_3 = b^\beta, 1 \le \beta \le 2m, \beta | 2m$*.*

Proof. The proof follows the general outline of the proof of Theorem 2. Regard F as the free product $F = \langle a \rangle \star \langle b \rangle$ together with the length L and order with respect to this factorization. We may assume $\{x_1, \dots, x_k\}$ is Nielsen reduced with $x_i \ne 1$ for all i. Further we may assume from the start that there is no Nielsen transformation from $\{x_1, \dots, x_k\}$ to a system $\{y_1, \dots, y_k\}$ with $[a^n, b^m] \in \langle y_1, \dots, y_{k-1} \rangle$, that is k is minimal with respect to this property.

As in the proof of Theorem 2, for this system we then have an equation

$$\prod_{k=1}^{q} x_{\nu_k}^{\epsilon_k} = [a^n, b^m] \tag{2}$$

where $\epsilon_k = \pm 1, \epsilon_k = \epsilon_{k+1}$ if $\nu_k = \nu_{k+1}$.

Among the equations as in (2) there is one for which q is minimal and let us assume that this is the case in Equation (2). Further we may also assume that each $x_i \ne 1$ and that each x_i occurs in (2). If some x_i occurs only once in (2) as either x_i or x_i^{-1} then case (1) of the theorem holds. Therefore for the rest of the proof we assume that case (1) does not hold.

Hence each x_i either occurs twice in Equation (2) with the same exponent $\epsilon = \pm 1$ or occurs in (2) exactly once with exponent $+1$ and once with exponent -1. In either case we always have

$$L\left(x_{\nu_k}^{\epsilon_k} \cdots x_{\nu_h}^{\epsilon_h}\right) \ge L\left(x_{\nu_l}^{\epsilon_l}\right)$$

for $1 \le k \le l \le h \le q$ and

$$L\left(x_{\nu_k}^{\epsilon_k} x_{\nu_{k+1}}^{\epsilon_{k+1}} x_{\nu_{k+2}}^{\epsilon_{k+2}}\right) \ge L(x_{\nu_k}) - L(x_{\nu_{k+1}}) + L(x_{\nu_{k+2}})$$

for $1 \le k \le q-2$. Especially, we have $L(x_i) \le 4$ for all i. Since we have only two cyclic factors and since $\{x_1, \dots, x_k\}$ is Nielsen reduced, for each x_i, which is not conjugate to a power of a or b, we have at least two places which are semistable letters for this x_i. This excludes the possibility $L(x_i) = 4$. The blockwise decription as in the proof of Theorem 2, together with $L([a^n, b^m]) = 4$ gives that there is at most one x_i which is not conjugate to a power of a or b, and which occurs in (2) exactly twice with the same exponent or exactly once with exponent $+1$ and once with exponent -1. Also, if there is such an x_i it must have length two.

Now suppose first there is an x_i which is not conjugate to a power of a or b, and which occurs in (2) exactly once with exponent $+1$ and once with exponent -1. Suppose that $x_i = a^{\alpha_1}b^{\beta_1}$. Then, since $L([a^n, b^m]) = 4$ the other x_j are powers of a or b. Recall that if we have, for instance, powers $a^{\alpha_1}, \dots, a^{\alpha_p}$, $p \geq 2$, then there is a Nielsen transformation from $\{a^{\alpha_1}, \dots, a^{\alpha_p}\}$ to $\{a^\gamma, 1, \dots, 1\}$ with $\gamma = \gcd(\alpha_1, \dots, \alpha_p)$. Hence by the minimality of k, it follows that no two of the x_i are powers of a and no two of the x_i are powers of b. A typical possible situation to consider is now, after some renumbering,

$$\begin{aligned} x_1^{\alpha_0}x_2x_3^{\beta_0}x_2^{-1}x_1^{\alpha_2}x_3^{\beta_2} &= a^{\gamma_1\alpha_0}a^{\alpha_1}b^{\beta_1}b^{\gamma_2\beta_0}b^{-\beta_1}a^{-\alpha_1}a^{\gamma_1\alpha_2}b^{\gamma_2\beta_2} \\ &= a^nb^ma^{-n}b^{-m} \end{aligned}$$

with $\alpha_0 \neq 0 \neq \alpha_2$. Then, necessarily $\gamma_1\alpha_0 + \alpha_1 = n$, $\gamma_2\beta_0 = m$, $\gamma_1\alpha_2 - \alpha_1 = -n$ and $\gamma_2\beta_2 = -m$. In particular we have $\alpha_0 = -\alpha_2$ and via a Nielsen transformation we may replace x_2 by $x_1^{\alpha_0}x_2 = a^{\gamma_1\alpha_0}a^{\alpha_1}b^{\beta_1} = a^nb^\beta$. But this contradicts the minimality of k. Therefore the above possible situation reduces to, again after some renumbering, to the situation

$$x_1 = a^nb^{\beta_1}, x_2 = b^{\gamma_2} \text{ and } x_1x_2^{\beta_0}x_1^{-1}x_2^{-\beta_0} = [a^n, b^m]$$

because $\beta_0 = -\beta_2$. We may reduce β_1 to a β with $1 \leq \beta < \gamma_2$ via a Nielsen transformation and obtain case (6) in the theorem.

An analogous case by case consideration gives that we obtain either case (5) or case (6) if there is an x_i which is not conjugate to a power of a or b and which occurs in (2) exactly once with exponent $+1$ and once with exponent -1. If each x_i is conjugate to a power of a or b we obtain cases (2), (3) or (4) since again no two of the x_i are powers of a and no two of the x_i are powers of b.

Finally suppose that one x_i is not conjugate to a power of a or b and occurs twice with the same exponent. Without loss of generality assume this exponent to be $+1$. Because of $L([a^n, b^m]) = 4$ and the blockwise description as given in the proof of Theorem 2, this x_i occurs exactly twice and has length 2 and recall that no other x_l occurs which is not a power of a or b. Let $x_i = a^\gamma b^\delta, \gamma \neq 0, \delta \neq 0$. Then after renumbering we get an equation

$$x_1^{\alpha_1}x_2x_3^{\beta_1}x_1^{\alpha_2}x_2x_3^{\beta_2} = [a^n, b^m]$$

with $x_1 = a^{\gamma_1}, x_3 = b^{\delta_1}, x_2 = a^\gamma b^\delta$. We may assume that $\gamma_1 \geq 1$ and $\delta_1 \geq 1$. We have necessarily

$$\gamma_1\alpha_1 + \gamma = n,\ \delta_1\beta_1 + \delta = m,\ \gamma_1\alpha_2 + \gamma = -n \text{ and } \delta_1\beta_2 + \delta = -m.$$

Then $\gamma_1|2n$ and $\delta_1|2m$, and $1 \le \gamma_1 \le 2n, 1 \le \delta_1 \le 2m$. If we replace x_2 by $y_2 = x_1^{\alpha_1} x_2 x_3^{\beta_1}$ this defines a Nielsen transformation and we obtain an equation

$$y_2 x_1^{\alpha_2-\alpha_1} y_2 x_3^{\beta_2-\beta_1} = [a^n, b^m]$$

with $y_2 = a^n b^m$, $x_1 = a^{\gamma_1}$, $1 \le \gamma_1 \le 2n, \gamma_1|2n$ and $x_3 = b^{\delta_1}$, $1 \le \delta_1 \le 2m, \delta_1|2m$. Since our system is Nielsen reduced we have that γ_1 does not divide n and δ_1 does not divide m. This gives case (7).

The case $x_i = b^\delta a^\gamma, \delta \ne 0, \gamma \ne 0$, cannot occur since the exponent is $+1$ and $[a^n, b^m]$ starts with a power of a and ends with a power of b. □

Using the theorem we first obtain the following corollaries. The first is due to Comerford and Edmonds [**C-E**] and the second due to Turner [**T**].

Corollary 3. *Let F be the free group on x, y and let $[x_1, x_2] = [x^n, y^m]$, $n, m \ge 1$. Then $\{x_1, x_2\}$ is Nielsen equivalent to a pair $\{y_1, y_2\}$ with either $y_1 = x^n$ and $y_2 = y^m x^\alpha, 0 \le \alpha < n$ or $y_1 = y^m$ and $y_2 = x^n y^\beta, 1 \le \beta < m$.*

Corollary 4. *The element $[x^n, y^m]$ is a test word in the free group of rank two on x, y for any $n, m \ge 1$.*

Recall that $\mathcal{U}_n$ is the variety generated by the laws $\mathcal{V}_n = \{[x, y], z^n\}, n = 0$ or $n \ge 2$. We let $\mathcal{L}_n$ be the variety generated by the laws $\mathcal{W}_n = \{[x^n, y^n]\}, n \ge 1$. We then obtain the following class of generic elements.

Corollary 5. *Let F be a free group of rank 2 on x, y. Then $[x^n, y^n]$ is $\mathcal{L}_n$-generic in F but for $n \ge 2$ it is not $\mathcal{U}_n$-generic in F.*

Corollary 6. *Let F be a free group of rank 2 on x, y. Then the element $[x^n, y^m]$, $n, m \ge 1$, is an APE if and only if $n = m = 1$.*

Recall that in general it is not true that being an APE implies being a test word. As mentioned earlier if $F = F(x, y)$ and $w = x^2 y x^{-1} y^{-1}$ then w is an APE but is not a test word. Since generic elements are test words this example shows further that APE does not imply generic in general. However using the same techniques as in Theorem 4 we can generalize the fact that the element w above is an APE to obtain further examples of APE's and testwords.

Theorem 5. *Let $F = \langle a, b; \rangle$ and let $X = \langle x_1, \dots, x_k \rangle \subset F$, $k \ge 1$. Suppose $a^n b a^{-1} b^{-1} \in X, n \ge 2$. Then there is a Nielsen transformation from $\{x_1, \dots, x_k\}$ to a basis $\{y_1, \dots, y_p\}$, $1 \le p \le k$, of X such that one of the following cases holds:*

(1) *$y_1 = a^n b a^{-1} b^{-1}$ or*
(2) *$y_1 = a, y_2 = b$.*

From this theorem and Theorem 1 we get the following corollary.

Corollary 7. *Let $F = \langle a, b; \rangle$. Then*

(1) *$a^n ba^{-1}b^{-1}$, $n \geq 2$, is an APE;*
(2) *$a^n ba^{-1}b^{-1}$, $n \geq 3$, is $\mathcal{U}_{n-1}$-generic;*
(3) *$a^n ba^{-1}b^{-1}$, $n \geq 3$, is a testword in F;*
(4) *$a^2 ba^{-1}b^{-1}$ is not a testword in F.*

Proof of Theorem 5. The proof follows the same outline as the proof of Theorem 4. Assume that $\{x_1, \dots, x_k\}$ is Nielsen reduced and k is minimal in the sense that $\{x_1, \dots, x_k\}$ is not Nielsen equivalent to a system $\{y_1, \dots, y_k\}$ with $a^n ba^{-1}b^{-1} \in \langle y_1, \dots, y_{k-1}\rangle$. Assume further that each x_i occurs at least twice in the freely reduced equation expressing $a^n ba^{-1}b^{-1}$ in terms of $x_1, \dots, x_k$. Assume that case (1) does not hold and assume that there is one x_i which is not conjugate to a power of a or b. Suppose first that this x_i occurs twice with the same exponent, without loss of generality say $+1$. As in the proof of Theorem 4, it follows from $L(a^n ba^{-1}b^{-1}) = 4$, the fact that the system is Nielsen reduced and the blockwise description that this x_i occurs exactly twice, has length 2, and no other x_l occurs which is not a power of a or b. Let $x_i = a^\gamma b^\delta, \alpha \neq 0 \neq \delta$. As in Theorem 4, $x_i = b^\delta a^\gamma$ cannot occur. Then after renumbering we obtain an equation

$$x_1^{\alpha_1} x_2 x_3^{\beta_1} x_1^{\alpha_2} x_2 x_3^{\beta_2} = a^n ba^{-1}b^{-1}$$

with $x_1 = a^{\gamma_1}, x_3 = b^{\delta_1}, x_2 = a^\gamma b^\delta$. We may assume that $1 \leq \delta_1, 1 \leq \gamma_1$, $1 \leq \gamma < \gamma_1$ and $1 \leq \delta < \delta_1$. Then necessarily $\delta + \beta_1\delta_1 = 1$ and $\delta + \beta_2\delta_1 = -1$. Hence $\delta_1 = 1$ or 2. Since the system is Nielsen reduced $\delta_1 \neq 1$ and hence $\delta_1 = 2$. Then $\delta = 1$ and hence $\beta_1 = 0$ and $\beta_2 = -1$. Thus we have $x_1^{\alpha_1} x_2 x_1^{\alpha_2} x_2 x_3^{-1} = a^n ba^{-1}b^{-1}$ contradicting the assumption that case (1) does not hold. It follows therefore that this x_i occurs exactly once with exponent $+1$ and once with exponent -1. Then as in the proof of Theorem 4 we must consider equations of the form

$$x_1^{\alpha_1} x_2 x_3^{\beta_1} x_2^{-1} x_1^{\alpha_2} x_3^{\beta_2} = a^n ba^{-1}b^{-1}$$

with $x_1 = a^{\gamma_1}, x_3 = b^{\delta_1}, x_2 = a^\gamma b^\delta$ and $0 \neq \gamma, 0 \neq \delta$. Without loss of generality let $1 \leq \gamma_1, 1 \leq \delta_1$, ($\gamma_1 = 0$ or $\delta_1 = 0$ cannot occur since $n \geq 2$). Then necessarily $\delta_1\beta_1 = 1$ and hence $\delta_1 = 1$. But this contradicts the fact that the system is Nielsen reduced since we can replace x_2 by $x_2 x_3^{-\delta} = a^\gamma b^\delta b^{-\delta} = a^\gamma$. The other possibilities are analogous. If each x_i is conjugate to a power of a or b then case (2) certainly holds if case (1) does not. This completes the proof of the theorem. □

The corollary now follows easily. If $w = a^n ba^{-1}b^{-1}$ is in any proper subgroup of F then condition (2) of the theorem cannot hold and hence condition (1) must hold, that is w is primitive. Therefore w is an APE. If $n \geq 3$ then $w \in \mathcal{V}_{n-1}(F)$ where $\mathcal{V}_n$ is the set of laws $\mathcal{V}_n = \{[x, y], z^n\}$. As before if $\mathcal{U}_n$ is the variety defined by this set of laws, then $\mathcal{U}_n$ is an non-trivial variety and it follows that w is an APE and that w is $\mathcal{U}_{n-1}$-generic and hence a testword. Finally part (4) comes from Turner [**T**].

5. A Result on Varieties and Primitive Elements.

The following result relates when the laws determined by a single element generate a trivial variety and being in a retract.

Theorem 6. *Let F be the free group on $x_1, \dots, x_n$ with $n \geq 2$ and let w be a freely reduced non-empty word in the generators of F which does not define a proper power of F. Then if the law $w = 1$ determines the trivial variety* (*consisting only of trivial groups*) *then w is a primitive in a retract of F.*

Proof. Notice first that if $\mathcal{B}$ is a non-trivial variety then since $\mathcal{B}$ is closed under subgroup formation it must contain non-trivial cyclic groups. Hence the intersection of $\mathcal{B}$ with the abelian variety $\mathcal{A}$ is not the trivial variety.

Now let $\mathcal{V}$ be the variety determined by the law $w = 1$. Let $\mathcal{E}$ stand for the trivial variety. Then $\mathcal{V} = \mathcal{E}$ is equivalent to the following two conditions being simultaneously satsified:

(1) $w \notin [F, F] = F'$ and
(2) If e_i is the exponent sum in w of x_i, $i = 1, \dots, n$, then $\gcd(e_1, \dots, e_n) = 1$.

To see this suppose that (1) and (2) are satisfied by w. Since $w \equiv x_1^{e_1} \cdots x_n^{e_n} (\mathrm{mod}[F, F])$ (1) is equivalent to $(e_1, \dots, e_n) \neq (0, \dots, 0)$. Let $m_1, \dots, m_n$ be integers such that $m_1 e_2 + \cdots + m_n e_n = 1$. Let G be an abelian group lying in $\mathcal{V}$. Then G must satisfy the law $x_1^{e_1} \cdots x_n^{e_n} = 1$. Let $x \in G$. Let $x_i = x^{m_i}$. Then from the law $x^{m_1 e_1 + \cdots + m_n e_n} = x = 1$. Therefore G is trivial. Hence $\mathcal{V}$ contains no non-trivial abelian groups and therefore it follows from the remark above that $\mathcal{V}$ is itself trivial.

Conversely suppose $\mathcal{V}$ is trivial. We show that conditions (1) and (2) must hold. Suppose (1) does not hold so that $w \in [F, F]$. Let $c_1, \dots, c_n$ be arbitrary integers. Then the infinite cyclic group $A = \langle a; \rangle$ lies in $\mathcal{V}$ since $w(a^{c_1}, \dots, a^{c_n}) \in [A, A] = 1$. This contradicts the triviality of $\mathcal{V}$ so therefore (1) must hold.

Now suppose $w \notin [F, F]$ but (2) is violated. Suppose $\gcd(e_1, \dots, e_n) = d > 1$. Then the finite cyclic group $B = \langle b; b^d = 1\rangle$ lies in $\mathcal{V}$ since for any integers $c_1, \dots, c_n$,

$$w(b^{c_1}, \dots, b^{c_n}) = (b^{c_1})^{e_1} \cdots (b^{c_n})^{e_n} = b^{c_1 e_1 + \cdots + c_n e_n} = 1$$

and $d | c_1 e_1 + \cdots + c_n e_n$. Since $d > 1$, B is non-trivial contradiciting the triviality of $\mathcal{V}$ so therefore (2) must also hold.

Now suppose w satisfies (1) and (2) and $m_1, \dots, m_n$ are integers such that $m_1 e_1 + \cdots + m_n e_n = 1$. Consider the map $F \to \langle w \rangle$ given by $x_1 \to$

$w^{m_1}, \dots, x_n \to w^{m_n}$. Suppose $w = x_{i_1}^{k_1} \cdots x_{i_l}^{k_l}$ where each k_j is a non-zero integer with $i_j \neq i_{j+1}$ for $j = 1, \dots, l-1$. Then under the above map

$$\begin{aligned} w \to w^{m_{i_1}k_1 + \cdots + m_{i_l}k_l} &= w^{m_1 \sum_{i_j=1} k_j + \cdots + m_n \sum_{i_j=n} k_j} \\ &= w^{m_1 e_1 = \cdots + m_n e_n} = w. \end{aligned}$$

It follows that $F \to \langle w \rangle$ is a retraction and clearly w is primitive in $\langle w \rangle$. $\square$

The above proof depended on the fact that if $\mathcal{A}$ is the abelian variety and $\mathcal{E}$ is the trivial variety then $\mathcal{B} \cap \mathcal{A} = \mathcal{E}$ implies that $\mathcal{B} = \mathcal{E}$. The next result completely characterizes the varieties such as $\mathcal{A}$ with this property. Recall that a variety $\mathcal{V}$ has **exponent** n if it satisfies the law $X^n = 1$. If $\mathcal{V}$ has no finite exponent it has **infinite exponent**.

Theorem 7. *Let $\mathcal{V}$ be a variety. Then $\mathcal{V}$ has the property that $\mathcal{B} \cap \mathcal{V} = \mathcal{E}$ implies that $\mathcal{B} = \mathcal{E}$ for an arbitrary variety $\mathcal{B}$ if and only if $\mathcal{V}$ has infinite exponent.*

Proof. Suppose $\mathcal{V}$ has infinite exponent. Therefore $\mathcal{V}$ contains infinite cyclic groups and since it is closed under the formation of quotients it contains cyclic groups of all possible finite orders. Since any non-trivial variety must contain cyclic groups of some order it follows that $\mathcal{V}$ will intersect non-trivially with any non-trivial variety.

Conversely suppose $\mathcal{V}$ has the stated property. If $\mathcal{V}$ has finite exponent n let m be an integer relatively prime to n and let $\mathcal{V}_1$ be a variety of finite exponent m. Since $\mathcal{V}$ satisfies the law $X^n = 1$ and $\mathcal{V}_1$ satisfies the law $X^m = 1$ and $(m, n) = 1$ it follows that their intersection satisfies the law $X = 1$. Hence only trivial groups are in their intersection. But $\mathcal{V}_1$ is non-trivial contradicting the stated property. Therefore $\mathcal{V}$ must have infinite exponent. $\square$

6. Extensions to Arbitrary Groups.

As pointed out by Turner the characterization and determination of test elements in arbitrary non-free groups is much more subtle and complicated than in free groups. First we show that there can exist test elements in non-free groups. The fact that $[x, y]$ is a test word in the free group of rank two on x, y followed from the following method of Nielsen: if u, v are elements of the free group of rank two on x, y and $[x, y] = [u, v]$ then the set $\{u, v\}$ is Nielsen equivalent to the set $\{x^{\pm 1}, y^{\pm 1}\}$. Exactly the same type of Nielsen transformation arguments can be applied in the free product of two cyclic groups (not excluding finite) provided that we allow an extended Nielsen transformation which replaces a generator x by x^d where x^d is also a generator of $\langle x \rangle$. In particular in a free product of cyclic groups with basis x, y the commutator is a test element.

Theorem 8. *Let $G = \langle x, y : x^p = y^q = 1\rangle \cong \mathbb{Z}_p \star \mathbb{Z}_q$ be the free product of two finite cyclic groups. Then the commutator $[x, y]$ is a test element.*

Much of the development on almost primitive elements and generic elements can be translated to more general situations. Let $\mathcal{U}$ be a variety defined by the set of laws $\mathcal{V}$ and G a group. Then we say that $\mathcal{U}$ is **efficient for** G if $\mathcal{V}(H) \neq H$ for any non-trivial subgroup of G. Recall that an element $g \in G$ is primitive if $G = \langle g\rangle \star G_1$ with $G_1 \neq G$ and g of infinite order. We then get the following.

Lemma 1. *Let g be primitive in G. Then $g \notin \mathcal{V}(g)$ for any non-trivial set of laws $\mathcal{V}$ unless $\mathcal{V}(G) = G$.*

Proof. If g is primitive in G and H is any group then any map g into H can be extended to a homomorphism $G \to H$. Let H be a $\mathcal{V}$-group. Then any $g_1 \in \mathcal{V}(G)$ goes to the identity. Therefore $g \notin \mathcal{V}(G)$ unless $\mathcal{V}(G) = G$. □

From this we can extend Theorem 1 almost exactly.

Theorem 9. *Let $\mathcal{U}$ be a variety defined by the laws $\mathcal{V}$ and suppose $\mathcal{U}$ is efficient for G. Let $g \in \mathcal{V}(G)$. Then if g is an APE in G it follows that g is $\mathcal{U}$-generic. Further if G is Hopfian then g is a test element.*

Proof. The proof is almost identical to the proof of Theorem 1. Suppose $g \in \mathcal{V}(G)$. Let $K \subset G$ be a proper subgroup. If $g \notin K$ then $g \notin \mathcal{V}(K)$. If $g \in K$ then $\mathcal{V}(K) \neq K$ since $\mathcal{U}$ is efficient for G. Since g is an APE it is primitive in K and hence from Lemma 1 $g \notin \mathcal{V}(K)$. Therefore g is $\mathcal{U}$-generic. If G is Hopfian then as before generic elements are test elements. □

As an example consider the Modular group $M = \mathbb{Z}_2 \star \mathbb{Z}_3$ the free product of a cyclic group of order two and a cyclic group of order three. Let $M = \langle x, y; x^2 = y^3 = 1\rangle$ and let $g = [x, y]$. Now $[x, y] = xyxy^2$ so the same proof as in the free group case shows that any Nielsen transformation will map this to a cyclic rewrite up to conjugation and exponent ± 1. It follows than that g is an APE. The abelian variety $\mathcal{A}$ is M-efficient so from Theorem 9, g is $\mathcal{A}$-generic. Since M is Hopfian this gives another proof that g is a test element.

The following straightforward propositions give some additional results.

Proposition 2. *Let F be a free group. $w \in F$ is a test word if and only if whenever $f : F \to F$ is an endomorphism with $f(w) = w_1$ with w_1 Whitehead related to w then f is an automorphism.*

Proof. Suppose $w \in F$ is a test word and suppose $f : F \to F$ is an endomorphism with $f(w) = w_1$ with w_1 Whitehead related to w. Then there is an automorphism $\alpha : F \to F$ with $\alpha(w_1) = w$. Then $f_1 = \alpha f$ is an endomorphism of F with $f_1(w) = w$. Since w is a testword it follows that f_1 is an automorphism. Therefore $f = \alpha^{-1} f_1$ is also an automorphism. The converse is clear. $\square$

Proposition 3. *Let w be a test word in the free group F and let N be a normal subgroup of F. Suppose that whenever $w \equiv w_1(N)$ it follows that there is a $w_2 \in w_1 N$ which is Whitehead related to w. Let $p : F \to F/N$ be the natural projection and let $g = p(w)$. Then g is a test element in $G = F/N$.*

Proof. Let $\phi : G \to G$ be an endomorphism with $\phi(g) = g$. Relative to a fixed generating system of G, ϕ can be lifted to an endomorphism $\phi^\star$ of the free group F. Let $w_1 = \phi^\star(w)$. Since $\phi(g) = g$ it follows that $p(w) = p(w_1)$ and hence $w \equiv w_1(N)$. From the condition we may assume that w is Whitehead related to w_1 and hence from Proposition 1 it follows that $\phi^\star$ is an automorphism of the free group and therefore ϕ is an automorphism of G. $\square$

7. The Test Rank of a Group.

If g is a test element of a group G then it is straightforward to see that this is equivalent to the fact that if $f(g) = \alpha(g)$ for some endomorphism f of G and some automorphism α of G then f must also be an automorphism. A **test set** in a group G consists of a set of elements $\{g_i\}$ with the property that if f is an endomorphism of G and $f(g_i) = \alpha(g_i)$ for some automorphism α of G and for all i then f must also be an automorphism. Any set of generators for G is a test set and if G posses a test element then this is a singleton test set. The **test rank** of a group is the minimal size of a test set. Clearly the test rank of any finitely generated group is finite and bounded above by the rank and below by 1. Further the test rank of any free group of finite rank is 1 since these contain test elements. For a free abelian group of rank n the test rank is precisely n.

Lemma 2. *If $G = \mathbb{Z}^n$, $n \geq 1$ is a free abelian group of rank n then its test rank is n.*

Proof. Let $G = \mathbb{Z} \times \mathbb{Z}$ be a free abelian group of rank 2. We show that G contains no test element. The proof in the general case that a set of k elements in a free abelian group of rank n with $k < n$ cannot be a test set is analogous.

Let x, y be a basis for G. We will write the group additively. Suppose $g = mx + ny$ with m, n integers is a test element. We will show that there exists a non-invertible endomorphism of G which fixes g. Any mapping

$$x \to ax + by$$
$$y \to cx + dy$$

determines an endomorphism of G to G. This homomorphism will be invertible and hence an automorphsim only if $ad - bc = \pm 1$. Suppose under this homomorphism $g \to g$. We thus have

$$mx + ny \to m(ax + by) + n(cx + dy) = g = mx + ny.$$

Considering a, b, c, d as integral unknowns we are then led to the sytem of two equations in four unknowns

$$ma + nc = m$$

$$mb + nd = n.$$

This has infinitely many integral solution with c dependent on a and b dependent on d. Choosing one such solution such that $ad - bc \neq \pm 1$ gives the desired homomorphism. □

Thus free abelian groups have maximal test rank while free groups of finite rank have minimal test rank. For given integers n and k with $k < n$ there always exist groups of rank n and test rank k.

Lemma 3. *Given integers n and k with $k < n$ there exist a group of rank n and test rank k.*

Proof. Let G_m stand for a free abelian group of rank m and F_d stand for a free group of rank d. Then the group $G = F_d \times G_m$ has rank $m + d$ and test rank $m + 1$. Given arbitrary n and $k < n$ choose m, d so that $m + 1 = k$ and $m + d = n$. The group G then has the desired property. □

We close with two questions on test rank.

(1) Given a finite presentation for a group G and given knowledge of the rank can one determine the test rank?

(2) Can one give an example of a group G with rank n and test rank $1 < k < n$ other than those of the type in the proof of Lemma 3 - that is not of the form $F_d \times G_m$.

References

[B-B-O] A.M. Brunner, R.G. Burns and S. Oates-Williams, *On almost primitive elements of free groups with an application to Fuchsian groups*, Can. J. Math., **45** (1993), 225-254.

[C] L.P. Comerford, *Generic elements of free groups*, Archiv der Math, to appear.

[C-E] L.P. Comerford and C.C. Edmonds, *Products of commutators and products of squares in a free group*, Int. J. of Algebra and Computation, **4** (1994), 469-480.

[Do] A. Dold, *Nullhomologous words in free groups which are not nullhomologous in any proper subgroup*, Arch. Math., **50** (1988), 564-569.

[D] V.G. Durnev, *The Mal'cev-Nielsen equation in a free metabelian group of rank two*, Math. Notes, **64** (1989), 927-929.

[F-R-S] B. Fine, G. Rosenberger and M. Stille, *Nielsen Transformations and Applications: A Survey*, Groups Korea, 1994, Kim/Johnson Eds., DeGruyter, (1995), 69-105.

[G-S] N. Gupta and V. Shpilrain, *Nielsen's commutator test for two-generator groups*, Arch. Math., **44** (1985), 1-14.

[H-P] P. Hill and S. Pride, *Commutators, generators and conjugacy equations in groups*, Math Proc. Camb. Phil. Soc., **114** (1993), 295-301.

[I] N. Isermann, *Generische Elemente und Testelemente in freien Gruppen*, Diplomarbeit - Universitat Dortmund, 1996.

[K-R] R.N. Kalia and G.Rosenberger, *Automorphisms of the Fuchsian groups of type* $(0, 2, 2, 2, q : 0)$, Comm. in Alg., **6(11)** (1978), 115-129.

[M-P] J. McCool and A. Pietrowski, *On free products with amalgamation of two infinite cyclic groups*, J. of Alg., **18** (1971), 377-383.

[Ne] H. Neumann, *Varieties of Groups*, Springer-Verlag, 1967.

[N] J. Nielsen, *Die Automorphisem der algemeinen unendlichen Gruppe mit zwei Erzeugenden*, Math. Ann., **78** (1918), 385-397.

[R1] G. Rosenberger, *Uber Darstellungen von Elementen und Untergruppen in freien Produkten*, Proc. of Groups Korea 1983: Springer Lecture Notes in Math., **1098** (1984), 142-160

[R2] ———, *Alternierende produkte in freien Gruppen*, Pac. J. Math., **78** (1978), 243-250.

[R3] ———, *Minimal generating systems for plane discontinuous groups and an equation in free groups*, Proc. of Groups Korea 1988: Springer Lecture Notes in Math., (1989), 170-186.

[S1] V. Shpilrain, *Test elements for endomorphisms of free groups and algebras*, preprint.

[S2] ———, *Recognizing automorphisms of the free groups*, Arch. Math., **62** (1994), 385-392.

[St] J. Stallings, *Problems about free quotients of groups*, preprint.

[T] E.C. Turner, *Test words for automorphisms of the free groups*, J. London Math. Soc., to appear.

[Z1] H. Zieschang, *Alternierende produkte in freien Gruppen*, Abh. Math. Sem. Univ. Hamburg, **27** (1964), 12-31.

[Z2] ———, *Automorphismen ebener discontinuerlicher Gruppen*, Math. Ann., **166** (1966), 148-167.

[Z3] ———, *Uber die Nielsensche Kurzungsmethode in freien Produkten mit Amalgam*, Invent. Math., **10** (1970), 4-37.

Received March 7, 1997 and revised May 18, 1998.

Fairfield University
Fairfield, CT 06430
E-mail address: fine@fair1.fairfield.edu

Fachbereich Mathematik Universitat
Dortmund
Federal Republic of Germany

St Joseph University
Philadelphia, PA 19131

Fachbereich Mathematik Universitat
Dortmund
Federal Republic of Germany

PACIFIC JOURNAL OF MATHEMATICS
Vol. 190, No. 2, 1999

LINEARLY UNRELATED SEQUENCES

Jaroslav Hančl

The paper deals with the so-called linearly unrelated sequences. The criterion and the application for irrational sequences and series is included too.

1. Introduction.

There are not many new results concerning the linear independence of numbers. Exceptions in the last decade are, e.g., the result of Sorokin [8] which proves the linear independence of logarithmus of special rational numbers, or that of Bezivin [2] which proves linear independence of roots of special functional equations.

The algebraic independence of numbers can be considered as a generalization of linear independence. One can find many results of this nature. For instance, in [4] Bundschuh proves that if the special series of rational numbers converges to infinity very fast then they are algebraically independent. In [7] I prove a similar result for continued fractions. In that paper the so-called continued fractional algebraic independence of sequences was also defined.

If we consider irrationality as a special case of linear independence then we can obtain many results. For instance, in [1] Apery proves the irrationality of $\zeta(3)$ and in [3] Borwein proves the irrationality of the sum $\sum_{n=1}^{\infty} 1/(q^n+r)$, where q and r are integers such that $q>1$ and $r\neq 0$.

In 1975 Erdös defined the so-called irrationality of sequences in [5] (we will consider a generalization of this definition in Section 3) and in the same paper he proves the irrationality of the sequence $\{2^{2^n}\}$. In 1993 in [6] I proved:

Theorem. *Let $\{r_n\}_{n=1}^{\infty}$ be a nondecreasing sequence of positive real numbers such that $\lim_{n\to\infty} r_n=\infty$, let B be a positive integer, and let $\{a_n\}_{n=1}^{\infty}$, $\{b_n\}_{n=1}^{\infty}$ be sequences of positive integers such that*

$$b_{n+1}\leq r_n^B$$

and

$$a_n\geq r_n^{2^n}$$

holds for every large n. Then the series

$$A = \sum_{n=1}^{\infty} b_n/a_n$$

and the sequence $\{a_n/b_n\}_{n=1}^{\infty}$ are irrational.

2. Linearly Unrelated Sequences.

Definition 2.1. Let $\{a_{i,n}\}_{n=1}^{\infty}$ $(i = 1, \dots, K)$ be sequences of positive real numbers. If for every sequence $\{c_n\}_{n=1}^{\infty}$ of positive integers the numbers $\sum_{n=1}^{\infty} 1/(a_{1,n}c_n)$, $\dots$, $\sum_{n=1}^{\infty} 1/(a_{K,n}c_n)$, and 1 are linearly independent, then the sequences $\{a_{i,n}\}_{n=1}^{\infty}$ $(i = 1, \dots, K)$ are linearly unrelated.

Theorem 2.1. *Let $\{a_{i,n}\}_{n=1}^{\infty}$, $\{b_{i,n}\}_{n=1}^{\infty}$ $(i = 1, \dots, K-1)$ be sequences of positive integers and $\epsilon > 0$ such that*

$$\frac{a_{1,n+1}}{a_{1,n}} \geq 2^{K^{n-1}}, a_{1,n}|a_{1,n+1} \quad (a_{1,n} \ \textit{divides} \ a_{1,n+1}) \tag{1}$$

$$b_{i,n} < 2^{K^{n-(\sqrt{2}+\epsilon)\sqrt{n}}}, \quad i = 1, \dots, K-1 \tag{2}$$

$$\lim_{n\to\infty} \frac{a_{i,n}b_{j,n}}{b_{i,n}a_{j,n}} = 0, \quad \textit{for all} \ j, i \in \{1, \dots, K-1\}, i > j \tag{3}$$

$$a_{i,n}2^{-K^{n-(\sqrt{2}+\epsilon)\sqrt{n}}} < a_{1,n} < a_{i,n}2^{K^{n-(\sqrt{2}+\epsilon)\sqrt{n}}}, i = 1, \dots, K-1 \tag{4}$$

hold for every sufficiently large natural number n. Then the sequences $\{a_{i,n}/b_{i,n}\}_{n=1}^{\infty}$ $(i = 1, \dots, K-1)$ are linearly unrelated.

Proof. We will prove that for every sequence $\{c_n\}_{n=1}^{\infty}$ of positive integers and for every $(K-1)$-tuple of integers $\alpha_1, \dots, \alpha_{K-1}$ (not all equal to zero) the sum

$$A = \sum_{j=1}^{K-1} \alpha_j \sum_{n=1}^{\infty} \frac{b_{j,n}}{a_{j,n}c_n}$$

is an irrational number. Suppose that A is a rational number. Let R be a maximal index such that $\alpha_R \neq 0$. Then we have

$$\begin{aligned} A &= \sum_{j=1}^{K-1} \alpha_j \sum_{n=1}^{\infty} \frac{b_{j,n}}{a_{j,n}c_n} = \sum_{n=1}^{\infty}\sum_{j=1}^{R} \alpha_j \frac{b_{j,n}}{a_{j,n}c_n} \\ &= \sum_{n=1}^{\infty} \frac{b_{R,n}}{a_{R,n}c_n} \left(\sum_{j=1}^{R-1} \alpha_j \frac{b_{j,n}a_{R,n}}{a_{j,n}b_{R,n}} + \alpha_R \right). \end{aligned}$$

Because of (3), there is a natural number N such that for every $n \geq N$ the number

$$\sum_{j=1}^{R-1} \alpha_j \frac{b_{j,n}a_{R,n}}{a_{j,n}b_{R,n}} + \alpha_R$$

and the number α_R have the same sign. Without loss of generality we may assume $\alpha_R > 0$ and (1)-(4) hold for every $n \geq N$. Thus, there are positive integers p and q such that

$$B = \frac{p}{q} = \sum_{n=N}^{\infty} \sum_{j=1}^{R} \alpha_j \frac{b_{j,n}}{a_{j,n} c_n}.$$

We reorder the sequences $\{a_{j,n} c_n\}_{n=N}^{\infty}$ to obtain the sequences $\{c_{j,n}\}_{n=N}^{\infty}$ $(j = 1, \ldots, R)$ so that $c_{1,N} \leq c_{1,N+1} \leq c_{1,N+2} \leq ....$ Thus, there is a map $\phi:\{n \geq N\} \to \{n \geq N\}$, such that $c_{1,n} = a_{1,\phi(n)} c_{\phi(n)}$ for $n \geq N$. It follows that

$$B = \frac{p}{q} = \sum_{n=N}^{\infty} \sum_{j=1}^{R} \alpha_j \frac{d_{j,n}}{c_{j,n}}, \tag{5}$$

where $d_{j,n} = b_{j,\phi(n)}$ for every $j = 1, \ldots, K-1$, $n = N, N+1, \ldots$. We will consider two cases.

1. First we assume that

$$\limsup_{n\to\infty} c_{1,n}^{1/K^n} = 2^V. \tag{6}$$

Then (1), (6), and the definition of the sequence $\{c_{1,n}\}_{n=1}^{\infty}$ imply that

$$V > 0.$$

Also, (6) implies that for every $\delta > 0$ there is a $n(\delta)$ such that for every $j > n(\delta)$

$$c_{1,j} < 2^{(V+\delta)K^j}, \tag{7}$$

and there are infinitely many M such that

$$c_{1,M} > 2^{(V-\delta)K^M}. \tag{8}$$

From $c_{1,n} = a_{1,\phi(n)} c_{\phi(n)} \leq 2^{(V+\delta)K^n}$, we get $a_{1,\phi(n)} \leq 2^{(V+\delta)K^n}$. Now, condition (1) gives

$$a_{1,\phi(n)} \geq a_{1,1} 2^{\frac{K^{\phi(n)-1}-1}{K-1}} \geq 2^{\frac{K^{\phi(n)-1}-1}{K-1}}.$$

Thus, $K^{\phi(n)-1} \leq 1 + (K-1)(V+\delta)K^n$ for all sufficiently large n. Hence,

$$\phi(n) - 1 \leq n + \frac{\log(V+\delta) + \log(K-1) + \log\left(1 + \frac{1}{(K-1)(V+\delta)K^n}\right)}{\log K},$$

and, $\phi(n) \leq n + \frac{\log(V+\delta)}{\log K} + 2$ for n sufficiently large. From the latter inequality, it follows from the fact that $x \to x - (\sqrt{2}+\epsilon)\sqrt{x}$ is increasing that

$$d_{j,n} < 2^{K^{n+\gamma-(\sqrt{2}+\epsilon)\sqrt{n}}}, j = 1, \ldots, R, \tag{9}$$

holds for every $n \geq N_1$, where $\gamma = \frac{\log(V+\delta)}{\log K} + 2$. For the same reason, and with the help of (4), we also obtain that

$$c_{j,n} 2^{-K^{n+\gamma-(\sqrt{2}+\epsilon)\sqrt{n}}} < c_{1,n} < c_{j,n} 2^{K^{n+\gamma-(\sqrt{2}+\epsilon)\sqrt{n}}}, j = 1, \dots, R \tag{10}$$

holds for every $n \geq N_2$. Now, (9) and (10) imply that

$$\sum_{n=M}^{\infty} \sum_{j=1}^{R} \alpha_j \frac{d_{j,n}}{c_{j,n}} \leq \sum_{n=M}^{\infty} \frac{2^{K^{n+\gamma-(\sqrt{2}+\epsilon)\sqrt{n}+3}}}{c_{1,n}} \tag{11}$$

for every sufficiently large M. Let $h \in N$ such that $\gamma + 1 \geq h > \gamma$. Now we will prove

$$T_M = \sum_{n=M}^{\infty} \frac{2^{K^{n+\gamma-(\sqrt{2}+\epsilon)\sqrt{n}+3}}}{c_{1,n}} \leq \frac{2^{K^{M+\beta-(\sqrt{2}+\epsilon)\sqrt{M}+4}}}{c_{1,M}} \tag{12}$$

for every sufficiently large M where $\beta = \gamma + h$. Also (1) yields $a_{1,n} \geq 2^{K^{n-2}}$. Thus $c_{1,n} \geq 2^{K^{n-2}}$. From this and (7) we have

$$\begin{aligned}
T_M &= \sum_{n=M}^{\infty} \frac{2^{K^{n+\gamma-(\sqrt{2}+\epsilon)\sqrt{n}+3}}}{c_{1,n}} \\
&= \sum_{n=M}^{M+h} \frac{2^{K^{n+\gamma-(\sqrt{2}+\epsilon)\sqrt{n}+3}}}{c_{1,n}} + \sum_{n=M+h+1}^{\infty} \frac{2^{K^{n+\gamma-(\sqrt{2}+\epsilon)\sqrt{n}+3}}}{c_{1,n}} \\
&\leq (h+1) \frac{2^{K^{M+\gamma-(\sqrt{2}+\epsilon)\sqrt{M}+3+h}}}{c_{1,M}} + \sum_{n=M+h+1}^{\infty} \frac{2^{K^{n+\gamma-(\sqrt{2}+\epsilon)\sqrt{n}+3}}}{c_{1,n}}
\end{aligned}$$

because $c_{1,M+j} \geq c_{1,M}$ for $j \geq 0$, and

$$\begin{aligned}
\sum_{n=M+h+1}^{\infty} \frac{2^{K^{n+\gamma-(\sqrt{2}+\epsilon)\sqrt{n}+3}}}{c_{1,n}} &\leq \sum_{n=M+h+1}^{\infty} \frac{2^{K^{n+\gamma-(\sqrt{2}+\epsilon)\sqrt{n}+3}}}{2^{K^{n-2}}} \\
&\leq 2 \frac{2^{K^{M+\gamma-(\sqrt{2}+\epsilon)\sqrt{M}+3+h}}}{2^{K^{M+h-1}}}.
\end{aligned}$$

So

$$T_M \leq (h+1) \frac{2^{K^{M+\gamma-(\sqrt{2}+\epsilon)\sqrt{M}+3+h}}}{c_{1,M}} + 2 \frac{2^{K^{M+\gamma-(\sqrt{2}+\epsilon)\sqrt{M}+3+h}}}{2^{K^{M+h-1}}}.$$

Now the inequality is proven if

$$\begin{aligned}
&\left(2^{K^{M+\gamma-(\sqrt{2}+\epsilon)\sqrt{M}+h+4}} - (h+1) 2^{K^{M+\gamma-(\sqrt{2}+\epsilon)\sqrt{M}+3+h}}\right) 2^{K^{M+h-1}} \\
&\geq c_{1,M} 2^{K^{M+\gamma-(\sqrt{2}+\epsilon)\sqrt{M}+3+h}+1}
\end{aligned}$$

which is true for M large by the choice of h, and the fact $c_{1,j} \leq 2^{(V+\delta)K^j}$ for all large j. The proof of inequality (12) is finished. It follows from (11) and (12) that

$$\sum_{n=M}^{\infty} \sum_{j=1}^{R} \alpha_j \frac{d_{j,n}}{c_{j,n}} \leq \frac{2^{K^{M+\beta-(\sqrt{2}+\epsilon)\sqrt{M}+4}}}{c_{1,M}} \tag{13}$$

for every sufficiently large natural number M. Hence, we have

$$\begin{aligned} B = \frac{p}{q} &= \sum_{n=N}^{\infty} \sum_{j=1}^{R} \alpha_j \frac{d_{j,n}}{c_{j,n}} \\ &= \sum_{n=N}^{M-1} \sum_{j=1}^{R} \alpha_j \frac{d_{j,n}}{c_{j,n}} + \sum_{n=M}^{\infty} \sum_{j=1}^{R} \alpha_j \frac{d_{j,n}}{c_{j,n}}. \end{aligned}$$

Thus

$$\begin{aligned} &p.lcm(c_{1,N}, \dots, c_{R,N}, c_{1,N+1}, \dots, c_{R,N+1}, \dots, c_{1,M-1}, \dots, c_{R,M-1}) \\ &= q.lcm(c_{1,N}, \dots, c_{R,M-1}) \sum_{n=N}^{M-1} \sum_{j=1}^{R} \alpha_j \frac{d_{j,n}}{c_{j,n}} \\ &\quad + q.lcm(c_{1,N}, \dots, c_{R,M-1}) \sum_{n=M}^{\infty} \sum_{j=1}^{R} \alpha_j \frac{d_{j,n}}{c_{j,n}}, \end{aligned}$$

where $lcm(x_1, \dots, x_n)$ denotes the least common multiple of numbers $x_1, ...,$ x_n. Thus, the number

$$C = q.lcm(c_{1,N}, \dots, c_{R,M-1}) \sum_{n=M}^{\infty} \sum_{j=1}^{R} \alpha_j \frac{d_{j,n}}{c_{j,n}}$$

is a positive integer. From this and (13) we obtain

$$\begin{aligned} C &= q.lcm(c_{1,N}, \dots, c_{R,M-1}) \sum_{n=M}^{\infty} \sum_{j=1}^{R} \alpha_j \frac{d_{j,n}}{c_{j,n}} \\ &\leq \frac{lcm(c_{1,N}, \dots, c_{R,M-1})}{c_{1,M}} 2^{K^{M+\beta+4-(\sqrt{2}+\epsilon)\sqrt{M}}} = \frac{D}{c_{1,M}} \end{aligned} \tag{14}$$

for every sufficiently large natural number M. From (1) and the definition of the sequence $\{c_{1,n}\}_{n=1}^{\infty}$ we have

$$\begin{aligned} D &= lcm(c_{1,N}, \dots, c_{R,M-1}) 2^{K^{M+\beta+4-(\sqrt{2}+\epsilon)\sqrt{M}}} \\ &\leq \left(\prod_{n=N}^{M-2} 2^{K^{n-2}} \right)^{-1} \left(\prod_{n=N}^{M-1} \prod_{j=1}^{R} c_{j,n} \right) 2^{K^{M+\beta+4-(\sqrt{2}+\epsilon)\sqrt{M}}}. \end{aligned}$$

From this, (7), (10), and the fact $\beta = \gamma + h$ we obtain

$$D \le 2^{\frac{-1}{K-1}(K^{M-3}-K^N)} \left(\prod_{n=N}^{M-1} \prod_{j=1}^{R} 2^{(V+\delta)K^n} 2^{K^{n+\beta+2-(\sqrt{2}+\epsilon)\sqrt{n}}} \right) \cdot 2^{K^{M+\beta+4-(\sqrt{2}+\epsilon)\sqrt{M}}} S(N_1, N_2, \delta),$$

where $S(N_1, N_2, \delta)$ does not depend on M. It follows that

$$\begin{aligned} D &\le 2^{-\frac{(K^{M-3}-K^N)}{K-1}} S(N_1, N_2, \delta) \left(\prod_{n=N}^{M-1} 2^{R(V+\delta)K^n} 2^{RK^{n+\beta+2-(\sqrt{2}+\epsilon)\sqrt{n}}} \right) \\ &\quad \cdot 2^{K^{M+\beta+4-(\sqrt{2}+\epsilon)\sqrt{M}}} \\ &\le 2^{-\frac{K^{M-3}-K^N}{K-1}} S(N_1, N_2, \delta) 2^{R(V+\delta)\frac{K^M-K^N}{K-1}} 2^{K^{M+\beta+5-(\sqrt{2}+\epsilon)\sqrt{M}+\log M}} \\ &\le 2^{-\frac{K^{M-3}-K^N}{K-1}} s(N_1, N_2, \delta) 2^{(V+\delta)K^M} 2^{K^{M+\beta+5-(\sqrt{2}+\epsilon)\sqrt{M}+\log M}}. \end{aligned}$$

Hence,

$$D \le 2^{(V+\delta-K^{-4})K^M}$$

for every sufficiently large M. This, (8), and (14) imply that

$$C = \frac{D}{c_{1,M}} \le 2^{(V+\delta-K^{-4})K^M} . 2^{-(V-\delta)K^M} = 2^{(2\delta-K^{-4})K^M}$$

for infinitely many natural numbers M. But this is impossible for a sufficiently small δ and a sufficiently large M.

2. Secondly, let us assume that

$$\limsup_{n\to\infty} c_{1,n}^{1/K^n} = \infty. \tag{15}$$

Let Q be a sufficiently large positive integer. Let the number of $c_{1,n}$ such that $c_{1,n} < 2^{K^Q}$ be Z. (The definition of the sequence $\{c_{1,n}\}_{n=N}^{\infty}$ and (1) imply that $Z - 1 < Q$.) Let $g(X, Y)$ be the number of $c_{1,n}$ satisfying $c_{1,n} \in [2^{K^Y}, 2^{K^X})$ and put $f(X, Y) = X - g(X, Y)$. Then (15) yields

$$\limsup_{X\to\infty} f(X, Y) = \infty \tag{16}$$

and

$$f(X + 1, Y) - f(X, Y) \le 1. \tag{17}$$

Because of (16) and (17) there is a least positive integer P such that

$$g(P, Q) = P - Q - Z - 2. \tag{18}$$

It follows that for every S $(Q \le S < P)$

$$g(P, S) \le P - S - 1. \tag{19}$$

(Otherwise $g(S,Q) = g(P,Q) - g(P,S) \leq P - Q - Z - 2 - (P - S) = S - Q - Z - 2$ and the number P would not be the least.) Now (18) and (19) imply that for every $j = 0, 1, \dots, P - Q - Z - 3$,

$$c_{1,P-Q-3-j+N} \leq 2^{K^{P-j-1}}.$$

Thus,

$$\prod_{c_{1,j}<2^{K^P}} c_{1,j} = \prod_{j=N}^{P-Q-3+N} c_{1,j} = \prod_{j=N}^{N+Z-1} c_{1,j} \prod_{j=N+Z}^{P-Q-3+N} c_{1,j} \tag{20}$$
$$< 2^{ZK^Q} \prod_{j=N+Z}^{P-Q-3+N} 2^{K^{Q+j-N+2}}$$
$$= 2^{ZK^Q} 2^{\frac{1}{K-1}(K^P - K^{Q+Z+2})} \leq 2^{\frac{1}{K-1}K^P}.$$

Now we define a sequence $\{S_n\}_{n=0}^{\infty}$ by induction in the following way. Let us put $S_0 = P$. Suppose that we have $S_0, S_1, \dots, S_{k-1}$. Because of (16) and (17) there is a least positive integer S_k such that

$$g(S_k, S_{k-1}) = S_k - S_{k-1} - 1. \tag{21}$$

Similarly (21) implies that for every S ($S_{k-1} \leq S \leq S_k$)

$$g(S_k, S) \leq S_k - S - 1. \tag{22}$$

The last inequality implies that for every $j = 1, \dots, S_k - S_{k-1} - 1$

$$c_{1,N+S_{k-1}-Q-2-k+j} \leq 2^{K^{S_{k-1}+j}}.$$

Hence, it follows that

$$\prod_{c_{1,j}\in(2^{K^{S_{k-1}}},2^{K^{S_k}})} c_{1,j} = \prod_{j=1}^{S_k-S_{k-1}-1} c_{1,N+S_{k-1}-Q-2-k+j} \tag{23}$$
$$\leq \prod_{j=1}^{S_k-S_{k-1}-1} 2^{K^{S_{k-1}+j}} = 2^{\frac{1}{K-1}(K^{S_k} - K^{S_{k-1}+1})}.$$

Now we will prove that there are infinitely many positive integers $T \geq P$ such that

$$lcm(c_{1,j}, c_{1,j} < 2^{K^T}) \leq 2^{\frac{1}{K-1}(K^T - K^{T-(\sqrt{2}+\frac{\epsilon}{4})\sqrt{T}})} \tag{24}$$

and

$$\prod_{c_{1,j}<2^{K^T}} c_{1,j} \leq 2^{\frac{1}{K-1}K^T}. \tag{25}$$

To prove this, we will consider three cases.

2.1. First, let us assume that

$$S_k - S_{k-1} < \sqrt{2S_k} \tag{26}$$

for infinitely many numbers k. Then (20), (23), and (26) yield

$$\begin{aligned}\prod_{c_{1,j}<2^{K^{S_k}}} c_{1,j} &= \left(\prod_{c_{1,j}<2^{K^P}} c_{1,j}\right)\left(\prod_{i=1}^{k}\prod_{c_{1,j}\in[2^{K^{S_{i-1}}},2^{K^{S_i}})} c_{1,j}\right)\\ &\le 2^{\frac{1}{K-1}K^P}.\prod_{i=1}^{k} 2^{\frac{1}{K-1}(K^{S_i}-K^{S_{i-1}+1})}\\ &= 2^{\frac{1}{K-1}(K^{S_0}+K^{S_1}-K^{S_0+1}+\dots+K^{S_k}-K^{S_{k-1}+1})}\\ &\le 2^{\frac{1}{K-1}(K^{S_k}-K^{S_{k-1}})} < 2^{\frac{1}{K-1}(K^{S_k}-K^{S_k-\sqrt{2S_k}})}.\end{aligned}$$

Thus (24) and (25) hold under condition (26).

2.2. Secondly, let us assume that for every positive integer k

$$S_k - S_{k-1} \ge \sqrt{2S_k}.$$

It follows that

$$S_k - \sqrt{2S_k} - S_{k-1} \ge 0.$$

Thus,

$$S_k \ge \left(\frac{1}{\sqrt{2}} + \sqrt{\frac{1}{2}+S_{k-1}}\right)^2 = 1 + S_{k-1} + \sqrt{1+2S_{k-1}}. \tag{27}$$

Now, by mathematical induction we prove that

$$S_k \ge \frac{1}{2}k^2. \tag{28}$$

For $k=0$ (28) holds. Suppose that (28) holds for $k-1$. Then (27) and (28) imply

$$\begin{aligned}S_k &\ge 1 + S_{k-1} + \sqrt{1+2S_{k-1}}\\ &\ge 1 + \frac{1}{2}(k-1)^2 + \sqrt{1+2\frac{1}{2}(k-1)^2}\\ &> 1 + \frac{1}{2}k^2 - k + \frac{1}{2} + (k-1) > \frac{1}{2}k^2.\end{aligned}$$

From (18) and (21) the number of $c_{1,j}$ such that $c_{1,j} < 2^{K^{S_k}}$ is equal to

$$g(S_k,0) = Z + g(S_0,Q) + \sum_{j=1}^{k} g(S_j,S_{j-1}) \tag{29}$$
$$= Z + P - Q - Z - 2 + \sum_{j=1}^{k}(S_j - S_{j-1} - 1)$$
$$= P - Q - 2 + S_k - S_0 - k = S_k - k - Q - 2.$$

Now, (28) and (29) imply that

$$g(S_k,0) = S_k - k - Q - 2 \tag{30}$$
$$\geq S_k - \sqrt{2S_k} - Q - 2 \geq S_k - \left(\sqrt{2} + \frac{\epsilon}{2}\right)\sqrt{S_k} + 2$$

for every sufficiently large k. Also (20), (23), and (30) yield

$$\prod_{c_{1,j}<2^{K^{S_k}}} c_{1,j} = \prod_{c_{1,j}<2^{K^P}} c_{1,j} \prod_{i=1}^{k} \prod_{c_{1,j}\in[2^{K^{S_{i-1}}},2^{K^{S_i}})} c_{1,j}$$
$$\leq 2^{\frac{1}{K-1}K^P} \prod_{i=1}^{k} 2^{\frac{1}{K-1}(K^{S_i} - K^{S_{i-1}+1})}$$
$$= 2^{\frac{1}{K-1}(K^P + \sum_{i=1}^{k}(K^{S_i} - K^{S_{i-1}+1}))} \leq 2^{\frac{1}{K-1}K^{S_k}}$$

for every sufficiently large k. From this, (1), (30), and the definition of the sequence $\{c_{1,n}\}_{n=N}^{\infty}$ it follows that

$$lcm(c_{1,j}, c_{1,j} < 2^{K^{S_k}}) \leq 2^{\frac{-1}{K-1}(K^{g(S_k,0)-1} - K^N)} . \prod_{c_{1,j}<2^{K^{S_k}}} c_{1,j}$$
$$\leq 2^{\frac{1}{K-1}(K^{S_k} - K^{S_k - (\sqrt{2}+\frac{\epsilon}{3})\sqrt{S_k}})}$$

for every sufficiently large k.

2.3. Third, let us assume that $S_k - S_{k-1} \leq \sqrt{2S_k}$, and $S_j - S_{j-1} \geq \sqrt{2S_j}$ for every $j > k$. Let us put $P' = S_k = S'_0$, and $S'_j = S_{k+j}$. We now proceed as in the second case with $\{S'_j\}_{j=0}^{\infty}$ in place of $\{S_j\}_{j=0}^{\infty}$. Thus (24) and (25) hold. Now let T be a positive integer such that (24) and (25) hold. Then we obtain from (5) that

$$B.q.lcm(c_{1,N}, \ldots, c_{1,N+g(T,0)-1}, c_{2N}, \ldots, c_{R,N+g(T,0)-1})$$
$$= q.lcm(c_{1,N}, \ldots, c_{R,N+g(T,0)-1}) \sum_{n=N}^{\infty} \sum_{j=1}^{R} \alpha_j \frac{d_{j,n}}{c_{j,n}}.$$

Thus, there is a positive integer E such that

$$E = q.lcm(c_{1,N}, \dots , c_{R,N+g(T,0)-1}) \sum_{n=N+g(T,0)}^{\infty} \sum_{j=1}^{R} \alpha_j \frac{d_{j,n}}{c_{j,n}}. \tag{31}$$

From (1), (4), the definition of the sequence $\{c_{1,n}\}_{n=N}^{\infty}$, (18), (21), (24), and (25) it follows that for infinitely many sufficiently large T

$$\begin{aligned} &lcm(c_{1,N}, \dots , c_{R,N+g(T,0)-1}) \\ &\le lcm(c_{1,N}, \dots , c_{1,N+g(T,0)-1}) \left(\prod_{j=N}^{N+g(T,0)-1} c_{1,j} 2^{K^{T+2-(\sqrt{2}+\epsilon)\sqrt{T}}} \right)^{K-2} \\ &= lcm(c_{1,j}, c_{1,j} < 2^{K^T}) \left(\prod_{c_{1,j}<2^{K^T}} c_{1,j} 2^{K^{T+2-(\sqrt{2}+\epsilon)\sqrt{T}}} \right)^{K-2} \\ &\le 2^{\frac{1}{K-1}\left(K^T - K^{T-(\sqrt{2}+\frac{\epsilon}{4})\sqrt{T}}\right)} \left(2^{\frac{1}{K-1}K^T} 2^{TK^{T+2-(\sqrt{2}+\epsilon)\sqrt{T}}} \right)^{K-2} \\ &= 2^{K^T - \frac{1}{K-1}K^{T-(\sqrt{2}+\frac{\epsilon}{4})\sqrt{T}} + T(K-2)K^{T+2-(\sqrt{2}+\epsilon)\sqrt{T}}} \le 2^{K^T - K^{T-(\sqrt{2}+\frac{\epsilon}{3})\sqrt{T}}}. \end{aligned} \tag{32}$$

On the other hand (1), (2), (4), the definition of the sequence $\{c_{1,n}\}_{n=N}^{\infty}$, (18), and (21) imply that

$$\begin{aligned} \sum_{n=N+g(T,0)}^{\infty} \sum_{j=1}^{R} \alpha_j \frac{d_{j,n}}{c_{j,n}} &\le \frac{T.K. \max_{j=1,\dots,R} |\alpha_j| . 2^{K^{T+2-(\sqrt{2}+\epsilon)\sqrt{T}}}}{2^{K^T}} \\ &\le 2^{K^{T-(\sqrt{2}+\frac{\epsilon}{2})\sqrt{T}} - K^T} \end{aligned} \tag{33}$$

for all sufficienly large T. Finally (31)-(33) imply that

$$\begin{aligned} E &\le q.2^{K^T - K^{T-(\sqrt{2}+\frac{\epsilon}{3})\sqrt{T}}} 2^{K^{T-(\sqrt{2}+\frac{\epsilon}{2})\sqrt{T}} - K^T} \\ &= q.2^{K^{T-(\sqrt{2}+\frac{\epsilon}{2})\sqrt{T}} - K^{T-(\sqrt{2}+\frac{\epsilon}{3})\sqrt{T}}} \end{aligned}$$

for infinitely many natural numbers T. But this is impossible for a positive integer E and a sufficiently large T. □

Example 1. Let $a_{j,n} = 2^{K^n}$, $b_{j,n} = (j+n)!$ $(j = 1, 2, \dots , K-1)$. Then the sequences $\{a_{j,n}/b_{j,n}\}_{n=1}^{\infty}$ are linearly unrelated.

3. Irrational Sequences.

Definition 3.1. Let $\{A_n\}_{n=1}^{\infty}$ be a sequence of positive real numbers. If for every sequence $\{c_n\}_{n=1}^{\infty}$ of positive integers the series

$$\sum_{n=1}^{\infty} \frac{1}{A_n c_n}$$

is irrational, then the sequence $\{A_n\}_{n=1}^{\infty}$ is irrational. If $\{A_n\}_{n=1}^{\infty}$ is not an irrational sequence, then it is a rational sequence.

Theorem 3.1. *Let $\epsilon > 0$, and let $\{a_n\}_{n=1}^{\infty}$ and $\{b_n\}_{n=1}^{\infty}$ be two sequences of positive integers such that*

$$a_n \geq 2^{2^n}$$

and

$$b_n \leq 2^{2^{n-(\sqrt{2}+\epsilon)\sqrt{n}}}.$$

Then the sequence $\left\{\frac{\prod_{i=1}^{n} a_i}{b_n}\right\}_{n=1}^{\infty}$ is irrational and the series $\sum_{n=1}^{\infty} \frac{b_n}{\prod_{i=1}^{n} a_i}$ is irrational too.

This theorem is an immediate consequence of Theorem 2.1. It is enough to put $K = 2$.

Example 2. The sequences $\{2^{2^n - n^2}\}_{n=1}^{\infty}$, $\{2^{2^n}/n\}_{n=1}^{\infty}$, and $\{2^{2^n - n}\}_{n=1}^{\infty}$ are irrational sequences.

Open Problem. Is the sequence $\left\{2^{[2^n(1-\frac{1}{n})]}\right\}_{n=1}^{\infty}$ irrational or not? ($[x]$ denotes the greatest integer less than or equal x.)

Remark. Let us put in Theorem 3.1 $a_n = 2^{2^n}$ and $b_n = 1$ for every natural number n. Then we obtain the very famous result of Erdös (see [5]) which states that the sequence $\{2^{2^n}\}_{n=1}^{\infty}$ is irrational.

From the last theorem we also obtain the following criterion for the so-called Cantor sequences.

Theorem 3.2. *Let $\epsilon > 0$ and let $\{b_n\}_{n=1}^{\infty}$ be a sequence of positive integers such that*

$$b_n \leq 2^{2^{n-(\sqrt{2}+\epsilon)\sqrt{n}}}.$$

Let us put

$$a_n = \left[2^{n\left(1-\frac{1}{n}\log_2\left(\frac{n}{\frac{\log_2 n}{n}+1}\right)\right)} \right].$$

Then the sequence $\{\frac{a_n!}{b_n}\}_{n=1}^{\infty}$ is irrational.

This theorem is an immediate consequence of Theorem 3.1.

Example 3. The sequences $\left\{2^{[n(1-\frac{1}{\sqrt{n}})]}!\right\}_{n=1}^{\infty}$ and $\left\{2^{[n(1-\frac{1}{\sqrt{n}})]}!/n!\right\}_{n=1}^{\infty}$ are irrational.

Thank you very much to reviewer and to Professor Carter for their correction of this paper.

References

[1] R. Apery, *Irrationality of $\zeta(2)$ and $\zeta(3)$*, Soc. Math. France, Asterisque, **61** (1979), 11-13.

[2] J.P. Bezivin, *Linear independence of the values of transcendental solutions of some functional equations*, Manuscripta Math., **61**(**1**) (1988), 103-129.

[3] P.B. Borwein, *On the irrationality of* $\sum(1/q^n+r)$, J. Num. Theory, **37** (1991), 253-259.

[4] P. Bundschuh, *A criterion for algebraic independence with some applications*, Osaka J. Math., **25** (1988), 849-858.

[5] P. Erdös, *Some problems and results on the irrationality of the sum of infinite series*, J. Math. Sci., **10** (1975), 1-7.

[6] J. Hančl, *Criterion for irrational sequences*, J. Num. Theory **43**(**1**) (1993), 88-92.

[7] ———, *Continued fractional algebraic independence of sequences*, Publ. Math., **46**(**1**-**2**) (1995), 27-31.

[8] V.N. Sorokin, *Linear independence of logarithm of some rational numbers*, Mat. Zametki, **46**(**3**) (1989), 74-79, 127, translated in Math. Notes, **46**(**3**-**4**) (1989), 727-730.

Received May 20, 1997 and revised February 26, 1998. This work was supported by Grant no. 201/97/0433 of the Czech Grant Agency.

University of Ostrava
701 03 Ostrava 1
Czech Republic
E-mail address: hancl@osu.cz

PACIFIC JOURNAL OF MATHEMATICS
Vol. 190, No. 2, 1999

LOOP SPACES OF H-SPACES WITH FINITELY GENERATED COHOMOLOGY

Yusuke Kawamoto

Suppose X is a simply connected mod p H-space such that the mod p cohomology $H^*(\Omega X)$ is a finitely generated algebra. We show that the loop space ΩX is homotopy equivalent to a finite product of Eilenberg-MacLane spaces $K(\mathbb{Z},1)$, $K(\mathbb{Z},2)$ and $K(\mathbb{Z}/p^i,1)$ for $i \geq 1$. This is a generalization of the result due to Lin, in which the same result was proved under the assumption that X is an A_p-space.

1. Introduction.

Let p be an odd prime. We assume that all spaces are completed at p in the sense of Bousfield-Kan [**2**], and the cohomologies are taken with $\mathbb{Z}/p$-coefficients unless otherwise specified. In this paper, we investigate the homotopy type for the loop space of an H-space whose cohomology is finitely generated as an algebra. In the case of the cohomology is finite dimensional, there is the following theorem due to Aguadé-Smith:

Theorem 1.1 ([**1**])**.** *If X is a simply connected mod p H-space such that $H^*(\Omega X)$ is finite dimensional, then ΩX has the homotopy type of a torus.*

The above theorem is known as the mod p torus theorem, and some generalizations of Theorem 1.1 are investigated by Hemmi [**8**] and McGibbon [**15**]. Hemmi showed that a connected finite quasi C_p-space has the homotopy type of a torus, where a quasi C_p-space is defined as an H-space which has certain higher homotopy associativity and commutativity (see [**8**, Def. 2.1]).

Our main result is stated as follows:

Theorem A. *If X is a simply connected mod p H-space such that $H^*(\Omega X)$ is finitely generated as an algebra, then ΩX is homotopy equivalent to a finite product of Eilenberg-MacLane spaces $K(\mathbb{Z},1)$, $K(\mathbb{Z},2)$ and $K(\mathbb{Z}/p^i,1)$ for $i \geq 1$.*

Theorem A generalizes Theorem 1.1 since $K(\mathbb{Z},2)$ and $K(\mathbb{Z}/p^i,1)$ for $i \geq 1$ do not have the finite cohomology. Our theorem also generalizes a result of Lin [**12**] who has shown Theorem A under the assumption that X is an A_p-space in the sense of Stasheff [**19**]. We owe much to the results in [**12**] and [**13**] (see §2). From the result of Hemmi, it may be possible to generalize our result to the case of quasi C_p-spaces instead of loop spaces on H-spaces.

For $p = 2$, there is the following more general result due to Slack and Broto-Crespo:

Theorem 1.2 ([**18**, Cor. 0.2], [**3**, Cor. 1.5])**.** *If X is a connected homotopy commutative* mod 2 *H-space such that the* mod 2 *cohomology $H^*(X)$ is finitely generated as an algebra, then X is homotopy equivalent to a finite product of Eilenberg-MacLane spaces $K(\mathbb{Z}, 1)$, $K(\mathbb{Z}, 2)$ and $K(\mathbb{Z}/2^i, 1)$ for $i \geq 1$.*

We remark that for the odd prime case, the corresponding result of Theorem 1.2 does not hold. In fact, Iriye-Kono [**9**] have shown that for an odd prime p, any mod p H-space possesses a multiplication which is homotopy commutative. Moreover, one may guess that a homotopy commutative mod p loop space which has the finitely generated cohomology is homotopy equivalent to a product of Eilenberg-MacLane spaces. However, we note that $Sp(2)$ for $p = 3$ and S^3 for $p \geq 5$ are counterexamples (see [**14**, Thm. 2]).

In the proof of Theorem A, we use a technique for H-fibrations introduced by Broto-Crespo [**3**]. Their observation was concentrated on the mod 2 case, and some parts of their proof have generalizations to the odd prime cases with simple modifications (see Proposition 3.3 and Proposition 3.6). We combine these results with the computations in §2 for the cohomology of ΩX to establish a proof of Theorem A (see §4).

Now we provide an outline of the proof of Theorem A so that the reader has an overview of the ideas and strategy.

For a mod p H-space X satisfying the assumption, we consider the three-connected cover $\tilde{X}$. Then we have a fibration

$$\Omega\tilde{X} \longrightarrow \Omega X \longrightarrow K,$$

where K is a finite product of Eilenberg-MacLane spaces of degrees 1 and 2. We see that $H^*(\Omega\tilde{X})$ is free commutative, finitely generated as an algebra which has generators in degrees $2p, 2p+1, 2p^2$ and $2p^2+1$ with certain Steenrod relations induced from $H^*(\Omega X)$ (see Proposition 2.3). For a generator x of degree $2p$, using the Lannes theory, we construct an H-map $\phi : B\mathbb{Z}/p \to \Omega\tilde{X}$ such that $\phi^*(x) = \omega^p$, where $\omega \in H^2(B\mathbb{Z}/p)$ denotes the generator. We construct an H-fibration

$$B\mathbb{Z}/p \xrightarrow{\phi} \Omega\tilde{X} \longrightarrow E_1,$$

where E_1 is an H-space given by the Borel construction for ϕ. By repeating this construction, we have a sequence of H-spaces and H-maps

$$\Omega\tilde{X} \longrightarrow E_1 \longrightarrow E_2 \longrightarrow \cdots ,$$

and if we set $Y = \varinjlim_s E_s$, then the three-connected cover $Y\langle 3\rangle \simeq \Omega\tilde{X}$, and the cohomology $H^*(Y)$ is related to $H^*(\Omega\tilde{X})$ in that $H^*(Y)$ has an additional three dimensional generator and one less $2p$-dimensional generators (see Proposition 4.1). Applying this procedure a finite number of times, we obtain a mod p H-space Z such that $Z\langle 3\rangle \simeq \Omega\tilde{X}$ and the cohomology $H^*(Z)$ has no $2p$-dimensional generator.

By using the same methods, we can knock off the $2p^2$-dimensional generators, and thus we obtain a mod p H-space W such that $W\langle 3\rangle \simeq \Omega\tilde{X}$ and the cohomology $H^*(W)$ is an exterior algebra with generators in degrees 3 and $2p+1$ (see Proposition 4.8).

By the localization theory due to Dror Farjoun and Neisendorfer, we can show that W is also the loop space on an H-space, and so W is contractible by Theorem 1.1. This implies that $\Omega\tilde{X}$ is also contractible and therefore $\Omega X \simeq K$. The ideas and strategy come from [**3**].

This paper is organized as follows: In §2, we prove Theorem A using Theorem 1.1, Proposition 2.6 and results for the localization theory due to Dror Farjoun [**6**] and Neisendorfer [**17**]. Here Proposition 2.6 is the key to the proof of Theorem A, and we postpone the proof until §4. In §3, we recall the Lannes theory and show some properties for H-fibrations. In §4, we prove Proposition 2.6 using the results of §3.

The author would like to thank Prof. Y. Hemmi, Prof. K. Ishiguro and Prof. M. Imaoka for their many helpful suggestions and conversations. We also appreciate the referee for many useful comments.

2. Proof of Theorem A.

In this section we prove Theorem A. Thus, throughout this section, the space X is always assumed to satisfy the hypothesis of Theorem A. First, we recall the following result due to Lin:

Theorem 2.1 ([**12**, Thm. A]). *$H^*(\Omega X)$ is free commutative, primitively generated on generators in degrees $1, 2, 2p, 2p+1, 2p^2$ and $2p^2+1$.*

Remark 2.2. Lin has proved that if X is a simply connected mod p H-space such that $H^*(\Omega X)$ is finitely generated as an algebra, then $H^*(\Omega X)$ is primitively generated on generators in degrees $1, 2, 2p, 2p+1, 2p^2$ and $2p^2+1$, and under the assumption that $H_*(X)$ is associative, $H^*(\Omega X)$ is free commutative. We note that his proof does not use this assumption to show that $H^*(\Omega X)$ is primitively generated (see [**12**, Cor. 2.2, Thm. 2.3]). But we see that the last statement also holds without this assumption. In fact, since ΩX is homotopy associative, homotopy commutative H-space and $H^*(\Omega X)$ is primitively generated, by a theorem of Browder [**4**, Thm. 8.15], $H^*(\Omega X)$ is free commutative.

By [12, Cor. 1.2, Thm. 2.1], $H^*(X)$ is generated by odd degree generators in degrees $2p^j+1$ for some $j \geq 0$ and even degree generators in degrees 2, $2p^j+2$ for some $j \geq 1$. We choose the basis of $H^2(X)$ and $H^3(X)$ as $B_2 = \{x_{j_0}\} \cup \{x_{j_1}\} \cup \cdots \cup \{x_{j_r}\}$ and $B_3 = \{\beta_1(x_{j_1})\} \cup \cdots \cup \{\beta_r(x_{j_r})\} \cup \{y_k\}$, respectively, where x_{j_0} are the mod p reductions of the integral classes and β_s denotes the s-th Bockstein operation. We define a generalized Eilenberg-MacLane space K as

$$K = \prod_{\{x_{j_0}\}} K(\mathbb{Z},2) \times \prod_{\{x_{j_1}\}} K(\mathbb{Z}/p,2) \times \cdots \times \prod_{\{x_{j_r}\}} K(\mathbb{Z}/p^r,2) \times \prod_{\{y_k\}} K(\mathbb{Z},3).$$

Let $f : X \to K$ be an H-map which represents the generators of the integral cohomology of dimension 2 and 3, and $\tilde{X}$ denote the homotopy fiber of f. Then, $\tilde{X}$ is an H-space and $2p$-connected. By the spectral sequence argument, we see that $H^*(\Omega\tilde{X})$ is finitely generated as an algebra, and so $\tilde{X}$ satisfies the same conditions as X.

Now we define an algebra A as

$$A = \mathbb{Z}/p[x_1, \ldots, x_m, y_1, \ldots, y_n] \otimes \Lambda(z_1, \ldots, z_{m+n}, w_1, \ldots, w_n),$$

where $|x_i| = 2p$ for $1 \leq i \leq m$, $|y_j| = 2p^2$ for $1 \leq j \leq n$, $|z_k| = 2p+1$ for $1 \leq k \leq m+n$, and $|w_l| = 2p^2+1$ for $1 \leq l \leq n$.

Then we can prove the following proposition:

Proposition 2.3. *$H^*(\Omega\tilde{X}) \cong A$ as algebras, and the following operations act on $H^*(\Omega\tilde{X})$:*

$$\begin{cases} \beta(x_i) = z_i & \text{for } 1 \leq i \leq m, \\ \beta(y_j) = \mathcal{P}^p(z_{m+j}) = w_j & \text{for } 1 \leq j \leq n. \end{cases} \tag{2.4}$$

For a mod p H-space Y, we denote the primitive and indecomposable modules of the Hopf algebra $H^*(Y)$ by $PH^*(Y)$ and $QH^*(Y)$, respectively. We need the following fact for the proof of Proposition 2.3.

Lemma 2.5 ([16, Thm. 4.21])**.** *If Y is a connected* mod p *H-space, then there is the following exact sequence:*

$$0 \to P(\xi H^*(Y)) \to PH^*(Y) \to QH^*(Y),$$

where $\xi : H^*(Y) \to H^*(Y)$ is a map defined as $\xi(x) = x^p$.

Proof of Proposition 2.3. By Theorem 2.1, $H^*(\Omega\tilde{X})$ is free commutative, and has generators x_i, y_j, z_k and w_l with $|x_i| = 2p$, $|y_j| = 2p^2$, $|z_k| = 2p+1$ and $|w_l| = 2p^2+1$ for $1 \leq i \leq m$, $1 \leq j \leq n$, $1 \leq k \leq q$ and $1 \leq l \leq r$, where generators are primitive.

Since $\beta(x_i) \in PH^{2p+1}(\Omega\tilde{X})$ for $1 \leq i \leq m$, we see that $\beta(x_i) \in QH^{2p+1}(\Omega\tilde{X})$ by Lemma 2.5. By [13, Cor. E], we have that $\beta(x_i) \neq 0$, and if $i_1 \neq i_2$, then $\beta(x_{i_1}) \neq \beta(x_{i_2})$. Thus, we can set $\beta(x_i) = z_i$ for $1 \leq i \leq m$. Similarly, we can set $\beta(y_j) = w_j$ for $1 \leq j \leq n$.

Since the suspension map $\sigma^* : QH^{2p^2+2}(\tilde{X}) \to PH^{2p^2+1}(\Omega\tilde{X})$ is an epimorphism, and $\beta : QH^{2p^2+1}(\tilde{X}) \to QH^{2p^2+2}(\tilde{X})$ is also an epimorphism by [**12**, Thm. 1.10], we have that $w_l \in \beta PH^{2p^2}(\Omega\tilde{X})$ for $1 \le l \le r$. Thus, we have that $w_l \in \beta QH^{2p^2}(\Omega\tilde{X})$ by Lemma 2.5, which implies that $r = n$.

Using [**12**, Thm. 1.9], the similar arguments show that $w_l \in \mathcal{P}^p QH^{2p+1}(\Omega\tilde{X})$ for $1 \le l \le n$. We can assume that $\mathcal{P}^p(z_{m+l}) = w_l$ for $1 \le l \le n$ since $\mathcal{P}^p(z_k) = \mathcal{P}^1\beta\mathcal{P}^{p-1}(x_k) + \beta\mathcal{P}^p(x_k) = 0$ for $1 \le k \le m$.

If we set

$$\mathcal{P}^p(z_{m+n+1}) = \sum_{l=1}^{n} \sigma_l w_l$$

for $\sigma_l \in \mathbb{Z}/p$, then for

$$\zeta = z_{m+n+1} - \sum_{l=1}^{n} \sigma_l z_{m+l},$$

we have that $\mathcal{P}^p(\zeta) = 0$. Since $\sigma^* : QH^{2p+2}(\tilde{X}) \to PH^{2p+1}(\Omega\tilde{X})$ is an epimorphism, $\zeta = \sigma^*(\mu)$ for some $\mu \in QH^{2p+2}(\tilde{X})$. Since $\sigma^*(\mathcal{P}^p(\mu)) = \mathcal{P}^p(\zeta) = 0$, by [**10**, Thm. B], there exists $\nu \in QH^{2p+1}(\tilde{X})$ such that $\mathcal{P}^p(\mu) = \beta\mathcal{P}^p(\nu)$ in $QH^{2p^2+2}(\tilde{X})$. Applying the Adem relation $\mathcal{P}^p\beta = \mathcal{P}^1\beta\mathcal{P}^{p-1} + \beta\mathcal{P}^p$ to ν, we have that $\mathcal{P}^p(\mu) = \mathcal{P}^p(\beta(\nu))$, which implies that $\mu = \beta(\nu)$ by [**12**, Thm. 1.9]. Then, $\zeta = \sigma^*(\mu) = \beta(\sigma^*(\nu)) \in \beta QH^{2p}(\Omega\tilde{X})$ by Lemma 2.5, which implies that

$$\zeta = \sum_{k=1}^{m} \tau_k z_k$$

for $\tau_k \in \mathbb{Z}/p$. Therefore, we have that

$$z_{m+n+1} = \sum_{k=1}^{m} \tau_k z_k + \sum_{l=1}^{n} \sigma_l z_{m+l},$$

which implies that $q = m + n$. This completes the proof. □

The following proposition is crucial for our study, which will be proved in §4 using the Lannes theory.

Proposition 2.6. *If Y is a* mod p *H-space with $H^*(Y) \cong A$ as algebras, and the operations* (2.4) *act on $H^*(Y)$, then there is a simply connected* mod p *finite H-space W such that $Y \simeq W\langle 3\rangle$, where $W\langle 3\rangle$ is the three-connected cover of W.*

Using Proposition 2.6, we can prove Theorem A as follows:

Proof of Theorem A. By Proposition 2.3 and Proposition 2.6, there exists a simply connected mod p finite H-space W such that $\Omega\tilde{X} \simeq W\langle 3\rangle$. Let L_g denote the localization functor with respect to a map g constructed by

Dror Farjoun [6]. For the constant map $c : B\mathbb{Z}/p \to *$, $L_c(\Omega\tilde{X}) \simeq W$ by the results due to Neisendorfer [17, Thm. 0.1]. Since $L_c(\Omega\tilde{X}) \simeq \Omega L_{\Sigma c}(\tilde{X})$ by [6, Thm. 3.A.1], and $L_{\Sigma c}$ preserves the H-structure, we see that the space W is the loop space of an H-space. By Theorem 1.1, W is contractible, and so $\Omega\tilde{X} \simeq W\langle 3\rangle$ is also contractible. Therefore, $\Omega X \simeq \Omega K$, and we have the required conclusion. This completes the proof of Theorem A. $\square$

3. Lannes T-functor and H-fibrations.

In this section we recall some results concerning the Lannes theory and the H-fibrations, which will be used in the next section.

Let $\mathcal{K}$ denote the category of unstable $\mathcal{A}_p$-algebras. The objects of $\mathcal{K}$ are called $\mathcal{K}$-algebras. It is known that $H^*(X)$ is a $\mathcal{K}$-algebra for any space X.

The Lannes T-functor $T : \mathcal{K} \to \mathcal{K}$ is a left adjoint of the functor $H^*(B\mathbb{Z}/p) \otimes -$, that is, there is the adjoint isomorphism $\mathrm{Hom}_{\mathcal{K}}(T(A), B) \cong \mathrm{Hom}_{\mathcal{K}}(A, H^*(B\mathbb{Z}/p) \otimes B)$ for $\mathcal{K}$-algebras A and B.

For a $\mathcal{K}$-map $f : A \to H^*(B\mathbb{Z}/p)$, its adjoint restricts to a $\mathcal{K}$-map $T(A)^0 \to \mathbb{Z}/p$, where $T(A)^0$ is the subalgebra of $T(A)$ of elements of degree 0. The connected component of $T(A)$ corresponding to f is defined by $T_f(A) = T(A) \otimes_{T(A)^0} \mathbb{Z}/p$, and there is the natural $\mathcal{K}$-map $\epsilon_f : A \to T_f(A)$.

The evaluation map $e : B\mathbb{Z}/p \times \mathrm{Map}(B\mathbb{Z}/p, X) \to X$ induces a $\mathcal{K}$-map e^*, and taking the adjoint of this yields a $\mathcal{K}$-map $\lambda : T(H^*(X)) \to H^*(\mathrm{Map}(B\mathbb{Z}/p, X))$. On the component level, for a map $\phi : B\mathbb{Z}/p \to X$, there is a $\mathcal{K}$-map $\lambda_{\phi^*} : T_{\phi^*}(H^*(X)) \to H^*(\mathrm{Map}(B\mathbb{Z}/p, X)_\phi)$. The composite $\lambda_{\phi^*}\epsilon_{\phi^*}$ is induced by the evaluation at the base point $e_\phi : \mathrm{Map}(B\mathbb{Z}/p, X)_\phi \to X$. The following theorem is due to Lannes:

Theorem 3.1 ([11, Thm. 3.2.1]). *Let X be a space and $\phi : B\mathbb{Z}/p \to X$ be a map. If $T_{\phi^*}(H^*(X))^1 = 0$, then $\lambda_{\phi^*} : T_{\phi^*}(H^*(X)) \to H^*(\mathrm{Map}(B\mathbb{Z}/p, X)_\phi)$ is an isomorphism.*

For the cohomology of an H-space, Dwyer-Wilkerson have proved the following:

Proposition 3.2 ([7, Thm. 3.2, Lemma 4.5]). *If X is a* mod p *H-space with finitely generated cohomology and $f : H^*(X) \to H^*(B\mathbb{Z}/p)$ is a $\mathcal{K}$-map, then $\epsilon_f : H^*(X) \to T_f(H^*(X))$ is an isomorphism.*

Recently, an important theory of H-fibrations using the Lannes theory was introduced by Broto-Crespo [3]. Their observation was concentrated on the mod 2 case. However, we also have the corresponding results for the odd prime case.

Proposition 3.3. *Let X be a* mod p *H-space with finitely generated cohomology, and $\phi : B\mathbb{Z}/p \to X$ be an H-map with $H^*(B\mathbb{Z}/p)$ is finitely* *generated $H^*(X)$-module induced by ϕ^*. If*

$$B\mathbb{Z}/p \xrightarrow{\phi} X \xrightarrow{\psi} Y$$

is a principal fibration, then Y is an H-space and ψ is an H-map.

Lemma 3.4. *Let $c : B\mathbb{Z}/p \times B\mathbb{Z}/p \to Y$ denote the constant map, where Y comes from Proposition 3.3. Then the base point evaluation map e_c : $\mathrm{Map}(B\mathbb{Z}/p \times B\mathbb{Z}/p, Y)_c \to Y$ is a homotopy equivalence.*

Proof. We have the following commutative diagram of fibrations:

$$\begin{tikzcd}
\mathrm{Map}(B\mathbb{Z}/p \times B\mathbb{Z}/p, B\mathbb{Z}/p)_S \arrow[r, "e_c"] \arrow[d] & B\mathbb{Z}/p \arrow[d, "\phi"] \\
\mathrm{Map}(B\mathbb{Z}/p \times B\mathbb{Z}/p, X)_c \arrow[r, "e_c"] \arrow[d] & X \arrow[d, "\psi"] \\
\mathrm{Map}(B\mathbb{Z}/p \times B\mathbb{Z}/p, Y)_c \arrow[r, "e_c"] & Y,
\end{tikzcd}$$

where $S = \{g : B\mathbb{Z}/p \times B\mathbb{Z}/p \to B\mathbb{Z}/p \mid \phi g \simeq c\}$ and e_c denote the base point evaluation maps.

Since X has the finitely generated cohomology, $e_c : \mathrm{Map}(B\mathbb{Z}/p \times B\mathbb{Z}/p, X)_c \to X$ is a homotopy equivalence by [7, Thm. 3.2]. It is known that $H^*(B\mathbb{Z}/p) \cong \Lambda(\theta) \otimes \mathbb{Z}/p[\omega]$ with $\beta(\theta) = \omega$. For a map $g : B\mathbb{Z}/p \times B\mathbb{Z}/p \to B\mathbb{Z}/p$ with $\phi g \simeq c$, there exists some $n \geq 1$ so that $g^*(\omega)^n = g^*(\omega^n) = 0$ since $H^*(B\mathbb{Z}/p)$ is finitely generated $H^*(X)$-module induced by ϕ^* and $g^*\phi^* = 0$, which implies that $g^*(\omega) = 0$. If we put $g^*(\theta) = a_1\theta_1 + a_2\theta_2$ for $a_1, a_2 \in \mathbb{Z}/p$, then $g^*(\omega) = \beta(g^*(\theta)) = a_1\omega_1 + a_2\omega_2 = 0$, and we must have $a_1 = a_2 = 0$, which implies that $g^*(\theta) = 0$. By a result of Lannes [11, Thm. 3.1.1], we obtain that $g \simeq c$.

Then we have that $S = \{c\}$, and thus $e_c : \mathrm{Map}(B\mathbb{Z}/p \times B\mathbb{Z}/p, B\mathbb{Z}/p)_S \to B\mathbb{Z}/p$ is a homotopy equivalence. Using the five lemma, $e_c : \mathrm{Map}(B\mathbb{Z}/p \times B\mathbb{Z}/p, Y)_c \to Y$ is a homotopy equivalence, and thus we have the required conclusion. □

For the proof of Proposition 3.3, we need the following fact which is known as the Zabrodsky lemma:

Lemma 3.5 ([21, Lemma 3.1])**.** *Let*

$$F \xrightarrow{i} E \xrightarrow{p} B$$

be a principal fibration, and Y be a space which satisfies that $e_c : \mathrm{Map}(F, Y)_c \to Y$ is a homotopy equivalence. Then the induced map $\mathrm{Map}(B, Y) \to \mathrm{Map}(E, Y)_S$ is a homotopy equivalence, where $S = \{g : E \to Y \mid gi \simeq c\}$.

Now we can prove Proposition 3.3 as follows:

Proof of Proposition 3.3. By Lemma 3.4, the evaluation map e_c : $\mathrm{Map}(B\mathbb{Z}/p \times B\mathbb{Z}/p, Y)_c \to Y$ is a homotopy equivalence. Then, applying Lemma 3.5 to a principal fibration

$$B\mathbb{Z}/p \times B\mathbb{Z}/p \xrightarrow{\phi\times\phi} X \times X \xrightarrow{\psi\times\psi} Y \times Y,$$

we have that $\mathrm{Map}(Y \times Y, Y) \simeq \mathrm{Map}(X \times X, Y)_S$, where $S = \{g : X \times X \to Y \mid g(\phi\times\phi) \simeq c\}$. If we denote the multiplication of the H-space X as μ_X, then there is a map $\mu_Y : Y \times Y \to Y$ so that $\psi\mu_X \simeq \mu_Y(\psi\times\psi)$. Using Lemma 3.5 again, we see that the map μ_Y gives an H-structure on Y. This completes the proof. □

Proposition 3.6. *Suppose that there is an H-fibration*

$$B\mathbb{Z}/p \xrightarrow{\phi_i} X_i \xrightarrow{\psi_i} X_{i+1} \tag{3.7}$$

for $i \geq 0$, and we put $Y = \varinjlim_i X_i$. If $H^*(Y)$ is finitely generated as an algebra, then the space Y has an H-structure.

Proof. We set $\mu = \varinjlim_i \mu_i : Y \times Y \to Y$ for the multiplication $\mu_i : X_i \times X_i \to X_i$ of the H-space X_i. Let $\iota_j : Y \to Y \times Y$ denote the inclusion map on the j-th factor for $j = 1, 2$. If we show that $\mu\iota_j \simeq 1_Y$ for $j = 1, 2$, then we have the required conclusion.

We denote the inclusion map as $\kappa_i : X_i \to Y$ for $i \geq 0$. Since μ_i is a multiplication for $i \geq 0$, we have that $\mu\iota_j\kappa_i \simeq \kappa_i\mu_i\iota^i_j \simeq \kappa_i$, where $\iota^i_j : X_i \to X_i \times X_i$ denotes the inclusion map on the j-th factor for $j = 1, 2$. By [**20**, Prop. 4], the obstruction to construct a homotopy between $\mu\iota_j$ and 1_Y lies in

$$\varprojlim_i{}^k \pi_k(\mathrm{Map}(X_i, Y)_{\kappa_i}) \tag{3.8}$$

for $k \geq 1$. Since $H^*(Y)$ is a finitely generated algebra, $\mathrm{Map}(B\mathbb{Z}/p, Y)_c \simeq Y$ by [**7**, Thm. 3.2]. Then, applying Lemma 3.5 to the fibration (3.7), we have that $\mathrm{Map}(X_i, Y)_{\kappa_i} \simeq \mathrm{Map}(X_{i+1}, Y)_{\kappa_{i+1}}$ for $i \geq 0$, and so the obstruction group (3.8) vanishes. This completes the proof. □

Now we introduce a result which is useful to compute the Serre spectral sequence for an H-fibration, which will be used in §4. Let X and Y be H-spaces and

$$X \longrightarrow Y \longrightarrow B^2\mathbb{Z}/p$$

be an H-fibration. We consider the Serre spectral sequence for the fibration whose E_2-term is given as

$$E_2^{*,*} = H^*(B^2\mathbb{Z}/p) \otimes H^*(X). \tag{3.9}$$

Then we see that the spectral sequence has a differential Hopf algebra structure, and for $r \geq 2$, if we put $A_r = E_r^{*,0}$ and $B_r = E_r^{0,*}$, then they have Hopf algebra structures induced from the E_2-term.

Proposition 3.10. (1) *If $d_r(B_r) \neq 0$, then the transgression $\tau : B_r^{r-1} \to P^r(A_r)$ is non-trivial.*

(2) *For $x \in B_r^q$, $d_r(x) \in P^r(A_r) \otimes B_r^{q-r+1}$, where $P^n(A_r)$ denotes the primitive module of A_r^n.*

We need the following lemma to show Proposition 3.10:

Lemma 3.11. (1) *For $r \geq 2$, the E_r-term is given as*

$$E_r^{*,*} \cong A_r \otimes B_r \otimes \Lambda(\alpha_1, \dots, \alpha_k),$$

where $\alpha_i \in E_r^{s_i,t_i}$ with $s_i < r$ and $|\alpha_i| = 2m_i + 1$ with $p|m_i$ for $1 \leq i \leq k$.

(2) *If $x \in P^{2s}(A_r)$ with $2s \geq r$, then x has the infinite height.*

Proof. We show (1) and (2) by induction. For $r = 2$, by (3.9) and since $H^*(B^2\mathbb{Z}/p)$ is free commutative, the results (1) and (2) hold. We assume that the results (1) and (2) have already shown for the E_r-term.

By a result of Browder [**5**, Thm. 5.8], the E_{r+1}-term is described as

$$E_{r+1}^{*,*} \cong A_{r+1} \otimes B_{r+1} \otimes \Lambda(\alpha_1, \dots, \alpha_k, \beta_1, \dots, \beta_l),$$

where α_i are elements stated in the lemma, $\beta_j \in E_{r+1}^{s_j,t_j}$ with $t_j < r - 1$ and $|\beta_j| = 2n_j - 1$ with $p|n_j$ for $1 \leq j \leq l$. From the proof of [**5**, Thm. 5.8], we see that $\beta_j = \{x_j \cdot d_r(x_j)^{p-1}\}$ for some $x_j \in E_r^{*,*}$ with $d_r(x_j) \in P(A_r)$. But by assumption, $d_r(x_j)$ has the infinite height, and so the element of the form β_j cannot occur, which shows (1).

For a non-trivial element $x \in P^{2s}(A_{r+1})$ with $2s \geq r + 1$, we assume that $x^{p^k} = 0$ for some $k \geq 1$, and obtain a contradiction from this assumption. By inductive hypothesis, $x^{p^k} \neq 0 \in A_r$, and then there exists an element $y \in E_r^{*,*}$ so that $d_r(y) = x^{p^k}$. By the form of the E_r-term, we have either a generator $z \in B_r^{r-1}$ with $d_r(z) = x^{p^{k_1}}$ for some $k_1 \leq k$ or a generator $\alpha \in E_r^{*,*}$ with $d_r(\alpha) = x^{p^{k_2}}$ for some $k_2 \leq k$. On the one hand, if $d_r(z) = x^{p^{k_1}}$, then $|x^{p^{k_1}}| = r < |x|$, which causes a contradiction. On the other hand, if $d_r(\alpha) = x^{p^{k_2}}$, then $|x^{p^{k_2}}| = 2m + 2$ for some $m \geq 1$ with $p|m$. This shows that $k_2 = 0$, and so $\{x\} = 0$ in the E_{r+1}-term, which also causes a contradiction. This completes the proof. ∎

Now we can prove Proposition 3.10 as follows:

Proof of Proposition 3.10. First we show (1). By assumption, there is an element $x \in B_r^q$ so that $d_r(x) \neq 0$. We can assume that if $y \in B_r^{\bar{q}}$ with $\bar{q} < q$, then $d_r(y) = 0$. If we set that

$$\Delta(x) = x \otimes 1 + 1 \otimes x + \sum_i \bar{x}_i \otimes \bar{\bar{x}}_i,$$

then $\Delta(d_r(x)) = d_r(\Delta(x)) = d_r(x) \otimes 1 + 1 \otimes d_r(x)$, and so $d_r(x) \in P(E_r^{r,q-r+1})$. By Lemma 3.11, the primitive elements of $E_r^{*,*}$ consist of $P(A_r)$, $P(B_r)$ and

$\alpha_i \in E_r^{s,t}$ with $s < r$, and then we have that $q = r-1$ and $d_r(x) \in P^r(A_r)$. This implies the required conclusion.

Next to show (2), let $x \in B_r^q$. By Lemma 3.11, we can set that

$$d_r(x) = \sum_{i \geq 1} a_i \cdot b_i \in E_r^{r,q-r+1},$$

where $a_i \in A_r \otimes \Lambda(\alpha_1, \dots, \alpha_k)$, $b_i \in B_r$ with $|a_i| + |b_i| = q+1$ for $i \geq 1$. We can assume that the elements b_i are linearly independent for $i \geq 1$. Then we have that

$$\begin{aligned}\Delta(d_r(x)) &= \sum_i \Delta(a_i)\Delta(b_i) \\ &= \sum_i \left(a_i \otimes 1 + 1 \otimes a_i + \sum_j \bar{a}_{i,j} \otimes \bar{\bar{a}}_{i,j}\right) \\ &\qquad \cdot \left(b_i \otimes 1 + 1 \otimes b_i + \sum_k \bar{b}_{i,k} \otimes \bar{\bar{b}}_{i,k}\right),\end{aligned}$$

where $0 < |\bar{a}_{i,j}|, |\bar{\bar{a}}_{i,j}| < |a_i|$ and $0 < |\bar{b}_{i,k}|, |\bar{\bar{b}}_{i,k}| < |b_i|$. On the other hand, we obtain that

$$\Delta(d_r(x)) = d_r(\Delta(x)) \in \bigoplus_{s+t=q} (E_r^{r,s-r+1} \otimes B_r^t) \oplus (B_r^s \otimes E_r^{r,t-r+1}).$$

For the dimensional reason, we see that $\sum_{i,j} \bar{a}_{i,j} b_i \otimes \bar{\bar{a}}_{i,j} = 0$, which implies that $\sum_j \bar{a}_{i,j} \otimes \bar{\bar{a}}_{i,j} = 0$ for $i \geq 1$ since b_i are linearly independent. This implies that

$$a_i \in P(A_r \otimes \Lambda(\alpha_1, \dots, \alpha_k)) = P(A_r) \oplus \{\alpha_1, \dots, \alpha_k\},$$

and then $a_i \in P^r(A_r)$ for $i \geq 1$. Thus we can conclude that $d_r(x) \in P^r(A_r) \otimes B_r^{q-r+1}$. This completes the proof. $\square$

Remark 3.12. We note that by Proposition 3.10, for $r \geq 2$, if either $P^r(A_r) = 0$ or $Q^{r-1}(B_r) = 0$, then $d_r(B_r) = 0$.

4. Proof of Proposition 2.6.

In this section we prove Proposition 2.6, and thus we assume that Y is a mod p H-space such that $H^*(Y) \cong A$, and the operations (2.4) act on $H^*(Y)$.

For $1 \leq t \leq m+1$, we set an algebra K_t as

$$\begin{aligned}K_t = \mathbb{Z}/p[x_t, \dots, x_m, y_1, \dots, y_n] \\ \otimes \Lambda(v_1, \dots, v_{t-1}, z_t, \dots, z_{m+n}, w_1, \dots, w_n)\end{aligned}$$

with x_i, y_j, z_k and w_l are as in A, $|v_q| = 3$ for $1 \leq q \leq t-1$. First, we prove the following proposition:

Proposition 4.1. *For $1 \le t \le m+1$, there is a* mod p *H-space Y_t such that $Y_t\langle 3\rangle \simeq Y$ and $H^*(Y_t) \cong K_t$ with the following operations:*

$$\begin{cases} \beta(x_i) = z_i & \text{for } t \le i \le m, \\ \beta(y_j) = w_j & \text{for } 1 \le j \le n, \\ \mathcal{P}^p(z_{m+l}) = w_l + \delta_l & \text{for } 1 \le l \le n, \end{cases}$$

where δ_l is some decomposable element of K_t for $1 \le l \le n$.

For $1 \le t \le m$, we set an algebra C_t as

$$C_t = \mathbb{Z}/p[u, x_{t+1}, \dots, x_m, y_1, \dots, y_n] \otimes \Lambda(v_1, \dots, v_t, z_{t+1}, \dots, z_{m+n}, w_1, \dots, w_n)$$

with x_i, y_j, z_k, w_l and v_q for $1 \le q \le t-1$ are as in K_t, $|u| = 2$ and $|v_t| = 3$.

An algebra A is said to be a $\mathcal{K}$-Hopf algebra if A is a $\mathcal{K}$-algebra and has a Hopf algebra structure compatible with the $\mathcal{K}$-structure, namely the diagonal map of A becomes a $\mathcal{K}$-map. It is known that for an H-space X, $H^*(X)$ is a $\mathcal{K}$-Hopf algebra. We see that if K_t and C_t have $\mathcal{K}$-Hopf algebra structures, then for the dimensional reason, v_q is primitive for $1 \le q \le t$.

Lemma 4.2. *Suppose that the algebras K_t and C_t are $\mathcal{K}$-Hopf algebras with the following operations:*

$$\begin{cases} \beta(u) = \lambda v_t & \text{for } \lambda = 0 \text{ or } 1, \\ \beta(x_i) = z_i & \text{for } t \le i \le m, \\ \beta(y_j) = w_j & \text{for } 1 \le j \le n, \\ \mathcal{P}^p(z_{m+l}) = w_l + \delta_l & \text{for } 1 \le l \le n, \end{cases} \tag{4.3}$$

where δ_l is some decomposable element of K_t for $1 \le l \le n$. Then the following hold:

(1) *There is a map of $\mathcal{K}$-Hopf algebra $f : K_t \to H^*(B\mathbb{Z}/p)$ such that $f(x_t) = \omega^p$ and $f = 0$ on the other generators of K_t, where $H^*(B\mathbb{Z}/p) \cong \Lambda(\theta) \otimes \mathbb{Z}/p[\omega]$ with $\beta(\theta) = \omega$.*

(2) *There is a map of $\mathcal{K}$-Hopf algebra $g : C_t \to H^*(B\mathbb{Z}/p)$ such that $g(u) = \omega$ and $g = 0$ on the other generators of C_t.*

Proof. We show only (2), since (1) is proved by similar arguments.

Let I denote the ideal of C_t generated by odd degree generators. For the dimensional reason, we see that I is a Hopf ideal of C_t. We show that I is closed under the action of $\mathcal{A}_p$.

For the dimensional reason, $\mathcal{P}^a(I) \subset I$ for $a \ge 1$, and using the relation $\beta\beta = 0$, we have that $\beta(z_k) = \beta(w_l) = 0$ for $t+1 \le k \le m$ and $1 \le l \le n$. Thus, it sufficies to show that $\beta(v_q), \beta(z_k) \in I$ for $1 \le q \le t$ and $m+1 \le k \le m+n$. We see that $\beta(v_q)$ is primitive since v_q is primitive, and so $\beta(v_q) = 0$ since $P^4(C_t) = 0$.

For the dimensional reason, we can put

$$\beta(z_k) = \kappa u^{p+1} + \sum_{i=t+1}^{m} \rho_i u x_i \mod I$$

for $\kappa, \rho_i \in \mathbb{Z}/p$.

If $\lambda = 1$, then using the relation $\beta\beta(z_k) = 0$, $\kappa = \rho_i = 0$ for $t+1 \le i \le m$, which implies that $\beta(z_k) \in I$.

When $\lambda = 0$, using the relation $\beta\beta(z_k) = 0$, we have that $\beta(z_k) = \kappa u^{p+1} \mod I$. For the dimensional reason, we have that $\beta(z_k)$ is primitive since $\beta(u) = \beta(v_q) = 0$, which implies that $\kappa = 0$, and so $\beta(z_k) \in I$.

From the above considerations, C_t/I is a $\mathcal{K}$-Hopf algebra, and the quotient map $\pi : C_t \to C_t/I$ becomes a map of $\mathcal{K}$-Hopf algebra. Since C_t/I is a polynomial algebra, there is a monomorphism of $\mathcal{K}$-Hopf algebra $\sigma : C_t/I \to H^*(BV)$ by [**1**], where V is a $(m+n-t+1)$-dimensional vector space over $\mathbb{Z}/p$. It is known that $H^*(BV) \cong \Lambda(\theta_1, \dots \theta_{m+n-t+1}) \otimes \mathbb{Z}/p[\omega_1, \dots \omega_{m+n-t+1}]$ with $\beta(\theta_k) = \omega_k$ for $1 \le k \le m+n-t+1$.

Taking a suitable basis of V, and by the argument of [**1**], we can assume that $\sigma(u) = \omega_1$, $\sigma(x_i) = \omega_{i-t+1}^p$ for $t+1 \le i \le m$, and $\sigma(y_j) = \omega_{m+j-t+1}^{p^2}$ for $1 \le j \le n$. If we define a map $g = (Bi)^*\sigma\pi$, where $i : \mathbb{Z}/p \to V$ is the inclusion on the first factor, then g is a map of $\mathcal{K}$-Hopf algebra which satisfies the required properties. This completes the proof. □

Proof of Proposition 4.1. We proceed by an induction on t. For $t = 1$, if we put $Y_1 = Y$, then $Y\langle 3\rangle \simeq Y$ since Y is 3-connected, and by assumption, $H^*(Y) \cong A \cong K_1$ with the operations (2.4). Now we assume that there exists an H-space Y_t with the required properties.

From now on, we construct an H-space Y_{t+1} satisfying the required properties. For the map f of Lemma 4.2, a result of Lannes [**11**, Thm. 3.1.1] implies that there is a map $\phi : B\mathbb{Z}/p \to Y_t$ such that $\phi^* = f$. We see that the evaluation map $e_\phi : \mathrm{Map}(B\mathbb{Z}/p, Y_t)_\phi \to Y_t$ becomes a homotopy equivalence by Theorem 3.1 and Proposition 3.2. Let $\iota : B\mathbb{Z}/p \to \mathrm{Map}(B\mathbb{Z}/p, Y_t)_\phi$ be the adjoint of $\phi\mu$, where μ is the multiplication of an H-structure of $B\mathbb{Z}/p$. Then we have the following commutative diagram of fibrations:

$$\begin{tikzcd} B\mathbb{Z}/p \arrow[r, equal] \arrow[d, "\phi"'] & B\mathbb{Z}/p \arrow[r] \arrow[d, "\iota"'] & EB\mathbb{Z}/p \arrow[r] \arrow[d] & B^2\mathbb{Z}/p \arrow[d, equal] \\ Y_t & \mathrm{Map}(B\mathbb{Z}/p, Y_t)_\phi \arrow[l, "e_\phi"', "\simeq"] \arrow[r, "\kappa"] & E_1 \arrow[r, "\zeta"] & B^2\mathbb{Z}/p, \end{tikzcd} \tag{4.4}$$

where $E_1 = EB\mathbb{Z}/p \times_{B\mathbb{Z}/p} \mathrm{Map}(B\mathbb{Z}/p, Y_t)_\phi$ denotes the Borel construction.

Since f is a map of $\mathcal{K}$-Hopf algebra, ϕ is an H-map, and so the bottom fibration becomes an H-fibration by Proposition 3.3. The E_2-term of the

Serre spectral sequence for this fibration is given as

$$E_2^{*,*} = H^*(B^2\mathbb{Z}/p) \otimes H^*(Y_t)$$

for $H^*(B^2\mathbb{Z}/p) \cong \mathbb{Z}/p[\eta, \beta\mathcal{P}^{\Delta_i}\beta(\eta) \mid i \geq 0] \otimes \Lambda(\beta(\eta), \mathcal{P}^{\Delta_i}\beta(\eta) \mid i \geq 0)$, where $\mathcal{P}^{\Delta_i} = \mathcal{P}^{p^i} \cdots \mathcal{P}^1$ and η denotes the fundamental class. Now we use the notations from Proposition 3.10.

For the dimensional reason and by Remark 3.12, we have that $E_{2p+1}^{*,*} \cong E_2^{*,*}$. The generator x_i is transgressive for $t \leq i \leq m$, and then by using the naturality of the diagram (4.4), and by Proposition 3.10, we obtain that

$$d_{2p+1}(x_i) = \begin{cases} \mathcal{P}^1\beta(\eta) & \text{for } i = t, \\ 0 & \text{for } t+1 \leq i \leq m. \end{cases} \tag{4.5}$$

By the Kudo transgression theorem, there are the following differentials:

$$\begin{cases} d_{2p^{k+1}+1}(x_t^{p^k}) = \mathcal{P}^{\Delta_k}\beta(\eta) & \text{for } k \geq 1, \\ d_{2p^k(p-1)+1}(\mathcal{P}^{\Delta_{k-1}}\beta(\eta) \otimes x_t^{p^{k-1}(p-1)}) = \beta\mathcal{P}^{\Delta_k}\beta(\eta) & \text{for } k \geq 1. \end{cases} \tag{4.6}$$

In particular, we see that

$$d_{2p(p-1)+1}(\mathcal{P}^1\beta(\eta) \otimes x_t^{p-1}) = \beta\mathcal{P}^{\Delta_1}\beta(\eta) \tag{4.7}$$

in the $E_{2p(p-1)+1}$-term. Since $H^1(Y_t) = 0$, $d_{2p+1}(z_k) = 0$ for $1 \leq k \leq m+n$. If $d_{2p+1}(y_j) \neq 0$, then we can replace the generator y_j so that $d_{2p+1}(y_j) = 0$ for $1 \leq j \leq n$. In fact, by Proposition 3.10, we can write

$$d_{2p+1}(y_j) = \mathcal{P}^1\beta(\eta) \otimes \sum_{s=0}^{p-1} b_s x_t^s,$$

where b_s are polynomials of generators of $H^*(Y_t)$ other than x_t for $0 \leq s \leq p-1$. If we put $\bar{y}_j$ as

$$\bar{y}_j = y_j - \sum_{s=0}^{p-2} \frac{1}{s+1} b_s x_t^{s+1},$$

then by (4.5), $d_{2p+1}(\bar{y}_j) = \mathcal{P}^1\beta(\eta) \otimes b_{p-1}x_t^{p-1}$, and applying the differential $d_{2p(p-1)+1}$ to $\{\mathcal{P}^1\beta(\eta) \otimes b_{p-1}x_t^{p-1}\} = 0$ in the $E_{2p(p-1)+1}$-term, we have that $\{\beta\mathcal{P}^{\Delta_1}\beta(\eta) \otimes b_{p-1}\} = 0$. This implies that $b_{p-1} = 0$, and so $d_{2p+1}(\bar{y}_j) = 0$. Similarly, we can replace the generators w_l so that $d_{2p+1}(w_l) = 0$ for $1 \leq l \leq n$. Then the E_{2p+2}-term of the spectral sequence is given as

$$E_{2p+2}^{*,*} \cong A_{2p+2} \otimes B_{2p+2} \otimes \Lambda(\mathcal{P}^1\beta(\eta) \otimes x_t^{p-1}),$$

where $A_{2p+2} \cong A_{2p+1}/(\mathcal{P}^1\beta(\eta))$ and B_{2p+2} is generated by the generators of B_{2p+1} other than x_t.

By (4.5), we have that

$$d_{2p+2}(z_k) = \begin{cases} \beta\mathcal{P}^1\beta(\eta) & \text{for } k = t, \\ 0 & \text{for } t+1 \le k \le m, \end{cases}$$

and for $m+1 \le k \le m+n$, if $d_{2p+2}(z_k) \neq 0$, then by Proposition 3.10, $d_{2p+2}(z_k) = a_k\beta\mathcal{P}^1\beta(\eta)$ for some $a_k \in \mathbb{Z}/p$. If we set $\bar{z}_k = z_k - a_k z_t$, then $d_{2p+2}(\bar{z}_k) = 0$. If $d_{2p+2}(y_j) \neq 0$, then by Proposition 3.10, we can write that

$$d_{2p+2}(y_j) = \beta\mathcal{P}^1\beta(\eta) \otimes (b_0 + b_1 z_t),$$

where b_s are polynomials of generators of B_{2p+2} other than z_t for $s = 0, 1$. If we set $\bar{y}_j = y_j - b_0 z_t$, then $d_{2p+2}(\bar{y}_j) = \beta\mathcal{P}^1\beta(\eta) \otimes b_1 z_t$, and applying d_{2p+2} to $d_{2p+2}(\bar{y}_j)$, we have that $(\beta\mathcal{P}^1\beta(\eta))^2 \otimes b_1 = 0$, which implies that $b_1 = 0$, and so $d_{2p+2}(\bar{y}_j) = 0$. By the same arguments, we can replace the generators w_l so that $d_{2p+2}(w_l) = 0$ for $1 \le l \le n$. Then we obtain that

$$E_{2p+3}^{*,*} \cong A_{2p+3} \otimes B_{2p+3} \otimes \Lambda(\mathcal{P}^1\beta(\eta) \otimes x_t^{p-1}),$$

where $A_{2p+3} \cong A_{2p+2}/(\beta\mathcal{P}^1\beta(\eta))$ and B_{2p+3} is generated by the generators of B_{2p+2} other than z_t. For the dimensional reason and by Remark 3.12, $E_{2p(p-1)+1}^{*,*} \cong E_{2p+3}^{*,*}$, and by (4.7), $E_{2p(p-1)+2}^{*,*} \cong A_{2p(p-1)+2} \otimes B_{2p(p-1)+2}$, where $A_{2p(p-1)+2} \cong A_{2p(p-1)+1}/(\beta\mathcal{P}^{\Delta_1}\beta(\eta))$ and $B_{2p(p-1)+2} \cong B_{2p(p-1)+1}$. Furthermore, for the dimensional reason and by Remark 3.12, $E_{2p^2+1}^{*,*} \cong E_{2p(p-1)+2}^{*,*}$, and so we conclude that

$$E_{2p^2+1}^{*,*} \cong A_{2p^2+1} \otimes B_{2p^2+1}$$

for

$$A_{2p^2+1} \cong H^*(B^2\mathbb{Z}/p)/(\beta\mathcal{P}^1\beta(\eta), \mathcal{P}^1\beta(\eta))$$

and

$$\begin{aligned} B_{2p^2+1} \cong\ & \mathbb{Z}/p[x_t^p, x_{t+1}, \dots, x_m, y_1, \dots, y_n] \\ & \otimes \Lambda(v_1, \dots, v_{t-1}, z_{t+1}, \dots, z_{m+n}, w_1, \dots, w_n). \end{aligned}$$

By iterating this process, we can compute the spectral sequence. In particular, the differentials are completely determined by (4.6), and so we have that for $k \ge 1$,

$$E_{2p^k+1}^{*,*} \cong A_{2p^k+1} \otimes B_{2p^k+1},$$

where

$$A_{2p^k+1} \cong H^*(B^2\mathbb{Z}/p)/\left(\beta\mathcal{P}^{\Delta_j}\beta(\eta), \mathcal{P}^{\Delta_j}\beta(\eta) \mid 0 \le j \le k-2\right)$$

and

$$B_{2p^k+1} \cong \mathbb{Z}/p\left[x_t^{p^{k-1}}, x_{t+1}, \dots, x_m, y_1, \dots, y_n\right] \otimes \Lambda(v_1, \dots, v_{t-1}, z_{t+1}, \dots, z_{m+n}, w_1, \dots, w_n).$$

This implies that $H^*(E_1) \cong C_t$ as algebras, where u and v_t represent the generators η and $\beta(\eta)$ in $H^*(B^2\mathbb{Z}/p)$. Since

$$\begin{cases} \kappa^*(x_i) = x_i & \text{for } t+1 \le i \le m, \\ \kappa^*(y_j) = y_j & \text{for } 1 \le j \le n, \\ \kappa^*(z_k) = z_k & \text{for } t+1 \le k \le m, \\ \kappa^*(z_k) = z_k - a_k z_t & \text{for } m+1 \le k \le m+n,\ a_k \in \mathbb{Z}/p, \\ \kappa^*(w_l) = w_l & \text{for } 1 \le l \le n \end{cases}$$

up to decomposable elements and $\mathcal{P}^p(z_t) = 0$, we can take the generators of $H^*(E_1)$ satisfying the condition (4.3) with $\lambda = 1$.

Next we apply same arguments to the H-space E_1. For the map g of Lemma 4.2, a result of Lannes [**11**, Thm. 3.1.1] implies that there is a map $\psi_1 : B\mathbb{Z}/p \to E_1$ such that $\psi_1^* = g$. The evaluation map $e_{\psi_1} : \mathrm{Map}(B\mathbb{Z}/p, E_1)_{\psi_1} \to E_1$ is a homotopy equivalence by Theorem 3.1 and Proposition 3.2. Let $\iota_1 : B\mathbb{Z}/p \to \mathrm{Map}(B\mathbb{Z}/p, E_1)_{\psi_1}$ be the adjoint of $\psi_1\mu$. Then, we have the following H-fibration by the same construction as above:

$$E_1 \xleftarrow{\simeq} \mathrm{Map}(B\mathbb{Z}/p, E_1)_{\psi_1} \xrightarrow{\kappa_1} E_2 \longrightarrow B^2\mathbb{Z}/p,$$

where $E_2 = (\mathrm{Map}(B\mathbb{Z}/p, E_1)_{\psi_1})_{hB\mathbb{Z}/p}$ denotes the Borel construction. Computing the spectral sequence for this fibration as above, we conclude that $H^*(E_2) \cong C_t$ with the operations (4.3) with $\lambda = 0$.

Iterating this process, we have the following sequence of H-spaces and H-maps:

$$Y_t \xrightarrow{\kappa} E_1 \xrightarrow{\kappa_1} E_2 \xrightarrow{\kappa_2} \cdots$$

satisfying $H^*(Y_t) \cong K_t$, $H^*(E_s) \cong C_t$ with the operations (4.3) with $\lambda = 1$ for $s = 1$ and $\lambda = 0$ for $s > 1$, $\kappa_s^*(u) = 0$ and

$$\kappa_s^* : H^*(E_{s+1})/(u) \longrightarrow H^*(E_s)/(u)$$

is an isomorphism for $s \ge 1$.

If we set $Y_{t+1} = \varinjlim_s E_s$, then there is the Milnor exact sequence

$$0 \to \varprojlim{}_s^1 H^{*+1}(E_s) \to H^*(Y_{t+1}) \to \varprojlim{}_s H^*(E_s) \to 0.$$

Since $\varprojlim{}_s^1 H^{*+1}(E_s) = 0$ by the Mittag-Leffler condition, we have that $H^*(Y_{t+1}) \cong \varprojlim{}_s H^*(E_s) \cong K_{t+1}$, and by Proposition 3.6, we see that Y_{t+1} has an H-structure. Let F be the homotopy fiber of the composite $E_1 \to Y_{t+1}$, then $H^*(F) \cong H^*(K(\mathbb{Z}, 2))$ by the spectral sequence argument, and this implies that $F \simeq K(\mathbb{Z}, 2)$. By the cohomology, E_1 is homotopy equivalent to the homotopy fiber of $[p]v_t : Y_{t+1} \to K(\mathbb{Z}, 3)$. Therefore, we have the following commutative diagram of fibrations:

$$\begin{tikzcd}
Y_t \arrow[r, "\kappa"] \arrow[d, equal] & E_1 \arrow[r, "\zeta"] \arrow[d] & B^2\mathbb{Z}/p \arrow[d] \\
Y_{t+1}\langle v_t\rangle \arrow[r] \arrow[d] & Y_{t+1} \arrow[r, "v_t"] \arrow[d, "{[p]v_t}"] & K(\mathbb{Z}, 3) \arrow[d, "{[p]}"] \\
* \arrow[r] & K(\mathbb{Z}, 3) \arrow[r, equal] & K(\mathbb{Z}, 3),
\end{tikzcd}$$

which implies that $Y_t \simeq Y_{t+1}\langle v_t\rangle$, where $Y_{t+1}\langle v_t\rangle$ denotes the homotopy fiber of the map $v_t : Y_{t+1} \to K(\mathbb{Z}, 3)$. By the induction hypothesis, $Y_t\langle 3\rangle \simeq Y$, and so we have that $Y_{t+1}\langle 3\rangle \simeq (Y_{t+1}\langle v_t\rangle)\langle 3\rangle \simeq Y$. This completes the proof. $\square$

Next, for $1 \leq t \leq n+1$, we set an algebra L_t as

$$\begin{aligned} L_t = {} & \mathbb{Z}/p[y_t, \dots , y_n] \\ & \otimes \Lambda(v_1, \dots , v_{m+t-1}, z_{m+t}, \dots , z_{m+n}, c_1, \dots , c_{t-1}, w_t, \dots , w_n) \end{aligned}$$

with y_j, z_k and w_l are as in A, $|v_q| = 3$ for $1 \leq q \leq m+t-1$, and $|c_r| = 2p+1$ for $1 \leq r \leq t-1$. Then we have the following proposition:

Proposition 4.8. *For $1 \leq t \leq n+1$, there is a mod p H-space Z_t such that $Z_t\langle 3\rangle \simeq Y$ and $H^*(Z_t) \cong L_t$ with the following operations:*

$$\begin{cases} \beta(y_j) = w_j & \text{for } t \leq j \leq n, \\ \mathcal{P}^1(v_{m+r}) = c_r & \text{for } 1 \leq r \leq t-1, \\ \mathcal{P}^p(z_{m+l}) = w_l + \delta_l & \text{for } t \leq l \leq n, \end{cases} \tag{4.9}$$

where δ_l is some decomposable element of L_t for $t \leq l \leq n$.

Proposition 4.8 is proved by same arguments as in Proposition 4.1, and so we give an outline of the proof.

We proceed by an induction on $1 \leq t \leq n+1$. For $t = 1$, if we set $Z_1 = Y_{m+1}$, then by Proposition 4.1, Z_1 satisfies the required properties. We assume that there exists an H-space Z_t with the conditions of Proposition 4.8, and construct an H-space Z_{t+1} satisfying the required properties.

We can construct a $\mathcal{K}$-Hopf algebra map $h : H^*(Z_t) \to H^*(B\mathbb{Z}/p)$ such that $h(y_t) = \omega^{p^2}$ and $h = 0$ on the other generators. By a result of Lannes, there is an H-map $\xi : B\mathbb{Z}/p \to Z_t$ such that $\xi^* = h$, and we see that the evaluation map $e_\xi : \mathrm{Map}(B\mathbb{Z}/p, Z_t)_\xi \to Z_t$ becomes a homotopy equivalence. For an H-structure μ of $B\mathbb{Z}/p$, if $\iota : B\mathbb{Z}/p \to \mathrm{Map}(B\mathbb{Z}/p, Z_t)_\xi$ denotes the adjoint of $\xi\mu$, then we have the following fibration:

$$Z_t \xleftarrow{\simeq} \mathrm{Map}(B\mathbb{Z}/p, Z_t)_\xi \longrightarrow F_1 \longrightarrow B^2\mathbb{Z}/p,$$

where F_1 is an H-space given by the Borel construction for ι.

For $1 \leq t \leq n+1$, we set an algebra D_t as

$$\begin{aligned} D_t = &\ \mathbb{Z}/p[u, y_{t+1}, \dots, y_n] \\ &\otimes \Lambda(v_1, \dots, v_{m+t}, z_{m+t+1}, \dots, z_{m+n}, c_1, \dots, c_t, w_{t+1}, \dots, w_n) \end{aligned}$$

with y_j, z_k, w_l and v_q for $1 \leq q \leq m+t-1$ are as in L_t, $|u| = 2$ and $|v_{m+t}| = 3$. Then, using the Serre spectral sequence, we have that $H^*(F_1) \cong D_t$ with the operations (4.9) and $\beta(u) = v_{m+t}$. Iterating this process, we have a sequence of H-spaces and H-maps

$$Z_t \longrightarrow F_1 \longrightarrow F_2 \longrightarrow \cdots$$

such that $H^*(Z_t) \cong L_t$, $H^*(F_s) \cong D_t$ with the operations (4.9). If we set $Z_{t+1} = \varinjlim_s F_s$, then Z_{t+1} has an H-structure, and using the Milnor exact sequence, we obtain that $H^*(Z_{t+1}) \cong L_{t+1}$ with the operations (4.9). We can show that the homotopy fiber $Z_{t+1}\langle v_{m+t}\rangle \simeq Z_t$, and so by the induction hypothesis, the three-connected cover $Z_{t+1}\langle 3\rangle \simeq Y$. This establishes the proof of Proposition 4.8.

Now we set $W = Z_{n+1}$. Then W is a simply connected mod p finite H-space such that

$$H^*(W) \cong \Lambda(v_1, \dots, v_{m+n}, c_1, \dots, c_n)$$

with $\mathcal{P}^1(z_{m+r}) = c_r$ for $1 \leq r \leq n$, and $Y \simeq W\langle 3\rangle$ by Proposition 4.8. This completes the proof of Proposition 2.6. $\square$

References

[1] J. Aguadé and L. Smith, *On the* mod p *Torus Theorem of John Hubbuck*, Math. Z., **191** (1986), 325-326.

[2] A. Bousfield and D. Kan, *Homotopy Limits, Completions and Localizations*, Springer Lecture Notes in Math., **304** (1972).

[3] C. Broto and J.A. Crespo, *H-spaces with Noetherian* mod *two cohomology algebra*, preprint.

[4] W. Browder, *Homotopy commutative H-spaces*, Ann. of Math., **75** (1962), 283-311.

[5] ———, *On differential Hopf algebras*, Trans. Amer. Math. Soc., **107** (1963), 153-178.

[6] E. Dror Farjoun, *Cellular Spaces, Null Spaces and Homotopy Lacalization*, Springer Lecture Notes in Math., **1622** (1996).

[7] W.G. Dwyer and C.W. Wilkerson, *Spaces of null homotopic maps*, Astérisque, **191** (1990), 97-108.

[8] Y. Hemmi, *Higher homotopy commutativity of H-spaces and the* mod p *Torus theorem*, Pacific J. Math., **149** (1991), 95-111.

[9] K. Iriye and A. Kono, Mod p *retracts of G-product spaces*, Math. Z., **190** (1985), 357-363.

[10] D. Kraines, *The kernel of the loop suspension map*, Illinois J. Math., **21** (1977), 91-108.

[11] J. Lannes, *Sur les espaces fonctionnels dont la source est le classifiant d'un p-groupe abélien élémentaire*, Publ. Math. Inst. HES, **75** (1992), 135-244.

[12] J.P. Lin, *Loops of H-spaces with finitely generated cohomology rings*, Topology Appl., **60** (1994), 131-152.

[13] ———, *Finitely generated cohomology Hopf algebras and torsion*, Pacific J. Math., **172** (1996), 215-221.

[14] C. McGibbon, *Homotopy commutativity in localized groups*, Amer. J. Math., **106** (1984), 665-687.

[15] ———, *Higher forms of homotopy commutativity and finite loop spaces*, Math. Z., **201** (1989), 363-374.

[16] J. Milnor and J.C. Moore, *On the structure of Hopf algebras*, Ann. of Math., **81** (1965), 211-264.

[17] J.A. Neisendorfer, *Localizations and connected covers of finite complexes*, Contemp. Math., **181** (1995), 385-390.

[18] M. Slack, *A classification theorem for homotopy commutative* mod 2 *H-spaces with finitely generated cohomology rings*, Mem. Amer. Math. Soc., **92** (1991).

[19] J.D. Stasheff, *Homotopy associativity of H-spaces*, I, II, Trans. Amer. Math. Soc., **108** (1963), 275-292, 293-312.

[20] Z. Wojtkowiak, *On maps from holim F to Z*, Springer Lecture Notes in Math., **1298** (1986), 227-236.

[21] A. Zabrodsky, *On phantom maps and a theorem of H. Miller*, Israel J. Math., **58** (1987), 129-143.

Received January 2, 1998 and revised August 5, 1998. The author was partially supported by JSPS Research Fellowships for Young Scientists.

Hiroshima University
Kagamiyama 1-3-1
Higashi-Hiroshima 739-8526
Japan
E-mail address: yusuke@top3.math.sci.hiroshima-u.ac.jp

PACIFIC JOURNAL OF MATHEMATICS
Vol. 190, No. 2, 1999

DUAL SPACES AND ISOMORPHISMS OF SOME DIFFERENTIAL BANACH $*$-ALGEBRAS OF OPERATORS

Edward Kissin and Victor S. Shulman

The paper continues the study of differential Banach *-algebras $\mathcal{A}_S$ and $\mathcal{F}_S$ of operators associated with symmetric operators S on Hilbert spaces H. The algebra $\mathcal{A}_S$ is the domain of the largest *-derivation δ_S of $B(H)$ implemented by S and the algebra $\mathcal{F}_S$ is the closure of the set of all finite rank operators in $\mathcal{A}_S$ with respect to the norm $\|A\| = \|A\| + \|\delta_S(A)\|$. When S is selfadjoint, $\mathcal{F}_S$ is the domain of the largest *-derivation of the algebra $C(H)$ implemented by S. If S is bounded, $\mathcal{F}_S = C(H)$ and $\mathcal{A}_S = B(H)$, so $\mathcal{A}_S$ is isometrically isomorphic to the second dual of $\mathcal{F}_S$. For unbounded selfadjoint operators S the paper establishes the full analogy with the bounded case: $\mathcal{A}_S$ is isometrically isomorphic to the second dual of $\mathcal{F}_S$. The paper also classifies the algebras $\mathcal{A}_S$ and $\mathcal{F}_S$ up to isometrical *-isomorphism and obtains some partial results about bounded but not necessarily isometrical *-isomorphisms of the algebras $\mathcal{F}_S$.

1. Introduction and preliminaries.

Extensive development of non-commutative geometry requires elaborating of the theory of differential Banach *-algebras, that is, dense *-subalgebras of C^*-algebras whose properties in many respects are analogous to the properties of algebras of differentiable functions.

Blackadar and Cuntz [**2**] and the authors [**12**] introduced and studied various classes of differential Banach *-algebras; the most interesting class consists of **D**-algebras, that is, dense *-subalgebras $\mathcal{A}$ of C^*-algebras $(\mathfrak{U}, \|\cdot\|)$ which, in turn, are Banach *-algebras with respect to another norm $\|\cdot\|_1$ and the norms $\|\cdot\|$ and $\|\cdot\|_1$ on $\mathcal{A}$ satisfy the inequality:

$$\|xy\| \le D(\|x\|\,\|y\|_1 + \|x\|_1\|y\|), \quad \text{for } x, y \in \mathcal{A}, \tag{1.1}$$

for some $D > 0$. This class contains, for example, the domains $D(\delta)$ of closed unbounded *-derivations δ of C^*-algebras $\mathfrak{U}$ where the norm $\|\cdot\|_1$ on $D(\delta)$ is defined, as usual, by the formula

$$\|A\|_1 = \|A\| + \|\delta(A)\|, \quad \text{for } A \in D(\delta).$$

Much work has been done on the investigation of properties of the differential Banach *-algebras (see Blackadar and Cuntz [2] and Kissin and Shulman [12, 13]) and the algebras $D(\delta)$ in particular (see, for example, Bratteli and Robinson [3] and Sakai [16]).

In many cases closed *-derivations of C^*-algebras $\mathfrak{U}$ of operators on Hilbert spaces are implemented by closed symmetric operators. In particular, Bratteli and Robinson [3] showed that if $\mathfrak{U}$ contains the ideal of all compact operators then any closed *-derivation of $\mathfrak{U}$ is implemented by a symmetric operator.

Any closed symmetric operator S on a Hilbert space H implements closed *-derivations of various C^*-algebras of operators on H. Among all these derivations there is the largest one - δ_S with domain $D(\delta_S)$ (which we denote by $\mathcal{A}_S$) containing the domains of all derivations implemented by S:

$$\begin{aligned}\mathcal{A}_S = \Big\{A \in B(H) : AD(S) \subseteq D(S),\ A^*D(S) \subseteq D(S) \quad \text{and} \\ (SA - AS)|_{D(S)} \ \text{extends to a bounded operator} \ \ A_S\Big\} \\ \text{and} \ \ \delta_S(A) = i\,\text{Closure}\,(SA - AS), \quad \text{for} \ \ A \in \mathcal{A}_S.\end{aligned}$$

The closure of $\mathcal{A}_S$ with respect to the norm $\|\cdot\|$ in $B(H)$ is the enveloping C^*-algebra which we denote by $\mathfrak{U}_S$.

The algebra $\mathcal{A}_S$ is a unital Banach *-algebra with respect to the norm

$$\|A\|_S = \|A\| + \|A_S\|. \tag{1.2}$$

If S implements a *-derivation δ of a C^*-algebra $\mathfrak{U}$ of operators on H then

$$D(\delta) \subseteq \mathcal{A}_S, \quad \mathfrak{U} \subseteq \mathfrak{U}_S \ \text{ and } \ \delta = \delta_S|\mathfrak{U}.$$

By $C(H)$ we denote the algebra of all compact operators on H. The *-algebras

$$\mathcal{K}_S = \mathcal{A}_S \cap C(H) \quad \text{and} \quad \mathcal{J}_S = \{A \in \mathcal{K}_S : \delta_S(A) \in C(H)\}$$

are dense in $C(H)$ and are the domains of the largest closed *-derivations from $C(H)$ into $B(H)$ and $C(H)$, respectively, implemented by S.

By $\mathcal{F}_S$ we denote the closure with respect to the norm $\|\cdot\|_S$ of the subalgebra of all finite rank operators in $\mathcal{A}_S$.

It was shown in [13] that $(\mathcal{K}_S, \|\cdot\|_S)$ and $(\mathcal{J}_S, \|\cdot\|_S)$ are semisimple Banach *-algebras, that $(\mathcal{F}_S, \|\cdot\|_S)$ is a simple Banach *-algebra and

$$\mathcal{F}_S \subseteq \mathcal{J}_S \subseteq \mathcal{K}_S \subseteq \mathcal{A}_S.$$

Furthermore, $\mathcal{F}_S$, $\mathcal{J}_S$ and $\mathcal{K}_S$ are closed two-sided ideals of $(\mathcal{A}_S, \|\cdot\|_S)$ and $\mathcal{F}_S$ is contained in any closed two-sided ideal of $(\mathcal{A}_S, \|\cdot\|_S)$. The relation between the ideals $\mathcal{F}_S$, $\mathcal{J}_S$ and $\mathcal{K}_S$ and the question of how the properties of the operator S are reflected in the structure of $\mathcal{K}_S$, $\mathcal{J}_S$ and $\mathcal{F}_S$ were investigated in [13]. In particular, it was established that $\overline{(\mathcal{K}_S)^2} = \overline{(\mathcal{J}_S)^2} = \mathcal{F}_S$, for all symmetric S, and that the ideals $\mathcal{J}_S$ and $\mathcal{F}_S$ have a bounded approximate identity if and only if S is selfadjoint. For selfadjoint S, it was also proved that $\mathcal{K}_S \neq \mathcal{J}_S = \mathcal{F}_S$.

In spite of the fact that the structure of the algebras $\mathcal{F}_S$, $\mathcal{J}_S$, $\mathcal{K}_S$, $\mathcal{A}_S$ and $\mathfrak{U}_S$ is comparatively simple, many important questions still remain open. In Section 2 we mainly study the structure of the algebras $\mathcal{A}_S$ and $\mathfrak{U}_S$ in the case when S is a selfadjoint operator. However, we also consider the case when S is a symmetric operator with at least one finite deficiency index and show that the algebras $\mathcal{A}_S$ and $\mathfrak{U}_S$ contain closed ideals of finite codimension.

If S is a bounded symmetric operator on H then $\mathcal{F}_S = C(H)$ and $\mathcal{A}_S = B(H)$, so $\mathcal{A}_S$ is isometrically isomorphic to the second dual of $\mathcal{F}_S$. In Section 3 we investigate the structure of the dual and the second dual spaces of the algebras $\mathcal{F}_S$ for unbounded symmetric operators S. In the case when S is selfadjoint we establish the full analogy with the bounded case: The algebra $\mathcal{A}_S$ is isometrically isomorphic to the second dual of $\mathcal{F}_S$.

In Section 4 we study the problem of classification of the algebras $\mathcal{F}_S$ and $\mathcal{A}_S$ up to *-isomorphism. For isometrical *-isomorphism this problem is completely solved in Theorem 4.4. For bounded but not necessarily isometrical *-isomorphism we obtain some interesting partial results in the case when S is selfadjoint.

2. Structure of the algebras $\mathcal{A}_S$ and the enveloping C^*-algebras $\mathfrak{U}_S$.

The main purpose of this section is to study the structure of the algebras $\mathcal{A}_S$ and $\mathfrak{U}_S$ in the case when S is a selfadjoint operator. However, we start the section by considering the case when S is a symmetric operator with at least one finite deficiency index. Making use of the existence of a J-symmetric representation of $\mathcal{A}_S$ on the deficiency space of S, we will show that the algebras $\mathcal{A}_S$ and $\mathfrak{U}_S$ contain closed ideals of finite codimension.

Let S be symmetric, S^* be the adjoint operator, let $N_-(S)$ and $N_+(S)$ be the deficiency spaces of S and

$$n_\pm(S) = \dim\left(N_\pm(S)\right)$$

be the deficiency indices of S. It is well known that $D(S^*)$ is a Hilbert space with respect to the scalar product

$$\langle x, y\rangle = (x, y) + (S^*x, S^*y), \quad \text{for} \ \ x, y \in D(S^*),$$

and it is the orthogonal sum of the closed subspaces $D(S), N_-(S)$ and $N_+(S)$:

$$D(S^*) = D(S)_{\langle + \rangle} N_-(S)_{\langle + \rangle} N_+(S).$$

Set $N(S) = N_-(S)_{\langle + \rangle} N_+(S)$ and let Q be the projection on $N(S)$ in $D(S^*)$. It was shown in [7] and [8] that

$$[x, y] = i(x, S^* y) - i(S^* x, y), \quad \text{for} \ \ x, y \in N(S),$$

is an indefinite non-degenerate sesquilinear form on $N(S)$, that

$$\pi_S(A) = QA|_{N(S)}, \quad \text{for} \ \ A \in \mathcal{A}_S,$$

is a bounded representation of $(\mathcal{A}_S, \|\cdot\|_S)$ on $N(S)$ and that it is J-*symmetric*:

$$[\pi_S(A)x, y] = [x, \pi_S(A^*)y], \quad \text{for} \ \ x, y \in N(S).$$

A subspace L in $N(S)$ is *neutral* if

$$[x, y] = 0, \quad \text{for all} \ \ x, y \in L.$$

The operator S is *well-behaved* if the representation π_S has no neutral invariant subspace.

Let $\kappa_S = \min(n_-(S), n_+(S))$ and assume that $0 < \kappa_S < \infty$. It was proved in [10] that the representation π_S has a κ_S-dimensional subrepresentation σ. Let ρ be an irreducible subrepresentation of σ. It was shown in [11] that ρ is bounded with respect to the operator norm $\|\cdot\|$ in $\mathcal{A}_S$ and, therefore, extends to a bounded *-representation of the enveloping C^*-algebra $\mathfrak{U}_S$. If S is well-behaved, it follows from Theorem 28.13 [14] that $\mathcal{K}_S \subseteq \mathrm{Ker}(\rho)$. This yields

Theorem 2.1. *Let S be a symmetric unbounded operator and $0 < \kappa_S < \infty$.*

(i) *There exists a closed two-sided ideal J in the Banach *-algebra $(\mathcal{A}_S, \|\cdot\|)$ such that the quotient algebra $\mathcal{A}_S/J$ is isomorphic to the full matrix algebra $M_n(\mathbb{C})$ with $0 < n \leq \kappa_S$.*

(ii) *The uniform closure $\overline{J}$ of J in $\mathfrak{U}_S$ is a closed two-sided ideal and the quotient algebra $\mathfrak{U}_S/\overline{J}$ is isomorphic to the full matrix algebra $M_n(\mathbb{C})$.*

(iii) *If S is well-behaved then $\mathcal{K}_S \subseteq J$ and $C(H) \subseteq \overline{J}$.*

Example 2.2. Let $H = L^2(0,1)$ and $S = i\frac{d}{dt}$ with domain $D(S)$ consisting of all absolutely continuous functions h such that $h' \in L^2(0,1)$ and $h(0) = h(1) = 0$. Then S is a symmetric operator and $n_-(S) = n_+(S) = 1$.

It was proved in [9] that S is well-behaved. Therefore it follows from Theorem 2.1 that there exists a closed two-sided ideal J in $(\mathcal{A}_S, \|\cdot\|)$ containing $\mathcal{K}_S$ such that $\dim(\mathcal{A}_S/J) = 1$ and that the uniform closure of J in $\mathfrak{U}_S$ is an ideal of codimension 1. $\square$

Let S be the same as in Example 2.2 and let $\mathrm{Lip}\,(0,1)$ be the algebra of all functions on $[0,1]$ satisfying a Lipshitz condition: $|g(t) - g(s)| \leq K_g|t - s|$ for some $K_g > 0$ and all $t, s \in [0,1]$. For $g \in \mathrm{Lip}\,(0,1)$, denote by M_g the operator of multiplication by g on $L^2(0,1)$ and set $\mathcal{B} = \{M_g : g \in \mathrm{Lip}\,(0,1)\}$. Then $M_g D(S) \subseteq D(S)$, $(M_g)^* D(S) = M_{\overline{g}} D(S) \subseteq D(S)$ and

$SM_g - M_gS$ extends to the operator $iM_{g'}$ which is bounded, since g' is essentially bounded on $[0,1]$. Thus $\mathcal{B} \subset \mathcal{A}_S$.

(The authors are grateful to the referee of the paper for pointing out an error in the definition of the algebra $\mathcal{B}$ in the first version of the paper.)

Problem 2.3. Is $\mathcal{A}_S = \mathcal{B} + \mathcal{K}_S$?

The assumption that a symmetric operator S is selfadjoint makes the task of studying the structure of the algebras $\mathcal{A}_S$ and $\mathfrak{U}_S$ easier. First of all, the structure of the ideals $\mathcal{K}_S$, $\mathcal{J}_S$ and $\mathcal{F}_S$ is simpler. While for arbitrary symmetric operators S it is only known (see [13]) that $\overline{(\mathcal{K}_S)^2} = \overline{(\mathcal{J}_S)^2} = \mathcal{F}_S$, where the closure is taken with respect to the norm $\|\cdot\|_S$, for selfadjoint operators S it was shown in [13] that $\mathcal{F}_S = \mathcal{J}_S \neq \mathcal{K}_S$. Secondly, in the selfadjoint case we can employ the Spectral Theorem to establish the structure of $\mathcal{A}_S$ and $\mathfrak{U}_S$.

Let
$$S = \int_{-\infty}^{\infty} \lambda \, dE_S(\lambda)$$
be the spectral decomposition of S. For every integer n, set

$$P_S(n) = E_S(n+1) - E_S(n) \quad \text{and} \quad [S] = \sum_{-\infty}^{\infty} nP_S(n). \tag{2.1}$$

Then $[S]$ is a selfadjoint operator, $\mathrm{Sp}\,([S]) \subseteq \mathbb{Z}$ and the operator $S - [S]$ is bounded. Therefore it follows that
$$\mathcal{A}_S = \mathcal{A}_{[S]}, \quad \mathcal{K}_S = \mathcal{K}_{[S]} \quad \text{and} \quad \mathcal{F}_S = \mathcal{F}_{[S]}$$
and the norms $\|\cdot\|_S$ and $\|\cdot\|_{[S]}$ are equivalent on $\mathcal{A}_S$. This reduces the problem of the description of the structure of the algebras $\mathcal{A}_S$ and $\mathfrak{U}_S$ to the case when $\mathrm{Sp}\,(S) \subseteq \mathbb{Z}$.

We denote by $\mathcal{S}_{\mathbb{Z}}$ the set of all selfadjoint operators S on H such that $\mathrm{Sp}\,(S) \subseteq \mathbb{Z}$ and set

$$H_S(n) = P_S(n)H, \quad \text{for } \ n \in \mathrm{Sp}\,(S). \tag{2.2}$$

Then

$$H = \sum_{n \in \mathrm{Sp}(S)} \oplus H_S(n). \tag{2.3}$$

We omit the proof of the following simple result.

Proposition 2.4. *Let $S, T \in \mathcal{S}_{\mathbb{Z}}$. If there exists a one-to-one mapping φ from $\mathrm{Sp}\,(T)$ onto $\mathrm{Sp}\,(S)$ such that $\dim(H_T(n)) = \dim(H_S(\varphi(n)))$, for $n \in \mathrm{Sp}\,(T)$, and*
$$\sup_{n \in \mathrm{Sp}(T)} |\varphi(n) - n| < \infty$$
then there exists a unitary operator U such that $\mathcal{A}_T = U\mathcal{A}_S U^*$.

Let $S \in \mathcal{S}_{\mathbb{Z}}$. Every operator A in $B(H)$ has a block-matrix form $A = (A_{ij})$, $i, j \in \mathrm{Sp}(S)$, with respect to decomposition (2.3). We denote by $\mathcal{D}_S$ the C^*-algebra of all block-diagonal operators $A = (A_{ij})$ in $B(H)$, that is, $A_{ij} = 0$ if $i \neq j$. By $\mathcal{R}$ we denote the subalgebra of all operators $A = (A_{ij})$ in $B(H)$ with only finite number of non-zero entries A_{ij}. Then, clearly,

$$\mathcal{D}_S \subseteq \mathcal{A}_S \quad \text{and} \quad \mathcal{R}_S \subseteq \mathcal{A}_S.$$

Let $\overline{\mathcal{R}}_S$ be the closure of $\mathcal{R}_S$ in $(\mathcal{A}_S, \|\cdot\|_S)$ and let $C_S(H)$ be the uniform closure of $\mathcal{R}_S$ in $B(H)$.

Lemma 2.5. *$\mathcal{D}_S + C_S(H)$ is a C^*-subalgebra of $\mathfrak{U}_S$ and $\mathcal{D}_S + \overline{\mathcal{R}}_S$ is a closed *-subalgebra of $(\mathcal{A}_S, \|\cdot\|_S)$.*

Proof. Let $\mathcal{L}$ be the uniform closure of $\mathcal{D}_S + \mathcal{R}_S$ in $B(H)$. Then $\mathcal{L}$ is a C^*-subalgebra of $\mathfrak{U}_S$. Since $\mathcal{R}_S$ is a two-sided ideal of the algebra $\mathcal{D}_S + \mathcal{R}_S$, the C^*-algebra $C_S(H)$ is a two-sided ideal of $\mathcal{L}$. Therefore it follows from Corollary 1.8.4 [4] that $\mathcal{D}_S + C_S(H)$ is a C^*-algebra, so $\mathcal{L} = \mathcal{D}_S + C_S(H)$.

For $A \in B(H)$, set

$$\phi(A) = \sum_{n \in \mathrm{Sp}(S)} P_S(n) A P_S(n) \quad \text{and} \quad \widetilde{A} = A - \phi(A).$$

Then ϕ is a conditional expectation from $B(H)$ onto $\mathcal{D}_S$ and

$$\|\phi(A)\| \leq \|A\| \quad \text{and} \quad \|\widetilde{A}\| \leq 2\|A\|. \tag{2.4}$$

If $A \in \mathcal{A}_S$ then $\widetilde{A} \in \mathcal{A}_S$ and $\mathrm{Closure}\,(SA - AS) = \mathrm{Closure}\,(S\widetilde{A} - \widetilde{A}S)$.

Assume that $\{A_n\}$ converge to A in $\mathcal{A}_S$ with respect to $\|\cdot\|_S$. Then

$$\|A - A_n\| \to 0 \;\text{ and }\; \|\mathrm{Closure}\,(S(A - A_n) - (A - A_n)S)\| \to 0, \;\text{ as }\; n \to \infty,$$

and therefore, by (1.2) and (2.4),

$$\begin{aligned} \|\widetilde{A} - \widetilde{A}_n\|_S &= \|\widetilde{A} - \widetilde{A}_n\| + \|\mathrm{Closure}\,(S(\widetilde{A} - \widetilde{A}_n) - (\widetilde{A} - \widetilde{A}_n)S)\| \\ &\leq 2\|A - A_n\| + \|\mathrm{Closure}\,(S(A - A_n) - (A - A_n)S)\| \to 0, \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{as } n \to \infty. \end{aligned} \tag{2.5}$$

Hence $\widetilde{A}_n$ converge to $\widetilde{A}$ with respect to $\|\cdot\|_S$.

Suppose now that $B \in \overline{\mathcal{R}}_S$. Then there are $\{B_n\}$ in $\mathcal{R}_S$ converging to B with respect to $\|\cdot\|_S$. It follows from (2.5) that $\widetilde{B}_n$ converge to $\widetilde{B}$ with respect to $\|\cdot\|_S$ and, since $\widetilde{B}_n$ belong to $\mathcal{R}_S$, we obtain that $\widetilde{B} \in \overline{\mathcal{R}}_S$.

Finally, let $C_n = A_n + B_n$ converge to C in $\mathcal{A}_S$ with respect to $\|\cdot\|_S$ where $A_n \in \mathcal{D}_S$ and $B_n \in \overline{\mathcal{R}}_S$. Then $\widetilde{C}_n = \widetilde{B}_n$ and, by (2.5), $\widetilde{B}_n$ converge to $\widetilde{C}$ with respect to $\|\cdot\|_S$. Since, by the above argument, all $\widetilde{B}_n$ belong to $\overline{\mathcal{R}}_S$, the operator $\widetilde{C}$ also belong to $\overline{\mathcal{R}}_S$. Hence $C \in \mathcal{D}_S + \overline{\mathcal{R}}_S$ and $\mathcal{D}_S + \overline{\mathcal{R}}_S$ is a closed *-subalgebra of $(\mathcal{A}_S, \|\cdot\|_S)$. ∎

Let $S \in \mathcal{S}_{\mathbb{Z}}$. We number the elements of $\mathrm{Sp}(S)$ in such a way that $\mathrm{Sp}(S) = \{n_i\}_{i \in I}$ is an increasing sequence,

$$0 \le n_i, \quad \text{for} \quad 0 \le i, \qquad \text{and} \quad 0 > n_i, \quad \text{for} \quad 0 > i.$$

Then $|i| \le |n_i|$ and, depending on S, the set I is either the set $\mathbb{Z}$ of all integers, or the set of all integers from $-\infty$ to some m, or from m to ∞. We consider the case when $I = \mathbb{Z}$. Two other cases can be considered similarly.

Set

$$\rho_S(k) = \left(\inf_{i \in \mathbb{Z}} |n_{i+k} - n_i| \right)^{-1}, \quad \text{for} \quad k \neq 0, \quad \text{and} \quad \rho_S(0) = 0.$$

Since $\inf_{i \in \mathbb{Z}} |n_{i+k} - n_i| \ge |k|$,

$$0 < \rho_S(k) \le \frac{1}{|k|}, \quad \text{for} \quad k \neq 0.$$

Proposition 2.6. *If*

$$\lim_{|i| \to \infty} (n_{i+1} - n_i) = \infty \tag{2.6}$$

$$\text{and} \quad \sum_{k \in \mathbb{Z}} \rho(k) \quad \text{converges} \tag{2.7}$$

then $\mathfrak{U}_S = \mathcal{D}_S + C_S(H)$.

Proof. Let $A = (A_{ij}) \in \mathcal{A}_S$, where A_{ij} are bounded operators from $H_S(n_j)$ into $H_S(n_i)$. Then the operator

$$B = SA - AS = (B_{ij}), \quad \text{where} \quad B_{ij} = (n_i - n_j) A_{ij},$$

is bounded. Set $b = \|B\|$. Since $\|B_{ij}\| \le \|B\|$, for all $i, j \in \mathbb{Z}$,

$$\|A_{ij}\| \le \frac{b}{|n_i - n_j|}, \quad \text{for} \quad i \neq j. \tag{2.8}$$

For $k \in \mathbb{Z} \setminus 0$ and $m > 0$, let

$$G_{ij}^{km} = A_{ij}, \quad \text{if} \quad j = i + k \quad \text{and} \quad -m \le i \le m, \quad \text{and} \quad G_{ij}^{km} = 0 \text{ otherwise.}$$

Then the operator $G^{km} = (G_{ij}^{km})$ belongs to $\mathcal{R}_S$. Taking into account (2.6) and (2.8), we obtain that the operators G^{km} converge uniformly in $B(H)$ to a bounded operator $G^k = (G_{ij}^k)$, as $m \to \infty$, where

$$G_{ij}^k = A_{ij}, \quad \text{if} \quad j = i + k, \quad \text{and} \quad G_{ij}^k = 0 \quad \text{otherwise.}$$

Therefore $G^k \in C_S(H)$ and, by (2.8),

$$\|G^k\| = \sup_i \|A_{ii+k}\| \le b \rho_S(k).$$

It follows from (2.7) that the operator $G = \sum_{k \in \mathbb{Z} \setminus 0} G^k$ belongs to $C_S(H)$. Since $A - G \in \mathcal{D}_S$, we obtain that $A \in \mathcal{D}_S + C_S(H)$, so that $\mathcal{A}_S \subseteq \mathcal{D}_S + C_S(H)$. It follows from Lemma 2.5 that $\mathfrak{U}_S = \mathcal{D}_S + C_S(H)$. $\square$

Corollary 2.7. *If there are $a>0$, $c>0$ and an integer N such that*
$$c|i|^a \le n_{i+1}-n_i \quad \text{for} \;\; N \le |i|$$
then $\mathfrak{U}_S = \mathcal{D}_S + C_S(H)$.

Proof. Condition (2.6), clearly, holds. Let $k>4N$. Then
$$\begin{aligned}\rho_S(k)^{-1} &= \inf_{i\in\mathbb{Z}} |n_{i+k}-n_i| = \inf_{i\in\mathbb{Z}} \sum_{p=1}^{k}(n_{i+p}-n_{i+p-1}) \\ &\ge c\sum_{m=N}^{[\frac{k}{2}]} m^a \ge \frac{c}{a+1}\left(\left[\frac{k}{2}\right]^{a+1} - (N-1)^{a+1}\right) \\ &\ge \frac{c}{a+1}\left(\frac{k}{4}\right)^{a+1}.\end{aligned}$$
Similarly, if $k<-2N$ then $\rho_S(k)^{-1} \ge \frac{c}{a+1}\left(\frac{|k|}{4}\right)^{a+1}$. Therefore condition (2.7) also holds and the result follows from Proposition 2.6. □

Suppose now that $\dim(H_S(n))=\infty$ for all $n\in\mathrm{Sp}(S)$ and let $n_0\in\mathrm{Sp}(S)$. Set $K=H_S(n_0)$. Then there exists a Hilbert space $\mathcal{H}$ with $\dim(\mathcal{H})=\infty$ such that the C^*-algebra $C_S(H)$ is isomorphic to the tensor product $B(K)\otimes C(\mathcal{H})$ where $C(\mathcal{H})$ is the C^*-algebra of all compact operators on $\mathcal{H}$. Choosing a basis $\{e_n\}_{n=1}^{\infty}$ in $\mathcal{H}$, we obtain that the algebra $\mathcal{D}_S$ is isomorphic to the von Neumann algebra tensor product $B(K)\overline{\otimes}\mathcal{L}$ of $B(K)$ and the W^*-algebra $\mathcal{L}$ of all operators on $\mathcal{H}$ diagonal with respect to $\{e_n\}_{n=1}^{\infty}$. From this and from Proposition 2.6 we obtain the following result.

Corollary 2.8. *Let $S\in\mathcal{S}_{\mathbb{Z}}$. If $\dim(H_S(n))=\infty$ for all $n\in\mathrm{Sp}(S)$ and conditions (2.6) and (2.7) hold then there exist Hilbert spaces K and $\mathcal{H}$ such that $\mathfrak{U}_S$ is isomorphic to $B(K)\overline{\otimes}\mathcal{L}+B(K)\otimes C(\mathcal{H})$, where $\mathcal{L}$ is the W^*-algebra of all operators on $\mathcal{H}$ diagonal with respect to some basis.*

Assume now that $\dim(H_S(n))<\infty$ for all $n\in\mathrm{Sp}(S)$. Then $C_S(H)$ coincides with the algebra $C(H)$ of all compact operators on H. Taking into account the definition of the ideal $\mathcal{K}_S$ and applying Proposition 2.6 we obtain the following result.

Corollary 2.9. *Let $S\in\mathcal{S}_{\mathbb{Z}}$ and $\dim(H_S(n))<\infty$ for all $n\in\mathrm{Sp}(S)$. If conditions (2.6) and (2.7) hold then $\mathfrak{U}_S=\mathcal{D}_S+C(H)$ and $\mathcal{A}_S=\mathcal{D}_S+\mathcal{K}_S$.*

Example 2.10. Let $\{e_i\}_{i=-\infty}^{\infty}$ be an orthonormal basis in H and let
$$Se_i = \mathrm{sgn}\,(i)|i|^{1+a}e_i, \quad \text{where} \;\; a>0.$$
Then $S\in\mathcal{S}_{\mathbb{Z}}$ and $n_i=\mathrm{sgn}\,(i)|i|^{1+a}$, so that
$$\lim_{|i|\to\infty}\frac{n_{i+1}-n_i}{\mathrm{sgn}\,(i)|i|^a} = 1+a.$$

Therefore, by Corollaries 2.7 and 2.9, $\mathfrak{U}_S = \mathcal{D}_S + C(H)$ and $\mathcal{A}_S = \mathcal{D}_S + \mathcal{K}_S$ where $\mathcal{D}_S$ is the algebra of all operators diagonal with respect to $\{e_i\}_{i=-\infty}^{\infty}$. Thus the quotient algebra $\mathcal{A}_S/\mathcal{K}_S$ is isomorphic to the commutative C^*-algebra $\mathcal{D}_S/\mathfrak{L}$ where $\mathfrak{L}$ is the algebra of all compact diagonal operators on H. □

Let $\{e_i\}_{i=-\infty}^{\infty}$ be an orthonormal basis in H and let

$$Se_i = ie_i \ \text{ and } \ Ue_i = e_{i+1}, \quad \text{for all } \ i \in \mathbb{Z}.$$

Then $S \in \mathcal{S}_{\mathbb{Z}}$ and U is the shift operator. We have that

$$UD(S) \subseteq D(S),\ U^*D(S) \subseteq D(S) \ \text{ and } \ (SU - US)|_{D(S)} \ \text{ extends to } \ U,$$

so that $U \in \mathcal{A}_S$. Hence $\mathfrak{U}_S$ contains the C^*-algebra $C(\mathcal{D}_S, U)$ generated by U and by the commutative algebra $\mathcal{D}_S$ of all operators diagonal with respect to $\{e_i\}_{i=-\infty}^{\infty}$.

Problem 2.11. Is $\mathfrak{U}_S = C(\mathcal{D}_S, U)$?

3. Dual and second dual spaces of the algebras $\mathcal{F}_S$.

Let S be a closed symmetric operator. Recall that $\mathcal{F}_S$ is the closure with respect to the norm $\|\cdot\|_S$ (see (1.2)) of the subalgebra of all finite rank operators in $\mathcal{A}_S$. If S is a bounded symmetric operator on H, it follows that $\mathcal{F}_S = C(H)$ and $\mathcal{A}_S = B(H)$, so that $\mathcal{A}_S$ is isometrically isomorphic to the second dual of $\mathcal{F}_S$. In this section we study the structure of the dual and the second dual spaces of the algebra $\mathcal{F}_S$ for unbounded symmetric operators S. In the case when S is selfadjoint we establish the full analogy with the bounded case: The algebra $\mathcal{A}_S$ is isometrically isomorphic to the second dual of $\mathcal{F}_S$.

By $T(H)$ we denote the Banach *-algebra of trace class operators on H with the norm

$$|A| = \sum_{i=1}^{\infty} s_i(A) = \mathrm{Tr}\left((A^*A)^{1/2}\right),$$

where $\{s_i(A)\}_{i=1}^{\infty}$ is the set of all eigenvalues of the positive compact operator $(A^*A)^{1/2}$.

It is well known that $T(H)$ can be identified with the dual space of the algebra $C(H)$: For any $T \in T(H)$,

$$F_T(A) = \mathrm{Tr}(AT), \quad A \in C(H),$$

is a bounded linear functional on $C(H)$ and $\|F_T\| = |T|$; and that $B(H)$ can be identified with the dual space of $T(H)$: For any $B \in B(H)$,

$$\theta_B(T) = \mathrm{Tr}\,(BT), \quad T \in T(H),$$

is a bounded linear functional on $T(H)$ and $\|\theta\| = \|B\|$.

Set $\widehat{B}(H) = B(H) \oplus B(H)$ and $\widehat{C}(H) = C(H) \oplus C(H)$. Then $\widehat{B}(H)$ and $\widehat{C}(H)$ are Banach spaces with the norm

$$\|A \oplus B\| = \|A\| + \|B\|.$$

Set $\widehat{T}(H) = T(H) \oplus T(H)$. It is a Banach space with the norm

$$|R \oplus T| = \max(|R|, |T|), \quad T, R \in T(H),$$

and it can be identified with the dual space of $\widehat{C}(H)$: For $R, T \in T(H)$,

$$F_{R\oplus T}(A \oplus B) = \mathrm{Tr}\,(AR) + \mathrm{Tr}\,(BT), \quad A \oplus B \in \widehat{C}(H), \tag{3.1}$$

is a bounded linear functional on $\widehat{C}(H)$ and $\|F_{R\oplus T}\| = |R \oplus T|$. Similarly, $\widehat{B}(H)$ can be identified with the dual space of $\widehat{T}(H)$: For $A, B \in B(H)$,

$$\theta_{A\oplus B}(R \oplus T) = \mathrm{Tr}\,(AR) + \mathrm{Tr}\,(BT), \quad R \oplus T \in \widehat{T}(H), \tag{3.2}$$

is a bounded linear functional on $\widehat{T}(H)$ and $\|\theta_{A\oplus B}\| = \|A \oplus B\|$.

Set

$$\widehat{\mathcal{A}}_S = \{A \oplus A_S : A \in \mathcal{A}_S\} \quad \text{and} \quad \widehat{\mathcal{F}}_S = \{A \oplus A_S : A \in \mathcal{F}_S\},$$

where $A_S = \mathrm{Closure}\,(SA - AS)$. Then $(\mathcal{A}_S, \|\cdot\|_S)$ and $(\widehat{\mathcal{A}}_S, \|\cdot\|)$, $(\mathcal{F}_S, \|\cdot\|_S)$ and $(\widehat{\mathcal{F}}_S, \|\cdot\|)$ are isometrically isomorphic, since

$$\|A\|_S = \|A\| + \|A_S\| = \|A \oplus A_S\|.$$

Therefore $\widehat{\mathcal{A}}_S$ is a closed subspace of $\widehat{B}(H)$ and $\widehat{\mathcal{F}}_S$ is a closed subspace of $\widehat{C}(H)$, since $A \in \mathcal{F}_S$ implies $A_S \in C(H)$.

Set

$$\mathfrak{T}_S = \Big\{T \in T(H) : TD(S) \subseteq D(S^*),\ T^*D(S) \subseteq D(S^*) \text{ and the operator}$$
$$(S^*T - TS)|_{D(S)} \text{ extends to a bounded trace class operator } \mathbb{T}\Big\}.$$

If $T \in \mathfrak{T}_S \cap \mathcal{A}_S$ then $\mathbb{T}_S = T_S$. In particular, if S is selfadjoint then $\mathbb{T}_S = T_S$ for all $T \in \mathfrak{T}_S$. Clearly, $\mathfrak{T}_S$ is a linear subspace in $T(H)$ and

$$\check{\mathfrak{T}}_S = \{\mathbb{T}_S \oplus T : T \in \mathfrak{T}_S\}$$

is a linear subspace in $\widehat{T}(H)$. For $T \in \mathfrak{T}_S$ and $z, u \in D(S)$,

$$-((\mathbb{T}_S)^*z, u) = -(z, \mathbb{T}_S u) = -(z, (S^*T - TS)u) = ((S^*T^* - T^*S)z, u),$$

so that

$$-(\mathbb{T}_S)^*|_{D(S)} = (S^*T^* - T^*S)|_{D(S)} = (\mathbb{T}^*)_S|_{D(S)}. \tag{3.3}$$

Therefore $T^* \in \mathfrak{T}_S$.

For $x, y \in H$, the rank one operator $x \otimes y$ on H is defined by the formula

$$(x \otimes y)z = (z, x)y. \tag{3.4}$$

It is easy to check that

$$\|x\otimes y\| = \|x\|\,\|y\|, \tag{3.5}$$
$$(x\otimes y)^* = y\otimes x,\ (x\otimes y)(u\otimes v) = (v,x)(u\otimes y),$$
$$R(x\otimes y) = x\otimes Ry,\ \text{ and }\ (x\otimes y)R\ \text{ extends to }\ (R^*x)\otimes y,$$

if R is a densely defined operator, $y\in D(R)$ and $x\in D(R^*)$. Let $\{e_j\}_{j=1}^\infty$ be a basis in H. Then

$$\begin{aligned}\operatorname{Tr}(x\otimes y) &= \sum_{j=1}^\infty((x\otimes y)e_j,e_j) = \sum_{j=1}^\infty (e_j,x)(y,e_j)\\ &= \left(y, \sum_{j=1}^\infty (x,e_j)e_j\right) = (y,x).\end{aligned} \tag{3.6}$$

Let $x,y\in D(S^*)$ and $T = x\otimes y$. By (3.4) and (3.5),

$$Tz = (z,x)y\in D(S^*) \tag{3.7}$$
$$T^*z = (y\otimes x)z = (z,y)x\in D(S^*),\quad \text{for }\ z\in H,$$
$$\text{and }\ \mathbb{T}_S = S^*T - TS = x\otimes S^*y - (S^*x)\otimes y\in T(H),$$

so that $T\in\mathfrak{T}_S$. By Φ_S we denote the set of all linear combinations of the operators $x\otimes y$, for $x,y\in D(S^*)$. Clearly, $\Phi\subset\mathfrak{T}_S$ and

$$\check{\Phi}_S = \{\mathbb{T}_S\oplus T : T\in\Phi_S\}$$

is a linear subspace of $\check{\mathfrak{T}}_S$.

Let X^* be the dual space of a Banach space X and Y be a linear subspace of X. The *annihilator*

$$Y^\perp = \{F\in X^* : F(y) = 0,\quad \text{for all }\ y\in Y\}$$

of Y in X^* is a closed subspace of X^* and from the general theory of Banach spaces (see [**5**] II.4.18 and [**15**] III, Problem 30) we have the following lemma.

Lemma 3.1. *The dual space Y^* of a closed subspace Y of X is isometrically isomorphic to the quotient space $X^*/Y^\perp$ and the second dual Y^{**} of Y is isometrically isomorphic to $Y^{\perp\perp}$ where*

$$Y^{\perp\perp} = \{\theta\in X^{**} : \theta(F) = 0,\quad \textit{for all }\ F\in Y^\perp\}.$$

Since $\widehat{\mathcal{F}}_S\subseteq\widehat{C}(H)$, the annihilator $(\widehat{\mathcal{F}}_S)^\perp$ is a closed subspace of the dual space $\widehat{C}(H)^* = \widehat{T}(H)$ and, since $\check{\Phi}_S\subseteq\check{\mathfrak{T}}_S\subseteq\widehat{T}(H)$, the annihilator $(\check{\Phi}_S)^\perp$ is a closed subspace of the dual space $\widehat{T}(H)^* = \widehat{B}(H)$.

Theorem 3.2. (i) *$\check{\mathfrak{T}}_S$ is a closed subspace in $\widehat{T}(H)$ and $(\widehat{\mathcal{F}}_S)^\perp = \check{\mathfrak{T}}_S$.*
(ii) $(\check{\mathfrak{T}}_S)^\perp\subseteq(\check{\Phi}_S)^\perp = \{A\oplus A_S : A\in\mathcal{A}_S\ \ \textit{and}\ \ AD(S^*)\subseteq D(S)\}\subseteq\widehat{\mathcal{A}}_S$.

Proof. Let $\mathbb{T}_S \oplus T \in \check{\mathfrak{T}}_S$ and $x, y \in D(S)$. Then $A = x \otimes y \in \mathcal{F}_S$ and, by (3.3) and (3.5),

$$\begin{aligned} &A_S = S(x \otimes y) - (x \otimes y)S = x \otimes Sy - (Sx) \otimes y, \\ &A_S T = (x \otimes Sy)T - ((Sx) \otimes y)T = (T^*x) \otimes Sy - (T^*Sx) \otimes y, \\ &A\mathbb{T}_S = (x \otimes y)\mathbb{T}_S = ((\mathbb{T}_S)^*x) \otimes y = ((T^*S - S^*T^*)x) \otimes y. \end{aligned} \tag{3.8}$$

Therefore, by (3.1), (3.6) and (3.8),

$$\begin{aligned} F_{\mathbb{T}_S \oplus T}(A \oplus A_S) &= \mathrm{Tr}\,(A\mathbb{T}_S) + \mathrm{Tr}(A_S T) \\ &= (y, (T^*S - S^*T^*)x) + (Sy, T^*x) - (y, T^*Sx) = 0. \end{aligned}$$

It follows from Lemma 3.1 [**13**] that any finite rank operator A in $\mathcal{F}_S$ has the form $A = \sum_{i=1}^n x_i \otimes y_i$ where $x_i, y_i \in D(S)$. Hence $F_{\mathbb{T}_S \oplus T}(A \oplus A_S) = 0$ for any finite rank operator $A \in \mathcal{F}_S$. Since, by definition of $\mathcal{F}_S$, finite rank operators are dense in $(\mathcal{F}_S, \|\cdot\|_S)$ and since $(\mathcal{F}_S, \|\cdot\|_S)$ and $(\widehat{\mathcal{F}}_S, \|\cdot\|)$ are isometrically isomorphic, the operators $A \oplus A_S$, where A are finite rank operators, are dense in $\widehat{\mathcal{F}}_S$. Since $F_{\mathbb{T}_S \oplus T}$ is continuous on $\widehat{C}(H)$, $F_{\mathbb{T}_S \oplus T}(A \oplus A_S) = 0$, for all $A \in \mathcal{F}_S$. Therefore $F_{\mathbb{T}_S \oplus T} \in (\widehat{\mathcal{F}}_S)^\perp$, so that $\check{\mathfrak{T}}_S \subseteq (\widehat{\mathcal{F}}_S)^\perp$.

Conversely, let $R \oplus T \in (\widehat{\mathcal{F}}_S)^\perp \subseteq \widehat{T}(H)$ and let $A = x \otimes y \in \mathcal{F}_S$, where $x, y \in D(S)$. From (3.1), (3.5), (3.6) and (3.8) it follows that

$$\begin{aligned} 0 = F_{R \oplus T}(A \oplus A_S) &= \mathrm{Tr}(AR) + \mathrm{Tr}(A_S T) \\ &= \mathrm{Tr}((R^*x) \otimes y) + \mathrm{Tr}[(T^*x) \otimes Sy - (T^*Sx) \otimes y] \\ &= (y, R^*x) + (Sy, T^*x) - (y, T^*Sx). \end{aligned}$$

Hence

$$(Sy, T^*x) = (y, (T^*S - R^*)x), \quad \text{for all } \; x, y \in D(S).$$

Therefore $T^*x \in D(S^*)$ and $S^*T^*x = (T^*S - R^*)x$. Thus $T^*D(S) \subseteq D(S^*)$ and

$$(Sx, Ty) = (T^*Sx, y) = (S^*T^*x, y) + (R^*x, y) = (x, TSy) + (x, Ry).$$

From this it follows that $Ty \in D(S^*)$ and $S^*Ty = TSy + Ry$. Hence

$$TD(S) \subseteq D(S^*) \quad \text{and} \quad R|_{D(S)} = S^*T|_{D(S)} - TS|_{D(S)}.$$

Therefore $T \in \mathfrak{T}_S$ and $R = \mathbb{T}_S$. Thus $(\widehat{\mathcal{F}}_S)^\perp \subseteq \check{\mathfrak{T}}_S$, so that $(\widehat{\mathcal{F}}_S)^\perp = \check{\mathfrak{T}}_S$. From this we also obtain that $\check{\mathfrak{T}}_S$ is a closed subspace of $\widehat{T}(H)$. Part (i) is proved.

Since $\check{\Phi}_S \subseteq \check{\mathfrak{T}}_S$, we have $(\check{\mathfrak{T}}_S)^\perp \subseteq (\check{\Phi}_S)^\perp$. Let now $A \oplus A_S \in \widehat{\mathcal{A}}_S$ and $AD(S^*) \subseteq D(S)$. It was shown in Lemma 3.1 [**13**] that

$$A_S|_{D(S^*)} = (S^*A - AS^*)|_{D(S^*)}.$$

For $x, y \in D(S^*)$, the operator $T = x \otimes y$ belongs to Φ_S and, taking the above equality into account, we obtain from (3.5) and (3.7) that

$$A_S T = x \otimes A_S y = x \otimes (S^*A - AS^*)y \quad \text{and}$$
$$A\mathbb{T}_S = A(x \otimes S^*y - (S^*x) \otimes y) = x \otimes AS^*y - (S^*x) \otimes Ay.$$

Therefore, by (3.2) and (3.6),

$$\begin{aligned}\theta_{A\oplus A_S}(\mathbb{T}_S \oplus T) &= \mathrm{Tr}(A\mathbb{T}_S) + \mathrm{Tr}(A_S T)\\ &= (AS^*y, x) - (Ay, S^*x) + (S^*Ay, x) - (AS^*y, x)\\ &= (S^*Ay, x) - (Ay, S^*x).\end{aligned}$$

Since $AD(S^*) \subseteq D(S)$, it follows that $S^*Ay = SAy$ and $(Ay, S^*x) = (SAy, x)$. Hence $\theta_{A\oplus A_S}(\mathbb{T}_S \oplus T) = 0$ and, by linearity, it holds for all $T \in \Phi_S$. Therefore

$$\{A \oplus A_S : A \in \mathcal{A}_S \ \text{ and } \ AD(S^*) \subseteq D(S)\} \subseteq \left(\check{\Phi}_S\right)^{\perp}. \tag{3.9}$$

Conversely, let $A \oplus B \in (\check{\Phi}_S)^{\perp}$. Then, for every $x, y \in D(S^*)$, $T = x \otimes y \in \Phi_S$ and

$$\theta_{A\oplus B}(\mathbb{T}_S \oplus T) = \mathrm{Tr}(AT_S) + \mathrm{Tr}(BT) = 0.$$

By (3.5), $BT = x \otimes By$ and, as above, $A\mathbb{T}_S = x \otimes AS^*y - (S^*x) \otimes Ay$. Hence, by (3.6),

$$0 = (AS^*y, x) - (Ay, S^*x) + (By, x).$$

Thus

$$(Ay, S^*x) = (AS^*y, x) + (By, x), \quad \text{for all } \ x, y \in D(S^*).$$

Therefore $Ay \in D(S^{**})$ and $S^{**}Ay = AS^*y + By$. Since S is closed, $S^{**} = S$ and we obtain that

$$AD(S^*) \subseteq D(S) \quad \text{and} \quad B|_{D(S^*)} = (SA - AS^*)|_{D(S^*)}. \tag{3.10}$$

Restricting (3.10) to D(S), we have

$$AD(S) \subseteq D(S) \quad \text{and} \quad B|_{D(S)} = (SA - AS)|_{D(S)}.$$

Making use of (3.10), we obtain that for $z \in D(S)$ and $u \in D(S^*)$,

$$(A^*z, S^*u) = (z, AS^*u) = (z, SAu) - (z, Bu) = (A^*Sz, u) - (B^*z, u).$$

Therefore $A^*z \in D(S^{**})$. Since $S^{**} = S$, we have $A^*D(S) \subseteq D(S)$. Thus $A \in \mathcal{A}_S$ and $B = A_S$, so $A \oplus B = A \oplus A_S \in \widehat{\mathcal{A}}_S$. Taking into account that $AD(S^*) \subseteq D(S)$, we obtain that

$$(\check{\Phi}_S)^{\perp} \subseteq \{A \oplus A_S : A \in \mathcal{A}_S \ \text{ and } \ AD(S^*) \subseteq D(S)\}.$$

Combining this with (3.9), we complete the proof of the theorem. □

Since the Banach spaces $(\mathcal{F}_S, \|\cdot\|_S)$ and $(\widehat{\mathcal{F}}_S, \|\cdot\|)$ and the Banach spaces $(\mathcal{A}_S, \|\cdot\|_S)$ and $(\widehat{\mathcal{A}}_S, \|\cdot\|)$ are isometrically isomorphic and since $(\widehat{\mathcal{F}}_S, \|\cdot\|)$ is a closed subspace of $\widehat{C}(H)$, Lemma 3.1 and Theorem 3.2 yield:

Corollary 3.3. *The dual space of the Banach *-algebra* $(\mathcal{F}_S, \|\cdot\|_S)$ *is isometrically isomorphic to the quotient space* $\widehat{T}(H)/\check{\mathfrak{T}}_S$ *and the second dual space of* $(\mathcal{F}_S, \|\cdot\|_S)$ *is isometrically isomorphic to a closed subspace of* $(\mathcal{A}_S, \|\cdot\|_S)$.

The following example shows that if S is not selfadjoint then, generally speaking, $(\check{\Phi}_S)^{\perp} \neq \widehat{\mathcal{A}}_S$, so that $(\mathcal{F}_S)^{\perp\perp} \neq \widehat{\mathcal{A}}_S$ and the second dual space of $(\mathcal{F}_S, \|\cdot\|_S)$ is isometrically isomorphic to a proper subspace of $(\mathcal{A}_S, \|\cdot\|_S)$.

Example 3.4. Let, as in Example 2.2, $H = L^2(0,1)$ and the operator $S = i\frac{d}{dt}$ with domain $D(S) = \{h(t) : h, h' \in L_2(0,1)$ and $h(0) = h(1) = 0\}$. Then S is a symmetric operator, non-selfadjoint and

$$D(S^*) = \{h(t) : h, h' \in L^2(0,1)\}.$$

Let $g(t)$ be a differentiable function on $[0,1]$ such that $g(0) \neq 0$ and let M_g be the bounded operator of multiplication by $g(t)$ on H. Then $M_g \in \mathcal{A}_S$. If $h(t) \in D(S^*)$ and $h(0) \neq 0$ then $(M_g h)(0) = g(0)h(0) \neq 0$, so that $M_g h \notin D(S)$. Thus $M_g \oplus (M_g)_S \notin \{A \oplus A_S : A \in \mathcal{A}_S$ and $AD(S^*) \subseteq D(S)\}$. Hence $(\check{\Phi}_S)^{\perp} \neq \mathcal{A}_S$.

Assume now that S is selfadjoint. Then $D(S^*) = D(S)$, $\mathbb{T}_S = T_S$, for $T \in \mathfrak{T}_S$, and

$$\mathfrak{T}_S = \{T \in T(H) \cap \mathcal{A}_S : T_S \in T(H)\} \subseteq \mathcal{A}_S.$$

It is well known (see, for example, [**5**] and [**6**]) that the algebra $T(H)$ is a two-sided ideal of $B(H)$ and if $A \in B(H)$ and $B \in T(H)$ then

$$|AB| \leq \|A\|\,|B|, \quad |B^*| = |B| \quad \text{and} \quad \|B\| \leq |B|. \tag{3.11}$$

We consider now two equivalent norms on $\mathfrak{T}_S$:

$$|T|_1 = |T| + |T_S| \ \text{ and } \ |T|_2 = \max(|T|, |T_S|), \quad \text{for } \ T \in \mathfrak{T}_S.$$

Since

$$\mathbb{T}_S = T \ \text{ and } \ |T|_2 = \max(|T|, |T_S|) = |\mathbb{T}_S \oplus T|, \quad \text{for } \ T \in \mathfrak{T}_S,$$

$(\mathfrak{T}_S, |\cdot|_2)$ is isometrically isomorphic to $\check{\mathfrak{T}}_S$.

Proposition 3.5. *Let S be selfadjoint. Then:*

(i) $\mathfrak{T}_S \subset \mathcal{F}_S$ *and* $(\mathfrak{T}_S, |\cdot|_2)$ *is a two-sided Banach* $\mathcal{A}_S$ *-module;*
(ii) $(\mathfrak{T}_S, |\cdot|_1)$ *is a Banach *-algebra and a* **D**-*subalgebra of* $C(H)$ (*see* (1.1)) *with* $D = 1$.

Proof. It was shown in [**13**] that if S is selfadjoint then $\mathcal{F}_S$ coincides with the algebra $\mathcal{J}_S = \{A \in \mathcal{A}_S : A$ and A_S belong to $C(H)\}$. Since $\mathfrak{T}_S \subset \mathcal{J}_S$, we obtain that $\mathfrak{T}_S \subset \mathcal{F}_S$.

In Theorem 3.2(i) it was shown that $\check{\mathfrak{T}}_S$ is a closed subspace of $\widehat{T}(H)$. Since $(\mathfrak{T}_S, |\cdot|_2)$ is isometrically isomorphic to $\check{\mathfrak{T}}_S$, it is a Banach space.

Since the norms $|\cdot|_1$ and $|\cdot|_2$ are equivalent, $(\mathfrak{T}_S, |\cdot|_1)$ is also a Banach space.

For $A, B \in \mathcal{A}_S$,

$$\begin{aligned}(AB)_S|_{D(S)} &= (SAB - ABS)|_{D(S)} \\ &= [(SA - AS)B + A(SB - BS)]|_{D(S)} = (A_S B + AB_S)|_{D(S)},\end{aligned}$$

so that

$$(AB)_S = A_S B + AB_S. \tag{3.12}$$

Let $T \in \mathfrak{T}_S$ and $A \in \mathcal{A}_S$. Then $T, T_S \in T(H)$. Since $\mathfrak{T}_S \subset \mathcal{A}_S$ and $T(H)$ is a two-sided ideal of $B(H)$, it follows that $AT \in T(H) \cap \mathcal{A}_S$ and, by (3.12),

$$(AT)_S = A_S T + AT_S \in T(H).$$

Therefore $AT \in \mathfrak{T}_S$. Making use of (3.11), we obtain that

$$\begin{aligned}|AT|_2 &= \max(|AT|, |(AT)_S|) \leq \max(\|A\|\,|T|, \|A_S\|\,|T| + \|A\|\,|T_S|) \\ &\leq (\|A\| + \|A_S\|)\max(|T|, |T_S|) = \|A\|_S |T|_2.\end{aligned}$$

Similarly, $TA \in \mathfrak{T}_S$ and $|TA|_2 \leq \|A\|_S |T|_2$. Thus $(\mathfrak{T}_S, |\cdot|_2)$ is a two-sided Banach $\mathcal{A}_S$-module. Part (i) is proved.

From (i) and from the fact that $\mathfrak{T}_S \subseteq \mathcal{A}_S$, we have that $\mathfrak{T}_S$ is an algebra. We also have that $T^* \in \mathfrak{T}_S$ and, since $\mathbb{T}_S = T_S$, it follows from (3.3) that $(T^*)_S = -(T_S)^* \in T(H)$. Taking this and (3.11) into account, we obtain that

$$|T^*|_1 = |T^*| + |(T^*)_S| = |T^*| + |-(T_S)^*| = |T| + |T_S| = |T|_1$$

and

$$\begin{aligned}|TR|_1 &= |TR| + |(TR)_S| = |TR| + |T_S R + TR_S| \\ &\leq \|T\|\,|R| + |T_S|\,\|R\| + \|T\|\,|R_S| \\ &\leq |T|\,|R| + |T_S|\,|R| + |T|\,|R_S| \leq |T|_1\,|R|_1,\end{aligned}$$

for $T, R \in \mathfrak{T}_S$. Hence $(\mathfrak{T}_S, |\cdot|_1)$ is a Banach *-algebra.

Clearly, $\mathfrak{T}_S$ is dense in $C(H)$. For $T, R \in \mathfrak{T}_S$, it follows from (3.11) that

$$\begin{aligned}|TR|_1 &= |TR| + |(TR)_S| = |TR| + |T_S R + TR_S| \\ &\leq \|T\|\,|R| + |T_S|\,\|R\| + \|T\|\,|R_S| \\ &\leq \|T\|(|R| + |R_S|) + (|T| + |T_S|)\|R\| \\ &= \|T\|\,|R|_1 + |T|_1\|R\|.\end{aligned}$$

Thus $(\mathfrak{T}_S, |\cdot|_1)$ is a **D**-subalgebra of $C(H)$ with the constant $D = 1$. $\square$

If S is selfadjoint, it follows from Theorem 3.2 that $(\check{\Phi}_S)^\perp = \widehat{\mathcal{A}}_S$ and

$$\left(\widehat{\mathcal{F}}_S\right)^{\perp\perp} = (\check{\mathfrak{T}}_S)^\perp \subseteq (\check{\Phi}_S)^\perp = \widehat{\mathcal{A}}_S.$$

In order to prove that $(\widehat{\mathcal{F}}_S)^{\perp\perp} = \widehat{\mathcal{A}}_S$ it suffices to show that $\check{\Phi}_S$ is dense in $\check{\mathfrak{T}}_S$. For this we need the following lemma which is a partial case of the general result obtained by Gohberg and Krein [**6**, Theorem 6.3] for symmetrically normable ideals.

Lemma 3.6. *Let $T \in T(H)$ and let Q_n be finite rank projections which converge to $\mathbf{1}_H$ in the strong operator topology. Then*

$$|T - Q_nT| \to 0 \quad and \quad |T - TQ_n| \to 0, \quad as \quad n \to \infty.$$

Proof. Let $A = x\otimes y$, $x, y \in H$. By (3.5), $A^*A = \|y\|^2(x\otimes x)$ and the operator $(A^*A)^{1/2} = \frac{\|y\|}{\|x\|}(x \otimes x)$ has only one non-zero eigenvalue $\lambda = \|x\|\,\|y\|$. Hence

$$|x \otimes y| = |A| = \mathrm{Tr}(A^*A)^{1/2} = \lambda = \|x\|\,\|y\|. \tag{3.13}$$

If $T = \sum_{i=1}^k x_i \otimes y_i$ is a finite rank operator then, by (3.5) and (3.13),

$$\begin{aligned} |T - Q_nT| &= \left|\sum_{i=1}^k x_i \otimes (y_i - Q_ny_i)\right| \le \sum_{i=1}^k |x \otimes (y_i - Q_ny_i)| \\ &= \sum_{i=1}^k \|x_i\|\,\|y_i - Q_ny_i\| \to 0, \end{aligned}$$

as $n \to \infty$. For any T in $T(H)$ and any $\varepsilon > 0$, there is a finite rank operator T_ε such that $|T - T_\varepsilon| < \varepsilon$. Making use of the inequality (3.11), we obtain that

$$\begin{aligned} |T - Q_nT| &\le |T - T_\varepsilon| + |T_\varepsilon - Q_nT_\varepsilon| + |Q_n(T - T_\varepsilon)| \\ &\le \varepsilon + |T_\varepsilon - Q_nT_\varepsilon| + \|Q_n\|\,|T - T_\varepsilon| \\ &\le 2\varepsilon + |T_\varepsilon - Q_nT_\varepsilon|. \end{aligned}$$

Since T_ε is a finite rank operator, by the above argument, there is n_ε such that $|T_\varepsilon - Q_nT_\varepsilon| \le \varepsilon$, for $n > n_\varepsilon$. Hence $|T - Q_nT| \le 3\varepsilon$ and $|T - Q_nT| \to 0$, as $n \to \infty$. Similarly, one can prove that $|T - TQ_n| \to 0$, as $n \to \infty$. □

Proposition 3.7. *Let S be selfadjoint. Then Φ_S is dense in $(\mathfrak{T}_S, |\cdot|_1)$.*

Proof. Let $[S]$ be the selfadjoint operator constructed in Section 2. Then $D(S) = D([S])$, so that $\Phi_S = \Phi_{[S]}$. Since $B = S - [S]$ is a bounded operator, $BT - TB \in T(H)$, for $T \in T(H)$. Therefore, taking into account that

$$(ST - TS)_{D(S)} = ([S]T - T[S])_{D(S)} + (BT - TB)_{D(S)},$$

we conclude that $\mathfrak{T}_S = \mathfrak{T}_{[S]}$ and $T_S = T_{[S]} + BT - TB$.

Making use of (3.11), we obtain that for any $T \in \mathfrak{T}_S$,

$$\begin{aligned} |T| + |T_S| &= |T| + \left|T_{[S]} + BT - TB\right| \\ &\le |T| + \left|T_{[S]}\right| + 2\|B\|\,|T| \\ &\le (1 + 2\|B\|)\left(|T| + \left|T_{[S]}\right|\right). \end{aligned}$$

Similarly, $|T|+|T_{[S]}| \le (1+2\|B\|)(|T|+|T_S|)$. Thus the norms $|\cdot|_1$ generated by the operators S and $[S]$ on $\mathfrak{T}_S$ are equivalent. Hence to obtain the proof we only have to show that $\Phi_{[S]}$ is dense in $(\mathfrak{T}_{[S]}, |\cdot|_1)$.

In every subspace $H_S(n)$ (see (2.2)) we choose an increasing sequence of finite-dimensional projections $\{Q_n^k\}_{k=1}^{\infty}$ converging to the projection $P_S(n)$ (see (2.1)) in the strong operator topology as $k \to \infty$. Set

$$Q^k = \sum_{n=-k}^{k} \oplus Q_n^k.$$

Then Q^k are finite-dimensional projections commuting with $[S]$. Hence $Q^k \in \Phi_{[S]}$. The projections Q^k converge to $\mathbf{1}_H$ in the strong operator topology. Let $T \in \mathfrak{T}_{[S]}$. Then $Q_nT \in \Phi_{[S]}$ and

$$[S]Q^kT - Q^kT[S] = Q^k[S]T - Q^kT[S] = Q^k([S]T - T[S]) = Q^kT_{[S]}.$$

Therefore $(Q^kT)_{[S]} = Q^kT_{[S]}$.

Since $T, T_{[S]} \in T(H)$, we obtain from Lemma 3.6 that

$$|T - Q^kT| \to 0 \ \text{ and } \ \left|T_{[S]} - (Q^kT)_{[S]}\right| = \left|T_{[S]} - Q^kT_{[S]}\right| \to 0, \ \text{ as } \ k \to \infty.$$

Hence

$$|T - Q^kT|_1 = |T - Q^kT| + \left|T_{[S]} - (Q^kT)_{[S]}\right| \to 0$$

as $k \to \infty$, so that $\Phi_{[S]}$ is dense in $(\mathfrak{T}_{[S]}, |\cdot|_1)$. □

Corollary 3.8. *Let S be a selfadjoint operator. Then:*

(i) *the Banach *-algebra $(\mathfrak{T}_S, |\cdot|_1)$ is simple;*
(ii) $(\check{\mathfrak{T}}_S)^{\perp} = (\check{\Phi}_S)^{\perp} = \widehat{\mathcal{A}}_S$;
(iii) *the dual space of $(\mathfrak{T}_S, |\cdot|_2)$ is isometrically isomorphic to the quotient space $\widehat{B}(H)/\widehat{\mathcal{A}}_S$.*

Proof. Let I be a closed two-sided ideal of $(\mathfrak{T}_S, |\cdot|_1)$ and $0 \neq T \in I$. Since $D(S)$ is dense in H, there is $x \in D(S)$ such that $Tx \neq 0$. Since S is selfadjoint, it follows from the definition of $\mathfrak{T}_S$ that $Tx \in D(S)$. From this and from the discussion before Lemma 3.1 we obtain that the rank one operators $y \otimes x$ and $Tx \otimes z$ belong to $\mathfrak{T}_S$ for any $y, z \in D(S)$. By (3.5), $T(y \otimes x) = (y \otimes Tx) \in I$ and

$$(Tx \otimes z)(y \otimes Tx) = \|Tx\|^2(y \otimes z) \in I.$$

Thus $y \otimes z \in I$ and, therefore, $\Phi_S \subseteq I$. Since I is closed, we obtain from Proposition 3.7 that $I = \mathfrak{T}_S$. Part (i) is proved.

Since the norms $|\cdot|_1$ and $|\cdot|_2$ are equivalent on $\mathfrak{T}_S$, it follows from Proposition 3.7 that Φ_S is dense in $(\mathfrak{T}_S, |\cdot|_2)$. Taking into account that $(\mathfrak{T}_S, |\cdot|_2)$ is isometrically isomorphic to the closed subspace $\check{\mathfrak{T}}_S$ of $\widehat{T}(H)$, we obtain that the linear subspace $\check{\Phi}_S$ is dense in $\check{\mathfrak{T}}_S$. From this and from Theorem 3.2(ii) we obtain $(\check{\mathfrak{T}}_S)^\perp = (\check{\Phi}_S)^\perp = \widehat{\mathcal{A}}_S$. Part (ii) is proved.

The dual space of $(\mathfrak{T}_S, |\cdot|_2)$ is isometrically isomorphic to the dual space of the closed subspace $\check{\mathfrak{T}}_S$ of $\widehat{T}(H)$. Since $\widehat{T}(H)^* = \widehat{B}(H)$, part (iii) follows from (ii) and from Lemma 3.1. $\square$

Theorem 3.9. *If S is a selfadjoint operator then $(\widehat{\mathcal{F}}_S)^{\perp\perp} = \widehat{\mathcal{A}}_S$ and the second dual space of the algebra $(\mathcal{F}_S, \|\cdot\|_S)$ is isometrically isomorphic to the algebra $(\mathcal{A}_S, \|\cdot\|_S)$.*

Proof. Combining Theorem 3.2(i) and Corollary 3.8(ii) yields $(\widehat{\mathcal{F}}_S)^{\perp\perp} = \widehat{\mathcal{A}}_S$. Therefore it follows from Lemma 3.1 that the second dual space of $(\widehat{\mathcal{F}}_S, \|\cdot\|)$ is isometrically isomorphic to $(\widehat{\mathcal{A}}_S, \|\cdot\|)$. Taking into account that $(\mathcal{F}_S, \|\cdot\|_S)$ is isometrically isomorphic to $(\widehat{\mathcal{F}}_S, \|\cdot\|)$ and that $(\mathcal{A}_S, \|\cdot\|_S)$ is isometrically isomorphic to $(\widehat{\mathcal{A}}_S, \|\cdot\|)$, we complete the proof. $\square$

4. Isomorphism of the algebras $\mathcal{F}_S$ and $\mathcal{A}_S$.

In this section we study the problem of classification of the algebras $\mathcal{F}_S$ and $\mathcal{A}_S$ up to *-isomorphism. For isometrical *-isomorphism this problem is completely solved in Theorem 4.4. As far as bounded but not necessarily isometrical *-isomorphism is concerned, we have obtained some partial results in Theorems 4.6 and 4.8 for the case when S is selfadjoint.

Banach *-algebras $(\mathcal{A}, \|\ \|_{\mathcal{A}})$ and $(\mathcal{B}, \|\ \|_{\mathcal{B}})$ are *-isomorphic if there is a bounded *-isomorphism φ from $\mathcal{A}$ onto $\mathcal{B}$. They are isometrically *-isomorphic if, in addition, $\|\varphi(A)\|_{\mathcal{B}} = \|A\|_{\mathcal{A}}$, for $A \in \mathcal{A}$.

Let $(\mathcal{A}, \|\ \|_{\mathcal{A}})$ and $(\mathcal{B}, \|\ \|_{\mathcal{B}})$ be Banach *-algebras of operators on Hilbert spaces H and $\mathcal{H}$ (the norms $\|\cdot\|_{\mathcal{A}}$ and $\|\cdot\|_{\mathcal{B}}$ do not, generally speaking, coincide with the operator norms in $B(H)$ and $B(\mathcal{H})$) and let φ be a bounded *-isomorphism from $\mathcal{A}$ onto $\mathcal{B}$. An isometry operator U from H into $\mathcal{H}$ implements φ if

$$\varphi(A) = UAU^*, \quad A \in \mathcal{A}.$$

Lemma 4.1. *Let R and T be symmetric operators on $\mathcal{H}$, S be a symmetric operators on H, U be an isometry operator from $\mathcal{H}$ onto H and $t \in \mathbb{R}$.*

(i) *If $\mathcal{F}_R = \mathcal{F}_T$ then the norms $\|\cdot\|_R$ and $\|\cdot\|_T$ on this algebra are equivalent, so that the Banach *-algebras $(\mathcal{F}_R, \|\cdot\|_R)$ and $(\mathcal{F}_T, \|\cdot\|_T)$ are *-isomorphic.*

(ii) *If $R = \pm T + t\mathbf{1}_{\mathcal{H}}$ then $\mathcal{F}_R = \mathcal{F}_T$ and the norms $\|\cdot\|_R$ and $\|\cdot\|_T$ coincide.*

(iii) *If $S = \lambda UTU^* + B$, where $0 \neq \lambda \in R$ and B is a bounded selfadjoint operator, then $A \to UAU^*$ is a bounded *-isomorphism from $(\mathcal{F}_T, \|\cdot\|_T)$ onto $(\mathcal{F}_S, \|\cdot\|_S)$. If $\lambda = \pm 1$ and $B = t\mathbf{1}_H$ then $A \to UAU^*$ is an isometric *-isomorphism.*

The same results hold for the algebras $\mathcal{A}_S$.

Proof. By Proposition 3.2 [**13**], the algebras $\mathcal{F}_R$ and $\mathcal{F}_T$ are semisimple. Hence if $\mathcal{F}_R = \mathcal{F}_T$ then it follows from Johnson's uniqueness of norm theorem that the norms $\|\cdot\|_R$ and $\|\cdot\|_T$ on this algebra are equivalent. Therefore the identity mapping is a bounded *-isomorphism from $(\mathcal{F}_R, \|\cdot\|_R)$ onto $(\mathcal{F}_T, \|\cdot\|_T)$.

Let $R = \pm T + t\mathbf{1}_{\mathcal{H}}$. Then $D(R) = D(T)$ and $A_T = A_R$ for any $A \in \mathcal{A}_T$. Hence $\|A\|_R = \|A\|_T$ and $\mathcal{A}_R = \mathcal{A}_T$. The sets of finite rank operators in the algebras $\mathcal{F}_R$ and $\mathcal{F}_T$ coincide and, since these algebras are the closures of these sets with respect to the norm $\|\cdot\|_T$, we obtain that $\mathcal{F}_S = \mathcal{F}_T$.

If $S = \lambda UTU^* + B$ then $D(S) = UD(T)$ and, for $A \in \mathcal{A}_T$,

$$UAU^*D(S) = UAD(T) \subseteq UD(T) = D(S) \quad \text{and}$$
$$SUAU^* - UAU^*S = \lambda U(TA - AT)U^* + (BA - AB),$$

so that $UAU^* \in \mathcal{A}_S$ and $(UAU^*)_S = \lambda UA_TU^* + (BA - AB)$. Thus $\mathcal{A}_S = U\mathcal{A}_TU^*$ and

$$\begin{aligned}\|UAU^*\|_S &= \|UAU^*\| + \|(UAU^*)_S\| = \|A\| + \|\lambda UA_TU^* + (BA - AB)\| \\ &\le \|A\| + \lambda\|A\| + 2\|B\|\,\|A\| \le \max(\lambda, 1 + 2\|B\|)\|A\|_T,\end{aligned}$$

so that $\psi(A) = UAU^*$ is a bounded *-isomorphism from $(\mathcal{A}_T, \|\cdot\|_T)$ onto $(\mathcal{A}_S, \|\cdot\|_S)$. If A is a finite rank operator in $\mathcal{A}_T$ then UAU^* is a finite rank operator in $\mathcal{A}_S$. Therefore $\mathcal{F}_S = \psi(\mathcal{F}_T)$. □

Let S be a symmetric operator with domain $D(S)$. It was shown in Lemma 3.1 [**13**] that a finite rank operator A belongs to $\mathcal{F}_S$ if and only if

$$A = \sum_{i=1}^{n} x_i \otimes y_i, \quad \text{where } x_i, y_i \in D(S). \tag{4.1}$$

Theorem 4.2. *Let S and T be symmetric operators on H and $\mathcal{H}$ and let $\mathcal{B}$ and $\mathcal{C}$ be closed *-subalgebras of $(\mathcal{A}_S, \|\cdot\|_S)$ and $(\mathcal{A}_T, \|\cdot\|_T)$, respectively, such that $\mathcal{F}_S \subseteq \mathcal{B}$ and $\mathcal{F}_T \subseteq \mathcal{C}$. Let ψ be a bounded *-isomorphism from $\mathcal{C}$ onto $\mathcal{B}$ and let $\varphi = \psi|\mathcal{F}_T$. Then:*

(i) *φ is a bounded *-isomorphism of $(\mathcal{F}_T, \|\cdot\|_T)$ onto $(\mathcal{F}_S, \|\cdot\|_S)$;*
(ii) *there is an isometry operator U from $\mathcal{H}$ onto H implementing ψ:*

$$\psi(A) = UAU^*, \quad \textit{for } A \in \mathcal{C},$$

and $D(S) = UD(T)$ and $\mathcal{F}_{UTU^*} = \mathcal{F}_S$.

Proof. For $x, y \in D(T)$, $x \neq 0$, $y \neq 0$, set $Y = \varphi(x \otimes y)$. If Y is not a rank one operator, there are $z, u \in D(S)$ such that $Yz \neq 0$, $Yu \neq 0$ and $Yz \perp Yu$. Since $Y \in \mathcal{A}_S$, we have that $Yz, Yu \in D(S)$, so that $Yz \otimes z \in \mathcal{F}_S$ and $u \otimes Yu \in \mathcal{F}_S$. By (3.5)

$$(Yz \otimes z)(u \otimes Yu) = (Yu, Yz)(u \otimes z) = 0. \tag{4.2}$$

Since $(z \otimes z)^* = z \otimes z$ and φ is a *-isomorphism, it follows from (3.5) that

$$\begin{aligned}\left(\varphi^{-1}(z \otimes z)x\right) \otimes y = (x \otimes y)\left[\varphi^{-1}(z \otimes z)\right]^* \\ = \varphi^{-1}(Y)\varphi^{-1}(z \otimes z) = \varphi^{-1}(z \otimes Yz) \neq 0.\end{aligned}$$

Thus $\varphi^{-1}(z \otimes z)x \neq 0$. Similarly, $\varphi^{-1}(u \otimes u)x \neq 0$. From this and from (3.5) and (4.2) it follows that

$$\begin{aligned}0 &= \varphi^{-1}((Yz \otimes z)(u \otimes Yu)) = \varphi^{-1}((z \otimes z)Y^*Y(u \otimes u)) \\ &= \varphi^{-1}(z \otimes z)\varphi^{-1}(Y^*)\varphi^{-1}(Y)\varphi^{-1}(u \otimes u) \\ &= \varphi^{-1}(z \otimes z)(y \otimes x)(x \otimes y)\varphi^{-1}(u \otimes u) \\ &= \varphi^{-1}(z \otimes z)\|y\|^2(x \otimes x)\varphi^{-1}(u \otimes u) \\ &= \|y\|^2([\varphi^{-1}(u \otimes u)x] \otimes [\varphi^{-1}(z \otimes z)x]) \neq 0.\end{aligned}$$

This contradiction shows that Y is a rank one operator. Hence $Y \in \mathcal{F}_S$ and, by (4.1), φ maps all finite rank operators in $\mathcal{F}_T$ into finite rank operators in $\mathcal{F}_S$. Since φ is bounded $\varphi(\mathcal{F}_T) \subseteq \mathcal{F}_S$. Similarly, $\varphi^{-1}(\mathcal{F}_S) \subseteq \mathcal{F}_T$, so that φ is a bounded *-isomorphism from $\mathcal{F}_T$ onto $\mathcal{F}$. Part (i) is proved.

Fix $x_0 \in D(T)$, $\|x_0\| = 1$. Since $x_0 \otimes x_0$ is a projection, $\varphi(x_0 \otimes x_0)$ is a one-dimensional projection in $\mathcal{F}_S$. By (4.1), we can choose ξ_0 in $D(S)$, $\|\xi_0\| = 1$, such that $\varphi(x_0 \otimes x_0) = \xi_0 \otimes \xi_0$. Let $y \in D(T)$. Making use of the equality $x_0 \otimes y = (x_0 \otimes y)(x_0 \otimes x_0)$, we obtain that

$$\begin{aligned}\varphi(x_0 \otimes y) = \varphi(x_0 \otimes y)\varphi(x_0 \otimes x_0) \\ = \varphi(x_0 \otimes y)(\xi_0 \otimes \xi_0) = \xi_0 \otimes \varphi(x_0 \otimes y)\xi_0.\end{aligned}$$

Since $\varphi(x_0 \otimes y) \in \mathcal{F}_S$, it follows from (4.1) that $\varphi(x_0 \otimes y)\xi_0$ belongs to $D(S)$.

Now $U : y \in D(T) \to \varphi(x_0 \otimes y)\xi_0$ is a linear mapping from $D(T)$ into $D(S)$ and $\varphi(x_0 \otimes y) = \xi_0 \otimes Uy$. Then

$$\begin{aligned}\varphi((y \otimes x_0)(x_0 \otimes y)) &= \|y\|^2\varphi(x_0 \otimes x_0) = \|y\|^2(\xi_0 \otimes \xi_0) \\ &= \varphi((x_0 \otimes y)^*)\varphi(x_0 \otimes y) \\ &= (Uy \otimes \xi_0)(\xi_0 \otimes Uy) = \|Uy\|^2(\xi_0 \otimes \xi_0).\end{aligned}$$

Thus $\|Uy\|^2 = \|y\|^2$, for $y \in D(T)$, and U extends to an isometry operator from $\mathcal{H}$ into H which we also denote by U. We have that, for $x, y \in D(T)$,

$$\begin{aligned}\varphi(x \otimes y) = \varphi((x_0 \otimes y)(x \otimes x_0)) = (\xi_0 \otimes Uy)(\xi_0 \otimes Ux)^* \\ = Ux \otimes Uy = U(x \otimes y)U^*.\end{aligned} \tag{4.3}$$

Similarly, there is an isometry operator V which maps $D(S)$ into $D(T)$ such that $\varphi^{-1}(\xi \otimes \eta) = V\xi \otimes V\eta$, for $\xi, \eta \in D(S)$. Hence

$$\xi \otimes \eta = \varphi(\varphi^{-1}(\xi \otimes \eta)) = \varphi(V\xi \otimes V\eta) = UV\xi \otimes UV\eta.$$

Thus $UV\xi = \lambda(\xi)\xi$ where λ is a function on $D(S)$ such that $|\lambda(\xi)| = 1$. Hence $UD(T) = D(S)$. Since $D(S)$ is dense in H and U is an isometry operator, we have $U\mathcal{H} = H$.

Let $A \in \mathcal{C}$ and set $R = U^*\psi(A)U$. Then $x \otimes y \in \mathcal{F}_T$, for any $x, y \in D(T)$, and, since $\mathcal{F}_T$ is an ideal of $\mathcal{A}_T$, we have $A(x \otimes y) = x \otimes Ay \in \mathcal{F}_T$. By (4.3),

$$\begin{aligned} R(x \otimes y) &= U^*\psi(A)U(x \otimes y) = U^*\psi(A)U(x \otimes y)U^*U \\ &= U^*\psi(A)\varphi(x \otimes y)U = U^*\psi(A)\psi(x \otimes y)U \\ &= U^*\psi(A(x \otimes y))U = U^*\varphi(x \otimes Ay)U = x \otimes Ay. \end{aligned}$$

Therefore $R(x \otimes y) = x \otimes Ry = x \otimes Ay$, so that $Ry = Ay$. Thus $R = A$ and

$$\psi(A) = UAU^*.$$

The operator $F = UTU^*$ is symmetric and $D(F) = UD(T) = D(S)$. By Lemma 4.1, $\mathcal{F}_F = U\mathcal{F}_T U^*$ and $A \to UAU^*$ is an isometric *-isomorphism from $(\mathcal{F}_T, \|\cdot\|_T)$ onto $(\mathcal{F}_F, \|\cdot\|_F)$. Hence

$$\varphi(U^*BU) = U(U^*BU)U^* = B, \quad \text{for } B \in \mathcal{F}_F,$$

is a bounded *-isomorphism from $\mathcal{F}_F$ onto $\mathcal{F}_S$. Therefore $\mathcal{F}_F = \mathcal{F}_S$. □

It was shown in Theorem 3.4 [**13**] that the algebra $(\mathcal{F}_S, \|\cdot\|_S)$ has a bounded approximate identity if and only if S is selfadjoint. Making use of this and of Theorem 4.2, we obtain the following result.

Corollary 4.3. *If the algebras $\mathcal{F}_S$ and $\mathcal{F}_T$ are *-isomorphic or the algebras $\mathcal{A}_S$ and $\mathcal{A}_T$ are *-isomorphic then the operators S and T are either selfadjoint or non-selfadjoint at the same time.*

Apart from the sufficient conditions of Lemma 4.1 and the necessary conditions of Corollary 4.3 for two algebras $\mathcal{F}_S$ and $\mathcal{F}_T$ to be *-isomorphic we do not know any other sufficient or necessary condition in the case when S and T are arbitrary symmetric operators. Later, in Theorem 4.6 and Corollary 4.8 we consider a particular case when the operators S and T are selfadjoint.

It follows from Theorem 4.2 that if $\mathcal{F}_S$ and $\mathcal{F}_T$ are *-isomorphic, they are unitary isomorphic. This, however, does not necessarily imply that they are isometrically isomorphic. In the following theorem we obtain necessary and sufficient conditions for algebras $\mathcal{F}_S$ and $\mathcal{F}_T$ to be *isometrically* *-isomorphic.

Theorem 4.4. *The algebras $(\mathcal{F}_S, \|\cdot\|_S)$ and $(\mathcal{F}_T, \|\cdot\|_T)$ are isometrically *-isomorphic if and only if there are $\lambda \in \mathbb{R}$ and an isometry operator U such that $S - \lambda\mathbf{1}_H = \pm UTU^*$. The same result holds for $(\mathcal{A}_S, \|\cdot\|_S)$ and $(\mathcal{A}_T, \|\cdot\|_T)$.*

Proof. From Lemma 4.1 it follows that the conditions of the theorem are sufficient. From Theorem 4.2 it follows that if these conditions are necessary for the algebras $(\mathcal{F}_S, \|\cdot\|_S)$ and $(\mathcal{F}_T, \|\cdot\|_T)$ to be isometrically *-isomorphic, they are also necessary for the algebras $(\mathcal{A}_S, \|\cdot\|_S)$ and $(\mathcal{A}_T, \|\cdot\|_T)$.

Let φ be an isometric *-isomorphism from $(\mathcal{F}_T, \|\cdot\|_T)$ onto $(\mathcal{F}_S, \|\cdot\|_S)$ and let U be the isometry operator as in Theorem 4.2 which implements φ:

$$\varphi(A) = UAU^*, \quad \text{for} \;\; A \in \mathcal{F}_T.$$

Set $F = UTU^*$. Then F is a symmetric operator on H, $D(S) = D(F) = UD(T)$ and $\mathcal{F}_S = \mathcal{F}_F$. Since φ is isometric, the norms $\|\cdot\|_S$ and $\|\cdot\|_F$ coincide.

We will show that there is $\lambda \in \mathbb{R}$ such that either $S - \lambda\mathbf{1}_H = F$ or $S - \lambda\mathbf{1}_H = -F$.

Step 1. Suppose that $z \in D(S)$ is not an eigenvector of S and $\|z\| = 1$. Set

$$s = (Sz, z), \quad t = (Fz, z), \quad R = S - s\mathbf{1}_H \quad \text{and} \;\; G = F - t\mathbf{1}_H.$$

Since S an F are symmetric, $s, t \in \mathbb{R}$, the operators R and G are symmetric and

$$D(R) = D(G), \quad Rz \neq 0 \quad \text{and} \;\; (Rz, z) = (Gz, z) = 0. \tag{4.4}$$

Set $D = D(R) = D(G)$. Since $\mathcal{F}_S = \mathcal{F}_F$ and the norms $\|\cdot\|_S$ and $\|\cdot\|_F$ coincide, it follows from Lemma 4.1 that $\mathcal{F}_R = \mathcal{F}_G$ and the norms $\|\cdot\|_R$ and $\|\cdot\|_G$ coincide.

Taking into account that R and G are symmetric, we obtain from (3.5) that

$$\begin{aligned}\|y \otimes x\|_R &= \|y \otimes x\| + \|y \otimes Rx - (Ry) \otimes x\| = \|y \otimes x\|_G \\ &= \|y \otimes x\| + \|y \otimes Gx - (Gy) \otimes x\|,\end{aligned}$$

for $x, y \in D$. Therefore

$$\|y \otimes Rx - (Ry) \otimes x\| = \|y \otimes Gx - (Gy) \otimes x\|. \tag{4.5}$$

Represent the elements Rx and Gx in the form

$$Rx = \alpha(x)x + x_R \quad \text{and} \;\; Gx = \beta(x)x + x_G, \tag{4.6}$$

where x_R and x_G are orthogonal to x. Then

$$\alpha(x)\|x\|^2 = (Rx, x) = (x, Rx) = \overline{\alpha(x)}\|x\|^2.$$

Thus $\alpha(x)$ is real, for $x \in D$. Therefore

$$\begin{aligned}x \otimes Rx - (Rx) \otimes x &= \alpha(x)(x \otimes x) + x \otimes x_R - \alpha(x)(x \otimes x) - x_R \otimes x \\ &= x \otimes x_R - x_R \otimes x.\end{aligned}$$

Since x and x_R are orthogonal, any $u \in H$ can be represented in the form $u = \nu x + \tau x_R + \widetilde{u}$, where $\nu, \tau \in \mathbb{C}$ and $\widetilde{u}$ is orthogonal to x and x_R. Therefore

$$\|u\| = |\nu|^2\|x\|^2 + |\tau|^2\|x\|^2 + \|\widetilde{u}\|^2$$

and, by (3.5),

$$\begin{aligned}\|(x\otimes x_R + x_R\otimes x)u\|^2 &= \|(u,x)x_R + (u,x_R)\,x\|^2\\ &= \big\|\nu\,\|x\|^2 x_R + \tau\|x_R\|^2 x\big\|^2\\ &= |\nu|^2\|x\|^4\|x_R\|^2 + |\tau|^2\|x_R\|^4\|x\|^2\\ &= \|x\|^2\|x_R\|^2(|\nu|^2\|x\|^2 + |\tau|^2\|x_R\|^2).\end{aligned}$$

Consequently,

$$\|x\otimes Rx - (Rx)\otimes x\|^2 = \|x\otimes x_R - x_R\otimes x\|^2 = \|x\|^2\|x_R\|^2.$$

Similarly, $\|x\otimes Gx - (Gx)\otimes x\|^2 = \|x\|^2\|x_G\|^2$ and it follows from (4.5) that

$$\|x_R\| = \|x_G\|, \quad \text{for } \ x\in D.$$

Therefore we obtain from (4.6) that for $x\in D$

$$\begin{aligned}\|x\|^2\|Rx\|^2 - |(Rx,x)|^2 &= \|x\|^2(|\alpha(x)|^2\|x\|^2 + \|x_R\|^2) - |\alpha(x)|^2\|x\|^4\\ &= \|x\|^2\|x_R\|^2 = \|x\|^2\|x_G\|^2\\ &= \|x\|^2\|Gx\|^2 - |(Gx,x)|^2.\end{aligned}$$

Hence

$$\|x\|^2(\|Rx\|^2 - \|Gx\|^2) = |(Rx,x)|^2 - |(Gx,x)|^2. \tag{4.7}$$

In particular, it follows from (4.4), (4.6) and (4.7) that

$$Rz = z_R, \quad Gz = z_G \quad \text{and} \ \ \|Rz\| = \|Gz\|. \tag{4.8}$$

Step 2. Set $D_Z^\perp = \{y\in D : y \text{ is orthogonal to } z\}$. Let $y\in D_Z^\perp$ and $x = y+\mu z$, $\mu\in\mathbb{C}$. Then $\|x\|^2 = \|y\|^2 + \|\mu z\|^2 = \|y\|^2 + |\mu|^2$ and, by (4.8),

$$\begin{aligned}\|Rx\|^2 - \|Gx\|^2 &= \|Ry\|^2 + \|\mu Rz\|^2 + 2\mathrm{Re}[\mu(Rz,Ry)]\\ &\quad - \|Gy\|^2 - \|\mu Gz\|^2 - 2\mathrm{Re}[\mu(Gz,Gy)]\\ &= A + 2\mathrm{Re}(\mu B),\end{aligned}$$

where

$$A = \|Ry\|^2 - \|Gy\|^2 \quad \text{and} \quad B = (Rz,Ry) - (Gz,Gy).$$

Since R is symmetric, it follows from (4.4) that

$$\begin{aligned}(Rx,x) &= (Ry,y) + (\mu Rz,y) + (Ry,\mu z) + (\mu Rz,\mu z)\\ &= (Ry,y) + 2\mathrm{Re}[\mu(Rz,y)].\end{aligned}$$

Similarly, $(Gx,x) = (Gy,y) + 2\mathrm{Re}[\mu(Gz,y)]$.

Let $\mu = re^{i\psi}$. Substituting all this in (4.7), we obtain that

$$\begin{aligned}&(\|y\|^2 + r^2)[A + 2r\mathrm{Re}(e^{i\psi}B)]\\ &\quad = \{(Ry,y) + 2r\mathrm{Re}[e^{i\psi}(Rz,y)]\}^2 - \{(Gy,y) + 2r\mathrm{Re}[e^{i\psi}(Gz,y)]\}^2.\end{aligned} \tag{4.9}$$

Set

$$C = (Ry,y)\mathrm{Re}[e^{i\psi}(Rz,y)] - (Gy,y)\mathrm{Re}[e^{i\psi}(Gz,y)] \qquad \text{and}$$
$$E = \{\mathrm{Re}[e^{i\psi}(Rz,y)]\}^2 - \{\mathrm{Re}[e^{i\psi}(Gz,y)]\}^2.$$

Since R and G are symmetric, (Ry,y) and (Gy,y) are real. Hence

$$C = \mathrm{Re}\{e^{i\psi}[(Ry,y)(Rz,y) - (Gy,y)(Gz,y)]\}.$$

Comparing the coefficients of the same powers of r in (4.9), we obtain that

$$\mathrm{Re}(e^{i\psi}B) = 0, \quad A = 4E \quad \text{and} \quad C = 0.$$

Taking into account that $\mathrm{Re}\,(e^{i\psi}K) = 0$, for $0 \le \psi < 2\pi$, implies $K = 0$, we obtain that $C = 0$ implies

$$(Ry,y)(Rz,y) - (Gy,y)(Gz,y) = 0. \tag{4.10}$$

Set $(Rz,y) = ae^{ib}$ and $(Gz,y) = ce^{id}$. Then

$$\begin{aligned} E &= a^2\left[\mathrm{Re}\left(e^{i(\psi+b)}\right)\right]^2 - c^2\left[\mathrm{Re}\left(e^{i(\psi+d)}\right)\right]^2 \\ &= a^2\cos^2(\psi+b) - c^2\cos^2(\psi+d). \end{aligned}$$

Since $A = 4E$ and since A does not depend on ψ, neither does E. Hence $a^2 = c^2$ and $d = b$ or $d = b + \pi$. Since $a \ge 0$ and $c \ge 0$, $a = c$. Thus

$$(Rz,y) = \pm(Gz,y), \quad \text{for } \ y \in D_Z^{\perp}. \tag{4.11}$$

Since D is dense in $\mathcal{H}$, $D_Z^{\perp}$ is dense in the subspace $\{\mathbb{C}z\}^{\perp}$. Hence (4.11) holds for all $y \in \{\mathbb{C}z\}^{\perp}$. From (4.9) it follows that $Rz = z_R \in \{\mathbb{C}z\}^{\perp}$. Substituting Rz for y in (4.11), we obtain $\|Rz\| = (Rz,Rz) = \pm(Gz,Rz)$. Let $Gz = \nu Rz + u$, where $\nu \in \mathbb{C}$ and u is orthogonal to Rz. Then

$$\|Rz\|^2 = \pm(Gz,Rz) = \pm\nu\|Rz\|^2.$$

Since $Rz \ne 0$ (see (4.4)), $\nu = \pm 1$. Taking (4.9) into account, we obtain

$$\begin{aligned} \|Rz\|^2 = \|Gz\|^2 &= (\nu Rz + u, \nu Rz + u) \\ &= |\nu|^2\|Rz\|^2 + \|u\|^2 = \|Rz\|^2 + \|u\|^2. \end{aligned}$$

Hence $u = 0$ and either $Rz = Gz$ or $Rz = -Gz$.

Step 3. Let $Rz = Gz$. Set $W = R - G$. Then W is symmetric, $Wz = 0$ and it follows from (4.10) that

$$[(Ry,y) - (Gy,y)](Rz,y) = (Wy,y)(Rz,y) = 0, \quad \text{for } \ y \in D_Z^{\perp}.$$

Any $x \in D$ can be represented in the form $x = y + \mu z$ where $\mu \in \mathbb{C}$ and $y \in D_Z^{\perp}$. Then $Wx = Wy$ and, since $(Rz,z) = 0$, we have $(Rz,x) = (Rz,y)$. Since $Wz = 0$,

$$\begin{aligned} (Wx,x)(Rz,x) &= (Wy, y + \mu z)(Rz,y) \\ &= [(Wy,y) + (y, \mu Wz)](Rz,y) = (Wy,y)(Rz,y) = 0. \end{aligned}$$

Therefore

$$(Wx,x)(Rz,x) = 0, \quad \text{for } x \in D. \tag{4.12}$$

Let $X = \{x \in H : (Rz,x) = 0\}$ be the orthogonal complement of the subspace $\{\mathbb{C}Rz\}$ in H. By (4.4), $Rz \neq 0$, so X has codimension 1. Set $\mathcal{D} = \{x \in D : x \notin X\}$. Since D is dense in H, $\mathcal{D}$ is also dense in H. For $x \in \mathcal{D}$, we have $(Rz,x) \neq 0$. Hence, by (4.12),

$$(Wx,x) = 0.$$

If $x, y \in \mathcal{D}$, there is $r > 0$ such that $x + re^{i\psi}y \in \mathcal{D}$, for all $0 \leq \psi < 2\pi$. Taking into account that W is symmetric, we obtain that

$$\begin{aligned} 0 &= (W(x+re^{i\psi}y), x+re^{i\psi}y) = (Wx,x) + 2r\mathrm{Re}[e^{i\psi}(Wy,x)] + r^2(Wy,y) \\ &= 2r\mathrm{Re}[e^{i\psi}(Wy,x)]. \end{aligned}$$

Hence $(Wy,x) = 0$. Since $\mathcal{D}$ is dense in H, we have $Wy = 0$, for $y \in \mathcal{D}$.

Let $u \in D \cap X$, so that $(Rz,u) = 0$. For $y \in \mathcal{D}$, $(Rz, y+u) = (Rz,y) \neq 0$. Hence $y + u \in \mathcal{D}$ and $0 = W(y+u) = Wy + Wu = Wu$. Thus $Wx = 0$, for all $x \in D$, so that $R = G$. Hence $S - s\mathbf{1}_H = F - t\mathbf{1}_H$. Setting $\lambda = s - t$, we obtain that

$$S - \lambda\mathbf{1}_H = F = UTU^*.$$

Similarly, in the case when $Rz = -Gz$ we obtain that $S - \lambda\mathbf{1}_H = -F = -UTU^*$ which concludes the proof of the theorem. □

In the rest of this section we study conditions for the algebras $\mathcal{F}_S$ and $\mathcal{F}_T$ to be *-isomorphic but not necessarily isometrically *-isomorphic in the case when S and T are selfadjoint operators. Taking Theorem 4.2(ii) into account, we may assume, without loss of generality, that $\mathcal{F}_S = \mathcal{F}_T$ and $D(S) = D(T)$.

In Example 4.7 we show that the coincidence of the domains of selfadjoint operators S and T even in the case when $\mathrm{Sp}(S) \subseteq \mathbb{Z}$, $\mathrm{Sp}(T) \subseteq \mathbb{Z}$ and S and T have the same sets of eigenvectors is not sufficient for $\mathcal{F}_S = \mathcal{F}_T$. In other words, the algebras $\mathcal{F}_S$ and $\mathcal{F}_T$ may be the closures of the same set of finite rank operators and, nevertheless, be non-isomorphic. Necessary and sufficient conditions for these algebras to be *-isomorphic will be obtained in Theorem 4.6.

Let $\mathfrak{H}$ be a Hilbert space with an orthogonal basis $\{e_i\}_{i=-\infty}^{\infty}$. Every operator T in $B(\mathfrak{H})$ has a matrix representation $T = (t_{ij})$, $-\infty < i,j < \infty$, where $t_{ij} = (Te_j, e_i)$. A matrix $M = (m_{ij})$, $-\infty < i,j < \infty$, is called a *Schur multiplier*, if, for any $T = (t_{ij}) \in B(\mathfrak{H})$, the matrix $M \circ T = (m_{ij}t_{ij})$ belongs to $B(\mathfrak{H})$. Then $T \to M \circ T$ is a bounded map of $B(\mathfrak{H})$ into itself; it will also be denoted by M and its norm by $|M|_{B(\mathfrak{H})}$.

Let $H = \sum_{i=-\infty}^{\infty} \oplus H_i$ be an orthogonal sum of Hilbert spaces H_i. Every operator A in $B(H)$ has a block-matrix representation $A = (A_{ij})$, $-\infty < i, j < \infty$, where A_{ij} are bounded operators from H_j into H_i.

Lemma 4.5. *Let $M = (m_{ij})$ be a Schur multiplier on $\mathfrak{H}$. It defines a bounded operator $\mathcal{M}$ on $B(H)$ by the formula*

$$\mathcal{M} \times A = (m_{ij} A_{ij}), \quad \text{where } A = (A_{ij}) \in B(H),$$

and $|\mathcal{M}|_{B(H)} = |M|_{B(\mathcal{H})}$.

Proof. Let $G = \{g_j\}_{j=-\infty}^{\infty}$ and $F = \{f_j\}_{j=-\infty}^{\infty}$ be sequences of elements such that $g_j, f_j \in H_j$ and $\|g_j\| = \|f_j\| = 1$. For $A = (A_{ij}) \in B(H)$, let $T^{G,F}(A) = \left(a_{ij}^{GF}\right)$, $-\infty < i, j < \infty$, be the matrix such that

$$a_{ij}^{GF} = (A_{ij} g_j, f_i) \in \mathbb{C}. \tag{4.13}$$

For $\alpha = \sum_{j=-\infty}^{\infty} \oplus \alpha_j e_j \in \mathfrak{H}$ and $\beta = \sum_{j=-\infty}^{\infty} \oplus \beta_j e_j \in \mathfrak{H}$, set

$$x_\alpha^G = \sum_{j=-\infty}^{\infty} \oplus \alpha_j g_j \quad \text{and} \quad y_\beta^F = \sum_{j=-\infty}^{\infty} \oplus \beta_j f_j.$$

Then $x_\alpha^G, y_\beta^F \in H$, $\left\|x_\alpha^G\right\| = \|\alpha\|$, $\left\|y_\beta^F\right\| = \|\beta\|$ and

$$\begin{aligned}\left(Ax_\alpha^G, y_\beta^F\right) &= \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \alpha_j \bar{\beta}_i (A_{ij} g_j, f_i) \\ &= \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \alpha_j \bar{\beta}_i a_{ij}^{GF} = \left(T^{G,F}(A)\alpha, \beta\right).\end{aligned}$$

Therefore $T^{G,F}(A) \in B(\mathfrak{H})$ and

$$\begin{aligned}\|A\| &= \sup_{\alpha,\beta,G,F} \frac{\left|\left(Ax_\alpha^G, y_\beta^F\right)\right|}{\|x_\alpha^G\| \left\|y_\beta^F\right\|} \\ &= \sup_{G,F} \left(\sup_{\alpha,\beta} \frac{|(T^{G,F}(A)\alpha, \beta)|}{\|\alpha\| \, \|\beta\|} \right) = \sup_{G,F} \left\|T^{G,F}(A)\right\|.\end{aligned} \tag{4.14}$$

It follows from (4.13) that $T^{G,F}(\mathcal{M} \times A) = M \circ T^{G,F}(A)$. Since M is a Schur multiplier, $M \circ T^{G,F}(A) \in B(\mathfrak{H})$ and, by (4.14),

$$\begin{aligned}\|\mathcal{M} \times A\| &= \sup_{G,F} \left\|T^{G,F}(\mathcal{M} \times A)\right\| = \sup_{G,F} \left\|M \circ T^{G,F}(A)\right\| \\ &\le \sup_{G,F} |M|_{B(\mathfrak{H})} \left\|T^{G,F}(A)\right\| = |M|_{B(\mathfrak{H})} \sup_{G,F} \left\|T^{G,F}(A)\right\| \\ &= |M|_{B(\mathfrak{H})} \|A\|.\end{aligned}$$

Hence $|\mathcal{M}|_{B(H)} \leq |M|_{B(\mathfrak{H})}$. On the other hand, it is easy to see that $|M|_{B(\mathfrak{H})} \leq |\mathcal{M}|_{B(H)}$. Thus $|\mathcal{M}|_{B(H)} = |M|_{B(\mathfrak{H})}$. □

Let S and T be selfadjoint operators on H and assume that $\mathrm{Sp}(S) \subseteq \mathbb{Z}$, $\mathrm{Sp}(T) \subseteq \mathbb{Z}$ and that

$$H = \sum_{i=-\infty}^{\infty} \oplus H_i \quad \text{where} \;\; S|_{H_i} = s_i \mathbf{1}_{H_i}, \quad T|_{H_i} = t_i \mathbf{1}_{H_i},$$
$$s_i \neq s_j \;\; \text{and} \;\; t_i \neq t_j \;\; \text{if } i \neq j.$$

Set

$$M = (m_{ij}) \;\; \text{where} \;\; m_{ij} = \frac{s_i - s_j}{t_i - t_j}, \;\; \text{for} \;\; i \neq j, \;\; \text{and} \;\; m_{ii} = 0, \quad \text{and}$$
$$N = (n_{ij}) \;\; \text{where} \;\; n_{ij} = \frac{t_i - t_j}{s_i - s_j}, \;\; \text{for} \;\; i \neq j, \;\; \text{and} \;\; n_{ii} = 0.$$

Theorem 4.6. *$\mathcal{F}_S = \mathcal{F}_T$ if and only if M and N are Schur multipliers.*

Proof. In every H_i we choose a non-decreasing sequence of finite-dimensional projections $\{Q_i^p\}_{p=1}^{\infty}$ which converge to $\mathbf{1}_{H_i}$ in the strong operator topology as $p \to \infty$. Set $Q_p = \sum_{i=-p}^{p} \oplus Q_i^p$. The finite-dimensional projections Q_p commute with S and T, belong to $\mathcal{F}_S \cap \mathcal{F}_T$ and converge to $\mathbf{1}_H$ in the strong operator topology. Therefore $\|Q_p\|_S = \|Q_p\|_T = \|Q_p\| = 1$.

For any $A = (A_{ij}) \in \mathcal{A}_S \cap \mathcal{A}_T$,

$$A_S = SA - AS = \left(A_{ij}^S\right) \;\; \text{and} \;\; A_T = TA - AT = \left(A_{ij}^T\right),$$

where $A_{ij}^S = (s_i - s_j)A_{ij}$ and $A_{ij}^T = (t_i - t_j)A_{ij}$. Set $B = A_T$. Then $A_S = \mathcal{M} \times B$,

$$\|A\|_S = \|A\| + \|A_S\| = \|A\| + \|\mathcal{M} \times B\| \qquad \text{and} \tag{4.15}$$
$$\|A\|_T = \|A\| + \|A_T\| = \|A\| + \|B\|.$$

We assume now that M and N are Schur multipliers and show that $\mathcal{F}_S = \mathcal{F}_T$. By Lemma 4.5 and (4.15),

$$\|A\|_S \leq \|A\| + |M|\,\|B\| \tag{4.16}$$
$$\leq \|A\| + |M|\,(\|A\|_T - \|A\|) \leq (|M| + 1)\|A\|_T.$$

Similarly,

$$\|A\|_T \leq (|N| + 1)\|A\|_S. \tag{4.17}$$

Let $A \in \mathcal{F}_S$. Then $Q_pA \in \mathcal{F}_S$ and, since Q_p commute with S,

$$(Q_pA)_S = \mathrm{Closure}\,(SQ_pA - Q_pAS) = \mathrm{Closure}\,Q_p(SA - AS) = Q_pA_S.$$

Since A and A_S are compact and since Q_p converge to $\mathbf{1}_H$ in the strong operator topology,

$$\|A - Q_pA\| \to 0 \;\; \text{and} \;\; \|A_S - (Q_pA)_S\| = \|A_S - Q_pA_S\| \to 0, \;\; \text{as} \;\; p \to \infty.$$

Hence $\|A - Q_pA\|_S \to 0$, so that $\{Q_p\}$ is a bounded approximate identity in $\mathcal{F}_S$. Similarly, it is a bounded approximate identity in $\mathcal{F}_T$.

Let $A \in \mathcal{F}_S$. For any p, $Q_pT = Q_pTQ_p = TQ_p$ is a finite rank operator. Hence

$$(Q_pAQ_p)_T = T(Q_pAQ_p) - (Q_pAQ_p)T = (TQ_p)AQ_p - Q_pA(Q_pT)$$

is a finite rank operator. Therefore $Q_pAQ_p \in \mathcal{F}_S \cap \mathcal{F}_T$ and, by (4.17),

$$\|Q_{p+k}AQ_{p+k} - Q_pAQ_p\|_T \leq (|N| + 1)\|Q_{p+k}AQ_{p+k} - Q_pAQ_p\|_S.$$

Since $\{Q_p\}$ is a bounded approximate identity in $\mathcal{F}_S$, the operators Q_pAQ_p converge to A with respect to $\|\cdot\|_S$. From the above inequality it follows that $\{Q_pAQ_p\}$ is a fundamental sequence with respect to $\|\cdot\|_T$. Hence there is $A_1 \in \mathcal{F}_T$ such that $\|A_1 - Q_pAQ_p\|_T \to 0$, as $p \to \infty$. Since $\|A - Q_pAQ_p\| \leq \|A - Q_pAQ_p\|_S \to 0$ and $\|A_1 - Q_pAQ_p\| \leq \|A_1 - Q_pAQ_p\|_T \to 0$, as $p \to \infty$, we obtain that $A = A_1$, so $\mathcal{F}_S \subseteq \mathcal{F}_T$. Similarly, $\mathcal{F}_T \subseteq \mathcal{F}_S$. Thus we conclude that $\mathcal{F}_S = \mathcal{F}_T$.

Suppose now that $\mathcal{F}_S = \mathcal{F}_T$. Choose elements $e_i \in H_i$ such that $\|e_i\| = 1$ and let $\mathfrak{H}$ be the subspace of H generated by all e_i, $-\infty < i < \infty$. Then $\mathfrak{H}$ is invariant for S and T, $Se_i = s_ie_i$ and $Te_i = t_ie_i$. By $S_{\mathfrak{H}}$ and $T_{\mathfrak{H}}$ we denote the restrictions of S and T to $\mathfrak{H}$. Since $\mathcal{F}_S = \mathcal{F}_T$,

$$\mathcal{F}_{S_{\mathfrak{H}}} = \mathcal{F}_{T_{\mathfrak{H}}}.$$

We shall show now that M and N are Schur multipliers on $\mathfrak{H}$.

The function $f(t) = i(\pi - t)$ on $[0, 2\pi]$ has Fourier coefficients $c_0 = 0$ and $c_n = \frac{1}{n}$, for $n = \pm 1, \pm 2, \dots$. Let $\mathcal{H}$ be a Hilbert space with an orthonormal basis $\{h_k\}_{k=-\infty}^{\infty}$ and $R = (r_{kl})$, $-\infty < k, l < \infty$, be a *Toeplitz* matrix such that $r_{kk} = 0$ and $r_{kl} = c_{k-l} = \frac{1}{k-l}$, $k \neq l$. Then $R \in B(\mathcal{H})$ and it follows from Theorem 8.1 [1] that R is a Schur multiplier and $|R| = \sup |f(t)| = \pi$.

Identifying e_i in $\mathfrak{H}$ with h_{t_i} in $\mathcal{H}$, we can consider $\mathfrak{H}$ as a subspace of $\mathcal{H}$. For $B = (b_{km}) \in B(\mathfrak{H})$, where $b_{km} = (Be_m, e_k)$, let $\widetilde{B} = (\widetilde{b}_{ij}) \in B(\mathcal{H})$ be such that $\widetilde{B}|_{\mathfrak{H}} = B$ and $\widetilde{B}|_{\mathfrak{H}^\perp} = 0$. Then $\|\widetilde{B}\| = \|B\|$,

$$\widetilde{b}_{t_kt_m} = \left(\widetilde{B}h_{t_m}, h_{t_k}\right) = (Be_m, e_k) = b_{km}, \qquad \text{and}$$
$$\widetilde{b}_{ij} = \left(\widetilde{B}h_j, h_i\right) = 0 \;\text{ if either }\; i \neq t_k \;\text{ or }\; j \neq t_m.$$

Since R is a Schur multiplier, the operator $\widetilde{C} = (\widetilde{c}_{ij}) = R \circ \widetilde{B}$ belongs to $B(\mathcal{H})$, where

$$\widetilde{c}_{t_kt_m} = r_{t_kt_m}\widetilde{b}_{t_kt_m} = (t_k - t_m)^{-1}b_{km}, \quad \text{if } \; k \neq m, \quad \text{and}$$
$$\widetilde{c}_{ij} = 0 \;\text{ if either }\; i \neq t_k \;\text{ or }\; j \neq t_m \;\text{ or }\; i = j = t_k.$$

Setting $C = \widetilde{C}|_{\mathfrak{H}}$, we obtain that $C = (c_{km}) \in B(\mathfrak{H})$, where

$$c_{km} = \widetilde{c}_{t_kt_m} = (t_k - t_m)^{-1}b_{km}, \;\text{ if }\; k \neq m, \;\text{ and }\; c_{kk} = 0,$$

that $\|\widetilde{C}\| = \|C\|$ and that $C = W \circ B$, where $W = (w_{km})$ is a matrix such that

$$w_{km} = (t_k - t_m)^{-1}, \quad k \neq m, \text{ and } w_{kk} = 0.$$

From this it follows that W is a Schur multiplier on $\mathfrak{H}$ and

$$\|W \circ B\| = \|C\| = \left\|\widetilde{C}\right\| = \left\|R \circ \widetilde{B}\right\| \leq |R| \left\|\widetilde{B}\right\| = |R| \, \|B\|.$$

Thus $|W| \leq |R| = \pi$.

Let P_n be the orthoprojections in $\mathfrak{H}$ on the subspaces $\sum_{j=-n}^{n} \oplus\{\mathbb{C}e_j\}$. Then P_n are finite rank operators commuting with operators $S_{\mathfrak{H}}$ and $T_{\mathfrak{H}}$ and $P_n\mathfrak{H} \subseteq D(S_{\mathfrak{H}})$. Hence $P_n \in \mathcal{F}_{S_{\mathfrak{H}}}$. For every $B \in B(\mathfrak{H})$, P_nBP_n are finite rank operators preserving $D(S_{\mathfrak{H}})$ and their adjoints $P_nB^*P_n$ also preserve $D(S_{\mathfrak{H}})$. Therefore

$$P_nBP_n \in \mathcal{F}_{S_{\mathfrak{H}}}. \tag{4.18}$$

Any $B = (b_{km}) \in B(\mathfrak{H})$ can be represented in the form $B = B_d + B_0$, where B_d is the diagonal operator such that $(B_d) = b_{kk}$. Then

$$\|B_d\| \leq \|B\| \quad \text{and} \quad \|B_0\| = \|B - B_d\| \leq 2\|B\|. \tag{4.19}$$

We have that

$$M \circ (P_nBP_n) = P_n(M \circ B)P_n. \tag{4.20}$$

Since $m_{kk} = 0$ in the matrix $M = (m_{km})$,

$$M \circ (P_nBP_n) = M \circ (P_nB_0P_n). \tag{4.21}$$

Set $A = W \circ B$. Since W is a Schur multiplier, $A \in B(\mathfrak{H})$ and, by (4.18), $P_nAP_n \in \mathcal{F}_{S_{\mathfrak{H}}}$. It is easy to check that

$$\begin{gathered} P_nB_0P_n = T_{\mathfrak{H}}(P_nAP_n) - (P_nAP_n)T_{\mathfrak{H}} = (P_nAP_n)_{T_{\mathfrak{H}}}, \quad \text{and} \\ M \circ (P_nB_0P_n) = S_{\mathfrak{H}}(P_nAP_n) - (P_nAP_n)S_{\mathfrak{H}} = (P_nAP_n)_{S_{\mathfrak{H}}}. \end{gathered} \tag{4.22}$$

Since $\mathcal{F}_{S_{\mathfrak{H}}} = \mathcal{F}_{T_{\mathfrak{H}}}$, it follows from Lemma 4.1(i) that the norms $\|\cdot\|_{S_{\mathfrak{H}}}$ and $\|\cdot\|_{T_{\mathfrak{H}}}$ are equivalent. Therefore there exists $D > 0$ such that $\|P_nAP_n\|_{S_{\mathfrak{H}}} \leq D\|P_nAP_n\|_{T_{\mathfrak{H}}}$. Hence we obtain from (4.19), (4.21) and (4.22) that

$$\begin{aligned} \|M \circ (P_nBP_n)\| &= \|M \circ (P_nB_0P_n)\| = \left\|(P_nAP_n)_{S_{\mathfrak{H}}}\right\| \\ &\leq \|P_nAP_n\|_{S_{\mathfrak{H}}} \leq D\|P_nAP_n\|_{T_{\mathfrak{H}}} \\ &= D\left(\|P_nAP_n\| + \left\|(P_nAP_n)_{T_{\mathfrak{H}}}\right\|\right) \\ &\leq D(\|A\| + \|P_nB_0P_n\|) \leq D(\|A\| + \|B_0\|) \\ &= D(\|W \circ B\| + \|B_0\|) \leq D(|R| \, \|B\| + 2\|B\|) = \rho. \end{aligned}$$

Thus all operators $M \circ (P_nBP_n)$, $1 \leq n < \infty$, lie in the ball $\mathbf{B}_\rho$ of $B(\mathfrak{H})$ of radius ρ. Compactness of $\mathbf{B}_\rho$ in the weak operator topology implies that the sequence $\{M \circ (P_n B P_n)\}_{n=1}^{\infty}$ has a cluster point $K \in B(\mathfrak{H})$. Therefore there is a subsequence $\{M \circ (P_{n_j} B P_{n_j})\}$ such that for all e_k and e_m,

$$(Ke_k, e_m) = \lim_{j \to \infty} (M \circ (P_{n_j} B P_{n_j}) e_k, e_m).$$

If $n_j \geq \max(|k|, |m|)$ then $P_{n_j} e_k = e_k$ and $P_{n_j} e_m = e_m$ and, by (4.20),

$$\left(M \circ \left(P_{n_j} B P_{n_j}\right) e_k, e_m\right) = \left(P_{n_j}(M \circ B) P_{n_j} e_k, e_m\right) = (M \circ B e_k, e_m).$$

Hence $(Ke_k, e_m) = ((M \circ B)e_k, e_m)$, $-\infty < k, m < \infty$. Thus $K = M \circ B$, so M is a Schur multiplier. Similarly, we obtain that N is also a Schur multiplier. $\square$

Example 4.7. Let

$$s_i = i \quad \text{and} \quad t_i = (-1)^i i$$

in Theorem 4.6. If $\mathcal{F}_S = \mathcal{F}_T$ then, by Theorem 4.6, M is a Schur multiplier and we have that $|m_{ij}| \leq |M|$ for all i and j. Let $i = 2k$ and $j = -2k + 1$. Then $s_i = t_i = 2k$ and $s_j = -t_j = -2k + 1$. Hence

$$m_{ij} = \frac{s_i - s_j}{t_i - t_j} = 4k - 1 \to \infty, \quad \text{as} \quad k \to \infty.$$

This shows that M is not a Schur multiplier and, therefore, $\mathcal{F}_S \neq \mathcal{F}_T$.

Making use of Theorem 4.6, we obtain the following result of a more general character.

Theorem 4.8. *Let S and T be selfadjoint operators on H and $\mathcal{H}$ respectively. If there exists a bijection φ of $\mathbb{Z}$ onto $\mathbb{Z}$ such that*

$$\dim(\mathcal{H}_T(\varphi(i))) = \dim(H_S(i)), \quad \textit{for all} \ \ i \in \mathbb{Z},$$

(see (2.2) for definition of $\mathcal{H}_T(i)$ and $H_S(i)$) and if

$$M = (m_{ij}) \ \ \textit{where} \ \ m_{ij} = \frac{\varphi(i) - \varphi(j)}{i - j}, \quad \textit{for} \ \ i \neq j, \quad \textit{and} \ \ m_{ij} = 0, \quad \textit{and}$$

$$N = (n_{ij}) \ \ \textit{where} \ \ n_{ij} = \frac{i - j}{\varphi(i) - \varphi(j)}, \quad \textit{for} \ \ i \neq j, \quad \textit{and} \ \ n_{ij} = 0$$

*are Schur multipliers then the algebras $\mathcal{F}_S$ and $\mathcal{F}_T$ are *-isomorphic.*

Proof. Consider the operators $[S]$ and $[T]$ (see (2.1)) and the corresponding decompositions

$$H = \sum_{i \in \mathbb{Z}} \oplus H_S(i) \quad \text{and} \quad \mathcal{H} = \sum_{i \in \mathbb{Z}} \oplus \mathcal{H}_T(i)$$

where $H_S(i) = P_S(i)H$ and $\mathcal{H}_T(i) = P_T(i)\mathcal{H}$ (see (2.3)). The operators $S - [S]$ and $T - [T]$ are bounded, so $\mathcal{F}_S = \mathcal{F}_{[S]}$ and $\mathcal{F}_T = \mathcal{F}_{[T]}$.

Consider the selfadjoint operator R on H such that all subspaces $H_S(i)$ are invariant for R and $R|_{H_S(i)} = \varphi(i)\mathbf{1}_{H_S(i)}$. Since M and N are Schur multipliers, it follows from Theorem 4.6 that $\mathcal{F}_R = \mathcal{F}_{[S]}$.

On the other hand, since $\dim(\mathcal{H}_T(\varphi(i))) = \dim(H_S(i))$, for all $i \in \mathbb{Z}$, there exists an isometry operator U from H onto $\mathcal{H}$ which maps $H_S(i)$ onto $\mathcal{H}_T(\varphi(i))$. Then $U^*[T]U = R$. By Lemma 4.1, the algebras $\mathcal{F}_R$ and $\mathcal{F}_{[T]}$ are *-isomorphic. Hence the algebras $\mathcal{F}_S$ and $\mathcal{F}_T$ are *-isomorphic. □

References

[1] G. Bennett, *Schur multipliers,* Duke Math. J., **44** (1977), 603-639.

[2] B. Blackadar and J. Cuntz, *Differential Banach algebra norms and smooth subalgebras of C^*-algebras*, J. Oper. Theory, **26** (1991), 255-282.

[3] O. Bratteli and D.W. Robinson, *Unbounded derivations of C^*-algebras*, I, Comm. Math. Phys., **42** (1975), 253-268.

[4] J. Dixmier, *Les C^*-algebres et leurs Representations*, Gauthier-Villars, Paris, 1969.

[5] N. Dunford and J.T. Schwartz, *Linear Operators*, Vol. I+II, Interscience Inc., New York, 1988.

[6] I.Ts. Gohberg and M.G. Krein, *Introduction to the theory of linear non-selfadjoint operators in Hilbert spaces*, Nauka, Moscow, 1965.

[7] P.E.T. Jorgensen and G.L. Price, *Index theory and second quantization of boundary value problems*, J. Funct. Anal., **104** (1992), 243-290.

[8] E. Kissin, *Symmetric operator extensions of unbounded derivations of C^*-algebras*, J. Funct. Anal., **81** (1988), 38-53.

[9] ______, *Indices of unbounded derivations of C^*-algebras*, Pacific J. Math., **152** (1992), 125-150.

[10] E. Kissin, A.I. Loginov and V.S. Shulman, *Derivations of C^*-algebras and almost Hermitian representations on Π_k-spaces*, Pacific J. Math., **174** (1996), 411-430.

[11] E. Kissin and V.S. Shulman, *Dense Q-subalgebras of Banach and C^*-algebras and unbounded derivations of Banach and C^*-algebras*, Proc. Edinburgh Math. Soc., **36** (1993), 261-276.

[12] ______, *Differential properties of some dense subalgebras of C^*-algebras*, Proc. Edinburgh Math. Soc., **37** (1994), 399-422.

[13] ______, *Differential Banach *-algebras of compact operators associated with symmetric operators*, J. Funct. Anal., **156** (1998), 1-29.

[14] ______, *Representations on Krein Spaces and Derivations of C^*-algebras*, Addison Wesley Longman, London, 1997.

[15] M. Reed and B. Simon, *Methods of Modern Mathematical Physics*, Academic Press, New York London, 1972.

[16] S. Sakai, *Operator Algebras in Dynamical Systems*, CUP, Cambridge, 1991.

Received December 10, 1997 and revised September 16, 1998.

University of North London
Great Britain
E-mail address: 11kissine@unl.ac.uk

Polytechnic Institute of Vologda
Vologda
Russia

PACIFIC JOURNAL OF MATHEMATICS
Vol. 190, No. 2, 1999

COMPLETELY REGULAR MULTIVARIATE STATIONARY PROCESSES AND THE MUCKENHOUPT CONDITION

S. Treil and A. Volberg

We are going to give necessary and sufficient conditions for a multivariate stationary stochastic process to be completely regular. We also give the answer to a question of V.V. Peller concerning the spectral measure characterization of such processes.

1. Introduction.

In this paper we shall give a necessary and sufficient condition for a multivariate stationary stochastic process to be *completely regular*. For the scalar case the description of completely regular processes was obtained by Helson an Sarason, see [**2**, **9**]. Almost none of the scalar methods is available in the vector situation. The explanation is simple. Our problem will be reduced to verifying L^2 weighted inequalities for a certain integral operator. The weight will be a matrix weight arising from the spectral measure of the process. All the pointwise estimates of integral operators become too crude for the vector valued case. For example, if a positive kernel is majorized by another one, and this second kernel gives the bounded operator in $L^2(\mu)$, then the original kernel obviously corresponds to a bounded operator in $L^2(\mu)$ too. But this is not the case if μ is a matrix measure even for scalar kernels.

The study of prediction theory for multivariate stationary stochastic processes was started by Kolmogorov and Wiener in the 50's, see, for example [**13**], [**14**], and [**4**]. It was later continued in works of I. Ibragimov, Yu. Rozanov, V. Solev, A. Yaglom, V. Peller, S. Khruschev, N.J. Young. An extensive bibliography can be found in [**6**] (for scalar processes) and in [**5**] (for vector ones).

Let us recall that a multivariate stationary stochastic process with discrete time is a sequence of d-tuples $x(n) = (x_1(n), x_2(n), \dots, x_d(n))$, $n \in \mathbb{Z}$ of scalar random variables such that $\mathbb{E}|x_j(n)|^2 < \infty$ and the *correlation matrix* $Q(n,k)$

$$Q(n,k) = \{Q(n,k)_{i,j}\}_{1\le i,j\le d} := \left\{\mathbb{E}x_i(n)\overline{x_j(k)}\right\}_{1\le i,j\le d}$$

depends only on the difference $n-k$; here $\mathbb{E}$ denotes mathematical expectation.

It is well known (see [8]) that there exists a matrix-valued nonnegative measure M on the unit circle $\mathbb{T}$ whose Fourier coefficients coincide with entries of the correlation matrix

$$Q(n,k) = \widehat{M}(n-k)\,, \qquad n,k \in \mathbb{Z}\,.$$

The measure M is called the *spectral measure* of the process $\{x(n)\}_{n\in\mathbb{Z}}$.

The random variables $x_j(n)$ can be treated as elements of Hilbert space $L^2(\Omega, dP)$, where Ω is the probability space and P is the probability, so $x(n)$ can be treated as elements of the $\mathbb{R}^d$-valued L^2 space $L^2_{\mathbb{R}^d}(\Omega, dP)$ For a moment n of time we can consider the past $\mathcal{X}_n$ and the future $\mathcal{X}^n$ of the process, which are defined as the subspaces

$$\mathcal{X}_n = \operatorname{span}\{x_j(k) \,:\, 1 \le j \le d,\, k < n\}$$
$$\mathcal{X}^n = \operatorname{span}\{x_j(k) \,:\, 1 \le j \le d,\, k \ge n\}$$

of $L^2(\Omega, dP)$.

A process is called *regular* if $\cap_{n\ge 0}\mathcal{X}^n = \{0\}$. In this case (see [8]) the spectral measure M of the process is absolutely continuous with respect to Lebesgue measure. Let W be the density of M with respect to Lebesgue measure. The matrix-valued function W is called the *spectral density* of the process.

A process $\{x(n)\}_{n\in\mathbb{Z}}$ is called *completely regular* if its past is asymptotically orthogonal to the future, namely if

$$\sup\left\{|\mathbb{E}(\xi\eta)| \,:\, \xi \in \mathcal{X}_0,\, \eta \in \mathcal{X}^n,\, \mathbb{E}|\xi|^2 \le 1,\, \mathbb{E}|\eta|^2 \le 1\right\} \longrightarrow 0 \qquad \text{as } n \to \infty\,.$$

Of course, complete regularity implies regularity. If the process is Gaussian (i.e. all random variables $x_j(k)$ have normal distribution) then the complete regularity means simply that past and future are almost independent. The problem we are dealing with is to characterize completely regular processes in terms of spectral measure.

It has been already mentioned (see again [8]) that if the process is completely regular, then its spectral measure is absolutely continuous, $dM = W dm$ where dm is the normalized ($m(\mathbb{T}) = 1$) Lebesgue measure on the unit circle $\mathbb{T}$.

The reader is referred to [8] once more to see that there exists $d_0 \le d$ (the rank of the process) such that the spectral density $W(t)$ has rank d_0 for almost all $t \in \mathbb{T}$. If $d_0 = d$ then the process $\{x(n)\}$ is said to be a *full rank*.

The study of processes of arbitrary rank can be easily reduced to the study of the processes of full rank, see [3]. So in this paper we shall consider only processes of full rank.

For the scalar case the description of completely regular processes was obtained by Helson an Sarason, see [2, 9]. To state their result we need a couple of definitions.

Let us recall that a function f on the unit circle $\mathbb{T}$ belongs to the space BMO (bounded mean oscillation) if

$$\sup_I \frac{1}{|I|}\int_I |f - f_I|dm = \|f\|_{\mathrm{BMO}} < \infty\,;$$

here f_I denotes the mean value of f on the interval I: $f_I := |I|^{-1}\int_I f dm$ and the *supremum* is taken over all subarcs I of $\mathbb{T}$.

The space VMO (vanishing mean oscillation) consists of all function $f \in$ BMO such that

$$\sup_I \frac{1}{|I|}\int_I |f - f_I|dm \longrightarrow 0 \qquad \text{as } |I| \to 0\,.$$

Theorem 1.1 (Helson, Sarason)**.** *Let w be the spectral density of a scalar stationary process. Then the process is completely regular if and only if w admits a representation*

$$w = |p|^2 e^{\varphi}\,,$$

where p is a polynomial with roots on the unit circle $\mathbb{T}$ and φ is a real-valued function in VMO.

It was conjectured by V. Peller in [5] that the same result holds for multivariate stationary processes. Namely he conjectured that a multivariate stationary process is completely regular if and only if its spectral density W admits the following representation

$$W = P^* e^{\Phi} P,$$

where P is a polynomial matrix whose determinant has roots on $\mathbb{T}$ and the matrix function $\Phi = \Phi^*$ belongs VMO.

In this direction he was able to prove the following theorem:

Theorem 1.2. *A multivariate stationary process is completely regular if and only if its spectral density W admits the factorization*

$$W = P^* W_1 P,$$

where P is a polynomial matrix whose determinant has roots on $\mathbb{T}$ and W_1 is the density of a completely regular stationary process such that $W_1^{-1} \in L^1$.

1.1. The main result. Let us recall that a measure μ on the unit disk $\mathbb{D}$ is called Carleson if

$$\sup_I \mu(Q(I)) \le C \cdot |I|$$

and is called the vanishing Carleson measure if

$$\limsup_{|I|\to 0} \mu(Q(I))/|I| = 0$$

where *limsup* is taken over all subarcs I of T. Here $Q(I)$ denotes the "Carleson square" for the arc I,

$$Q(I) = \{z \in \mathbb{D}\,:\, z/|z| \in I, 1 - |I| \le |z| < 1\}.$$

For a function F on the unit circle let $F(\lambda)$, $\lambda \in \mathbb{D}$, denote its harmonic extension at the point λ.

The main result of the paper is the following theorem.

Theorem 1.3. *Let the density W of a stationary process satisfy $W^{-1} \in L^1$. Then the the following are equivalent:*

1) *The process is completely regular;*
2) *W^{-1} is the spectral density of a completely regular process;*
3) $\limsup_{|I|\to 0} \left\| \Big(\frac{1}{|I|}\int_I W dm\Big)^{1/2} \Big(\frac{1}{|I|}\int_I W^{-1} dm\Big)^{1/2} \right\| = 1$*; here* supremum *is taken over all subarcs I of $\mathbb{T}$;*
4) $\limsup_{|\lambda|\to 1} \left\| \Big(W(\lambda)\Big)^{1/2} \Big(W^{-1}(\lambda)\Big)^{1/2} \right\| = 1$, *where $W(\lambda)$ and $W^{-1}(\lambda)$ are harmonic extensions of functions $W\big|\mathbb{T}$ and $W^{-1}\big|\mathbb{T}$ respectively at point $\lambda \in \mathbb{D}$.*
5) $\limsup_{|\lambda|\to 1} \Big\{ \det\Big(W(\lambda)\Big) \exp\Big(-\big[\log\det W\big](\lambda)\Big)\Big\} = 1$, *where $W(\lambda)$ and $\big[\log\det W\big](\lambda)$ are harmonic extensions of functions $W\big|\mathbb{T}$ and $\log\det W\big|\mathbb{T}$ respectively at point $\lambda \in \mathbb{D}$.*
6) *The measures*

$$\left\| W(z)^{-1/2}\left(\frac{\partial}{\partial x}W(z)\right)W(z)^{-1/2}\right\|^2 (1-|z|^2)dxdy$$

and

$$\left\| W(z)^{-1/2}\left(\frac{\partial}{\partial y}W(z)\right)W(z)^{-1/2}\right\|^2 (1-|z|^2)dxdy$$

are vanishing Carleson measures.

Together with Theorem 1.2 the above theorem yields the complete description of completely regular stationary processes.

Theorem 1.4. *A stationary process with spectral density $\mathcal{W}$ is completely regular if and only if $\mathcal{W}$ admits the representation*

$$\mathcal{W} = P^* W P,$$

where P is a polynomial matrix whose determinant has roots on $\mathbb{T}$ and the matrix-function W satisfies $W^{-1} \in L^1$ and one of equivalent conditions 3-6 of Theorem 1.3.

Let us discuss the main result (Theorem 1.3) a little bit. First of all it is not difficult to show directly that in the scalar case the conditions 3-6 of Theorem 1.3 are equivalent to $W = e^{\varphi}$, $\varphi \in \mathrm{VMO}$. We are leaving this as an exercise for the reader.

Usually in probability only real valued stationary processes are considered. In that case the spectral density of a process should satisfy $W(\overline{z}) = W(z)$, and only such functions can be realized as densities of stationary processes.

If one allow complex-valued processes, any nonnegative matrix function is the spectral density of some stationary process.

Our theorem deals with arbitrary nonnegative matrix-functions and can be applied to complex-valued processes (as well as to real-valued).

2. Scheme of the proof of the main result.

The diagram of the proof will be the following: $1 \Longrightarrow 4 \Longrightarrow 5 \Longrightarrow 6 \Longrightarrow 1$. Then we will show that $1 \Longrightarrow 2$ and so automatically $2 \Longrightarrow 1$.

And in this section we will show that $3 \Longleftrightarrow 4$.

Lemma 2.1. *For a scalar weight w the following conditions are equivalent:*

1) $\limsup\limits_{|I|\to 0}\Big(\frac{1}{|I|}\int_I w\Big)\Big(\frac{1}{|I|}\int_I w^{-1}\Big) = 1$*;*
2) $\limsup\limits_{|\lambda|\to 1} w(\lambda)w^{-1}(\lambda) = 1$*, where $w(\lambda)$ and $w^{-1}(\lambda)$ denote the harmonic extensions of w and w^{-1} respectively at the point λ;*
3) $w = e^{\varphi}$*, where $\varphi \in$ VMO.*

Proof. First of all let us rewrite condition 1. Let $\varphi := \log w$. For a function f let f_I denote its average over the arc I, $f_I := |I|^{-1}\int_I f$. Then clearly

$$w_I \cdot (w^{-1})_I = \big[w_I \exp(-\varphi_I)\big] \cdot \big[(w^{-1})_I \exp(\varphi_I)\big].$$

By Jensen inequality (geometric mean $\le$ arithmetic mean) the expressions in brackets are at least 1, so the condition 1 splits into the following 2 conditions

$$\limsup_{|I|\to 0}\big[w_I \exp(-\varphi_I)\big] = 1, \qquad \text{and} \qquad \limsup_{|I|\to 0}\big[(w^{-1})_I \exp(\varphi_I)\big] = 1.$$

Let f_+ denote the positive part of the function f, $f_+(x) := \max(f(x), 0)$. Then the inequality

$$x \le e^x - 1 \qquad \text{for } x \ge 0$$

implies

$$\begin{aligned}\frac{1}{|I|}\int_I(\varphi - \varphi_I)_+ &\le \frac{1}{|I|}\int_I\big(\exp(\varphi - \varphi_I) - 1\big)\\ &= w_I \exp(-\varphi_I) - 1 \to 0 \quad \text{as } |I| \to 0.\end{aligned}$$

Since $\int_I |\varphi - \varphi_I| = 2\int_I (f - f_I)_+$, one can conclude that $\varphi \in$ VMO.

Similarly, using Poisson averages instead of averages over intervals one can get from condition 2 of the lemma that harmonic extension of $|\varphi - \varphi(\lambda)|$ at the point λ tends to 0 as $\lambda \to 1$. But that is an equivalent definition of VMO, so the condition 2 also implies that $\varphi \in$ VMO.

On the other hand, if $\varphi \in$ VMO, John-Nirenberg Theorem (see [**1**, Chapter VI]) claims that the measure of the set $\{t \in I : |\varphi(t) - \varphi_I| > a\}$ is estimated from above by Ce^{-Ka}, where $K = K_I \to \infty$ as $|I| \to 0$. Therefore for $x > 1$ the measure of the set $\{t \in I : \exp(\varphi(t) - \varphi_I) > x\}$ is estimated from above by Cx^{-K}. Integrating this distribution function one can get that $\limsup_{|I|\to 0} w_I \exp(-\varphi_I) \le 1$ (in fact, it is 1, because by Jensen inequality $w_I \exp(-\varphi_I) \ge 1$). Similarly, $\limsup_{|I|\to 0} (w^{-1})_I \exp(\varphi_I) = 1$. Multiplying the above two inequalities one gets condition 1.

The proof that 3 $\implies$ 2 is similar. For a point $\lambda \in \mathbb{D}$ let I_λ be an interval with center at $\lambda/|\lambda|$ of length $(1 - |\lambda|)^{1/3}$. Since the Poisson Kernel $P_\lambda(z) = (1 - |\lambda|^2) \cdot |1 - \overline{\lambda} z|^{-2}$ satisfies $\sup_{z \in \mathbb{T}\setminus I_\lambda} P_\lambda(z) \to 0$ as $|\lambda| \to 1$, the distribution inequality for φ on I_λ implies that $w(\lambda) \cdot \exp(-\varphi(\lambda)) \to 1$ as $|\lambda| \to 1$, and therefore the condition 2 of the lemma. □

The following Lemma is probably well known and can be easily from the distribution function inequality for VMO (John-Nirenberg Theorem).

Lemma 2.2. *For $\lambda \in \mathbb{D}$ let I_λ be an interval centered at $\lambda/|\lambda|$ of length $1 - |\lambda|$. If $\varphi \in$ VMO, then $\varphi_{I_\lambda} - \varphi(\lambda) \to 0$ as $|\lambda| \to 1$.*

Corollary 2.3. *Let $\varphi \in$ VMO and let $w = e^\varphi$. Then for I_λ as in the above lemma we have*

$$\lim_{|\lambda|\to 1} \frac{w(\lambda)}{w_{I_\lambda}} = 1.$$

Proof. By the above lemma $\lim_{|\lambda|\to 1} \exp(\varphi(\lambda))/\exp(\varphi_{I_\lambda}) = 1$. On the other hand it follows from the proof of Lemma 2.1 that

$$\lim_{|\lambda|\to 1} w(\lambda)/\exp(\varphi(\lambda)) = 1 \qquad \text{and} \qquad \lim_{|I|\to 0} w_I/\exp(\varphi_I) = 1.$$

Taking the ration of the last 2 identities (with $I = I_\lambda$) we get the statement we need. □

Now to show equivalence of condition 3 and 4 of Theorem 1.3 is enough to show that these conditions imply that for a fixed vector $e \in \mathbb{C}^d$ scalar weight $w(z) = (W(z)e, e)$ satisfies conditions 1 and 2 of Lemma 2.1. Then Corollary 2.3 implies that the averages W_{I_λ} and $W(\lambda)$ are equivalent, the same holds for W^{-1}, and we are done.

It remains now to show that the scalar weight $w(z) = (W(z)e, e)$ satisfies condition 1 (equivalently 2) of Lemma 2.1. The easiest way to do that is to recall where the Muckenhoupt condition (A_2) came from, see [**10**].

Recall that the quantity $\big\|[W_I]^{1/2}[(W^{-1})_I]^{1/2}\big\|$ is just the norm of the averaging operator $f \mapsto f_I \cdot \chi_I$ in the weighted space $L^2(W)$, see [**10**, Lemma 2.1]. Then $[w_I]^{1/2}[(w^{-1})_I]^{1/2}$ is the norm of the restriction of the above averaging operator onto the subspace of $L^2(W)$ consisting of functions of form fe where f is a scalar function. Therefore

$$1 \le [w_I]^{1/2}[(w^{-1})_I]^{1/2} \le \left\|[W_I]^{1/2}[(W^{-1})_I]^{1/2}\right\|$$

so the weight w satisfies condition 1 of the lemma.

Similarly, the quantity $\left\|W(\lambda)^{1/2}W^{-1}(\lambda)^{1/2}\right\|$ is just the norm of another averaging operator $\left(f \mapsto \int_{\mathbb{T}} fk_\lambda\right)k_\lambda$, where k_λ is the normalized reproducing kernel of H^2, $k_\lambda(z) = (1-|\lambda|^2)^{1/2}(1-\overline{\lambda}z)^{-1}$, see [**10**, Lemma 2.1], so condition 4 of the theorem implies condition 2 of the lemma for the weight w.

3. Eliminating probability.

The problem of description of completely regular processes can be now stated without mentioning any probability theory at all.

First of all notice that without loss of generality we can assume that the process is complex-valued. Namely, if we have a real stationary process $\{x(n)\}_{n\in\mathbb{Z}}$ we can consider its comlexification, namely the same process but in the complex Hilbert space $L^2_{\mathbb{C}^d}(\Omega, dP)$. Consider the comlexificated past $(\mathcal{X}_n)_{\mathbb{C}}$ and future $(\mathcal{X}^n)_{\mathbb{C}}$

$$(\mathcal{X}_n)_{\mathbb{C}} = \operatorname{span}\{x_j(k) \,:\, 1 \le j \le d,\, k < n\}$$
$$(\mathcal{X}^n)_{\mathbb{C}} = \operatorname{span}\{x_j(k) \,:\, 1 \le j \le d,\, k \ge n\}$$

where span now means the closed linear span in the complex Hilbert space $L^2_{\mathbb{C}^d}(\Omega, dP)$. It is easy to see that

$$\begin{aligned}\sup\left\{|\mathbb{E}(\xi\eta)| \,:\, \xi \in \mathcal{X}_0,\, \eta \in \mathcal{X}^n,\, \mathbb{E}|\xi|^2 \le 1,\, \mathbb{E}|\eta|^2 \le 1\right\}\\ = \sup\left\{|\mathbb{E}(\xi\bar{\eta})| \,:\, \xi \in (\mathcal{X}_0)_{\mathbb{C}},\, \eta \in (\mathcal{X}^n)_{\mathbb{C}},\, \mathbb{E}|\xi|^2 \le 1,\, \mathbb{E}|\eta|^2 \le 1\right\},\end{aligned}$$

so a process and its comlexification are completely regular simultaneously. So we indeed can assume from the beginning that our process is complex valued.

Consider now the vector space $L^2(W)$ of $\mathbb{C}^d$-valued functions on the unit circle with the norm

$$\|f\|^2_{L^2(W)} = \int_{\mathbb{T}} (W(\xi)f(\xi), f(\xi))_{\mathbb{C}^d}\, dm(\xi)$$

(of course we have to take the quotient space over the functions of norm 0). The mapping $x_j(k) \mapsto z^k e_j$, where e_j, $j = 1, ..., d$ is the standard orthonormal basis in $\mathbb{C}^d$, is an isometric isomorphism between $\operatorname{span}\{x_j(k) : 1 \le j \le d, k \in \mathbb{Z}\}$ and $L^2(W)$.

The past $\mathcal{X}_n$ and future $\mathcal{X}^n$ are mapped to the spaces X_n and X^n of $L^2(W)$

$$X_n = \operatorname{span}\{z^k\mathbb{C}^d \,:\, k < n\} \tag{3.1}$$

$$X^n = \operatorname{span}\{z^k\mathbb{C}^d \,:\, k \geq n\}\,. \tag{3.2}$$

So the problem of describing completely regular stationary processes can be reformulated as follows: *Describe all matrix weights W such that the spaces X_0 and X^n are asymptotically* (*as $n \to \infty$*) *orthogonal to each other,*

$$\rho_n = \sup\Big\{|(\xi,\eta)_{L^2(W)}| \,:\, \xi \in X_0,\, \eta \in X^n,\\ \|\xi\|_{L^2(W)} \leq 1,\, \|\eta\|_{L^2(W)} \leq 1\Big\} \longrightarrow 0, \tag{3.3}$$

as $n \to \infty$.

4. Necessity (1 $\Longrightarrow$ 4).

In this section we are going to prove the implication 1 $\Longrightarrow$ 4 (see Theorem 4.1 below) and the equivalence 1 $\Longleftrightarrow$ 2 (see Lemma 4.4).

For a function F defined on the unit circle $\mathbb{T}$ let $F(\lambda)$ denote its harmonic extension at the point $\lambda \in \mathbb{D}$.

Theorem 4.1. *Let W be a matrix valued weight such that $W^{-1} \in L^1$. Suppose the "past" X_0 and "future" X^n defined by* (3.1), (3.2) *are asymptotically orthogonal, which is*

$$\rho_n = \sup\Big\{|(\xi,\eta)_{L^2(W)}| \,:\, \xi \in X_0,\, \eta \in X^n,\\ \|\xi\|_{L^2(W)} \leq 1,\, \|\eta\|_{L^2(W)} \leq 1\Big\} \longrightarrow 0$$

as $n \to \infty$. Then

$$\limsup_{|\lambda|\to 1} \left\| \Big(W(\lambda)\Big)^{1/2}\Big(W^{-1}(\lambda)\Big)^{1/2} \right\| = 1\,.$$

Proof. First of all let us show that if W^{-1} is completely regular and $W^{-1} \in L^1$ then W satisfies the Muckenhoupt (A_2) condition

$$\sup_{\lambda\in\mathbb{D}} \left\| \Big(W(\lambda)\Big)^{1/2}\Big(W^{-1}(\lambda)\Big)^{1/2} \right\| < \infty\,. \tag{A_p}$$

Recall that $\left\| \Big(W(\lambda)\Big)^{1/2}\Big(W^{-1}(\lambda)\Big)^{1/2} \right\|$ is exactly the norm of the operator $f \mapsto (f, k_\lambda)k_\lambda$ in the weighted space $L^2(W)$; here k_λ denotes the normalized reproducing kernel for H^2,

$$k_\lambda(z) := \frac{(1-|\lambda|^2)^{1/2}}{1-\bar{\lambda}z}\,, \qquad \lambda \in \mathbb{D},$$

$\|k_\lambda\|_2 = 1$. Note that $k_0 \equiv 1$. So if $W^{-1} \in L^1$ the operator $f \mapsto (f,1)1$ is bounded in $L^2(W)$, and therefore by translation invariance the operators $f \mapsto (f, z^n)z^n = \hat{f}(n)z^n$ are bounded as well (they all have the same norm).

We know that the spaces X_0 and X^n are asymptotically orthogonal, so we can say that for large enough N the operator P_+ restricted onto $\operatorname{span}\{X_0, X^N\} = \operatorname{span}\{z^n\mathbb{C}^d : n \notin [0,N]\}$ is bounded, say by 2,

$$\|P_+ f\|_{L^2(W)} \le 2\|f\|_{L^2(W)}, \ \forall f \in \operatorname{span}\{X_0, X^N\} = \operatorname{span}\{z^n\mathbb{C}^d : n \notin [0,N]\}.$$

Since $f - \sum_{n=0}^N \hat{f}(n)z^n \in \operatorname{span}\{X_0, X^N\} = \operatorname{span}\{z^n\mathbb{C}^d : n \notin [0,N]\}$, one can conclude that the operator P_+ is bounded in $L^2(W)$, and so the weight satisfies the Muckenhoupt condition (A_2).

We will need the following simple lemma about Muckenhoupt weights.

Lemma 4.2. *If w is a scalar Muckenhoupt weight, then its harmonic extension $w(\lambda)$ cannot decay too fast near the boundary of the disk. Namely, if the Muckenhoupt norm of w is at most M there is a function $\alpha = \alpha_M$, $\alpha : [0,1) \to (0,\infty)$, $\alpha(t) \searrow 0$ as $t \to 1+$ such that*

$$\frac{(1-|\lambda|^2)w(0)}{w(\lambda)} \le \alpha(|\lambda|).$$

Proof of the lemma. For an arc $I \subset \mathbb{T}$ and $k > 0$ let kI denote the arc of length $k|I|$ with the same center as I.

We are going to show that for a Muckenhoupt weight w with the Muckenhoupt norm at most M

$$w_{2^n I} \le M^2(2-\varepsilon)^n w_I\,, \qquad \varepsilon = \varepsilon(M) > 0\,. \tag{4.1}$$

Applying this formula in the case $2^n I = \mathbb{T}$ and using the trivial estimate

$$w(\lambda) \ge C w_{I_\lambda}$$

where I_λ is the arc with center at the point $\lambda/|\lambda|$, $|I_\lambda| = 1 - |\lambda|^2$ and C is an absolute constant, we can get from there (recall that $|I_\lambda| = 1 - |\lambda|^2 = 2^{-n}$)

$$w(\lambda) \ge c(2-\varepsilon)^{-n}\cdot w(0) = c(2-\varepsilon)^{\log_2(1-|\lambda|^2)}\cdot w(0) = c\cdot(e-\delta)^{\log(1-|\lambda|^2)}\cdot w(0),$$

where $\delta = \delta(\varepsilon) > 0$; here e is the base of the natural logarithm, not a vector in $\mathbb{C}^d$. This estimate implies the conclusion of the lemma with $\alpha(t) = c^{-1}(1-t^2)\cdot(e-\delta)^{-\log(1-t^2)}$.

To prove (4.1) we notice the since the weight w^{-1} is the Muckenhoupt (A_2) weight with the same Muckenhoupt norm as w, it is doubling and therefore

$$(w^{-1})_{2I} \ge (2-\varepsilon)^{-1}(w^{-1})_I\,,$$

where ε depends only on the Muckenhoupt norm of w. Iterating this inequality n times we get

$$(w^{-1})_{2^n I} \ge (2-\varepsilon)^{-n}(w^{-1})_I.$$

The last estimate and the Muckenhoupt condition imply

$$w_{2^n I} \le M/(w^{-1})_{2^n I} \le M \cdot (2-\varepsilon)^n/(w^{-1})_I \le M^2(2-\varepsilon)^n w_I$$

and that is exactly what we need. $\square$

Corollary 4.3. *If a matrix weight W satisfies the Muckenhoupt condition (A_2) with the Muckenhoupt norm at most M then for any $e \in \mathbb{C}^d$*

$$(1-|\lambda|^2)\cdot \frac{(W(0)e,e)_{\mathbb{C}^d}}{(W(\lambda)e,e)_{\mathbb{C}^d}} \le \alpha(|\lambda|) \to 0 \qquad \text{as } |\lambda| \to 1,$$

where $\alpha = \alpha_M$ is the function from Lemma 4.2.

Proof of the corollary. The proof follows immediately from the fact that the scalar weight w, $w(\xi) = \big(W(\xi)e,e\big)_{\mathbb{C}^d}$ is the Muckenhoupt (A_2) weight with the Muckenhoupt norm at most M (see [**11**], proof of Corollary 2.4). $\square$

We now return to the proof of the theorem.

The condition $W^{-1} \in L^1$ implies that $\int_{\mathbb{T}} \log\det W(\xi) dm(\xi) > -\infty$, hence (see [**7**]) there exists a factorization of W of the form $W = F^*F$, where F is an outer matrix function in H^2.

Take $e \in \mathbb{C}^d$ and let us compute the distance

$$\operatorname{dist}_{L^2(W)}\{z^{-1}e, \operatorname{span}\{z^n\mathbb{C}^d : n \ge 0\}\} = \operatorname{dist}_{L^2(W)}\{e, \operatorname{span}\{z^n\mathbb{C}^d : n > 0\}\}.$$

By the vectorial version of the Szegö theorem (see [**7**]) this distance is exactly $\|F(0)e\|$. Using the Möbius transformation of the disk one can get from there

$$\operatorname{dist}_{L^2(W)}\left\{\frac{(1-|\lambda|^2)^{1/2}}{z-\lambda}e, \operatorname{span}\{z^n\mathbb{C}^d : n \ge 0\}\right\} = \|F(\lambda)e\|_{\mathbb{C}^d}.$$

Writing the Fourier series expansion of $\frac{(1-|\lambda|^2)^{1/2}}{z-\lambda}$

$$\frac{(1-|\lambda|^2)^{1/2}}{z-\lambda} = (1-|\lambda|^2)^{1/2}\sum_{n=0}^{\infty}\lambda^n z^{-(n+1)}$$

one can see that for any fixed $N > 0$ the function $\frac{(1-|\lambda|^2)^{1/2}}{z-\lambda}e$ is almost in the "past" X_{-N} as $|\lambda| \to 1$. Namely,

$$\begin{aligned} f_\lambda &= \frac{(1-|\lambda|^2)^{1/2}}{z-\lambda}e \\ &= (1-|\lambda|^2)^{1/2}\sum_{n=0}^{N-1}\lambda^n z^{-(n+1)}e + (1-|\lambda|^2)^{1/2}\sum_{n=N}^{\infty}\lambda^n z^{-(n+1)}e \\ &= f_\lambda^1 + f_\lambda^2, \end{aligned}$$

where $f_\lambda^2 \in X_{-N}$, and f_λ^1 is small,

$$\frac{\|f_\lambda^1\|_{L^2(W)}}{\|f_\lambda\|_{L^2(W)}} \le \frac{(1-|\lambda|^2)^{1/2} N \cdot \|e\|_{L^2(W)}}{\big(W(\lambda)e, e\big)_{\mathbb{C}^d}^{1/2}}$$
$$= \frac{(1-|\lambda|^2)^{1/2} N \cdot \big(W(0)e, e\big)_{\mathbb{C}^d}^{1/2}}{\big(W(\lambda)e, e\big)_{\mathbb{C}^d}^{1/2}} \le N\alpha(|\lambda|)^{1/2} \to 0,$$

as $|\lambda| \to 1$, where $\alpha(.)$ is as in Lemma 4.2 and Corollary 4.3.

Since X_0 and X^N are asymptotically orthogonal, the shift invariance implies that the subspaces X_{-N} and X^0 are asymptotically orthogonal as well. Taking $|\lambda| \to 1$ and then $N \to \infty$ we can conclude that

$$\begin{aligned}
&\|F(\lambda)e\|_{\mathbb{C}^d} / \|W(\lambda)^{1/2} e\|_{\mathbb{C}^d} \\
&= \operatorname{dist}_{L^2(W)} \left\{ \frac{(1-|\lambda|^2)^{1/2}}{z-\lambda} e,\ \operatorname{span}\{z^n \mathbb{C}^d \,:\, n \ge 0\} \right\} / \, \|f_\lambda\|_{L^2(W)} \\
&\ge 1 - \beta(|\lambda|) \to 1,
\end{aligned}$$

where $\beta(.)$ depends only on the Muckenhoupt norm of W and $\beta(|\lambda|) \to 0$ as $|\lambda| \to 1$.

The last inequality implies

$$\|W(\lambda)^{1/2} F(\lambda)^{-1}\| \le (1 - \beta(|\lambda|))^{-1}. \tag{4.2}$$

Note that since $\|F(\lambda)e\|_{\mathbb{C}^d} / \|W(\lambda)^{1/2} e\|_{\mathbb{C}^d} \le 1$ for all $e \in \mathbb{C}^d$, we have

$$\|W(\lambda)^{1/2} F(\lambda)^{-1}\| \ge 1.$$

We will show a little later that under assumptions of the theorem the subspaces X_0 and X^N in the weighted space $L^2(W^{-1})$ are asymptotically orthogonal as well. The factorization $W = F^* F$ yields the factorization $W^{-1} = F^{-1}(F^{-1})^*$ of W^{-1}. Similarly to the previous case

$$\begin{aligned}
&\operatorname{dist}_{L^2(W^{-1})} \left\{ \frac{(1-|\lambda|^2)^{1/2}}{1-\overline{\lambda} z} e,\ \operatorname{span}\{z^n \mathbb{C}^d \,:\, n \ge 0\} \right\} \\
&= \|F^{-1}(\lambda)^* e\|_{\mathbb{C}^d} = \|F(\lambda)^{-1*} e\|_{\mathbb{C}^d}.
\end{aligned}$$

Acting as before we get

$$\|W^{-1}(\lambda)^{1/2} F(\lambda)^*\| \le (1 - \beta_1(|\lambda|))^{-1} \tag{4.3}$$

where $\beta_1(|\lambda|) \to 0$ as $|\lambda| \to 1$.

Combining (4.2) and (4.3) we get

$$\|W(\lambda)^{1/2} W^{-1}(\lambda)^{1/2}\| \le (1 - \beta(|\lambda|))^{-1} (1 - \beta_1(|\lambda|))^{-1} \to 1 \qquad \text{as } |\lambda| \to 1\,.$$

So, we completed the proof modulo the following lemma. □

This lemma also gives us the equivalence 1 $\iff$ 2.

Lemma 4.4. *Under assumptions of Theorem 4.1 the weight W^{-1} is a spectral density of a completely regular process, i.e., the spaces X_0 and X^N are asymptotically orthogonal (as $N \to \infty$) in the weighted space $L^2(W^{-1})$.*

Proof. It is enough to show that

$$\|P_+ \bigm| \operatorname{span}\{X_0, X^N\}\|_{L^2(W^{-1})\to L^2(W^{-1})} \to 1 \qquad \text{as } N \to \infty.$$

The later is true because

$$\begin{aligned} &\|P_+ \bigm| \operatorname{span}\{X_0, X^N\}\|_{L^2(W^{-1})\to L^2(W^{-1})} \\ &= \|W^{-1/2}\big(P_+ \bigm| \operatorname{span}\{X_0, X^N\}\big)W^{1/2}\|_{L^2\to L^2} \\ &= \|W^{1/2}\big(P_+ \bigm| \operatorname{span}\{X_0, X^N\}\big)W^{-1/2}\|_{L^2\to L^2} \\ &= \|P_+ \bigm| \operatorname{span}\{X_0, X^N\}\|_{L^2(W)\to L^2(W)} \end{aligned}$$

and

$$\|P_+ \bigm| \operatorname{span}\{X_0, X^N\}\|_{L^2(W)\to L^2(W)} \to 1 \qquad \text{as } N \to \infty$$

(since X_0 and X^N are asymptotically orthogonal in $L^2(W)$). $\square$

5. Vanishing Carleson measures.

Recall that $W(\lambda)$ and $W^{-1}(\lambda)$ denote harmonic extensions at the point $\lambda \in \mathbb{D}$ of the weights W and W^{-1} respectively.

Lemma 5.1. *Let a matrix weight W satisfy*

$$\lim_{|\lambda|\to 1} \|W(\lambda)^{1/2}\big(W^{-1}\big)(\lambda)^{1/2}\| = 1.$$

Then

$$\limsup_{|\lambda|\to 1} \left\{ \det\Big(W(\lambda)\Big) \exp\Big(-\big[\log\det W\big](\lambda)\Big)\right\} = 1\,.$$

Proof. First of all let us notice that the assumption of the lemma implies that $W, W^{-1} \in L^1(\mathbb{T})$, therefore $\log(\det W) \in L^1(\mathbb{T})$. Therefore there exists a factorization $W = F^*F$ a.e. on $\mathbb{T}$, where F is an outer function in $H^2(M_{d\times d})$.

Since F is an outer function in H^2, $\det F$ is an outer function in $H^{2/d}$. Therefore

$$|\det F(z)| = \exp\left\{(\log|\det F|)(z)\right\} = \exp\left\{\frac{1}{2}\,(\log\det W)\,(z)\right\}. \tag{5.1}$$

It is well known fact that $F^*(z)F(z) \le W(z)$ for any $z \in D$, where $\le$ means the inequality for quadratic forms. There are many proofs of this fact, for example it admits a very simple operator-theoretic interpretation which is in fact hidden in the proof of Theorem 4.1. Explanation that we present here is more function-theoretic: Direct computation shows that

$$\Delta\left(F(z)^*F(z)\right) = 4\left(\bar\partial F(z)^*\right)\left(\partial F(z)\right) = 4\left(\partial F(z)\right)^*\left(\partial F(z)\right) \geq 0\,,$$

so for any $e \in \mathbb{C}^d$ the function $\|F(z)e\|^2$ is subharmonic and coincide with $(W(\xi)e, e)$ on $\mathbb{T}$.

We can do the same factorization for W^{-1}. Namely, let G be an outer matrix-valued function in $H^2(M_{d\times d})$ such that $W^{-1} = G^*G$ on $\mathbb{T}$. We should point out to the reader that in general G does not necessarily coincide with F^{-1}. However, applying (5.1) to G one can conclude that

$$|\det G(z)| = \exp\left\{\frac{1}{2}\left(\log\det W^{-1}\right)(z)\right\} = |\det F(z)|^{-1}. \tag{5.2}$$

Now we are in position to prove the lemma. By the assumption

$$\lim_{|z|\to 1}\left\|W(z)^{1/2}(W^{-1})(z)^{1/2}\right\| = 1, \tag{5.3}$$

and therefore,

$$\lim_{|z|\to 1}\left|\det(W(z))\det\left((W^{-1})(z)\right)\right| = 1.$$

Using (5.2) one can rewrite the last identity as

$$\lim_{|z|\to 1}\left\{\left[\det W(z)/|\det F(z)|^2\right]\left[\det W^{-1}(z)/|\det G(z)|^2\right]\right\} = 1.$$

Since $F(z)^*F(z) \leq W(z)$ and $G(z)^*G(z) \leq W^{-1}(z)$, expressions in brackets are at least 1, so, taking into account (5.1) we get

$$\lim_{|z|\to 1}\left[\det W(z)/\exp\left\{(\log\det W)(z)\right\}\right] = 0$$

or equivalently

$$\lim_{|z|\to 1}\log\left\{\det(W(z))\right\} - \left(\log\det W\right)(z) = 0\,. \tag{5.4}$$

□

Theorem 5.2. *A matrix weight W satisfies*

$$\limsup_{|\lambda|\to 1}\left\{\det\Big(W(\lambda)\Big)\exp\Big(-\big[\log\det W\big](\lambda)\Big)\right\} = 1$$

if and only if the measures

$$\left\|W(z)^{-1/2}\left(\frac{\partial}{\partial x}W(z)\right)W(z)^{-1/2}\right\|^2(1-|z|^2)dxdy$$

and

$$\left\|W(z)^{-1/2}\left(\frac{\partial}{\partial y}W(z)\right)W(z)^{-1/2}\right\|^2(1-|z|^2)dxdy$$

are vanishing Carleson measures.

The implication 3 $\Longrightarrow$ 4 of Theorem 1.3 follows immediately from Theorem 5.2 and Lemma 5.1.

To prove the theorem we need the following well known description of vanishing Carleson measures:

Lemma 5.3. *A measure μ in the unit disk $\mathbb{D}$ is a vanishing Carleson measure if and only if*

$$\lim_{|\lambda|\to 1}\int_{\mathbb{D}}\frac{1-|\lambda|^2}{|1-\bar{\lambda}z|^2}d\mu(z)=0.$$

We also need the following lemma that was proved in [**11**], see Lemma 3.1 there.

Lemma 5.4. *Let W be a harmonic function of n variables with values in the space of strictly positive $d\times d$ matrices ($W(x)=W(x)^*>0\ \forall x$). Then*

$$\Delta\left(\log(\det W)\right)=-\sum_{j=1}^{n}\operatorname{trace}\left(\left(W^{-1/2}\frac{\partial W}{\partial x_j}W^{-1/2}\right)^2\right).$$

Proof of Theorem 5.2*.* The proof below follows the lines of the proof of Theorem 3.2 of [**11**].

By Green's formula and Lemma 5.4

$$\begin{aligned}
&\log\{\det(W(s))\}-(\log\det W)(s)\\
&=-\frac{1}{2\pi}\iint_{\mathbb{D}}\log\left|\frac{1-\bar{s}z}{z-s}\right|\Delta\log\{\det(W(z))\}\,dxdy\\
&=\frac{1}{4\pi}\iint_{\mathbb{D}}\left\{\operatorname{trace}\left(W(z)^{-1/2}\frac{\partial W(z)}{\partial x}W(z)^{-1/2}\right)^2\right.\\
&\quad\left.+\operatorname{trace}\left(W(z)^{-1/2}\frac{\partial W(z)}{\partial y}W(z)^{-1/2}\right)^2\right\}\log\left|\frac{1-\bar{s}z}{z-s}\right|^2 dxdy.
\end{aligned}$$

Using an elementary inequality $\log(1/a)\ge 1-a$ for $0<a\le 1$ and the fact that $\|A\|\le\operatorname{trace}A$ for a nonnegative matrix A, the last integral is at least

$$\begin{aligned}
&\frac{1}{4\pi}\iint_{\mathbb{D}}\left\|W(z)^{-1/2}\frac{\partial W(z)}{\partial x}W(z)^{-1/2}\right\|^2\log\left|\frac{1-\bar{s}z}{z-s}\right|^2 dxdy\\
&\ge\frac{1}{4\pi}\iint_{\mathbb{D}}\left\|W(z)^{-1/2}\frac{\partial W(z)}{\partial x}W(z)^{-1/2}\right\|^2\left(1-\left|\frac{z-s}{1-\bar{s}z}\right|^2\right)dxdy\\
&=\iint_{\mathbb{D}}\left\|W(z)^{-1/2}\frac{\partial W(z)}{\partial x}W(z)^{-1/2}\right\|^2\cdot\frac{(1-|s|^2)(1-|z|^2)}{|1-\bar{s}z|^2}\,dxdy.
\end{aligned}$$

Together with (5.4) this imply

$$\lim_{|s|\to 1} \iint_{\mathbb{D}} \frac{(1-|s|^2)}{|1-\overline{s}z|^2} \cdot \left\| W(z)^{-1/2} \frac{\partial W(z)}{\partial x} W(z)^{-1/2} \right\|^2 (1-|z|^2)\, dxdy = 0$$

that yields that the measure $\left\| W(z)^{-1/2} \left(\frac{\partial}{\partial x} W(z)\right) W(z)^{-1/2} \right\|^2 (1-|z|^2)\, dxdy$ is a vanishing Carleson measure.

The measure $\left\| W(z)^{-1/2} \left(\frac{\partial}{\partial y} W(z)\right) W(z)^{-1/2} \right\|^2 (1-|z|^2)\, dxdy$ is treated similarly.

To prove the opposite implication, let us estimate the integral

$$\iint_{\mathbb{D}} \operatorname{trace} \left(W(z)^{-1/2} \frac{\partial W(z)}{\partial x} W(z)^{-1/2} \right)^2 \log \left| \frac{1-\overline{s}z}{z-s} \right|^2 \, dxdy$$

(the integral with $\partial W/\partial y$ can be estimated absolutely the same way). Denote by b_s a Blaschke factor with zero at the point s, $b_s(z) = (z-s)(1-\overline{s}z)^{-1}$.

First of all, we can estimate the trace by $d \cdot \|\cdot\|$, where d is dimension of the space. So we can estimate the integral by

$$\begin{aligned} &C \iint\limits_{\mathbb{D}} \left\| W(z)^{-1/2} \frac{\partial W(z)}{\partial x} W(z)^{-1/2} \right\|^2 \log |b_s(z)|^{-2} dxdy \\ &= \iint\limits_{|b_s(z)|<\varepsilon} \cdots + \iint\limits_{|b_s(z)|\ge\varepsilon} \cdots . \end{aligned}$$

To estimate the second integral we notice that

$$\log |b_s(z)|^{-2} dxdy \le C(\varepsilon) \frac{(1-|s|^2)(1-|z|^2)}{|1-\overline{s}z|^2}$$

for $|b_s(z)| \ge \varepsilon$, and since the measure is a vanishing Carleson measure we can make the integral as small as we want when $|s| \to 1$.

To estimate the first integral let make a trivial observation: If $w \in L^1(\mathbb{T})$, $w \ge 0$ and $w(z)$ denotes its harmonic extension at the point z, then for all z such that $|z| \le 1/2$ (and therefore for all z such that $|z| < \varepsilon \le 1/2$)

$$\frac{\partial}{\partial x} w(z) \le C w(0),$$

where C is an absolute constant. Combining this observation with the Harnack inequality $w(0) \le C' w(z)$, $|z| \le 1/2$, and applying it to functions $w(.) = \big(W(\cdot)e, e\big)_{\mathbb{C}^d}$ we get the inequality for quadratic forms

$$\frac{\partial}{\partial x} W(z) \le C\varepsilon W(0) \le C_1 W(z).$$

This implies

$$\left\| W(z)^{-1/2} \Big(\frac{\partial}{\partial x} W(z) \Big) W(z)^{-1/2} \right\| \le C_1, \qquad \forall z : \ |z| < \varepsilon \le 1/2.$$

Using the Möbius transformation $z \mapsto b_s(z)$ we get

$$\left\| W(z)^{-1/2}\Big(\frac{\partial}{\partial x}W(z)\Big)W(z)^{-1/2}\right\| \le C\varepsilon\,, \qquad \forall z:\ |b_s(z)| < \varepsilon \le 1/2\,.$$

Since

$$\iint_{|b_s(z)|\le\varepsilon} \log|b_s(z)|^{-2}dxdy \le C\varepsilon^2 \log\frac{1}{\varepsilon}\,,$$

we can estimate the first integral by $C\varepsilon^2\log(1/\varepsilon)$; we can make this number as small as we want by picking sufficiently small ε.

□

6. Embedding theorem and equivalent norms.

By analogy with the scalar case (see [**12**]) we will say that a matrix weight W satisfies the *invariant* A_∞ *condition* if

$$(invA_\infty) \qquad \sup_{s\in\mathbb{D}}\Big\{\det\Big(W(s)\Big)\exp\Big(-\big[\log\det W\big](s)\Big)\Big\} < \infty\,.$$

The *supremum* is called the invariant A_∞ norm of W.

Theorem 5.2 implies that if the measures

$$\left\| W(z)^{-1/2}\left(\frac{\partial}{\partial x}W(z)\right)W(z)^{-1/2}\right\|^2 (1-|z|^2)dxdy$$

and

$$\left\| W(z)^{-1/2}\left(\frac{\partial}{\partial y}W(z)\right)W(z)^{-1/2}\right\|^2 (1-|z|^2)dxdy$$

are vanishing Carleson measures then the weight W satisfies the invariant A_∞ condition.

Literally repeating the proof of Theorem 5.2 one can obtain that the weight W satisfies the invariant A_∞ condition if and only if the above measures are Carleson.

We will need the following "embedding theorem". More general result was proved in [**11**], Lemma 4.1.

Lemma 6.1. *Let W be a matrix weight satisfying the invariant A_∞ condition, and let μ be a Carleson measure with the Carleson norm $\|\mu\|_C$. Then for any analytic (or antianalytic) vector-function f, the following inequality holds,*

$$\iint_{\mathbb{D}}(W(z)f(z),f(z))\,d\mu(z) \le C\|\mu\|_C\int_{\mathbb{T}}(W(\xi)f(\xi),f(\xi))dm(\xi),$$

where the constant C depends the dimension d and the invariant A_∞ norm of W.

Proof. The invariant A_∞ condition implies that $\log\det W \in L^1$, so there exists (see [7]) an outer function $F \in H^2(M_{d\times d})$ such that $W = F^*F$. It is well known (see again [7]) that

$$|\det F(z)| = \exp\left\{\frac{1}{2}\big[\log\det W\big](z)\right\}.$$

It is well known and it was already shown it in the proof of Lemma 5.1 that $F(z)^*F(z) \le W(z)$. Hence

$$\|W(z)^{1/2}F(z)^{-1}e\| \ge \|e\|, \qquad e \in \mathbb{C}^d. \tag{6.5}$$

Since

$$\left|\det\left\{W(z)^{1/2}F(z)^{-1}\right\}\right| = \left\{\det\Big(W(\lambda)\Big)\exp\Big(-\big[\log\det W\big](\lambda)\Big)\right\}^{1/2} \le C$$

we can estimate

$$\|W(z)^{1/2}F(z)^{-1}e\| \le C.$$

Together with (6.5) it implies that $(W(z)e, e)$ and $\|F(z)e\|^2$ are equivalent in a sense of two-sided estimate. Therefore

$$\begin{aligned}
&\iint_{\mathbb{D}} (W(z)f(z), f(z))\, d\mu(z)\\
&\le C\iint_{\mathbb{D}} (F(z)f(z), F(z)f(z))\, d\mu(z)\\
&\le C\|\mu\|_C \int_{\mathbb{T}} (F(\xi)f(\xi), F(\xi)f(\xi))dm(\xi)\\
&= C\|\mu\|_C \int_{\mathbb{T}} (W(\xi)f(\xi), f(\xi))dm(\xi).
\end{aligned}$$

□

We also need the following simple lemma.

Lemma 6.2 (equivalence of weighted norms)**.** *Let W be a matrix weight satisfying the invariant A_∞ condition. There exist a constant C such that for any analytic or antianalytic vector-function f in $L^2(W)$ satisfying $f(0) = 0$*

$$\frac{1}{C}\int_{\mathbb{T}} (Wf, f)dm \le \iint_{\mathbb{D}} (W(z)f'(z), f'(z))\log\frac{1}{|z|}\, dxdy \le C\int_{\mathbb{T}} (Wf, f)dm.$$

Proof. Let us recall the the operators ∂ and $\overline{\partial}$ are defined as

$$\partial f = \frac{1}{2}\left(\frac{\partial f}{\partial x} - i\frac{\partial f}{\partial y}\right), \qquad \overline{\partial} f = \frac{1}{2}\left(\frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y}\right).$$

Recall that for analytic functions $\partial f = f'$ and $\overline{\partial} f = 0$.

Let f be an analytic function, $f(0)=0$. Using the Green's formula and taking into account that $f(0)=0$ and $\Delta = 4\partial\overline{\partial} = 4\overline{\partial}\partial$ we get

$$\begin{aligned}\int_{\mathbb{T}} \big(Wf,f\big)dm &= \frac{1}{2\pi}\iint_{\mathbb{D}} \Delta\big(W(z)f(z),f(z)\big)\log\frac{1}{|z|}dxdy\\ &= \frac{2}{\pi}\iint_{\mathbb{D}}\big(\overline{\partial}W(z)f'(z),f(z)\big)\log\frac{1}{|z|}dxdy\\ &\quad+\frac{2}{\pi}\iint_{\mathbb{D}}\big(\partial W(z)f(z),f'(z)\big)\log\frac{1}{|z|}dxdy\\ &\quad+\frac{2}{\pi}\iint_{\mathbb{D}}\big(W(z)f'(z),f'(z)\big)\log\frac{1}{|z|}dxdy\\ &= \frac{2}{\pi}(\mathcal{I}_1+\mathcal{I}_2+\mathcal{I}_3).\end{aligned}$$

The last integral $\mathcal{I}_3$ is exactly the integral we want to estimate. Let us denote $A^2 := \int_{\mathbb{T}}\big(Wf,f\big)dm$, $B^2 := \mathcal{I}_3$. We want to show that $A \asymp B$ in a sense of two sided estimate. Let us estimate $\mathcal{I}_1$:

$$\begin{aligned}&|\mathcal{I}_1|\\ &= \Big|\iint_{\mathbb{D}}\big(W(z)^{-1/2}\overline{\partial}W(z)W(z)^{-1/2}W(z)^{1/2}f'(z),W(z)^{1/2}f(z)\big)\\ &\qquad\qquad\cdot\log\frac{1}{|z|}dxdy\Big|\\ &\le \Big|\iint_{\mathbb{D}}\big\|W(z)^{-1/2}\overline{\partial}W(z)W(z)^{-1/2}\big\|\cdot\big\|W(z)^{1/2}f'(z)\big\|\cdot\big\|W(z)^{1/2}f(z)\big\|\\ &\qquad\qquad\cdot\log\frac{1}{|z|}dxdy\Big|\\ &\le \Big(\iint_{\mathbb{D}}\big\|W(z)^{-1/2}\overline{\partial}W(z)W(z)^{-1/2}\big\|^2\big(W(z)f(z),f(z)\big)_{\mathbb{C}^d}\log\frac{1}{|z|}dxdy\Big)^{1/2}\\ &\qquad\qquad\cdot\Big(\iint_{\mathbb{D}}\big(W(z)f'(z),f'(z)\big)_{\mathbb{C}^d}\log\frac{1}{|z|}dxdy\Big)^{1/2}.\end{aligned}$$

The measure $\big\|W(z)^{-1/2}\overline{\partial}W(z)W(z)^{-1/2}\big\|^2\log\frac{1}{|z|}dxdy$ is Carleson, so by Lemma 6.1 the first term in the product is estimated by KA (K is a constant). The second term is just B so $|\mathcal{I}_1|\le KAB$. Similarly $|\mathcal{I}_2|\le KAB$. So

$$A^2 = B^2+\mathcal{I}_1+\mathcal{I}_2\,,$$

where

$$|\mathcal{I}_1|,\ |\mathcal{I}_2|\le KAB\,.$$

This immediately implies

$$\frac{1}{C}A\le B\le CA$$

for an appropriate choice of C. $\square$

7. Proof of the implication 6 $\Longrightarrow$ 1.

To prove the implication 6 $\Longrightarrow$ 1 we need to estimate $\int_{\mathbb{T}}(Wz^n f, g)dm$, $f \in X^0$, $g \in X_0$, $\|f\|_{L^2(W)} = \|g\|_{L^2(W)} = 1$.

Using the Green's formula and taking into account that $g(0) = 0$ and $\Delta = 4\partial\overline{\partial} = 4\overline{\partial}\partial$ we get

$$\begin{aligned}
\int_{\mathbb{T}}(Wz^n f, g)dm &= \frac{1}{2\pi}\iint_{\mathbb{D}} \Delta\big(W(z)z^n f(z), g(z)\big)_{\mathbb{C}^d} \log\frac{1}{|z|}dxdy \\
&= \frac{2}{\pi}\iint_{\mathbb{D}}\big(\overline{\partial}W(z)\partial(z^n f(z)), g(z)\big)_{\mathbb{C}^d}\log\frac{1}{|z|}dxdy \\
&\quad + \frac{2}{\pi}\iint_{\mathbb{D}}\big(\overline{\partial}W(z)(z^n f(z)), \overline{\partial}g(z)\big)_{\mathbb{C}^d}\log\frac{1}{|z|}dxdy \\
&= \frac{2}{\pi}(\mathcal{I}_1 + \mathcal{I}_2).
\end{aligned}$$

The second integral is easy to estimate:

$$\begin{aligned}
&|\mathcal{I}_2| \\
&= \Big|\iint_{\mathbb{D}}\big(W(z)^{-1/2}\overline{\partial}W(z)W(z)^{-1/2}W(z)^{1/2}(z^n f(z)), W(z)^{1/2}\overline{\partial}g(z)\big)_{\mathbb{C}^d} \\
&\qquad\qquad \cdot\log\frac{1}{|z|}dxdy\Big| \\
&\le \iint_{\mathbb{D}}\|W(z)^{-1/2}\overline{\partial}W(z)W(z)^{-1/2}\|\cdot\|W(z)^{1/2}(z^n f(z))\|\cdot\|W(z)^{1/2}\overline{\partial}g(z)\| \\
&\qquad\qquad \cdot\log\frac{1}{|z|}dxdy \\
&\le \Big(\iint_{\mathbb{D}}|z|^{2n}\cdot\|W(z)^{-1/2}\overline{\partial}W(z)W(z)^{-1/2}\|^2\cdot\big(W(z)f(z), f(z)\big)_{\mathbb{C}^d} \\
&\qquad\qquad \cdot\log\frac{1}{|z|}dxdy\Big)^{1/2} \\
&\quad \cdot\Big(\iint_{\mathbb{D}}\big(W(z)\overline{\partial}g(z), \overline{\partial}g(z)\big)_{\mathbb{C}^d}\log\frac{1}{|z|}dxdy\Big)^{1/2}.
\end{aligned}$$

The last term is equivalent to the norm $\|g\|_{L^2(W)}$ (see Lemma 6.2), so by Lemma 6.1

$$\begin{aligned}
|\mathcal{I}_2| &\le \|f\|_{L^2(W)}\cdot\|g\|_{L^2(W)} \\
&\qquad \cdot\Big\||z|^{2n}\cdot\|W(z)^{-1/2}\overline{\partial}W(z)W(z)^{-1/2}\|\log\frac{1}{|z|}dxdy\Big\|_C^{1/2}.
\end{aligned}$$

Since the measure $\|W(z)^{-1/2}\overline{\partial}W(z)W(z)^{-1/2}\| \log \frac{1}{|z|} dxdy$ is a vanishing Carleson measure, the Carleson norm

$$\left\| |z|^{2n} \cdot \|W(z)^{-1/2}\overline{\partial}W(z)W(z)^{-1/2}\| \log \frac{1}{|z|} dxdy \right\|_C^{1/2} \to 0$$

as $n \to \infty$. So $|\mathcal{I}_2| \to 0$ as $n \to \infty$.

To estimate $\mathcal{I}_1$ we pick $r < 1$ close to 1 and split the integral into two: $\mathcal{I}_1 = \iint_{r\mathbb{D}} \dots + \iint_{\mathbb{D}\setminus r\mathbb{D}} \dots$. Acting as with $\mathcal{I}_2$ we can estimate

$$\begin{aligned} &\left| \iint_X \dots \right| \\ &\le \left(\iint_X \cdot \|W(z)^{-1/2}\overline{\partial}W(z)W(z)^{-1/2}\|^2 \cdot \big(W(z)g(z), g(z)\big)_{\mathbb{C}^d} \cdot \log \frac{1}{|z|} dxdy \right)^{1/2} \\ &\quad \cdot \left(\iint_X \big(W(z)\partial\big(z^n f(z)\big), \partial\big(z^n f(z)\big)\big)_{\mathbb{C}^d} \log \frac{1}{|z|} dxdy \right)^{1/2}, \end{aligned}$$

where X is either $r\mathbb{D}$ or $\mathbb{D}\setminus r\mathbb{D}$. Note that both terms are uniformly bounded.

We can say even more. If $X = r\mathbb{D}$ the second term can be made as small as we wish by picking sufficiently large n.

Let now $X = \mathbb{D}\setminus r\mathbb{D}$. The measure $\|W(z)^{-1/2}\overline{\partial}W(z)W(z)^{-1/2}\| \log \frac{1}{|z|} dxdy$ is a vanishing Carleson measure, so for r sufficiently close to 1 its restriction onto $\mathbb{D} \setminus r\mathbb{D}$ has the Carleson norm as small as we want. So by Lemma 6.1 the first term is as small as we want if r is sufficiently close to 1.

8. A counterexample to Peller's conjecture.

In this section we are going to construct a weight W, such that $W^{-1} \in L^1$, $\log W \in \mathrm{VMO}$, but the corresponding stationary process is not completely regular (i.e., the weight W does not satisfy any of the conditions 1–6 of Theorem 1.3).

Let

$$W = U^* \begin{pmatrix} 1 & 0 \\ 0 & \delta(z) \end{pmatrix} U, \qquad U = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}.$$

Here

$$\delta(e^{it}) = 1/\log(1/|t|), \qquad -1/4 \le t \le 1/4,$$

and δ is a continuous function bounded away from 0 and ∞ on the rest of the circle, and

$$\alpha(e^{it}) = (t/|t|)\delta(e^{it})^{1/4}, \qquad -1/4 \le t \le 1/4,$$

and again α is continuous on the rest of the circle.

Then

$$\log W = U^* \begin{pmatrix} 0 & 0 \\ 0 & \log\delta \end{pmatrix} U = \begin{pmatrix} \sin^2\alpha\log\delta & \sin\alpha\cos\alpha\log\delta \\ \sin\alpha\cos\alpha\log\delta & \cos^2\alpha\log\delta \end{pmatrix},$$

and this matrix clearly belongs to VMO: $\log\delta = \log\log 1/|t|$ (considered only in a neighborhood of 0) is a "typical" unbounded function in VMO, so $\cos^2\alpha\log\delta \in \mathrm{VMO}$, and all other entries of the matrix are continuous.

Let us now show that the weight W does not even satisfies the Muckenhoupt condition (A_2). Direct computations show that

$$W = \begin{pmatrix} \cos^2\alpha & -\sin\alpha\cos\alpha \\ -\sin\alpha\cos\alpha & \sin^2\alpha \end{pmatrix} + \delta \begin{pmatrix} \sin^2\alpha & \sin\alpha\cos\alpha \\ \sin\alpha\cos\alpha & \cos^2\alpha \end{pmatrix}$$

and

$$W^{-1} = \begin{pmatrix} \cos^2\alpha & -\sin\alpha\cos\alpha \\ -\sin\alpha\cos\alpha & \sin^2\alpha \end{pmatrix} + \delta^{-1} \begin{pmatrix} \sin^2\alpha & \sin\alpha\cos\alpha \\ \sin\alpha\cos\alpha & \cos^2\alpha \end{pmatrix}.$$

If we pick I to be a symmetric arc $[e^{-i\varepsilon}, e^{i\varepsilon}]$ ($\varepsilon > 0$ is small), then off-diagonal entries of W_I and $(W^{-1})_I$ equal 0, and so we can estimate

$$W_I \ge C \begin{pmatrix} \cos^2\alpha(\varepsilon) & 0 \\ 0 & \sin^2\alpha(\varepsilon) \end{pmatrix},$$

$$(W^{-1})_I \ge C \begin{pmatrix} \delta(\varepsilon)^{-1}\sin^2\alpha(\varepsilon) & 0 \\ 0 & \delta(\varepsilon)^{-1}\cos^2\alpha(\varepsilon) \end{pmatrix}.$$

Therefore

$$\big\|[W_I]^{1/2}[(W^{-1})_I]^{1/2}\big\| \ge C\delta(\varepsilon)^{-1}\sin\alpha(\varepsilon)\cos\alpha(\varepsilon) \to \infty \qquad \text{as} \quad \varepsilon \to 0\,.$$

References

[1] J.B. Garnett, *Bounded Analytic Functions*, Acad. Press, NY, 1981.

[2] H. Helson and D. Sarason, *Past and future*, Math. Scand., **21** (1967), 5-16.

[3] I.A. Ibragimov, *Completely regular multidimensional stationary processes with discrete time*, Proc. Steklov Inst. Math., **111** (1970), 269-301.

[4] P. Masani and N. Wiener, *On bivariate stationary processes and the factorization of matrix-valued functions*, Theor. Probability Appl., **4** (1959), 300-308.

[5] V.V. Peller, *Hankel operators and multivariate stationary processes*, Operator theory: Operator algebras and applications, Part 1, (Durham, NH, 1988), 357-371, Proc. Sympos. Pure Math., **51**(**1**), Amer. Math. Soc., Providence, RI, 1990.

[6] V.V. Peller and S.V. Khruschev, *Hankel operators, best approximation, and stationary Gaussian processes*, Russian Math. Surveys, **37** (1982), 53-124.

[7] M. Rosenblum and J. Rovnyak, *Hardy Classes and Operator Theory*, (Oxford Mathematical Monographs) Oxford Science Publications, The Clarendon Press, Oxford University Press, New York, 1985.

[8] Yu.A. Rozanov, *Stationary Stochastic Processes*, Holden-Day, SF, 1967.

[9] D. Sarason, *An addendum to "Past and future"*, Math. Scand., **30** (1972), 62-64.

[10] S. Treil and A. Volberg, *Wavelets and the angle between past and future*, Journal of functional analysis, **143**(**2**) (1997), 269-308.

[11] ______, *Continuous frame decomposition and a vector Hunt-Muckenhoupt-Wheeden Theorem*, Arkiv för Matematik, **35**(**2**) (1997), 363-386.

[12] S. Treil, A. Volberg and D. Zheng, *Hilbert transform, Toeplitz operators and Hankel operators, and invariant A_∞ weights*, to appear in Rev. Mat. Iberoamericana.

[13] N. Wiener and P. Masani, *The prediction theory of multivariate stochastic processes,* I. *The regularity conditions*, Acta Math., **98** (1957), 111-150.

[14] ______, *The prediction theory of multivariate stochastic processes,* II. *The linear predicator*, Acta Math., **99** (1958), 93-137.

Received February 2, 1998. This research was partially supported by the NSF grant DMS 9622936, binational Israeli-USA grant BSF 00030, and research program at MSRI in the Fall of 1997.

Michigan State University
East Lansing, Michigan 48824
E-mail address: treil@math.msu.edu

Michigan State University
East Lansing, Michigan 48824
E-mail address: volberg@math.msu.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 190, No. 2, 1999

LOWER BOUND ESTIMATES OF THE FIRST EIGENVALUE FOR COMPACT MANIFOLDS WITH POSITIVE RICCI CURVATURE

DaGang Yang

We present some new lower bound estimates of the first eigenvalue for compact manifolds with positive Ricci curvature in terms of the diameter and the lower Ricci curvature bound of the manifolds. For compact manifolds with boundary, it is assumed that, with respect to the outward normal, it is of nonnegative second fundamental form for the first Neumann eigenvalue and the mean curvature of the boundary is nonnegative for the first Dirichlet eigenvalue.

1. Introduction.

For a smooth n-dimensional closed Riemannian manifold M^n whose Ricci curvature satisfies

$$\mathrm{Ric}(M^n) \geq (n-1)K > 0 \tag{1.1}$$

for some positive constant K, it has been shown by A. Lichnerowicz [**6**] in 1958 (see also [**7**]) that the first positive eigenvalue λ of the manifold M has a lower bound

$$\lambda \geq nK. \tag{1.2}$$

The aim of this paper is to give some new lower bound estimates in terms of the lower Ricci curvature bound $(n-1)K$ and the diameter d of the manifold M. The main results of this paper are summarized in the following two theorems.

Theorem 1. *Let M^n be a closed Riemannian manifold with* $\mathrm{Ric}(M^n) \geq (n-1)K \geq 0$ *and diameter d. Then the first positive eigenvalue λ on M^n satisfies the lower bound*

$$\lambda \geq \frac{1}{4}(n-1)K + \frac{\pi^2}{d^2}. \tag{1.3}$$

Theorem 2. *Let M^n be a compact manifold with nonempty boundary and with* $\mathrm{Ric}(M^n) \geq (n-1)K \geq 0$.

(a) *Assume that the boundary ∂M is weakly convex, that is, the second fundamental form with respect to the outward normal is nonnegative. Then the first positive Neumann eigenvalue λ on M^n satisfies the same lower bound (1.3).*

(b) *Assume that the mean curvature with respect to the outward normal of the boundary ∂M is nonnegative. Then the first positive Dirichlet eigenvalue λ on M^n satisfies the lower bound estimate*

$$\lambda \geq \frac{1}{4}\left\{(n-1)K + \frac{\pi^2}{r^2}\right\} \tag{1.4}$$

where r is the inscribed radius for M.

These results generalize the Li-Yau [**5**] and Yang-Zhong [**11**] (cf. [**4**], [**9**]) estimates where they proved that the first positive eigenvalue satisfies $\lambda \geq \frac{\pi^2}{d^2}$ for closed manifolds with nonnegative Ricci curvature. Notice that for manifolds with small diameter, Theorem 1 is better than the estimate (1.2) by A. Lichnerowicz. P. Li has conjectured that the first positive eigenvalue should satisfy the lower bound

$$\lambda \geq (n-1)K + \frac{\pi^2}{d^2}. \tag{1.5}$$

A proof of this conjecture would unify the Li-Yau and Yang-Zhong estimate for manifolds with nonnegative Ricci curvature with the Lichnerowicz estimate (1.1). It is my pleasure to thank Professor P. Li for raising to me this interesting problem.

Theorems 1 and 2 follow from Theorems 4.1 and 4.2, which are more precise statements of our results. Our ideas are based on the gradient estimate technique for eigenfunctions which was developed by P. Li and S.T. Yau [**3**], [**5**]. Some preliminary lemmas are proved in Section 2 and the gradient estimates of eigenfunctions are presented in Section 3. These estimates introduce a higher order term associated with the positive lower bound on the Ricci curvature in the gradient estimates of Li-Yau and Yang-Zhong. When $K = 0$, our gradient estimates reduce to the estimates derived by Li-Yau and Yang-Zhong. The proof of Theorems 4.1 and 4.2 are presented in Section 4.

2. Some Preliminary Lemmas.

Throughout this paper, M will be a compact n-dimensional Riemannian manifold with or without smooth boundary with $\mathrm{Ric}(M) \geq (n-1)K \geq 0$.

Let v be a normalized eigenfunction of a positive eigenvalue λ on M with either Dirichlet or Neumann boundary condition if $\partial M \neq \emptyset$, that is,

$$\Delta v = -\lambda v, \tag{2.1}$$

such that

$$\min v(x) \geq -1 \quad \text{and} \quad \max v(x) = 1 \tag{2.2}$$

if $\partial M \neq \emptyset$ and v is a Dirichlet eigenfunction, and

$$\min v(x) = m_1 - 1 \quad \text{and} \quad \max v(x) = m_1 + 1 \tag{2.3}$$

if $\partial M \neq \emptyset$ and v is a Neumann eigenfunction or $\partial M = \emptyset$, where $0 \leq m_1 < 1$ is the median of v.

To give a unified presentation of the three different cases, we set $m_1 = 0$ if $\partial M \neq \emptyset$ and v is a Dirichlet eigenfunction. Let $u = v - m_1$. Given a constant $0 < s < 1$, consider the function

$$w = su = sv - m \tag{2.4}$$

where $0 \leq m = sm_1 < s < 1$. Thus $\max |w(x)| = s < 1$ and $\lim_{s \to 1^-} w = u$.

Lemma 2.1. *Let $h(t)$ be a smooth positive function defined on the open interval $(-1, 1)$. Assume that $\mathrm{Ric}(M) \geq (n-1)K$ and there is a point $p \in M$ such that the smooth function*

$$H = |\nabla w|^2 - 2h(w) \tag{2.5}$$

satisfies the conditions that

$$H(p) = \max H(x) = 0 \quad \textit{and} \quad \nabla H(p) = 0. \tag{2.6}$$

Then, at $t = w(p)$, the function h satisfies the inequality

$$h'^2 + \lambda(t+m)h' + 2h\{(n-1)K - \lambda - h''\} \leq 0. \tag{2.7}$$

Proof. Let $t = w(p)$. Since $|w| \leq s < 1$, $H(p) = 0$, and h is a positive smooth function on $(-1, 1)$, we have

$$|\nabla w(p)|^2 = 2h(t) > 0. \tag{2.8}$$

Choose a normal orthonormal frame $e_1, e_2, \ldots, e_n$ on a neighborhood of p such that $e_i w(p) = 0$ for $i > 1$. For any smooth function f, we shall adopt the notation that $f_i = e_i f(p)$ and $f_{ij} = e_j e_i f(p)$ for $i, j = 1, 2, \ldots, n$. Then $w_i = 0$ for $i > 1$ and (2.8) implies that $w_1^2 = 2h(t) > 0$. Since $\nabla H(p) = 0$, we have

$$0 = H_j = 2\sum_{i=1}^{n} w_i w_{ij} - 2h'(t)w_j = 2(w_1 w_{1j} - h'(t)w_j) \tag{2.9}$$

for $j = 1, 2, \ldots, n$. In particular,

$$w_{11} = h'(t). \tag{2.10}$$

Since H attains its maximum at p and $\nabla H(p) = 0$, the maximum principle applies to give

$$\Delta H(p) \leq 0. \tag{2.11}$$

It follows from Bochner's formula

$$\frac{1}{2}\Delta|\nabla w|^2 = \sum_{i,j=1}^{n} w_{ij}^2 + \nabla w \cdot \nabla(\Delta w) + \mathrm{Ric}(\nabla w, \nabla w), \tag{2.12}$$

(2.8), (2.10), (2.11), and $\Delta w = s\Delta v = -\lambda(w+m)$ that

$$\begin{aligned} 0 &\geq \frac{1}{2}\Delta H(p) \\ &= \sum_{ij=1}^{n} w_{ij}^2 - \lambda|\nabla w|^2 + \mathrm{Ric}(\nabla w, \nabla w) - \{h'(t)\Delta w + h''(t)|\nabla w|^2\} \\ &\geq h'^2(t) + \lambda(t+m)h'(t) + 2h(t)\{(n-1)K - \lambda - h''(t)\}. \end{aligned}$$

If the maximum value of H is attained at an interior point p in M, the condition that $\nabla H(p) = 0$ in Lemma 2.1 is automatically satisfied. When $\partial M \neq \emptyset$ and $p \in \partial M$, the following lemma assures that $\nabla H(p) = 0$ remains to be true if suitable convexity conditions are imposed on the boundary. Thus the maximum principle still applies even the maximum value of H is attained on the boundary.

Lemma 2.2. *Let $h(t)$ be a smooth positive function on the open interval $(-1,1)$, Assume that $\partial M \neq \emptyset$ and the maximum value 0 of the function (2.5) is attained at a boundary point $p \in \partial M$. Then $\nabla H(p) = 0$ in either of the following two situations.*

(a) *v is a Neumann eigenfunction and ∂M is weakly convex in the sense that the second fundamental form S in the outward normal direction is nonnegative definite.*
(b) *v is a Dirichlet eigenfunction, ∂M has nonnegative mean curvature $tr\, S \geq 0$ in the outward normal direction, and $h(t)$ is an even function.*

Proof. Let $p \in \partial M$ and $H(p) = \max H(x) = 0$. We first consider the case where v is a Neumann eigenfunction and assume that ∂M is weakly convex. Let e_n be the unit outward normal vector field on ∂M. Then $e_n w = 0$ on ∂M since $w = sv - m$ and v is a Neumann eigenfunction. Let $e_1, e_2, \ldots, e_{n-1}$ be a local orthonormal frame tangent to ∂M on a neighborhood of p in ∂M such that $w_i = e_i w(p) = 0$ for $i > 1$. Extend $e_1, e_2, \ldots, e_n$ to an orthonormal frame in a neighborhood of p in M by parallel translation along the geodesics $\exp_{\partial M}(te_n)$. Thus $e_n = \frac{d}{dt}\exp_{\partial M}(te_n)$ and $D_{e_n} e_i = 0$ for $i = 1, 2, \cdots, n$, where D is the covariant differential operator of the Riemannian manifold M. Moreover

$$\nabla w(p) = w_1 e_1(p) \neq 0 \tag{2.13}$$

since $|\nabla w(p)|^2 = 2h(w(p)) > 0$ and $w_i = 0$ for $i > 1$. Since ∂M is smooth, $e_n w(p) = 0$, and H has a maximum at $p \in \partial M$, we have $H_i = e_i H(p) = 0$

for $i < n$ and

$$
\begin{aligned}
(2.14) \quad 0 \le H_n &= e_n|\nabla w(p)|^2 = 2w_1 e_n e_1 w(p) \\
&= 2w_1\{e_1 e_n w(p) + (D_{e_n} e_1)w(p) - (D_{e_1} e_n)w(p)\} \\
&= -2w_1^2 \langle D_{e_1} e_n, e_1\rangle
\end{aligned}
$$

where the last equality follows from the facts that $e_n w = 0$ on ∂M, e_1 is tangent to ∂M, $D_{e_n} e_1 = 0$, and $w_i = 0$ for $i > 1$.

On the other hand, since ∂M is weakly convex, that is, the second fundamental form S satisfies $S(V,V) = \langle D_V e_n, V\rangle \ge 0$ for all tangent vector V to ∂M, we obtain

$$
(2.15) \qquad 0 \le H_n = -2w_1^2 S(e_1, e_1) \le 0.
$$

Hence $H_n = 0$ and $\nabla H(p) = \sum_{i=1}^n H_i e_i(p) = 0$.

Now let h be an even function and let v be a Dirichlet eigenfunction on M with nonnegative mean curvature $\operatorname{tr} S \ge 0$. Extend e_n to a local orthonormal frame $e_1, e_2, \dots, e_n$ on a neighborhood of p in M such that $D_{e_n} e_i = 0$ for $i = 1, 2, \dots, n$. Recall that for Dirichlet boundary condition, we have $m_1 = 0$, thus $w = sv$ and $w|_{\partial M} = sv_{\partial M} = 0$. Therefore $e_i w|_{\partial M} = 0$ for $i < n$ and $\nabla h(w)|_{\partial M} = 0$ since h is an even function. Since H attains its maximum value at $p \in \partial M$, we have $H_i = 0$ for $i < n$ and

$$
(2.16) \qquad 0 \le H_n = e_n|\nabla w(p)|^2 = 2\sum_{i=1}^n w_i e_n e_i w(p) = 2w_n e_n^2 w(p).
$$

Since $w(p) = 0$ and $D_{e_n} e_n = 0$, it follows from the definition of the Laplace operator that

$$
\begin{aligned}
(2.17) \qquad 0 &= -\lambda w(p) = \Delta w(p) = \sum_{i=1}^n (e_i^2 - D_{e_i} e_i) w(p) \\
&= e_n^2 w(p) + \overline{\Delta} w(p) - w_n \sum_{i=1}^{n-1} \langle e_n, D_{e_i} e_i\rangle
\end{aligned}
$$

where $\overline{\Delta}$ is the Laplace operator on ∂M with the induced Riemannian metric. Since $w|_{\partial M} = 0$ and the mean curvature

$$
(2.18) \qquad \operatorname{tr} S = \sum_{i=1}^{n-1} \langle D_{e_i} e_n, e_i\rangle = -\sum_{i=1}^{n-1} \langle e_n, D_{e_i} e_i\rangle
$$

is nonnegative, we obtain $\overline{\Delta} w(p) = 0$ and

$$
(2.19) \qquad 0 \le H_n = 2w_n e_n^2 w(p) = -2w_n^2 \operatorname{tr} S(p) \le 0.
$$

Therefore $H_n = 0$ and $\nabla H(p) = 0$.

We shall also need the following lower bound estimate of the first eigenvalue, which is due to A. Lichnerowicz [**6**] when M is a compact manifold without boundary. For completeness sake, a proof is enclosed.

Lemma 2.3. *Assume that* $\mathrm{Ric}(M) \geq (n-1)K > 0$*. Let* λ *be the first positive eigenvalue on* M *(with either Dirichlet or Neumann boundary condition if* $\partial M \neq \emptyset$*). If* $\partial M \neq \emptyset$*, we also assume that* ∂M *is of nonnegative mean curvature* $\mathrm{tr}\, S \geq 0$ *if* λ *is a Dirichlet eigenvalue and* ∂M *is of nonnegative definite second fundamental form* $S \geq 0$ *if* λ *is a Neumann eigenvalue. Then*

$$\lambda \geq nK. \tag{2.20}$$

Proof. Let v be an eigenfunction of the eigenvalue λ. The lower bound (2.20) follows from integrating Bochner's formula for ∇v

$$\frac{1}{2}\Delta|\nabla v|^2 = \sum_{i,j=1}^{n} v_{ij}^2 - \lambda|\nabla v|^2 + \mathrm{Ric}(\nabla v, \nabla v) \tag{2.21}$$

on M and applying the boundary conditions. More specifically, using the Schwarz inequality

$$\sum_{i,j=1}^{n} v_{ij}^2 \geq \sum_{i=1}^{n} v_{ii}^2 \geq \frac{1}{n}(\Delta v)^2 = \frac{1}{n}\lambda^2 v^2, \tag{2.22}$$

and the lower bound on the Ricci curvature, integrate (2.21) over M yields

$$\frac{1}{2}\int_{\partial M} e_n|\nabla v|^2 = \frac{1}{2}\int_M \Delta|\nabla v|^2 \geq \int_M \left\{\frac{1}{n}\lambda^2 v^2 + [(n-1) - \lambda]|\nabla v|^2\right\}. \tag{2.23}$$

Since $\Delta v = -\lambda v$, multiply by v and integrate over M and use the boundary conditions yield $\int_M |\nabla v|^2 = \lambda \int_M v^2$. Hence

$$\frac{1}{2}\int_{\partial M} e_n|\nabla v|^2 \geq \frac{n-1}{n}\lambda(nK - \lambda)\int_M v^2. \tag{2.24}$$

If $\partial M = \emptyset$, then (2.20) follows from (2.24) immediately. Otherwise, we show that $e_n|\nabla v|^2 \leq 0$ pointwisely on ∂M for either of the two boundary conditions. Indeed, for any $p \in \partial M$, choose an orthonormal frame $e_1, e_2, \ldots, e_n$ as in the proof of Lemma 2.2. For Neumann boundary condition, similar computations as in (2.14), (2.15), and the convexity condition $S \geq 0$ yield

$$e_n|\nabla v|^2 = -2v_1^2 S(e_1, e_1) \leq 0. \tag{2.25}$$

For Dirichlet boundary condition, similar computations as in (2.17), (2.18), (2.19), and $\mathrm{tr}\, S \geq 0$ yield

$$e_n|\nabla v|^2 = -2v_n^2 \mathrm{tr}\, S \leq 0. \tag{2.26}$$

In any case, we have

$$\int_{\partial M} e_n|\nabla v|^2 \leq 0. \tag{2.27}$$

Thus the lower bound estimate (2.20) follows from (2.24) and (2.27).

Notice that if we have a test function h which satisfies the conditions in Lemma 2.1, then we get a gradient estimate

$$|\nabla w|^2 \leq 2h(w). \tag{2.28}$$

To construct a suitable test function h, the following function z, which was introduced by H.C. Yang and J.Q. Zhong [**11**] to estimate the first eigenvalue for manifolds with nonnegative Ricci curvature, is especially useful.

Lemma 2.4. *The function*

$$z(t) = \frac{2}{\pi}\left(\arcsin t + t\sqrt{1-t^2}\right) - t \tag{2.29}$$

is a continuous odd function on $[-1,1]$*. Furthermore, on the open interval* $(-1,1)$*,* z *is smooth and satisfies*

$$(1-t^2)z'' + tz' + t = 0, \tag{2.30}$$

$$\frac{2}{5}t^2(1-t^2) \leq |z(t)| < \frac{1}{4}(1-t^4), \tag{2.31}$$

$$z'^2 - 2zz'' > \frac{1}{4}(t - tz' + 2z)^2, \tag{2.32}$$

(2.33)
$$2(1-t^2)(3+t^2)(z'^2 - 2zz'' + z') > \left\{6tz + (1-t^2)\left(\frac{6}{\pi}\sqrt{1-t^2} - 1\right)\right\}^2.$$

Proof. It follows from the definition (2.29) for $z(t)$ that

$$z'(t) = \frac{4}{\pi}\sqrt{1-t^2} - 1, \tag{2.34}$$

$$z''(t) = -\frac{4}{\pi}t(1-t^2)^{-1/2}. \tag{2.35}$$

Thus the identity (2.30) is clearly true. Furthermore, we have

$$t - tz' + 2z = \frac{4}{\pi}\arcsin t, \tag{2.36}$$

(2.37)
$$z'^2 - 2zz'' = (1-t^2)^{-1/2}\left\{(1-t^2)^{1/2}\left(1 + \frac{16}{\pi^2}\right) - \frac{8}{\pi} + \frac{16}{\pi^2}t\arcsin t\right\},$$

$$z'^2 - 2zz'' + z' = 4(1-t^2)^{-1/2}\left\{\frac{4}{\pi^2}\left[(1-t^2)^{1/2} + t\arcsin t\right] - \frac{1}{\pi}(1+t^2)\right\}. \tag{2.38}$$

For the inequalities, we first notice that z is an odd function. Hence, all of the functions involved in the inequalities are even functions. Therefore, we need only to verify them on the interval $[0,1)$.

Let

$$\phi(t) = z(t) - \frac{2}{5}t^2(1-t^2), \tag{2.39}$$

$$\phi_1(t) = 1 - t^4 - 4z(t), \tag{2.40}$$

$$\phi_2(t) = (1-t^2)^{1/2}\left(1+\frac{16}{\pi^2}\right) - \frac{8}{\pi} + \frac{16}{\pi^2}t\arcsin t - \frac{4}{\pi^2}(1-t^2)^{1/2}(\arcsin t)^2, \tag{2.41}$$

$$\begin{aligned}\phi_3(t) &= 8(1-t^2)^{1/2}(3+t^2)\left\{\frac{4}{\pi^2}\left[(1-t^2)^{1/2} + t\arcsin t\right] - \frac{1}{\pi}(1+t^2)\right\} \\ &\quad - \left\{\frac{12}{\pi}t\arcsin t + \frac{6}{\pi}(1-t^2)^{1/2}(1+t^2) - 1 - 5t^2\right\}^2. \end{aligned}\tag{2.42}$$

Then the inequalities (2.31), (2.32), and (2.33) are equivalent to $\phi \geq 0$ and $\phi_i > 0$ for $i = 1,2,3$ on $[0,1)$. Since all of the functions are explicit elementary functions, it is easy to give a rigorous proof of these inequalities. However, it will take a few pages to do so. Instead, it is a much simpler matter to combine calculus with a graphing utility to verify these inequalities. The details will therefore be left to the readers.

3. Gradient Estimates of Eigenfunctions.

In this section, we prove the following gradient estimates.

Theorem 3.1. *Let M be a compact n-dimensional Riemannian manifold without boundary with Ric$(M) \geq (n-1)K \geq 0$. Let v be a normalized eigenfunction on M with median m_1 of a positive eigenvalue λ. Let $u = v - m_1$, $a = \frac{(n-1)K}{2\lambda}$, and let z be the function defined by (2.29). Then, the gradient of u satisfies the inequality*

$$|\nabla u|^2 \leq \lambda\{(1-u^2)[1-a(1-u^2)] + 2m_1 z(u)\}. \tag{3.1}$$

Theorem 3.2. *Let M be a compact n-dimensional Riemannian manifold with nonempty boundary and with* $\mathrm{Ric}(M) \geq (n-1)K \geq 0$ *and let $a = \frac{(n-1)K}{2\lambda}$.*

(a) *Assume that the boundary is weakly convex, that is, the second fundamental form S in the outward normal direction is nonnegative definite. Let v be a normalized Neumann eigenfunction on M with median m_1 of a positive eigenvalue λ. Then, the gradient of $u = v - m_1$ satisfies the same inequality* (3.1).

(b) *Assume that the boundary is of nonnegative mean curvature $trS \geq 0$ in the outward normal direction. Let v be a normalized Dirichlet eigenfunction on M of a positive eigenvalue λ. Then the gradient of v satisfies the inequality*

$$|\nabla v|^2 \leq \lambda(1-v^2)\{1-a(1-v^2)\}. \tag{3.2}$$

Notice that since $\lim_{s\to 1^-} m = m_1$ and $\lim_{s\to 1^-} w = v - m_1$ ($m_1 = 0$ for Dirichlet eigenfunction), to show the gradient estimates (3.1) and (3.2), it suffices to show the corresponding estimates for w. We shall use Lemma 2.1 twice. First, we show a gradient estimate for w in Lemma 3.3 which is a slight variation of the Yang-Zhong [**11**] estimate for compact manifolds without boundary (see also [**4**] and [**9**]).

Lemma 3.3. *Assume that* $\mathrm{Ric}(M) \geq 0$*. If $\partial M \neq \emptyset$, we also assume that either the second fundamental form S is nonnegative definite if v is a Neumann eigenfunction or the mean curvature* $\mathrm{tr}\, S$ *is nonnegative if v is a Dirichlet eigenfunction. Let $w = sv - m$ be as in Section* 2*. Then, for all* $0 < s < 1$*, the gradient of w satisfies the inequality*

$$|\nabla w|^2 \leq \lambda(1 - w^2 + 2mz(w)). \tag{3.3}$$

Proof. Since $|w| \leq s < 1$ and $0 \leq m = sm_1 < 1$, the inequality (2.31) implies that $1 - w^2 + 2mz(w)$ is a positive smooth function on M. Thus, there exists a positive constant β such that the smooth function

$$Q = |\nabla w|^2 - \beta(1 - w^2 + 2mz(w)) \tag{3.4}$$

has 0 as its maximum value. Thus, the inequality (3.3) will follow if $\beta \leq \lambda$.

Let

$$h(t) = \frac{\beta}{2}(1 - t^2 + 2mz(t)). \tag{3.5}$$

Notice that if v is a Dirichlet eigenfunction, then $m = sm_1 = 0$ and h is an even function. Let $p \in M$ be a point where the function Q attains its maximum value 0. The convexity conditions $S \geq 0$ or $\mathrm{tr}\, S \geq 0$ and Lemma 2.2 implies that $\nabla Q(p) = 0$. It follows from Lemma 2.1 that, at $t = w(p) \in (-1, 1)$, the function h defined by (3.5) satisfies the inequality (2.7) with $K = 0$, namely,

$$0 \geq h'^2 + \lambda(t + m)h' - 2h(\lambda + h''). \tag{3.6}$$

Since $h' = \beta(mz' - t)$, $h'' = \beta(mz'' - 1)$, and $\beta > 0$, divide the inequality (3.6) by β and simplify yield

$$\begin{aligned} 0 \ \geq \ & (\beta-\lambda)\left\{1 + m(t+2z-tz') + m^2(z'^2 - 2zz'')\right\} \\ & +\lambda m^2\left\{z'^2 - 2zz'' + z'\right\} - m\beta\{(1-t^2)z'' + tz' + t\}. \end{aligned} \tag{3.7}$$

The last term is 0 because of the identity (2.30). Completing the square in the first term yields

$$\begin{aligned} 0 \ \geq \ & (\beta-\lambda)\left\{1 + \frac{m}{2}(t+2z-tz')\right\}^2 + \lambda m^2\{z'^2 - 2zz'' + z'\} \\ & +(\beta-\lambda)m^2\left\{z'^2 - 2zz'' - \frac{1}{4}(t+2z-tz')^2\right\}. \end{aligned} \tag{3.8}$$

If $\beta > \lambda$, then, it follows from Lemma 2.4 that all of the three terms on the right side of the inequality (3.8) is nonnegative. Moreover, the first term is positive if $m = 0$ and the last two terms are both positive if $m \neq 0$. That is certainly not possible since the left side of the inequality (3.8) is 0. Hence, we must have $\beta \leq \lambda$.

The rest of this section will be devoted to the proof of Theorem 3.1 and 3.2. If $K = 0$, Theorem 3.1 and 3.2 follows from Lemma 3.3. So assume that $\mathrm{Ric}(M) \geq K > 0$. As already been noticed, we need only show that there exists a constant $\alpha \geq a = \frac{(n-1)K}{2\lambda}$ such that w satisfies the inequality

$$|\nabla w|^2 \leq \lambda\{(1-w^2)[1-\alpha(1-w^2)] + 2mz(w)\}. \tag{3.9}$$

It follows from Lemma 3.3 that there exists a nonnegative constant α such that the function

$$G = |\nabla w|^2 - \lambda\{(1-w^2)[1-\alpha(1-w^2)] + 2mz(w)\} \tag{3.10}$$

has 0 as its maximum value since G is a strictly increasing linear function in α and

$$G = |\nabla w|^2 - \lambda(1 - w^2 + 2mz(w)) + \lambda\alpha(1-w^2)^2 \leq \lambda\alpha(1-s^2)^2 < 0 \tag{3.11}$$

if $\alpha < 0$.

Suppose, on the contrary, that $\alpha < a$. By Lemma 2.3, we have $\lambda \geq nK$. Thus

$$0 \leq \alpha < a = \frac{(n-1)K}{2\lambda} \leq \frac{n-1}{2n} < \frac{1}{2}. \tag{3.12}$$

It follows from the inequality (2.31) and (3.12) that the new test function

$$h(t) = \frac{\lambda}{2}\{(1-t^2)[1-\alpha(1-t^2)] + 2mz(t)\} \tag{3.13}$$

is a positive smooth function on $(-1, 1)$.

Let $p \in M$ be a point where the smooth function G attains its maximum value 0. As in the proof of Lemma 3.3, the convexity condition $S \geq 0$ or $\operatorname{tr} S \geq 0$ and Lemma 2.2 implies that $\nabla G(p) = 0$. It follows from Lemma 2.1 that, at $t = w(p) \in (-1, 1)$, the function defined by (3.13) satisfies the inequality

$$h'^2 + \lambda(t+m)h' + 2h\{(n-1)K - \lambda - h''\} \leq 0. \tag{3.14}$$

Since

$$h'(t) = \lambda\{t[2\alpha(1-t^2) - 1] + mz'(t)\}, \tag{3.15}$$

$$h''(t) = \lambda\{2\alpha - 1 - 6\alpha t^2 + mz''(t)\}, \tag{3.16}$$

and $(n-1)K = 2a\lambda$, divide the inequality (3.14) by λ^2 and then simplify it using the identity (2.30) yield

$$\begin{aligned} 0 \geq\ & 2(a-\alpha)\{(1-t^2)[1-\alpha(1-t^2)] + 2mz\} \\ & +2\alpha t^2(1-t^2)[2-\alpha(1-t^2)] \\ & +2m\alpha t\left\{6tz + (1-t^2)\left[\frac{6}{\pi}\sqrt{1-t^2} - 1\right]\right\} + m^2(z'^2 - 2zz'' + z'). \end{aligned} \tag{3.17}$$

It follows from the inequalities (2.31), (2.33), and (3.12) that the first term on the right side of the inequality (3.17) is positive while the second and the fourth terms are nonnegative, thus $m > 0$ and $\alpha > 0$. Furthermore, it follows from $1 > 2a > 2\alpha > 0$, (2.31), and (3.17) that

$$\begin{aligned} 0 &> \alpha\Big\{t^2(1-t^2)(3+t^2) + 2mt\left[6tz + (1-t^2)\left(\frac{6}{\pi}\sqrt{1-t^2} - 1\right)\right] \\ &\qquad + 2m^2(z'^2 - 2zz'' + z')\Big\} \\ &\geq \alpha(1-t^2)^{-1}(3+t^2)^{-1}\Big\{t(1-t^2)(3+t^2) \\ &\qquad + m\left[6tz + (1-t^2)\left(\frac{6}{\pi}\sqrt{1-t^2} - 1\right)\right]\Big\}^2 \\ &\quad + m^2\alpha\Big\{2(z'^2 - 2zz'' + z') \\ &\qquad - (1-t^2)^{-1}(3+t^2)^{-1}\left[6tz + (1-t^2)\left(\frac{6}{\pi}\sqrt{1-t^2} - 1\right)\right]^2\Big\}. \end{aligned} \tag{3.18}$$

Since $m > 0$ and $\alpha > 0$, the inequality (3.18) apparently contradicts with the inequality (2.33) in Lemma 2.4. Therefore, we have proved that there exists a constant $\alpha \geq a$ such that w satisfies the gradient estimate (3.9) for each constant $0 < s < 1$. Taking the limit to the inequality (3.9) by letting $s \to 1^-$ now yields the inequalities (3.1) and (3.2). This completes the proof of Theorems 3.1 and 3.2.

4. Lower Bound Estimates of the First Positive Eigenvalue.

In this section, we apply the gradient estimates obtained in the previous section to derive some new lower bound estimates of the first positive eigenvalue on compact Riemannian manifolds whose Ricci curvature satisfies $\mathrm{Ric}(M) \geq (n-1)K \geq 0$.

Theorem 4.1. *Let M be a compact n-dimensional Riemannian manifold without boundary whose Ricci curvature satisfies $\mathrm{Ric}(M) \geq (n-1)K \geq 0$. Let d be the diameter of M. Let v be the normalized eigenfunction of the first positive eigenvalue λ so that*

$$\inf v(x) = m_1 - 1 \quad \textit{and} \quad \max v(x) = m_1 + 1 \tag{4.1}$$

where $0 \leq m_1 < 1$ is the median of v. Then

$$\begin{aligned}\lambda \geq \min\Bigg\{&(n-1)K + \frac{\pi^2}{d^2}, (n-1)K/4 + \frac{\pi^2}{d^2}\left[1+0.09m_1^2\right]^2\Bigg[1 \\ &+ \sum_{k=1}^{\infty} \frac{(4k-1)!!(2k-1)!!}{(4k)!!(2k)!!}\left(\frac{(n-1)Kd^2}{4\pi^2+3(n-1)Kd^2}\right)^{2k}\Bigg]^2\Bigg\}.\end{aligned} \tag{4.2}$$

Proof. Let $u = v - m_1$. It follows from Theorem 3.1 that

$$|\nabla u|^2 \leq \lambda\{(1-u^2)[1-a(1-u^2)] + 2m_1 z(u)\} \tag{4.3}$$

where $a = \frac{(n-1)K}{2\lambda}$. Hence

$$\frac{|\nabla u|}{(1-u^2)^{1/2}[1-a(1-u^2)]^{1/2}[1+2m_1 z(u)(1-u^2)^{-1}[1-a(1-u^2)]^{-1}]^{1/2}} \leq \lambda^{1/2}. \tag{4.4}$$

By (4.1), there exist two points p, $q \in M$ such that

$$u(p) = -1 \quad \text{and} \quad u(q) = 1. \tag{4.5}$$

Let $\gamma(t)$ be a minimal geodesic from p to q in M and let

$$\theta(x) = \arcsin u(x) \in [-\pi/2, \pi/2]. \tag{4.6}$$

Integrate (4.4) along γ yields

$$\begin{aligned}\lambda^{1/2}d &\geq \int_\gamma \lambda^{1/2}dt \geq \int_{-1}^{1}(1-u^2)^{-1/2}[1-a(1-u^2)]^{-1/2}\{1 \\ &\quad + 2m_1 z(u)(1-u^2)^{-1}[1-a(1-u^2)]^{-1}\}^{-1/2}du \\ &= \int_{-\pi/2}^{\pi/2}\left[1-\frac{a}{2}-\frac{a}{2}\cos 2\theta\right]^{-1/2}\left\{1+2m_1 z(\sin\theta)\sec^2\theta\left[1\right.\right. \\ &\quad \left.\left. -\frac{a}{2}-\frac{a}{2}\cos 2\theta\right]^{-1}\right\}^{-1/2}d\theta. \end{aligned} \tag{4.7}$$

Let $b=\frac{a}{2-a}=\frac{(n-1)K}{4\lambda-(n-1)K}$. Then $0\leq b<1/3$ since $\lambda\geq nK$. Thus

$$\left[1-\frac{a}{2}-\frac{a}{2}\cos 2\theta\right]^{-1/2}=(1-a/2)^{-1/2}(1-b\cos 2\theta)^{-1/2}\geq 1. \tag{4.8}$$

The inequality (2.31) in Lemma 2.4 implies that

$$2m_1|z(\sin\theta)|\sec^2\theta\left[1-\frac{a}{2}-\frac{a}{2}\cos 2\theta\right]^{-1}\leq m_1<1 \tag{4.9}$$

since $a<1/2$.

So we can apply the binomial series expansion

$$(1-y)^{-1/2}=1+\sum_{k=1}^{\infty}\frac{(2k-1)!!}{(2k)!!}y^k \tag{4.10}$$

for

$$y=-2m_1 z(\sin\theta)\sec^2\theta\left[1-\frac{a}{2}-\frac{a}{2}\cos 2\theta\right]^{-1} \tag{4.11}$$

and notice that (4.11) is an odd function in θ. It follows from (4.7), (4.8), (4.10), and (4.11) that

$$\begin{aligned}\lambda^{1/2}d &\geq (1-a/2)^{-1/2}\int_{-\pi/2}^{\pi/2}[1-b\cos 2\theta]^{-1/2}\left[1\right. \\ &\quad \left. +\sum_{k=1}^{\infty}\frac{(2k-1)!!}{(2k)!!}y^k\right]d\theta \\ &= 2(1-a/2)^{-1/2}\int_0^{\pi/2}[1-b\cos 2\theta]^{-1/2}\left[1\right. \\ &\quad \left. +\sum_{k=1}^{\infty}\frac{(4k-1)!!}{(4k)!!}y^{2k}\right]d\theta \\ &\geq 2(1-a/2)^{-1/2}\int_0^{\pi/2}[1-b\cos 2\theta]^{-1/2}\left[1+\frac{3}{8}y^2\right]d\theta. \end{aligned} \tag{4.12}$$

By the inequalities (2.31) and (4.8), we have

$$y^2 \geq \left[\frac{4}{5}m_1 \sin^2\theta\right]^2 = \frac{16}{25}m_1^2 \sin^4\theta. \tag{4.13}$$

Now expand $(1-b\cos 2\theta)^{-1/2}$ in (4.12) and integrate term by term yield

$$\begin{aligned} &d\lambda^{1/2}\left(1-\frac{a}{2}\right)^{1/2} \\ &\geq 2\int_0^{\pi/2}\left(1+\frac{6}{25}m_1^2\sin^4\theta\right)(1-b\cos 2\theta)^{-1/2}d\theta \\ &= 2\int_0^{\pi/2}\left(1+\frac{6}{25}m_1^2\sin^4\theta\right)\left[1+\sum_{k=1}^{\infty}\frac{(2k-1)!!}{(2k)!!}b^k\cos^k 2\theta\right]d\theta \\ &\geq \pi(1+0.09m_1^2)\left[1+\sum_{k=1}^{\infty}\frac{(4k-1)!!(2k-1)!!}{(4k)!!(2k)!!}b^{2k}\right]. \end{aligned} \tag{4.14}$$

Since $a = \frac{(n-1)K}{2\lambda}$, we obtain

$$\lambda \geq \frac{n-1}{4}K + \frac{\pi^2}{d^2}(1+0.09m_1^2)^2\left[1+\sum_{k=1}^{\infty}\frac{(4k-1)!!(2k-1)!!}{(4k)!!(2k)!!}b^{2k}\right]^2 \tag{4.15}$$

where $b = \frac{(n-1)K}{4\lambda-(n-1)K}$.

So either $\lambda \geq \frac{\pi^2}{d^2}+(n-1)K$ or else $b \geq \frac{(n-1)Kd^2}{4\pi^2+3(n-1)Kd^2}$ and

$$\begin{aligned} \lambda \geq{}& \frac{n-1}{4}K + \frac{\pi^2}{d^2}(1+0.09m_1^2)^2 \\ &\cdot\left[1+\sum_{k=1}^{\infty}\frac{(4k-1)!!(2k-1)!!}{(4k)!!(2k)!!}\left(\frac{(n-1)Kd^2}{4\pi^2+3(n-1)Kd^2}\right)^{2k}\right]^2. \end{aligned} \tag{4.16}$$

This completes the proof of Theorem 4.1.

Theorem 4.2. *Let M be a compact n-dimensional Riemannian manifold with nonempty boundary ∂M. Assume that the Ricci curvature satisfies* $\mathrm{Ric}(M) \geq (n-1)K \geq 0$. *Let $d = \mathrm{diam}(M)$ be the diameter of M and let $r = \sup\{d(x,\partial M) | x \in M\}$ be the inscribed radius of M. Then:*

(a) *If the second fundamental form of the boundary in the outward normal direction is nonnegative definite, then the first positive Neumann eigenvalue for M satisfies the same inequality (4.2).*
(b) *If the mean curvature of the boundary in the outward normal direction is nonnegative, then the first positive Dirichlet eigenvalue for M*

satisfies the lower bound

(4.17)
$$\lambda \geq \min\left\{(n-1)K + \frac{\pi^2}{4r^2}, \frac{n-1}{4}K + \frac{\pi^2}{4r^2}\left[1+\sum_{k=1}^{\infty}\frac{(4k-1)!!(2k-1)!!}{(4k)!!(2k)!!}\left(\frac{(n-1)Kr^2}{\pi^2+3(n-1)Kr^2}\right)^{2k}\right]^2\right\}.$$

Proof. Since the Neumann eigenfunction satisfies the same gradient estimate (3.1), the proof of the lower bound (4.2) for the first positive Neumann eigenvalue is identical with the proof of Theorem 4.1. The proof of the lower bound (4.17) for the first Dirichlet eigenvalue is also similar to the proof of Theorem 4.1.

Let v be the normalized first Dirichlet eigenfunction such that $0 \leq v \leq \max v(x) = 1$. Let $q \in M$ and $p \in \partial M$ be two points such that $v(q) = 1$ and $d(p,q) = d(q, \partial M)$. By the definition of the inscribed radius r, we have $d(p,q) \leq r$. It follows from the inequality (3.2) that

$$\frac{|\nabla v|}{(1-v^2)^{1/2}[1-a(1-v^2)]^{1/2}} \leq \lambda^{1/2}. \tag{4.18}$$

Integrate the inequality (4.18) along a minimal geodesic from p to q as in the proof of Theorem 4.1 yields the desired lower bound (4.17).

References

[1] J. Cheeger, *A lower bound for the smallest eigenvalue of the Laplacian*, Problems in Analysis, a symposium in honor of S. Bochner, Princeton Univ. Press, Princeton, 1970.

[2] J. Escobar, *Uniqueness theorems on conformal deformation of metrics, Sobolev inequalities, and an eigenvalue estimate.*

[3] P. Li, *A lower bound for the first eigenvalue of the Laplacian on a compact Riemannian manifold*, Indiana Univ. Math. J., **28** (1979), 1013-1019.

[4] ———, *Lecture Notes on Geometric Analysis*, Lecture Notes Series, #6, Seoul National University, Seoul, Korea.

[5] P. Li and S.T. Yau, *Eigenvalues of a compact Riemannian manifold*, AMS Proc. Symp. Pure Math., **36** (1980), 205-239.

[6] A. Lichnerowicz, *Géometrie des groupes de transformations*, Dunod, Paris, 1958.

[7] M. Obata, *Certain conditions for a Riemannian manifold to be isometric to the sphere*, J. Math. Soc. Japan, **14** (1962), 333-340.

[8] R. Reilly, *Applications of the Hessian operator in a Riemannian manifold*, Indiana Univ. Math. J., **26** (1977), 459-472.

[9] R. Schoen and S.T. Yau, *Lectures on Differential Geometry*, Conference Proceedings and Lecture Notes in Geometry and Topology, Vol. I, International Press, 1994.

[10] S.T. Yau, *Isoperimetric constants and the first eigenvalue of a compact Riemannian manifold*, Ann. Scient. Éc. Norm. Sup., **4** (1985), 487-507.

[11] H.C. Yang and J.Q. Zhong, *On the estimate of the first eigenvalue of a compact Riemannian manifold*, Sci. Sinica Ser., **A27** (1984), 1265-1273.

Received September 30, 1997. This research was partially supported by NSF grant 9404312.

Tulane University
New Orleans, LA 70118
E-mail address: dgy@math.tulane.edu

CONTENTS

Volume 190, no. 1 and no. 2

Guidelines for Authors

Authors may submit manuscripts at pjm.math.berkeley.edu/about/journal/submissions.html and choose an editor at that time. Exceptionally, a paper may be submitted in hard copy to one of the editors; authors should keep a copy.

By submitting a manuscript you assert that it is original and is not under consideration for publication elsewhere. Instructions on manuscript preparation are provided below. For further information, visit the web address above or write to pacific@math.berkeley.edu or to Pacific Journal of Mathematics, University of California, Los Angeles, CA 90095–1555. Correspondence by email is requested for convenience and speed.

Manuscripts must be in English, French or German. A brief abstract of about 150 words or less in English must be included. The abstract should be self-contained and not make any reference to the bibliography. Also required are keywords and subject classification for the article, and, for each author, postal address, affiliation (if appropriate) and email address if available. A home-page URL is optional.

Authors are encouraged to use LaTeX, but papers in other varieties of TeX, and exceptionally in other formats, are acceptable. At submission time only a PDF file is required; follow the instructions at the web address above. Carefully preserve all relevant files, such as LaTeX sources and individual files for each figure; you will be asked to submit them upon acceptance of the paper.

Bibliographical references should be listed alphabetically at the end of the paper. All references in the bibliography should be cited in the text. Use of BibTeX is preferred but not required. Any bibliographical citation style may be used but tags will be converted to the house format (see a current issue for examples).

Figures, whether prepared electronically or hand-drawn, must be of publication quality. Figures prepared electronically should be submitted in Encapsulated PostScript (EPS) or in a form that can be converted to EPS, such as GnuPlot, Maple or Mathematica. Many drawing tools such as Adobe Illustrator and Aldus FreeHand can produce EPS output. Figures containing bitmaps should be generated at the highest possible resolution. If there is doubt whether a particular figure is in an acceptable format, the authors should check with production by sending an email to pacific@math.berkeley.edu.

Each figure should be captioned and numbered, so that it can float. Small figures occupying no more than three lines of vertical space can be kept in the text ("the curve looks like this:"). It is acceptable to submit a manuscript will all figures at the end, if their placement is specified in the text by means of comments such as "Place Figure 1 here". The same considerations apply to tables, which should be used sparingly.

Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

PACIFIC JOURNAL OF MATHEMATICS

Volume 190 No. 2 October 1999
