View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Kyoto University Research Information Repository

Kyoto University Research Information Repository

KYOTO UNIVERSITY

Title	Diagonalization modulo norm ideals : spectral method and modulus of continuity (Spectral and Scattering Theory and Related Topics)
Author(s)	KURODA, S.T.
Citation	数理解析研究所講究録別冊 = RIMS Kokyuroku Bessatsu (2010), B16: 101-126
Issue Date	2010-04
URL	http://hdl.handle.net/2433/176843
Right	
Type	Departmental Bulletin Paper
Textversion	publisher

Kyoto University

RIMS Kôkyûroku Bessatsu
B16 (2010), 101–126

Diagonalization modulo norm ideals; spectral method and modulus of continuity

By

S. T. Kuroda*

Abstract

This is a half expository paper on whether a commutative N-tuple $\boldsymbol{A}$ of selfadjoint operators can be simultaneously diagonalized by perturbations belonging to a certain norm ideal. Along with $\boldsymbol{A}$, operators $f(\boldsymbol{A})$ defined by functional calculus are considered. After a short review of previous developments the spectral theoretical method given in Voigt [21] will be exploited. The method suit handling $f(\boldsymbol{A})$ well, though it cannot reach the best possible result for a tuple $\boldsymbol{A}$ by Voiculescu ([16], [2]). The result is expressed in terms of modulus of continuity of f, a résumé of which is also included (Section 3). Section 3 and Subsection 5.1 concerning a construction of CONS of Haar type can be read independently of other parts.

§ 1. Introduction

In the theory of diagonalization modulo norm ideals one asks if a given selfadjoint operator can be made diagonalizable by a perturbation belonging to a class of operators called norm ideals. A little more precisely, given a norm ideal $\boldsymbol{\mathcal{C}}$, one asks whether

$$\text{(D)} \quad \begin{cases} \text{for any selfadjoint operator } A \text{ there exists a selfadjoint } K \in \boldsymbol{\mathcal{C}} \text{ so that} \\ A + K \text{ has a pure point spectrum.} \end{cases}$$

This problem goes back to classical works of H. Weyl ([23, 1909]) and J. von Neumann ([11, 1935]). In particular, von Neumann showed that statement (D) holds for the Hilbert-Schmidt class $\boldsymbol{\mathcal{C}} = \boldsymbol{\mathcal{C}}_2$. Later, the author proved the same for any $\boldsymbol{\mathcal{C}} \neq \boldsymbol{\mathcal{C}}_1$, where $\boldsymbol{\mathcal{C}}_1$ is the trace class([9, 1958]). According to the famous Kato–Rosenblum theroem (D) does not hold for $\boldsymbol{\mathcal{C}}_1$. Thus, among norm ideals $\boldsymbol{\mathcal{C}}_p$ (for $\boldsymbol{\mathcal{C}}_p$ see Example 2.2) the borderline for being diagonalizable is $p = 1$.

Received May 11, 2009. Revised September 18, 2009.
2000 Mathematics Subject Classification(s): 26A15, 26B05, 47A10, 47A55, 47B10
Key Words: Modulus of continuity, Diagonalization, Norm ideal
*Emeritus, Gakushuin University and Tokyo University,
3-7-5-604 Ebisuminami, Shibuya-ku, Tokyo, 150-0022, Japan.
e-mail: `kuroda-s@khaki.plala.or.jp`
© 2010 Research Institute for Mathematical Sciences, Kyoto University. All rights reserved.

Following works for normal operators (cf. I. D. Berg ([3, 1971]), W. Sikonia ([14, 1971]), J. Weidmann ([22])) similar problems for commutative N-tuple of selfadjoint operators have been extensively investigated by J. Voigt ([21, 1977]), D. Voiculescu (e.g., [16, 1979]), and more recently by J. Xia (e.g., [24, 1997]). Let us call (D_N) a statement similar to (D) for N-tuples (cf. 2.2.2). Voigt proved that (D_N) holds for $\boldsymbol{\mathcal{C}}_p$ with $p > N$. This suggests that the borderline for the validity of (D_N) among $\boldsymbol{\mathcal{C}}_p$ may be $p \sim N$.[1] Voiculescu's theory showed that the borderline is in fact $\boldsymbol{\mathcal{C}}_N^-$ which is a little smaller than $\boldsymbol{\mathcal{C}}_N$ (for $\boldsymbol{\mathcal{C}}_N^-$ see Example 2.3). This is a decisive result.

The condition in (D) may be written as $A - (A + K) \in \boldsymbol{\mathcal{C}}$. Then, as the next step one may examine the condition of the type $f(A) - f(A + K) \in \boldsymbol{\mathcal{C}}$, where f belongs to a suitable class of functions. This is a simultaneous diagonalization with respect to f. One can formulate similar condition for N-tuples. Such problems were taken up by J. Xia ([24]) and examined for Lipschitz continuous f (cf. 2.2.2, 4.3).

Voiculescu's theory is based on various fields of functional analysis, in particular operator algebraic arguments. It is deep and difficult. On the other hand, the method evolved through [3], [14], [22], and culminating in its most general and systematic form in [21] is of more spectral theoretic nature and, though it cannot reach the best possible result of Voiculescu, the method may be apt for investigating problems with f. We call this method Berg-Sikonia-Weidmann-Voigt method (in short BSWV method). Main purpose of the present article is to examine how far we can go by BSWV method. In fact, without introducing much new idea, we can push the way considerably further to the direction of simultaneous diagonalization with respect to f. This is our main result (Theorem 4.2). We shall formulate the problem in terms of spectral measures, which is equivalent to the formulation by N-tuples. Let E be a spectral measure in $\mathbf{R}^N$ with compact support and let f be a function continuous on the support of E. The theorem will give a rather general criteria (cf. (4.10)) for pairs $(f(E), \boldsymbol{\mathcal{C}})$ in order that $f(E)$ is diagonalizable modulo $\boldsymbol{\mathcal{C}}$ simultaneously. In criteria (4.10) we use the concept of modulus of continuity of f, more general than Lipschitz continuity, while conditions imposed on $\boldsymbol{\mathcal{C}}$ is stronger than those of Voiculescu and Xia and cannot even include $\boldsymbol{\mathcal{C}}_N$.

The origin of the present work is a lecture by the author at the symposium of which this volume is the proceedings. The title of the lecture was "Diagonalization modulo norm ideals, a review and some remarks". Since a considerable improvement in the formulation (use of modulus of continuity) came after the symposium, the reviewing part was reduced. Yet, the article still retains the original character of the lecture of being expository. Thus, explanations of all preliminary materials, such as norm ideals,

[1] An apology. A review of [21] was written by the author (MathSciNet MR0451011 (56#9301)). He realized recently that in the review statement of Theorem 2.1 of [21] the crucial condition $p > m$ (in the present notation $p > N$) was missing. The author would like to take this opportunity to express his apology, especially to Professor Voigt, for this mistake.

spectral measures, modulus of continuity, are leisurely included. This work is a part of recent interest of the author in N-dimensional spectral measures. In this regard we quote [10].

The present article is organized as follows. After making a short review of symmetrically normed ideals in 2.1, a brief history of the development roughly described above will be presented in 2.2. In Section 3 we shall summarize some elementary facts about the modulus of continuity. Among them Proposition 3.4, which states that any function $\omega(\delta; f)$ expressing the modulus of continuity of f is equivalent to a concave function, may be of some interest and is included even though the proposition is not really needed in this article. 4.1 and 4.2 will be devoted to a short account on spectral measures. Our main result (Theorem 4.2) is presented in 4.3. Section 5 is devoted to the proof. In 5.1 we shall present an abstract generalization of Voigt's construction of complete orthonormal system based originally on a dyadic decomposition of a cube.

In this article we shall always work in a separable infinite dimensional Hilbert space. The following are lists of notations and abbreviations to be used throughout the article without further comment.

List of notations.

$\mathfrak{H}$	a separable infinite-dimensional Hilbert space
$\mathcal{L}$, $\mathcal{C}_\infty$, $\mathcal{R}$, or $\mathcal{L}(\mathfrak{H})$ etc.	the set of all bounded, compact, finite rank linear operators in $\mathfrak{H}$ respectively. $\\| \\|$ is the norm in $\mathcal{L}$
$\mathcal{P}$	the set of all orthogonal projections in $\mathfrak{H}$
$P_{\mathfrak{M}}$	the orthogonal projection onto a closed subspace $\mathfrak{M}$ of $\mathfrak{H}$
$\mathbb{N}$, $\mathbf{R}$, $\mathbf{C}$	the set of all natural, real, and complex numbers, respectively.
$\|A\|$	the cardinality of a set A
χ_ω	the characteristic function of a set ω
$r(A)$	the rank of an operator A of finite rank

List of abbreviations:

Abbreviation	For	Definition at
ONS	orthonormal system	
CONS	complete orthonormal system	
PPS	pure point spectrum	2.2.1, §4.1
LGD	league of graded decompositions	end of 5.1.1
LGDD	league of graded dyadic decompositions	beginning of 5.1.3

§ 2. A short review

§ 2.1. Symmetrically normed ideals

In this subsection we shall make a short review of the theory of norm ideals of operators. A standard reference is [6].

For a compact operator $A \in \boldsymbol{\mathcal{C}}_\infty$ we denote by $s_j = s_j(A)$, $j = 1, 2, \cdots$, the sequence of singular values of A (i.e. positive eigenvalues of $(A^*A)^{1/2}$) arranged in a nonincreasing order with repetitions according to the multiplicity. When $A \in \mathcal{R}$, the sequence $s_j(A)$ terminates at $j = r(A)$. We then let $s_j(A) = 0$ for $j > r(A)$ and denote the resulting sequence by $\boldsymbol{s}(A) = \{s_j(A)\}_{j\in\mathbb{N}}$.

Definition 2.1. Let $\boldsymbol{\mathcal{C}} \subset \mathcal{L}$ be a linear subspace of $\mathcal{L}$. When the following conditions (i)–(iv) are satisfied, $\boldsymbol{\mathcal{C}}$ is called a *symmetrically normed ideal* (abbr. *s.n. ideal*).

(i) $\boldsymbol{\mathcal{C}}$ is a two-sided ideal of $\mathcal{L}$ (i.e. $A \in \boldsymbol{\mathcal{C}}$ and $B, C \in \mathcal{L}$ implies $BAC \in \boldsymbol{\mathcal{C}}$).

(ii) $\boldsymbol{\mathcal{C}}$ is equipped with a norm $\|\cdot\|_{\boldsymbol{\mathcal{C}}}$ with respect to which $\boldsymbol{\mathcal{C}}$ is complete.

(iii) $\|UAV\|_{\boldsymbol{\mathcal{C}}} = \|A\|_{\boldsymbol{\mathcal{C}}}$ for any $A \in \boldsymbol{\mathcal{C}}$ and unitary operators U and V,
or equivalently,

(iii′) $\|BAC\|_{\boldsymbol{\mathcal{C}}} \le \|B\|\,\|C\|\,\|A\|_{\boldsymbol{\mathcal{C}}}$, $A \in \boldsymbol{\mathcal{C}}$ and $B, C \in \mathcal{L}$.[2]

(iv) $\|P\|_{\boldsymbol{\mathcal{C}}} = 1$ for a one-dimensional orthogonal projection P.

We call a norm on $\boldsymbol{\mathcal{C}}$ a *symmetric norm* when that norm satisfies (iii′).

We note that a two-sided ideal of $\mathcal{L}$ is always contained in $\boldsymbol{\mathcal{C}}_\infty$ (Calkin's theorem; [6, Theorem 1.1 of Chapter III]).

Example 2.2. Let $1 \le p < \infty$ and define $\boldsymbol{\mathcal{C}}_p$ and $\|A\|_p$ as

$$\boldsymbol{\mathcal{C}}_p = \Big\{ A \in \boldsymbol{\mathcal{C}}_\infty \;\Big|\; \sum_{j=1}^{\infty} s_j(A)^p < \infty \Big\}, \quad \|A\|_p = \|A\|_{\boldsymbol{\mathcal{C}}_p} = \Big(\sum_{j=1}^{\infty} s_j(A)^p \Big)^{1/p}.$$

$\boldsymbol{\mathcal{C}}_p$ is an s.n. ideal. For $1 < p < q < \infty$ we have $\boldsymbol{\mathcal{C}}_1 \subsetneq \boldsymbol{\mathcal{C}}_p \subsetneq \boldsymbol{\mathcal{C}}_q \subsetneq \boldsymbol{\mathcal{C}}_\infty$. $\boldsymbol{\mathcal{C}}_1$ and $\boldsymbol{\mathcal{C}}_2$ are called the *trace class* and the *Hilbert-Schmidt class*, respectively.

It is known that $\boldsymbol{\mathcal{C}} \supset \boldsymbol{\mathcal{C}}_1$ for any s.n. ideal $\boldsymbol{\mathcal{C}}$.

Example 2.3. Let

$$\boldsymbol{\mathcal{C}}_p^- = \Big\{ A \in \boldsymbol{\mathcal{C}}_\infty \;\Big|\; \|A\|_{\boldsymbol{\mathcal{C}}_p^-} = \sum_{j=1}^{\infty} s_j(A) j^{-1+1/p} < \infty \Big\}.$$

[2] The equivalence of (iii) and (iii′) is something to be proved (see, e.g., [6, Corollary 3.1 of Chapter III]). The implication (iii′)⇒ (iii) is easy, but the converse is not.

$\mathcal{C}_p^-$ is an s.n. ideal and satisfies $\mathcal{C}_q \subsetneq \mathcal{C}_p^- \subsetneq \mathcal{C}_p$, $1 \le q < p$. Obviously, $\mathcal{C}_1^- = \mathcal{C}_1$. $\mathcal{C}_p^-$ plays an important role in Voiculescu's theory to be explained below.

In the general theory of s.n. ideals one introduces a class of functions Φ on sequence spaces called *symmetric norming functions* and relates s.n. norms to these functions (cf. [6]). In this article we do not refer to these functions.

Proposition 2.4. *If $s_j(A) \le s_j(B)$, $\forall j$, then $\|A\|_{\mathcal{C}} \le \|B\|_{\mathcal{C}}$.* (cf. [6, p. 71])

For an operator A of finite rank $\|A\|_{\mathcal{C}}$ is estimated by the operator norm $\|A\|$ and the rank $r = r(A)$ of A. Let $\mathcal{C}$ be an s.n. ideal. By (iii) of Definition 2.1 the $\mathcal{C}$-norm of a finite-dimensional orthogonal projection $P_r \in \mathcal{P}$ of rank r is determined by r and does not depend on the choice of P_r. Namely, the following quantity $\nu_{\mathcal{C}(r)}$ is well-defined by the formula

$$\nu_{\mathcal{C}}(r) = \|P_r\|_{\mathcal{C}}, \quad P_r \in \mathcal{P}, \quad r(P_r) = r. \tag{2.1}$$

Example 2.5. $\nu_{\mathcal{C}_p}(r) = r^{1/p}$, in particular $\nu_{\mathcal{C}_1}(r) = r$. $\nu_{\mathcal{C}_p^-} = \sum_{j=1}^r j^{-1+1/p}$.

Proposition 2.6. *Let $\mathcal{C}$ be an s.n. ideal and let $A \in \mathcal{R}$ be of finite rank. Then,*

$$\|A\|_{\mathcal{C}} \le \nu_{\mathcal{C}}(r(A))\|A\|. \tag{2.2}$$

Proof. Let $r = r(A)$ and $P_r \in \mathcal{R} \cap \mathcal{P}$ with $r(P_r) = r$. Then, for $1 \le j \le r$ we have $s_j(A) \le s_1(A) = \|A\| = s_j(\|A\|P_r)$ and for $j > r$ we have $s_j(A) = s_j(\|A\|P_r) = 0$. Hence, $s_j(A) \le s_j(\|A\|P_r)$, $\forall j$, and (2.2) follows from Proposition 2.4. $\square$

Remark. Let P_r be as in (2.1). (a) $\nu_{\mathcal{C}}(P_1) = 1$ for any $\mathcal{C}$ by (iv) of Definition 2.1. (b) The equality holds in (2.2) for $A = P_r$, because $\|P_r\| = 1$. (c) $\nu_{\mathcal{C}}(r) \le r$ because P_r is a sum of r one-dimensional projections. (d) $\nu_{\mathcal{C}}(r) = o(r)$ if $\mathcal{C} \ne \mathcal{C}_1$ (cf. [9]), a small observation which was a key in [9]. In this article only a weaker assertion (2.2) will be used.

§ 2.2. Brief history

Brief descriptions of the development of the problem can be found in the literature, say in [24], and [28]. We follow the tradition and present a brief history [3].

2.2.1. Single operator A selfadjoint operator A in $\mathfrak{H}$ is said to have *pure point spectrum* if there exists a CONS $\{\varphi_k\}_{k\in\mathbb{N}}$ of $\mathfrak{H}$ consisting of eigenvectors of A, i.e.,

[3] Most of 2.2 was written before the author found [28]. Some inevitable similarities are left as they are. [28] contains more about recent developments.

$A = \sum_{k=1}^{\infty} \lambda_k(\cdot, \varphi_k)\varphi_k$. We use an abbreviation PPS for "pure point spectrum" and use expressions such as "A is PPS", "a PPS selfadjoint operator", etc.

A classical problem going back to H. Weyl is to examine whether any selfadjoint operator can be made PPS by adding an operator from a given s.n. ideal $\boldsymbol{\mathcal{C}}$.

1. Compact perturbation. H. Weyl (1909, [23]) proved that for any bounded selfadjoint operator A there exists a selfadjoint $K \in \boldsymbol{\mathcal{C}}_\infty$ such that $A + K$ is PPS.

2. Perturbation by Hilbert-Schmidt class. J. von Neumann (1935, [11]) proved the same with $K \in \boldsymbol{\mathcal{C}}_\infty$ replaced by $K \in \boldsymbol{\mathcal{C}}_2$ and proved moreover that the Hilbert-Schmidt norm $\|K\|_{\boldsymbol{\mathcal{C}}_2}$ of K can be made arbitrarily small. (A need not be bounded.)

3. Perturbation by trace class. M. Rosenbum and T. Kato (1957, [13], [8]) proved that the addition of a trace class operator does not change the absolutely continuous part. More precisely, the absolutely continuous parts of A and $A + K$ with K being selfadjoint and belonging to $\boldsymbol{\mathcal{C}}_1$ are unitarily equivalent. Thus, in 2 above the Hilbert-Schmidt class $\boldsymbol{\mathcal{C}}_2$ cannot be replaced by the trace class $\boldsymbol{\mathcal{C}}_1$.

4. Perturbation by s.n. ideal not equal to the trace class. The present author (1958, [9]) proved that in 2 the Hilbert-Schmidt class $\boldsymbol{\mathcal{C}}_2$ can be replaced by any s.n. ideal $\boldsymbol{\mathcal{C}}$ which is not equal to the trace class $\boldsymbol{\mathcal{C}}_1$.

5. We say that a selfadjoint operator A is *diagonalizable modulo* $\boldsymbol{\mathcal{C}}$ if there exists a selfadjoint $K \in \boldsymbol{\mathcal{C}}$ such that $A + K$ is PPS. Statement 3 above says that a necessary condition for a selfadjoint A to be diagonalizable modulo the trace class $\boldsymbol{\mathcal{C}}_1$ is that A has no absolutely continuous part. That this condition is also sufficient and that $\|K\|_{\boldsymbol{\mathcal{C}}_1}$ can be made arbitrarily small were proved by W. Carey and J. D. Pincus (1976, [5]).

With these results one may say that the problem for a single operator has been pretty much settled.

2.2.2. **N-tuple of commutative selfadjoint operators** For normal operators Weyl type theorem was obtained by Berg ([3, 1971]) and Sikonia ([14, 1971]). Exploiting methods used in these works, Weidmann ([22]) proved the diagonalizability result for $\boldsymbol{\mathcal{C}}_p$ with $p > 2$.

A normal operator N can be written as $N = A_1 + iA_2$, where A_1, A_2 are two commutative (i.e. $A_1A_2 - A_2A_1 = 0$) selfadjoint operators. Then the problem is naturally generalized to a similar problem for a commutative N-tuple of selfadjoint operators. Let $N \in \mathbb{N}$. We say that $\boldsymbol{A} = (A_1, \cdots, A_N)$ is a commutative N-tuple of selfadjoint operators when each $A_j \in \mathcal{L}$ is selfadjoint and $[A_j, A_k] = A_jA_k - A_kA_j = 0$, $\forall\, j,\, k = 1, \cdots, N$. When each A_j has a pure point spectrum (PPS), we say that $\boldsymbol{A}$ is a commutative N-tuple of PPS selfadjoint operators.

In order to simplify the exposition we introduce the following condition (D_N), where $N \in \mathbb{N}$. The condition (D_N) is a condition to be imposed on an s.n. ideal $\boldsymbol{\mathcal{C}}$.

$$(\mathrm{D}_N)\ \begin{cases} \text{For any commutative } N\text{-tuple } \boldsymbol{A} = \{A_1, \cdots, A_N\} \text{ of selfadjoint opera-} \\ \text{tors and } \varepsilon > 0 \text{ there exists a commutative } N\text{-tuple } \boldsymbol{D} = \{D_1, \cdots, D_N\} \\ \text{of PPS self-adjoint operators such that } A_j - D_j \in \boldsymbol{\mathcal{C}} \text{ and } \|A_j - D_j\|_{\boldsymbol{\mathcal{C}}} < \varepsilon, \\ j = 1, \cdots, N. \end{cases}$$

It is not difficult to see that, if D_j, $j = 1, \cdots, N$, are PPS and satisfy $[D_j, D_k] = 0$, then they are simultaneously diagonalizable. In other words, there exists a CONS $\{\psi_k\}$ and $\lambda_k^j \in \mathbf{R}$ such that $D_j = \sum_{k=1}^{\infty} \lambda_k^j (\cdot, \psi_k)\psi_k$. Thus, (D_N) says that, given N-tuple $\boldsymbol{A}$, all A_j, $j = 1, \cdots, N$, are simultaneously diagonalizable modulo arbitrarily small perturbation belonging to $\boldsymbol{\mathcal{C}}$.

As far as we are aware of, the first substantial work about N-tuples was done by J. Voigt (1977, [21]). (We mention though that some works forerunning to Voiculescu's work ([16]) to be mentioned below were being done in 1970's (e.g., [15]).)

Theorem 2.7. (Voigt[21, 1977]) *Let $N \in \mathbb{N}$ and let p be such that $N < p \leq \infty$. Then $\boldsymbol{\mathcal{C}}_p$ satisfies* (D_N).

Theorem 2.7 suggests that, in the case of N-tuples, the borderline for the diagonalizability may be $p \sim N$. Starting with [16] (and its forerunners), D. Voiculescu developed a general and extensive theory of diagonalization of N-tuples modulo $\boldsymbol{\mathcal{C}}$ and proved that the borderline is in fact $p \sim N$. The borderline, however, is not $\boldsymbol{\mathcal{C}}_N$, but is a little smaller ideal $\boldsymbol{\mathcal{C}}_N^-$ introduced in Example 2.3.

The following theorem is a decisive result.

Theorem 2.8. (Voiculescu[16, 1979], Bercovici-Voiculescu[2, 1989]) *An s.n. ideal $\boldsymbol{\mathcal{C}}$ satisfies* (D_N) *if and only if $\boldsymbol{\mathcal{C}} \setminus \boldsymbol{\mathcal{C}}_N^- \neq \emptyset$.*

Only if part and the statement that $\boldsymbol{\mathcal{C}}_N$ and hence any ideal $\boldsymbol{\mathcal{C}} \supset \boldsymbol{\mathcal{C}}_N$ satisfy (D_N) were established in [16]. The remaining gap was filled in [2].

Remark. **1.** Since $\boldsymbol{\mathcal{C}}_1^- = \boldsymbol{\mathcal{C}}_1$, Theorem 2.8 adapted to the case $N = 1$ is exactly the same as 4 of Subsubsection 2.2.1 as statement. However, it is mentioned in [18] that the method used to prove Theorem 2.8 is only effective for $N \geq 2$.

2. In the course of the proof of "only if part" of Theorem 2.8 it is proved that $\boldsymbol{A}$ is diagonalizable modulo $\boldsymbol{\mathcal{C}}_N^-$ if and only if the absolutely continuous part of $\boldsymbol{A}$ is absent (cf. [16, Proposition 4.1, Theorem 4.5]). This corresponds to the result of Carey-Pincus ($N = 1$) mentioned in 5 of Subsubsection 2.2.1.

Voigt's approach (BSWV method) is of spectral theoretic nature, while Voiculescu's theory is of more operator algebraic nature. In the latter, probably, the method is as important as the result. Let $\mathcal{R}^+ = \{A \in \mathcal{R}(\mathfrak{H}) \,|\, A \text{ is selfadjoint}, 0 \leq A \leq I\}$. Let $\boldsymbol{A} = \{A_1, \cdots, A_N\}$ be a commutative N-tuple of selfadjoint operators and let $\boldsymbol{\mathcal{C}}$ be a norm ideal. In Voiculescu's theory an invariant defined as

$$k_{\mathcal{C}}(\boldsymbol{A}) = \liminf_{R \in \mathcal{R}^+} \{ \max_{1 \le j \le N} \{ \|[R, A_j]\|_{\mathcal{C}} \} \}$$

(in Voiculescu's notation k_Φ) plays an important role not only in the diagonalization theory, but in various further developments. We only quote here the following crucial link between $k_{\mathcal{C}}(\boldsymbol{A})$ and the diagonalization ([16, Corollary 2.6]).

"A commutative N-tuple $\boldsymbol{A}$ is diagonalizable in the sense of $(\mathrm{D}_N) \Longleftrightarrow k_{\mathcal{C}}(\boldsymbol{A}) = 0$" The germ of this relation may be traced back to a classical work of P. Halmos ([7]).

We cannot write any review of further developments. The reader is referred to [16]–[20] and references cited there. As one of recent related works, we also quote [12].

2.2.3. Diagonalization theorems under functional calculus The relation between $\boldsymbol{A}$ and $\boldsymbol{D}$ in (D_N) may be written in abbreviation as $\boldsymbol{A} - \boldsymbol{D} \in \boldsymbol{C}$. In "diagonalization theorems under functional calculus", a phrase borrowed from [24], one is interested in the simultaneous diagonalization modulo $\mathcal{C}$ of $f(\boldsymbol{A})$, i.e., $f(\boldsymbol{A}) - f(\boldsymbol{D}) \in \mathcal{C}$, and asks for which class of functions f this relation holds. Since 1997 J. Xia has been developing a broad theory on this subject and published a series of works, e.g., [24]–[27].

In [24] Xia introduced the concept of $\mathcal{C}$-discreteness. Apart from some requirement on eigenvectors, the statement that $\boldsymbol{A}$ is $\mathcal{C}$-discrete is essentially equivalent, the author believes, to the statement that there exists a PPS $\boldsymbol{D}$ such that $f(\boldsymbol{A}) - f(\boldsymbol{D}) \in \mathcal{C}$ for all Lipschtz continuous f. Note that $A_j = f_j(\boldsymbol{A})$ with $f_j(\lambda_1, \cdots, \lambda_N) = \lambda_j$ so that $\mathcal{C}$-discreteness of $\boldsymbol{A}$ implies that $\boldsymbol{A}$ is diagonalizable modulo $\mathcal{C}$. Then, though not very explicitly stated, Corollary 7.2 of [24] generalizes the "if part" of Theorem 2.8 to a diagonalization theorem under functional calculus.

Our Theorem 4.2 will turn out to be a partial generalization of the above mentioned Xia's result in that the condition on f is relaxed while the condition on $\mathcal{C}$ is strengthened and excludes even $\mathcal{C}_N$. It might be possible that one can reach $\mathcal{C}_N^-$ version of Theorem 4.2 by some extension of Xia's method in [24], [27]. But due to heavy technicality in Xia's works we have not been able to investigate it so far. Our Theorem 4.2 is proved by a direct extension of BSWV method and our intention in the present article is to emphasize the simplicity of BSWV method.

So far, we have been concerned only with the problem of diagonalization. This is the problem of instability of continuous spectra under qualitatively "large" perturbation. The theory of wave operators and the stability of absolutely continuous spectra under qualitatively "small" perturbation were also investigated by Voigt ([21]) and Voiculescu ([17]) for commutative N-tuples. In this article we do not touch upon these topics and hope we shall be able to come back to them in another occasion.

§ 3. An essay on modulus of continuity

§ 3.1. Definition and basic properties

Throughout this subsection we fix a compact convex set $Q \subset \mathbf{R}^n$. We denote by $d(Q)$ the diameter of $Q : d(Q) = \operatorname{diam}(Q) = \sup\{\, |x-y| \;\big|\; x, y \in Q \,\}$.

Let $C(Q)$ be the set of all continuous functions on Q. For $f \in C(Q)$ we put

$$\omega(\delta) = \omega(\delta; f) = \sup\{\, |f(x) - f(y)| \;\big|\; x, y \in Q,\ |x-y| \le \delta \,\}, \quad \delta \ge 0, \tag{3.1}$$

and call $\omega(\delta; f)$ the *modulus of continuity* of f. (In the present paper we stick to this basic definition of modulus of continuity and do not argue about higher order version.)

In what follows we shall summarize useful properties of $\omega(\delta; f)$. All these are elementary. But, for reader's convenience we shall write down a proof of propositions in 3.2.

The following properties of $\omega(\delta)$ are immediate consequences of definition (3.1):

$$\omega(\delta) \text{ is non-decreasing;} \quad \omega(0) = 0; \quad \omega(\delta) = \omega(d(Q)) \quad \text{for} \quad \delta \ge d(Q).$$

Proposition 3.1. *$\omega(\delta; f)$ is a continuous function of δ, i.e., $\omega(\cdot; f) \in C([0, \infty))$.*

Remark. Definition 3.1 works for a general Q. As seen from the proof, $\omega(\delta)$ is right continuous for a general compact Q, but it is not necessarily left continuous. The convexity of Q is one of convenient sufficient conditions for the left continuity.

In order to characterize functions $\omega(\delta)$ we introduce the following classes of functions:

$$w \in J_Q \overset{\text{def}}{\iff} \begin{cases} w \in C([0,\infty)); \quad w(t) \ge 0; \quad w \text{ is non-decreasing,} \\ w(0) = 0, \quad w(t) = \text{const for } t \ge d(Q), \end{cases}$$

$$w \in K_Q \overset{\text{def}}{\iff} w \in J_Q \text{ and } w \text{ is concave,}$$

$$w \in M_Q \overset{\text{def}}{\iff} w \in J_Q \text{ and } w \text{ satisfies the subadditivity relation:}$$

$$w(t+s) \le w(s) + w(t), \quad t,\, s \ge 0. \tag{3.2}$$

Remark. It is easily seen that $K_Q \subset M_Q$, but the converse inclusion does not hold. An example is $f(\delta) = \delta$, $0 \le \delta \le 1$; $= 1$, $1 \le \delta \le 2$; $= \delta - 1$, $2 \le \delta \le 3$; $= 2$, $3 \le \delta$.

Proposition 3.2. *Let $w \in C([0, \infty))$. Then, we have*

$$\Big[\, \exists f \in C(Q) \quad \text{s. t.} \quad w(\delta) = \omega(\delta; f) \,\Big] \iff w \in M_Q.$$

For characterizing the degree of continuity such as Hölder continuity of order θ it is more convenient to use equivalence classes of functions in J_Q rather than functions themselves. Letting w_1, $w_2 \in J_Q$, we introduce relations $\prec$ and $\sim$ into J_Q and hence into M_Q as follows:

$$w_1 \prec w_2 \overset{\text{def}}{\Longleftrightarrow} \Big[\exists\, c > 0 \quad \text{s. t.} \quad w_1(t) \le c w_2(t), \quad \forall\, t \ge 0\Big],$$

$$w_1 \sim w_2 \overset{\text{def}}{\Longleftrightarrow} \Big[w_1 \prec w_2 \quad \text{and} \quad w_2 \prec w_1\Big].$$

Obviously, the relation $\sim$ is an equivalence relation. We denote by $\mathcal{M}_Q$ the set of all equivalence classes in $M_Q : \mathcal{M}_Q = M_Q/\sim$. An equivalence class in $\mathcal{M}_Q$ is represented by a boldface letter, so that we write like $w \in \boldsymbol{w} \in \mathcal{M}_Q$, where $w \in M_Q$. We note that the relation $\prec$ in M_Q induces a similar relation in $\mathcal{M}_Q$.

It is this $\mathcal{M}_Q$ whcich will be used as a convenient tool to classify the degree of continuity. We call $\boldsymbol{w} \in \mathcal{M}_Q$ an *equivalence class of modulus of continuity.* For $f \in C(Q)$ we denote by $\boldsymbol{\omega}(f)$ the equivalence class of modulus of continuity to which $\omega(\delta; f)$ belongs: $\omega(\cdot; f) \in \boldsymbol{\omega}(f) \in \mathcal{M}_Q$.

Let $f \in C(Q)$ and $\boldsymbol{w} \in \mathcal{M}_Q$. f is said to be $\boldsymbol{w}$*-continuous*, if $\boldsymbol{\omega}(f) \prec \boldsymbol{w}$. The set of all $\boldsymbol{w}$-continuous functions is denoted by $C^{\boldsymbol{w}}(Q)$.

Example 3.3. Let $0 < \theta \le 1$ and let $w_\theta(\delta) = \delta^\theta$, $0 \le \delta \le d(Q)$, and $w_\theta(\delta) = d(Q)^\theta$, $\delta \ge d(Q)$. Then, $w_\theta \in M_Q$. Let $\boldsymbol{w}_\theta = \boldsymbol{\omega}(w_\theta)$. $f \in C(Q)$ is $\boldsymbol{w}_\theta$-continuous if and only if there exists c such that $\omega(\delta; f) = \sup\limits_{|x-y|\le\delta} |f(x) - f(y)| \le c w_\theta(\delta) = c\delta^\theta$. This is equivalent to $|f(x) - f(y)| \le c|x-y|^\theta$, the Hölder continuity of order θ of f. (A function Hölder continuous of order θ is Hölder continuous of order θ' for any $\theta' \le \theta$. This is the reason why we used $\boldsymbol{\omega}(f) \prec \boldsymbol{w}$ rather than $\boldsymbol{\omega}(f) = \boldsymbol{w}$ in the definition of $\boldsymbol{w}$-continuity.)

It is convenient to have some qualitative measures in discussing $\boldsymbol{w}$-continuity. Take $\boldsymbol{w} \in \mathcal{M}_Q$ and fix one representative $w \in \boldsymbol{w}$. Let $f \in C(Q)$ and suppose that f is $\boldsymbol{w}$-continuous. Since $\boldsymbol{\omega}(f) \prec \boldsymbol{w}$ by definition, there exist $c > 0$ such that $\omega(\delta; f) \le c w(\delta)$, $\forall\, \delta \ge 0$. We define $|f|_w$ as the infimum of such c, i.e.,

$$|f|_w = \inf\{\, c \,|\, \omega(\delta; f) \le c w(\delta),\ \forall\, \delta \ge 0\,\}, \quad w \in \boldsymbol{w}, \quad \boldsymbol{\omega}(f) \prec \boldsymbol{w}. \tag{3.3}$$

It is readily seen that

$$|f|_w = \sup_{x,y\in Q,\ x\ne y} \frac{|f(x) - f(y)|}{w(|x-y|)}. \tag{3.4}$$

In particular, in Example 3.3 $|f|_{w_\theta}$ is equivalent to the ordinary $C^{0,\theta}(Q)$ norm of f in a bounded domain Q.

These are about all we need later. For its own interest we add two more propositions. In Remark just before Proposition 3.2 we noted that $K_Q \subsetneqq M_Q$. However, for any $w \in M_Q$ there exists $\widetilde{w} \in K_Q$ such that $w \sim \widetilde{w}$. In other words any $\boldsymbol{w} \in \mathcal{M}_Q$ has a concave representative. More precisely, we have the following proposition.

Proposition 3.4. *Let $w \in M_Q$ and let*

$$\widetilde{W} = \{ u \in K_Q,\ w(s) \le u(s),\ \forall\, s \ge 0 \}, \quad \widetilde{w}(t) = \inf\{ u(t) \mid u \in \widetilde{W} \}.$$

Then, $\widetilde{W} \neq \emptyset$, $\widetilde{w} \in K_Q$, and

$$\frac{1}{2}\widetilde{w}(t) \le w(t) \le \widetilde{w}(t), \quad \forall\, t \ge 0. \tag{3.5}$$

We remark that an exposition relevant to Proposition 3.4 can be found in [1].

The following proposition shows that, for classifying the degree of continuity of f, only the values of $\omega(\delta, f)$ at $\delta \sim 0$ matter.

Proposition 3.5. Let w_1, $w_2 \in M_Q$. If there exists a natural number J such that

$$w_1(2^{-j}) = w_2(2^{-j}), \qquad \forall\, j \ge J, \tag{3.6}$$

then $w_1 \sim w_2$, so that w_1, w_2 belong to the same equivalence class in $\mathcal{M}_Q$.

The same statement would be true if 2^{-j} is replaced by a^{-j}, $a > 0$. We do not know yet if 2^{-j} in (3.6) can be replaced by an arbitrary sequence t_j such that $t_j \searrow 0$.

§ 3.2. Proof of propositions

Proof of Proposition 3.1. Proof of the right continuity. Suppose on the contrary that $d \equiv \lim_{\delta' \downarrow \delta} \omega(\delta') - \omega(\delta) > 0$, where $\omega(\delta) = \omega(\delta; f)$, and hence that $\omega(\delta + 1/n) \ge \omega(\delta) + d$ $\forall\, n \in \mathbb{N}$. Then, there would exist x_n, $y_n \in Q$ such that $|x_n - y_n| \le \delta + 1/n$ and $|f(x_n) - f(y_n)| \ge \omega(\delta) + d - 1/n$. By extracting subsequence for which x_n and y_n converge, we would readily reach a contradiction. (Here only the compactness of Q and the continuity of f are used.)

Proof of the left continuity. Let $\delta > 0$ be fixed. We may assume that $\omega(\delta) > 0$, as $\omega(\delta) = 0$ would imply that $f(x)$ and hence $\omega(\delta)$ are identically zero. Fix an ε such that $0 < \varepsilon < \omega(\delta)$ and take x, $y \in Q$ such that $|x - y| \le \delta$ and $|f(x) - f(y)| \ge \omega(\delta) - \varepsilon/2$. Note that the last relation implies $x \neq y$. Let $\overline{(xy)}$ be the open line segment joining x and y. Since f is continuous, we can take $x' \in \overline{(xy)}$ such that $|f(x') - f(x)| < \varepsilon/2$. Put $\eta = |x' - x|$. Since Q is convex, $\overline{(xy)} \subset Q$ so that $x' \in Q$. Furthermore, $|x' - y| = |x - y| - \eta \le \delta - \eta$ and $|f(x') - f(y)| \ge |f(x) - f(y)| - |f(x) - f(x')| \ge \omega(\delta) - \varepsilon$.

These imply that $\omega(\delta-\eta) \geq \omega(\delta)-\varepsilon$. Since $\omega(\delta)$ is non-decreasing, this proves the left continuity. $\square$

Proof of Proposition 3.2. Proof of $\Longrightarrow$ Write as $w(t)=\omega(t;f)$. It is obvious that $w \in J_Q$. We verify (3.2). Let $\varepsilon>0$. Then, there exist x, $y \in Q$ such that $|x-y| \leq t+s$ and $|f(x)-f(y)| \geq w(t+s)-\varepsilon$. $|x-y| \leq t+s$ implies the existence of $z \in \overline{xy}$ such that $|x-z| \leq t$, $|z-y| \leq s$. Since Q is convex, $z \in Q$. Hence we see that $w(t+s)-\varepsilon \leq |f(x)-f(z)|+|f(z)-f(y)| \leq w(t)+w(s)$. Since ε is arbitrary, (3.2) is verified.

Proof of $\Longleftarrow$ Since Q is compact, there exist a, $b \in Q$ such that $|a-b|=d(Q)$. Introduce a new coordinate system into $\mathbf{R}^N$ by taking a as the origin and the direction of the vector $\overrightarrow{ab}$ as the positive direction of the x_1 axis. Let x' be the coordinate in the plane perpendicular to $\overrightarrow{ab}$ and express $x \in \mathbf{R}^N$ as $x=(x_1,x')$. Then, $a=(0,0)$ and $b=(d(Q),0)$. Note that Q sits between two hyperplanes $x_1=0$ and $x_1=d(Q)$.

Given $w \in M_Q$, we define $f \in C(Q)$ as $f(x)=f(x_1,x')=w(x_1)$ and prove that $w(\delta)=\omega(\delta;f)$. It suffices to prove this for $0 \leq \delta \leq d(Q)$. We have $\omega(\delta;f) \geq w(\delta)$, because $\omega(\delta;f) \geq |f(\delta,0)-f(0,0)| = w(\delta)$. Next, let x, $y \in Q$ and suppose that $|x-y| \leq \delta$. We may assume that $x_1 \leq y_1$ Then, $0 \leq y_1-x_1 \leq \delta$ and $|f(x)-f(y)| = w(y_1)-w(x_1) \overset{(*)}{\leq} w(y_1-x_1) \leq w(\delta)$, where the inequality marked $^{(*)}$ follows from (3.2). This proves that $\omega(\delta;f) \leq w(\delta)$. $\square$

Proof of Proposition 3.4. That $\widetilde{W} \neq \emptyset$ and $\widetilde{w} \in K_Q$ are immediately seen. The second inequality of (3.5) is obvious. To prove the first it suffices to show that

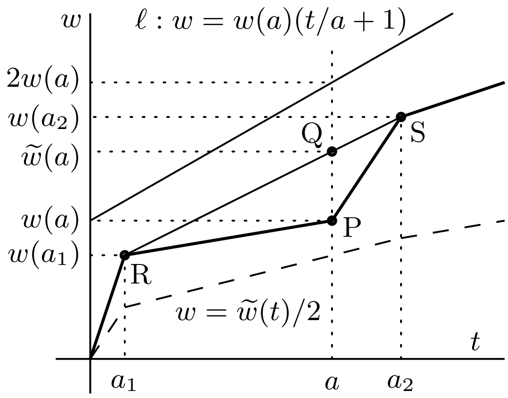
$$\big[\, w(a) < \widetilde{w}(a), \quad 0 < a < d(Q) \,\big] \Longrightarrow \widetilde{w}(a) < 2w(a).$$

We first show that

$$w(t) \leq w(a)\Big(\frac{t}{a}+1\Big), \quad t \geq a. \tag{3.7}$$

In fact, writing as $t=Na+r$, $N \in \mathbb{N} \cup \{0\}$, $0 \leq r < a$, and using (3.2) repeatedly, we see that $w(t) \leq w((N+1)a) \leq (N+1)w(a) = (t/a-r/a+1)w(a) \leq (t/a+1)w(a)$.

Since $w(a) < \widetilde{w}(a)$, we can find maximum a_1 such that $a_1 < a$ and $w(a_1)=\widetilde{w}(a_1)$ and minimum a_2 such that $a < a_2$ and $w(a_2)=\widetilde{w}(a_2)$. In the t-w plane put P$=(a,w(a))$, Q$=(a,\widetilde{w}(a))$, R$=(a_1,w(a_1))=(a_1,\widetilde{w}(a_1))$, S$=(a_2,w(a_2))=(a_2,\widetilde{w}(a_2))$ (see Figure). In $[a_1,a_2]$ the curve RPS represents the graph of $w=w(t)$ and the line segment $\overline{\text{RS}}$ represents the graph of $w=\widetilde{w}(t)$. Q must lie on $\overline{\text{RS}}$.



Let l be the line represented by $w = w(a)(t/a+1)$. Then, by (3.7) we see that S lies below l. Since $w(t)$ is nondecreasing, R also lies below l (see Figure). Hence, $\mathrm{Q} \in \overline{\mathrm{RS}}$ must lie below l, which implies that $\widetilde{w}(a) < 2w(a)$. $\square$

Proof of Proposition 3.5. For $t \geq 2^{-J}$ we have $w_2(2^{-J})/w_1(d(Q)) \leq w_2(t)/w_1(t) \leq w_2(d(Q))/w_1(2^{-J})$. For $2^{-(j+1)} \leq t \leq 2^{-j}$, $j \geq J$ we see that

$$\frac{w_1(2^{-(j+1)})}{w_1(2^{-j})} = \frac{w_2(2^{-(j+1)})}{w_1(2^{-j})} \leq \frac{w_2(t)}{w_1(t)} \leq \frac{w_2(2^{-j})}{w_1(2^{-(j+1)})} = \frac{w_1(2^{-j})}{w_1(2^{-(j+1)})}$$

Since $w_1(2^{-j}) = w_1(2^{-(j+1)} + 2^{-(j+1)}) \leq 2w_1(2^{-(j+1)})$ by (3.2), we see that the right hand side is not greater than 2 and the left hand side is not smaller than 1/2. Thus, we are done. $\square$

§ 4. How far one can go by BSWV method

§ 4.1. Spectral measures

Let $(\Omega, \mathcal{B})$ be a measurable space. A mapping E from $\mathcal{B}$ to $\mathcal{P}$ is called a *spectral measure* on $(\Omega, \mathcal{B})$, if (i) E is strongly σ-additive, that is,

$$\big[\omega_j \in \mathcal{B},\ j \in \mathbb{N},\ \omega_j \cap \omega_k = \emptyset,\ \ j \neq k\big] \implies \Big[E(\cup_{j=1}^{\infty}\omega_j)u = \sum_{j=1}^{\infty} E(\omega_j)u,\ \forall\, u \in \mathfrak{H}\Big],$$

and (ii) $E(\Omega) = I$. (I is the identity operator in $\mathfrak{H}$.) When we need to specify $\mathfrak{H}$, we call E an $\mathfrak{H}$-spectral measure. (The reader is referred to [4, Chapter 5] for a systematic exposition of spectral measures on general measurable space.)

Spectral measures which have pure point spectrum (PPS) will be an important ingredient in the present paper. A spectral measure E on $(\Omega, \mathcal{B})$ is said to be a *PPS spectral measure* if there exists a sequence $\xi_k \in \Omega$ and a CONS $\{\varphi_k\}$ of $\mathfrak{H}$ such that $E(\omega)$, $\omega \in \mathcal{B}$, is the orthogonal projection onto the closed subspace of $\mathfrak{H}$ spanned by all φ_k such that ξ_k belongs to ω, that is

$$E(\omega) = P_{\mathrm{c.l.h.}\{\varphi_k \,|\, \xi_k \in \omega\}}, \tag{4.1}$$

where c.l.h.$\mathfrak{S}$ denotes the closed linear hull of a set $\mathfrak{S} \subset \mathfrak{H}$. (Note that the sequence ξ_k may contain the same point repeatedly.) Formally, (4.1) may be written as

$$E = \sum_k \delta_{\xi_k} P_{\{\alpha\varphi_k | \alpha \in \mathbf{C}\}}. \tag{4.1$'$}$$

Functional calculus with respect to a spectral measure For a spectral measure E on $(\Omega, \mathcal{B})$ and a function f on Ω the operator $f(E)$ is defined as in the case of spectral measures on $\mathbf{R}$. In the present paper we only consider the case that f is bounded. The following is a quick review.

Let $L^\infty(\Omega)$ be the set of all $\mathcal{B}$-measurable bounded functions on Ω and put $\|f\|_\infty = \text{ess-}\sup_{\lambda\in\Omega} |f(\lambda)|$. For fixed $u,\ v \in \mathfrak{H}$,

$$\rho_{u,v}(\omega) = \rho_{u,v;E}(\omega) = (E(\omega)u, v), \quad \omega \in \mathcal{B}, \tag{4.2}$$

defines a complex measure on $(\Omega, \mathcal{B})$. We write as $\rho_u = \rho_{u,u}$. ρ_u is a bounded measure on Ω. $\rho_{u,v}$ satisfies $|\rho_{u,v}(\omega)| \le \|u\|\|v\|$ and it follows that $\left|\int_\Omega f(\lambda)d\rho_{u,v}(\lambda)\right| \le \|f\|_\infty \|u\|\,\|v\|$. Therefore, for $f \in L^\infty(\Omega)$, the formula

$$(f(E)u, v) = \int_\Omega f(\lambda)d\rho_{u,v}(\lambda), \qquad u,\ v \in \mathfrak{H}, \tag{4.3}$$

defines an operator $f(E) \in \mathcal{L}$ and $\|f(E)\| \le \|f\|_\infty$. It is easily verified that the mapping $f \longmapsto \mathcal{L}(\mathfrak{H})$ satisfies the basic formulas of functional calculus:

$$(\alpha f + \beta g)(E) = \alpha f(E) + \beta g(E), \quad (fg)(E) = f(E)g(E), \quad \bar{f}(E) = f(E)^*. \tag{4.4}$$

By taking $f = \chi_\omega$, where χ_ω is the characteristic function of ω, and noting (4.2), we see that $(\chi_\omega(E)u, v) = \int_\Omega \chi_\omega(\lambda)d\rho_{u,v}(\lambda) = \rho_{u,v}(\omega) = (E(\omega)u, v)$, and hence

$$\chi_\omega(E) = E(\omega). \tag{4.5}$$

§ 4.2. Decomposition into direct sum of simple spectral measures

This subsection is a short description of the direct sum decomposition similar to the so-called Hellinger–Hahn decomposition in the spectral theory of selfadjoint operators.

Simple spectral measure Let $(\Omega, \mathcal{B}, \rho)$ be a finite measure space and let $\mathfrak{H} = L^2(\Omega, \rho)$. For $f \in L^\infty(\Omega)$ we denote by M_f the operator of multiplication by f : $M_f u(\lambda) = f(\lambda)u(\lambda)$, $\forall u \in L^2(\Omega, \rho)$. Let E_χ be the family of operators defined as

$$E_\chi(\omega) = M_{\chi_\omega}, \quad \omega \in \mathcal{B}.$$

Then, it is easily verified that E_χ is a $L^2(\Omega, \rho)$-spectral measure on $(\Omega, \mathcal{B})$.

An $\mathfrak{H}$-spectral measure E on $(\Omega, \mathcal{B})$ is said to be a *simple spectral measure* if there exists a measure ρ on $(\Omega, \mathcal{B})$ such that E is unitarily equivalent to the $L^2(\Omega, \rho)$-spectral measure E_χ on $(\Omega, \mathcal{B})$.

Simple part of E generated by u. Let E be an $\mathfrak{H}$-spectral measure. Fixing $u \in \mathfrak{H}$, $u \neq 0$, we put

$$\mathfrak{H}(u; E) = \text{c.l.h.}\{f(E)u \mid f \in L^\infty(\Omega)\}.$$

From (4.3) and (4.4) it follows that

$$\|f(E)u\|^2 = (|f|^2(E)u, u) = \int_\Omega |f(\lambda)|^2 d\rho_u(\lambda). \tag{4.6}$$

Since $L^\infty(\Omega)$ is dense in $L^2(\Omega, \rho_u)$, we see by (4.6) that the mapping

$$L^\infty(\Omega) \ni f \longmapsto f(E)u \in \mathfrak{H}(u; E)$$

can be extended uniquely to a unitary operator

$$U: \; L^2(\Omega, \rho_u) \longrightarrow \mathfrak{H}(u; E).$$

Since $E(\omega)f(E) = (\chi_\omega f)(E)$ by (4.4) and (4.5), $u \in \mathfrak{H}(u; E)$ implies that $E(\omega)u \in \mathfrak{H}(u; E)$. Thus, the spectral measure E is reduced by $\mathfrak{H}(u; E)$. This means that the restriction of E to $\mathfrak{H}(u; E)$ is an $\mathfrak{H}(u; E)$-spectral measure on $(\Omega, \mathcal{B})$. We denote this spectral measure by E_u. Then we have

$$E_u(\omega) = UM_{\chi_\omega}U^{-1}, \quad E_u = UE_\chi U^{-1}. \tag{4.7}$$

Hence, E_u in $\mathfrak{H}(u; E)$ is unitarily equivalent to E_χ in $L^2(\Omega, \rho_u)$, so that E_u is a simple spectral measure. We call the pair $(\mathfrak{H}(u; E), E_u)$ a *spectrally simple part* of $(\mathfrak{H}, E)$ generated by u.

Decomposition into a direct sum of spectrally simple parts Take $u_1 \in \mathfrak{H}$, $u_1 \neq 0$, and put

$$\mathfrak{H}_1 = \mathfrak{H}(u_1; E), \qquad E_1 = E_{u_1} = E\restriction_{\mathfrak{H}_1}.$$

By (4.7) E_1 in $\mathfrak{H}_1$ is unitarily equivalent to E_χ in $L^2(\Omega, \rho_{u_1})$, or $(\mathfrak{H}_1, E_1)$ is a spectrally simple part of $(\mathfrak{H}, E)$ generated by u_1.

If $\mathfrak{H}_1 = \mathfrak{H}$, our construction of the decomposition ends here. Otherwise, take $u_2 \in \mathfrak{H} \ominus \mathfrak{H}_1$, $u_2 \neq 0$. Since E is reduced by $\mathfrak{H}_1$ and hence by $\mathfrak{H} \ominus \mathfrak{H}_1$, we can repeat the same construction as before to obtain $\mathfrak{H}_2 = \mathfrak{H}(u_2, E\restriction_{\mathfrak{H}\ominus\mathfrak{H}_1})$, $E_2 = E\restriction_{\mathfrak{H}_2}$. The spectral measure E_2 in $\mathfrak{H}_2$ is unitarily equivalent to the spectral measure E_χ in $L^2(\Omega, \rho_{u_2})$.

Continuing this process, we obtain the following proposition.

Proposition 4.1. *Let $\mathfrak{H}$ be a separable Hilbert space and let E be an $\mathfrak{H}$-spectral measure on $(\Omega, \mathcal{B})$. Then, $\mathfrak{H}$ is decomposed as*

$$\mathfrak{H} = \sum_{n=1}^{M} \oplus \mathfrak{H}_n = \sum_{n=1}^{M} \oplus \mathfrak{H}(u_n, E), \quad \text{either } M \in \mathbb{N} \text{ or } M = \infty. \tag{4.8}$$

Each $\mathfrak{H}_n = \mathfrak{H}(u_n, E)$ *reduces* E *and*

$$E_n \equiv E\restriction_{\mathfrak{H}_n} \text{ in } \mathfrak{H}_n \sim E_\chi \text{ in } L^2(\Omega; \rho_{u_n}) \qquad \text{(unitary equivalence).}$$

In short, $(\mathfrak{H}, E)$ *is decomposed into a direct sum of* (*mutually orthogonal*) *spectrally simple parts* $(\mathfrak{H}_n, E_n)$.

Remark. In the Hellinger-Hahn decomposition ($N = 1$) it is further required that the support of ρ_{u_n} is non-increasing. In the present paper we do not do that.

Equivalence of spectral measures on $\mathbb{R}^n$ and commutative N-tuples Let E be a spectral measure on $\mathbb{R}^n$ with compact support. Then

$$\boldsymbol{A} = (A_1, \cdots, A_N), \quad A_j = \int_{\mathbb{R}^n} \lambda_j dE(\lambda) \tag{4.9}$$

is a commutative N-tuple of bounded selfadjoint operators. Conversely, suppose that commutative bounded selfadjoint operators $A_j = \displaystyle\int_{\mathbb{R}} \lambda dE_j(\lambda)$, $j = 1, \cdots N$, are given. Then, E_j's are mutually commutative and one can construct the product spectral measure E of $E_1, \cdots, E_N$ (cf. [4]). With this E relation (4.9) holds. Thus, the problem of diagonalization of commutative N tuples is equivalent to the problem of the diagonalization of spectral measures.

§ 4.3. Results

The subject of our investigation is a spectral measure on $\mathbf{R}^N$, $N \geq 1$. In the present paper we consider only spectral measures with a compact support. This is for simplicity and generalization to general spectral measures will be immediate.

By an immediate extension of the proof of Theorem 2.7 we see that $f(\boldsymbol{A})$ is diagonalizable modulo $\boldsymbol{\mathcal{C}}_p$, $p > N$, if f is Lipschitz continuous, a special case of Xia's result mentioned in 2.2.3. Similarly, we can easily show that $f(\boldsymbol{A})$ is diagonalizable modulo $\boldsymbol{\mathcal{C}}_p$, $p > N/\theta$, if f is Hölder continuous of order θ, $0 < \theta \leq 1$. This suggests that as the degree of the continuity of f becomes weaker we need a bigger $\boldsymbol{\mathcal{C}}$ for the diagonalization modulo $\boldsymbol{\mathcal{C}}$. We use the modulus of continuity to express the degree of continuity of f and the quantity $\nu_{\boldsymbol{\mathcal{C}}}$ difined by (2.1) the bigness of $\boldsymbol{\mathcal{C}}$. A typical condition is

$$\sum_{j=1}^{\infty} \nu_{\boldsymbol{\mathcal{C}}}(2^{jN}) w(2^{-j}) < \infty.$$

As we shall strive for a broad range of simultaneous diagonalizability and a uniform estimate by the quantity $|f|_w$ introduced by (3.3), our theorem will have a little complicated appearance.

Main theorem

Theorem 4.2. *Let E be a spectral measure in $\mathbf{R}^n$ with* $\operatorname{supp}(E) \subset Q$, *where Q is a closed cube in $\mathbf{R}^n$. Let $\{\eta_j\}$, $\eta_j > 0$, be a sequence of positive numbers such that $\sum_{j=1}^{\infty} \eta_j < \infty$. Let $\varepsilon > 0$. Then there exists a PPS spectral measure E_0 in $\mathbf{R}^n$, depending only on E, $\{\eta_j\}$, and ε, which has the following property.*

If an s.n. ideal $\boldsymbol{\mathcal{C}}$ and a representative $w \in \boldsymbol{w}$ of an equivalence class $\boldsymbol{w}$ of modulus of continuity satisfy

$$\nu_{\boldsymbol{\mathcal{C}}}(2^{jN})w(2^{-j}) \leq \eta_j, \tag{4.10}$$

then $f(E) - f(E_0) \in \boldsymbol{\mathcal{C}}$ for all $\boldsymbol{w}$-continuous function f and the estimate

$$\|f(E) - f(E_0)\|_{\boldsymbol{\mathcal{C}}} \leq \varepsilon |f|_w, \qquad \forall\, f \in C^{\boldsymbol{w}}(Q), \tag{4.11}$$

holds, where $|f|_w$ is defined by (3.3).

Let $a_j > 0$, $b_j > 0$ be two sequences of positive numbers. We write $a_j \lesssim b_j$ if there exists $c > 0$ such that $a_j \leq c b_j$, $\forall\, j \in \mathbb{N}$.

Corollary 4.3. *Let E and $\{\eta_j\}$ be as in Theorem 4.2. Then, there exists a PPS spectral measure E_0 in $\mathbf{R}^n$, depending only on E and $\{\eta_j\}$, which has the following property.*

If an s.n. ideal $\boldsymbol{\mathcal{C}}$ and an equivalence class of modulus of continuity $\boldsymbol{w} \in \mathcal{M}_Q$ satisfy

$$\nu_{\boldsymbol{\mathcal{C}}}(2^{jN})w(2^{-j}) \lesssim \eta_j \quad \text{for one or equivalently all } w \in \boldsymbol{w}, \tag{4.12}$$

then $f(E) - f(E_0) \in \boldsymbol{\mathcal{C}}$ for all $\boldsymbol{w}$-continuous functions f.

Proof of Corollary. Take $\varepsilon = 1$ in Theorem 4.2 and construct E_0. Assumption (4.12) implies that certain representatives w of $\boldsymbol{w}$ satisfy (4.10). Apply the theorem to conclude that $f(E) - f(E_0) \in \boldsymbol{\mathcal{C}}$. □

Example 4.4. Let us consider the pair $(\boldsymbol{\mathcal{C}}_p, \boldsymbol{w}_\theta)$ with $1 \leq p < \infty$, $0 < \theta \leq 1$. (For $\boldsymbol{\mathcal{C}}_p$ and $\boldsymbol{w}_\theta$ see Examples 2.2, 3.3.) As a representative of $\boldsymbol{w}_\theta$ we take $w_\theta(\delta) = \delta^\theta$. Then, $C^{\boldsymbol{w}_\theta}$ is equal to $C^{0,\theta}$, the set of all Hölder continuous functions of order θ. For $f \in C^{0,\theta}(Q)$ we put $\|f\|_\theta = |f|_{w_\theta}$. As was mentioned after (3.4) $\|f\|_\theta$ is equivalent to the $C^{0,\theta}$-norm of f.

We have $\nu_{\boldsymbol{\mathcal{C}}_p}(2^{jN})w_\theta(2^{-j}) = 2^{-j(\theta - N/p)}$. If p and θ satisfies $\theta - N/p \geq \beta > 0$, then the right hand side is majorized by $2^{-\beta j}$. Therefore, by taking $\eta_j = 2^{-\beta j}$ and applying Theorem 4.2, we conclude that the following statement holds.

For any $\varepsilon > 0$ and β, $0 < \beta < 1$, there exists a PPS spectral measure E_0 such that

$$f(E) - f(E_0) \in \boldsymbol{\mathcal{C}}_p, \quad \|f(E) - f(E_0)\|_{\boldsymbol{\mathcal{C}}_p} \leq \|f\|_\theta \, \varepsilon$$

as long as $p \geq N/(\theta - \beta) > 0$.

In particular, if $p > N/\theta$, we can find a PPS spectral measure E_0 such that $f(E) - f(E_0) \in \boldsymbol{\mathcal{C}}_p$ for all f which is Hölder continuous of order θ. This result with $\theta = 1$ is the case of Lipschitz continuous f and corresponds to Theorem 2.7.

§ 5. Proof of Theorem 4.2

§ 5.1. CONS associated to graded decompositions

We shall mimic the proof of Theorem 2.1 of [21]. In that proof the use of a CONS associated to dyadic decompositions of the unit cube in $\mathbb{R}^N$ plays a crucial role. In this subsection we shall reproduce construction of [21] in an abstract setting (cf. 5.1.1 and 5.1.2) and recapture Voigt's CONS associated to dyadic decomposition in 5.1.3. One slight difference is that the first grade ($\mathcal{G}_0$ in the following notation) consists of a single element in [21] while here it is not necessarily so. This small change will facilitate the following discussion.

5.1.1. Graded decompositions of Ω. For a while we shall work in a measurable space $(\Omega, \mathcal{B})$. Suppose that there is given a sequence of decompositions $\mathcal{G}_0$, $\mathcal{G}_1$, $\cdots$ of Ω into a finite number of subsets so that $\mathcal{G}_{j+1}$ is a sub-decomposition of $\mathcal{G}_j$. More precisely, we suppose that for every $j \in \mathbb{N} \cup \{0\}$ there is given a collection

$$\mathcal{G}_j = \{Q_k^j \in \mathcal{B} \,|\, k = 1, 2, \cdots, r_j\}, \quad r_j = |\mathcal{G}_j| \in \mathbb{N}, \quad Q_k^j \neq \emptyset,$$

of finite number of subsets of Ω such that for each fixed $j \in \mathbb{N} \cup \{0\}$ the set Ω is a disjoint union of Q_k^j, i.e.

$$\Omega = \bigcup_{k=1}^{r_j} Q_k^j, \quad Q_k^j \cap Q_l^j = \emptyset, \quad k \neq l. \tag{5.1}$$

(For later convenience we let the suffix j start from $j = 0$ rather than $j = 1$.) Moreover, we suppose that $\mathcal{G}_{j+1}$ is a sub-decomposition of $\mathcal{G}_j$, $j \in \mathbb{N} \cup \{0\}$, i.e., each Q_k^j is a (disjoint) union of some of Q_m^{j+1}. More precisely, this can be expressed as follows. (i) $r_{j+1} \geq r_j$ and (ii) $\{1, 2, \cdots, r_{j+1}\}$ is divided into a disjoint union of r_j non-empty subsets I_m^{j+1} so that

$$\{1, 2, \cdots, r_{j+1}\} = \bigcup_{m=1}^{r_j} I_m^{j+1}, \quad I_m^{j+1} \neq \emptyset, \quad I_m^{j+1} \cap I_l^{j+1} = \emptyset, \quad m \neq l,$$

$$Q_k^j = \bigcup_{m \in I_k^{j+1}} Q_m^{j+1}. \qquad \text{(disjoint union)}$$

We write

$$\mathcal{G} = \bigcup_{j=0}^{\infty} \mathcal{G}_j$$

and call $\mathcal{G}$ a *league of graded decompositions* of Ω. In abbreviation we write LGD for league of graded decompositions.

5.1.2. Construction of a CONS associated to an LGD Let $(\Omega, \mathcal{B}, \mu)$ be a finite measure space ($0 < \mu(\Omega) < \infty$) and let $\mathcal{G} = \bigcup_j \mathcal{G}_j$ be an LGD of $(\Omega, \mathcal{B})$. Following [21], we shall construct an ONS of $L^2(\Omega, \mathcal{B}, \mu) = L^2(\Omega)$ associated to $\mathcal{G}$.

$\mathcal{G}_0$ plays a special role and subcubes $Q_k^0 \in \mathcal{G}_0$ appear twice in the construction. First we consider the family of functions $\{\chi_{Q_k^0}\}_{k=1,\cdots,r_0}$. From this family we omit those χ which are 0 in $L^2(\Omega)$ and normalize others as $\chi_{Q_k^0}/\|\chi_{Q_k^0}\|$. Since $\chi_{Q_k^0}$ are mutually orthogonal, we thus obtain an ONS which we denote by Θ^0. (Clearly, $\Theta^0 \neq \emptyset$.)

We next construct an ONS Θ_k^j, $j \in \mathbb{N} \cup \{0\}$, $k = 1, \cdots, r_j$, associated to Q_k^j. (Note that $j = 0$ appears again.) We start from the family of functions

$$\{\chi_{Q_k^j}\} \cup \{\chi_{Q_m^{j+1}} \,|\, m \in I_k^{j+1}\}, \tag{5.2}$$

namely, the set of the characteristic function of Q_k^j itself and those of all Q_m^{j+1} that are members of the decomposition of Q_k^j in $\mathcal{G}_{j+1}$. Let s_k^j be the dimension of the linear subspace of $L^2(\Omega)$ spanned by functions appearing in (5.2). Since $\chi_{Q_k^j} = \sum_{m \in I_k^{j+1}} \chi_{Q_m^{j+1}}$ and $\chi_{Q_m^{j+1}}$ are mutually orthogonal for different m, s_k^j is the number of non-zero (in L^2) functions among $\chi_{Q_m^{j+1}}$ with $m \in I_k^{j+1}$.

We set

$$\Theta_k^j = \emptyset, \qquad \text{if} \quad s_k^j = 0 \quad \text{or} \quad 1.$$

When $s_k^j \geq 2$ we apply the Schmidt orthogonalization process to the family of all functions appearing in (5.2) in such a way that the first element of the resulting ONS is $\chi_{Q_k^j}/\|\chi_{Q_k^j}\|$. Remove $\chi_{Q_k^j}/\|\chi_{Q_k^j}\|$ from that ONS and call the remaining family Θ_k^j.

We now put

$$\Theta = \Theta^0 \bigcup \left(\bigcup_{j \in \mathbb{N} \cup \{0\}} \bigcup_{k=1}^{r_j} \Theta_k^j \right).$$

The following three properties are immediate consequences of the construction.

(A) $\varphi \in \Theta^0$ implies $\operatorname{supp} \varphi \subset Q_k^0$ for some k and $\varphi \in \Theta_k^j$ implies $\operatorname{supp} \varphi \subset Q_k^j$.

(B) Any $\varphi \in \Theta_k^j$ is orthogonal to any function which is constant on Q_k^j.

(C) Let $j \in \mathbb{N}$. Then, $\varphi \in \Theta^0$ or $\varphi \in \Theta_l^{j'}$, $j' < j$, implies φ is constant on Q_k^j.

Proposition 5.1. *Let $\mathcal{M}$ be the closed subspace of $L^2(\Omega)$ spanned by the set $\{\chi_{Q_k^j} \,|\, j \in \mathbb{N} \cup \{0\},\ k = 1, \cdots, r_j\}$. Then, Θ is a CONS of $\mathcal{M}$.*

Proof. (i) Orthogonality. $\Theta_{k_1}^j \perp \Theta_{k_2}^j$, $k_1 \neq k_2$, is obvious by (A). $\Theta^0 \perp \Theta_k^j$ and $\Theta_k^{j_1} \perp \Theta_m^{j_2}$, $j_1 \neq j_2$, are immediate consequences of (B) and (C).

(ii) Completeness. Let $\mathcal{M}'$ be the closed subspace spanned by all $\varphi \in \Theta$. Then, it suffices to prove that $\mathcal{M} = \mathcal{M}'$. It is clear from the construction that any $\varphi \in \Theta$ is a linear combination of functions in $\{\chi_{Q_k^j}\}$. Hence, $\mathcal{M}' \subset \mathcal{M}$.

To prove $\mathcal{M} \subset \mathcal{M}'$ it suffices to show that

$$\chi_{Q_k^j} \in \text{l.h.}\{\Theta\}, \qquad \forall j \in \mathbb{N} \cup \{0\}, \quad \forall k = 1, \cdots, r_j, \tag{5.3}$$

where l.h.$\{\Theta\}$ is the set of all finite linear combinations of functions in Θ. We shall prove (5.3) by induction in j.

For $j = 0$ (5.3) is obvious because either $\chi_{Q_k^0} = 0$ in $L^2(\Omega)$ or $\chi_{Q_k^0}/\|\chi_{Q_k^0}\| \in \Theta^0$. Now, suppose that (5.3) has been proved for a certain j. Take a Q_k^{j+1}. Since $\mathcal{G}_{j+1}$ is a sub-decomposition of $\mathcal{G}_j$, there exists a unique m such that $Q_k^{j+1} \subset Q_m^j$. Recall that $\Theta_m^j \cup \{\chi_{Q_m^j}/\|\chi_{Q_m^j}\|\}$ is constructed by means of the Schmidt orthogonalization from $\{\chi_{Q_{k'}^{j+1}}\}_{k' \in I_m^{j+1}} \cup \{\chi_{Q_m^j}\}$, of which $\chi_{Q_k^{j+1}}$ is a member. This means that $\chi_{Q_k^{j+1}} \in$ l.h.$\{\Theta_m^j \cup \{\chi_{Q_m^j}\}\}$. Since $\chi_{Q_m^j} \in$ l.h.$\{\Theta\}$ by induction hypothesis, we conclude that $\chi_{Q_k^{j+1}} \in$ l.h.$\{\Theta\}$. This completes the proof of (5.3) and hence the proof of Proposition 5.1. □

5.1.3. Graded dyadic decompositions of the cube Q_0 and associated CONS

We now consider the case that Ω is a unit cube in $\mathbb{R}^N : \Omega = Q_0 = [0,1]^N$, and construct decompositions which may be called a *league of graded dyadic decompositions* (abbr. LGDD). We agree that all cubes are cubes whose sides are parallel to the coordinate axis. For a cube Q we denote by $l(Q)$ the side length of Q.

As the zeroth step we put $\mathcal{G}_0 = \{Q_0\}$. In the first step we bisect each side of Q_0. Then, Q_0 is decomposed in an obvious way into 2^N subcubes with $l(Q) = 2^{-1}$. (In this decompositions subcubes are half open cubes except those meeting the boundary of Q_0, to which some portion of the boundary is attached to ensure the validity of (5.1). It would not be necessary to elaborate this point.) In the second step we apply the same procedure to each subcube obtained in the first step. Repeating this procedure, Q_0 is decomposed at the jth step into 2^{jN} subcubes with $l(Q) = 2^{-j}$.

We let $\mathcal{G}_j$ be the set of all these subcubes Q_k^j obtained at the jth step and put

$$\mathcal{G} = \bigcup_{j=0}^{\infty} \mathcal{G}_j, \quad \mathcal{G}_j = \{Q_k^j \,|\, k = 1, 2, \cdots, 2^{jN}\}. \tag{5.4}$$

It is clear that $\mathcal{G}$ is an LGD of Q_0. We call subcubes Q_k^j *dyadic subcubes* of grade j. In this example we have $r_j = |\mathcal{G}_j| = 2^{jN}$.

More generally, we shall use decompositions which starts from dyadic cubes whose side length is 2^{-J}. Let $\mathcal{G}_j$ be as above. Forming the union from $j = J \in \mathbb{N} \cup \{0\}$, we put

$$\mathcal{G}^{(J)} = \bigcup_{j=J}^{\infty} \mathcal{G}_j, \quad \text{i.e.} \quad \mathcal{G}_j^{(J)} = \mathcal{G}_{J+j}. \tag{5.5}$$

$\mathcal{G}^{(J)}$ is also an LGD of Q_0. $\mathcal{G}^{(0)}$ coincides with $\mathcal{G}$ of (5.4). (In $\mathcal{G}_j^{(J)}$ the superfix distinguishes different LGD's, while the suffix indicates the grade of a decomposition in $\mathcal{G}^{(J)}$.)

CONS associated to $\mathcal{G}^{(J)}$. We next suppose that we have a finite measure μ on Q_0 and construct CONS $\Theta^{(J)}$ of $L^2(Q_0, \mu)$ associated to $\mathcal{G}^{(J)}$ following the recipe given in 5.1.2. First, for each Q_k^j of (5.4) we construct an ONS Θ_k^j as in 5.1.2. Since each Q_k^j is decomposed into 2^N subcubes of grade $j+1$, it follows from the construction that $|\Theta_k^j| \le 2^N - 1$. Put

$$\Theta^{J,0} = \left\{ \|\chi_{Q_k^J}\|^{-1} \chi_{Q_k^J} \;\middle|\; k \in \{1, \cdots, 2^{JN}\},\ \chi_{Q_k^J} \ne 0 \text{ in } L^2(Q_0; \mu) \right\}.$$

Then $\Theta^{(J)}$ associated to $\mathcal{G}^{(J)}$ is written as

$$\Theta^{(J)} = \Theta^{J,0} \bigcup \left(\bigcup_{j=J}^{\infty} \bigcup_{k=1}^{r_j} \Theta_k^j \right).$$

Since $\{\chi_{Q_k^j} \mid j \ge J,\, k = 1, \cdots, 2^{jN}\}$ spans a dense linear subspace of $L^2(Q_0, \mu)$, we see that $\Theta^{(J)}$ is a CONS of $L^2(Q_0, \mu)$. We note that $\Theta^{(0)}$ is the CONS given in [21].

Remark. Compare two CONS $\Theta^{(J)}$, $J \ge 1$, and $\Theta^{(0)}$. They have $\bigcup_{j=J}^{\infty} \bigcup_{k=1}^{r_j} \Theta_k^j$ in common. Therefore $\Theta^{J,0}$ and $\Phi^{(J)} \equiv \Theta^{0,0} \bigcup \left(\bigcup_{j=0}^{J-1} \bigcup_{k=1}^{r_j} \Theta_k^j \right)$ should span the same space. As an illustration let us verify that $|\Theta^{J,0}| = |\Phi^{(J)}|$, assuming that none of $\chi_{Q_k^j}$ is 0 in L^2. In fact, the left hand side is equal to 2^{JN}, while the right hand side is equal to $1 + \sum_{j=0}^{J-1} (2^N - 1) 2^{jN} = 2^{JN}$.

Projections associated to $\Theta^{J,0}$ and Θ_k^j. Let $\mathcal{M}_k^j$ be the subspace spanned by Θ_k^j and $P_k^j = P_{\mathcal{M}_k^j} \in \mathcal{P}$ the projection onto $\mathcal{M}_k^j$. Denoting elements of Θ_k^j by $\varphi_m^{j,k}$, we write as $\Theta_k^j = \{\varphi_1^{j,k}, \varphi_2^{j,k}, \cdots, \varphi_{|\Theta_k^j|}^{j,k}\}$, where $|\Theta_k^j| \le 2^N - 1$. When $|\Theta_k^j| < 2^N - 1$, we add $\varphi_m^{j,k} = 0$, $|\Theta_k^j| < m \le 2^N - 1$. Then, P_k^j is expressed as

$$P_k^j = \sum_{m=1}^{2^N - 1} (\cdot, \varphi_m^{j,k}) \varphi_m^{j,k} \equiv \sum_{m=1}^{2^N - 1} P_m^{j,k}, \quad \varphi_m^{j,k} = 0 \ \mu\text{-a.e. in } Q_0 \setminus Q_k^j. \tag{5.6}$$

Similarly, let $\mathcal{M}^{J,0}$ be the subspace spanned by $\Theta^{J,0}$ and let $P^{J,0} = P_{\mathcal{M}^{J,0}}$. Then, in a similar way we can express $P^{J,0}$ as

$$(5.7) \qquad P^{J,0} = \sum_{m=1}^{2^{JN}} (\cdot, \phi_m^J)\phi_m^J \equiv \sum_{m=1}^{2^{JN}} P_m^{J,0}, \quad \phi_m^J = 0 \ \mu\text{-a.e. in } Q_0 \setminus Q_m^J.$$

The main difference between (5.6) and (5.7) is that in (5.7) each ϕ_m^J is supported by the cube Q_m^J, which is mutually disjoint, while in (5.6) the support of each summand overlaps each other but $P_k^j u$ as a whole is supported by Q_k^j.

Since $\Theta^{(J)}$ is a CONS we have

$$(5.8) \qquad I = \sum_{m=1}^{2^{JN}} P_m^{J,0} + \sum_{j=J}^{\infty}\sum_{k=1}^{2^{jN}} P_k^j = \sum_{m=1}^{2^{JN}} P_m^{J,0} + \sum_{j=J}^{\infty}\sum_{k=1}^{2^{jN}}\sum_{m=1}^{2^N-1} P_m^{j,k}.$$

§ 5.2. Proof of Theorem 4.2

5.2.1. Reduction We may assume that $Q \subset Q_0 = [0,1]^N$. Then, in view of Proposition 4.1 the problem is reduced to the case that

$$(5.9) \qquad \mathfrak{H} = \sum_{n=1}^{\infty} \oplus L^2(Q_0; \mu_n), \quad E = \sum_{n=1}^{\infty} \oplus E_n,$$

where E_n is an $L^2(Q_0; \mu_n)$-spectral measure determined by the multiplication of characteristic functions: $E_n(\omega)u(\lambda) = \chi_\omega(\lambda)u(\lambda)$, $u \in L^2(Q_0; \mu_n)$. Here, we restricted our attention to the case that $M = \infty$ in (4.8). The proof for the case that $M \in \mathbb{N}$ is simpler as it does not involve an infinite direct sum.

The PPS spectral measure E_0 will be constructed as a direct sum $E_0 = \sum_{n=1}^{\infty} \oplus E_{0,n}$. Recall LGDD $\mathcal{G}^{(J)}$of Q_0 introduced in 5.1.3 (cf. (5.5)). The idea is to use finer decompositions (i.e. bigger J) as n in (5.9) increases. For that purpose we look up the sequence η_j appearing in Theorem 4.2. Since $\sum_{j=1}^{\infty} \eta_j < \infty$, we can choose J_n in such a way that

$$(5.10) \qquad \sum_{j=J_n}^{\infty} \eta_j < c_N 2^{-n}\varepsilon,$$

where c_N is a constant, depending only on N, to be determined later. Fixing such a sequence $\{J_n\}_{n\in\mathbb{N}}$, we introduce into $L^2(Q_0, \mu_n)$ the LGDD $\mathcal{G}^{(J_n)}$, the corresponding CONS $\Theta^{(J_n)}$, and the corresponding decomposition (5.8) of I. Here $\mathcal{G}^{(J_n)}$ does not depend on the measure μ_n but the CONS $\Theta^{(J_n)}$ and the associated projections do depend on μ_n. To avoid excessive complications, we do not indicate this dependence on μ_n in the notation.

We now choose and fix $\xi_k^j \in Q_k^j$ for each j, k. The family $\{\xi_k^j\}$ will be used common to all n. On the basis of (5.9), (5.8) we define $E_{0,n}$ and E_0 as follows:

$$E_0 = \sum_{n=1}^{\infty} \oplus E_{0,n}, \quad E_{0,n}(\omega) = \sum_{\{m | \xi_m^{J_n} \in \omega\}} P_m^{J_n,0} + \sum_{j=J_n}^{\infty} \sum_{\{k | \xi_k^j \in \omega\}} P_k^j. \tag{5.11}$$

It is obvious that $E_{0,n}$ is a PPS $L^2(Q_0;\mu_n)$-spectral measure and E_0 is a PPS $\mathfrak{H}$-spectral measure. Furthermore, we have $f(E) - f(E_0) = \sum_{n=1}^{\infty} \oplus (f(E_n) - f(E_{0,n}))$. It is now clear that the proof of Theorem 4.2 is reduced to the proof of the following proposition.

Proposition 5.2. *Suppose that* $\boldsymbol{\mathcal{C}}$ *and* $w \in \boldsymbol{w} \in \mathcal{M}_Q$ *satisfy* (4.10). *Then, for any* $\boldsymbol{w}$*-continuous* $f \in C^{\boldsymbol{w}}(Q)$ *we have* $f(E_n) - f(E_{0,n}) \in \boldsymbol{\mathcal{C}}$ *and*

$$\|f(E_n) - f(E_{0,n})\|_{\boldsymbol{\mathcal{C}}} \leq 2^{-n}\varepsilon |f|_w, \quad \forall n \in \mathbb{N}. \tag{5.12}$$

5.2.2. Proof of Proposition 5.2

From (5.8) and (5.11) it follows that

$$\begin{aligned} f(E_n) - f(E_{0,n}) &= \sum_{m=1}^{2^{J_n N}} (f(E) - f(\xi_m^{J_n}))P_m^{J_n,0} + \sum_{j=J_n}^{\infty} \sum_{k=1}^{2^{jN}} (f(E) - f(\xi_k^j))P_k^j \\ &\equiv A_n + B_n. \end{aligned} \tag{5.13}$$

Noting that each $P_m^{J_n,0}u$ is supported by $Q_m^{J_n}$ and hence $P_m^{J_n,0}u$ and $f(E)P_m^{J_n,0}u$ are mutually orthogonal for different m, we have for any $u \in L^2(Q_0;\mu_n)$

$$\begin{aligned} \|A_n u\|^2 &= \sum_{m=1}^{2^{J_n N}} \left\|(f(E) - f(\xi_m^{J_n}))P_m^{J_n,0}u\right\|^2 \\ &= \sum_{m=1}^{2^{J_n N}} \int_{Q_m^{J_n}} |f(\xi) - f(\xi_m^{J_n})|^2 d\|E(\xi)P_m^{J_n,0}u\|^2. \end{aligned} \tag{5.14}$$

On the right hand side of (5.14) ξ and $\xi_m^{J_n}$ both belong to $Q_m^{J_n}$ so that $|\xi - \xi_m^{J_n}| \leq N^{1/2} l(Q_m^{J_n}) = N^{1/2}2^{-J_n}$. Put $d_N = [N^{1/2}]+1$, where $[a]$ denotes the largest integer not exceeding a. Since f is $\boldsymbol{w}$ continuous and $w \in \boldsymbol{w}$, it follows from (3.4) that

$$|f(\xi) - f(\xi_m^{J_n})| \leq |f|_w w(N^{1/2}2^{-J_n}) \leq |f|_w w(d_N 2^{-J_n}) \leq d_N |f|_w w(2^{-J_n}), \tag{5.15}$$

where the last inequality is a consequence of (3.2) which w satisfies. Inserting (5.15) in (5.14), we see that

$$\|A_n u\|^2 \leq d_N^2 |f|_w^2 w(2^{-J_n})^2 \sum_{m=1}^{2^{J_n N}} \int_{Q_m^{J_n}} d\|E(\xi)P_m^{J_n,0}u\|^2 \leq d_N^2 |f|_w^2 w(2^{-J_n})^2 \|u\|^2,$$

and hence that $\|A_n\| \le d_N|f|_w w(2^{-J_n})$. The operator A_n is of finite rank with the rank at most 2^{J_nN}. Therefore by Proposition 2.6 we finally conclude that

$$\begin{aligned}\|A_n\|_{\mathcal{C}} &\le \nu_{\mathcal{C}}(2^{J_nN})\|A_n\| \le d_N\nu_{\mathcal{C}}(2^{J_nN})|f|_w w(2^{-J_n}) \\ &\le d_N|f|_w\eta_{J_n} \le d_N|f|_w \sum_{j=J_n}^{\infty}\eta_j,\end{aligned} \tag{5.16}$$

where the third inequality follows from the assumption (4.10).

We shall next estimate B_n appearing in (5.13). The summation over j will be handled later by the triangle inequality. Using (5.6), the sum over k can be written as

$$\begin{aligned}&\sum_{k=1}^{2^{jN}}(f(E)-f(\xi_k^j))P_k^j \\ &\quad= \sum_{k=1}^{2^{jN}}\sum_{m=1}^{2^N-1}(f(E)-f(\xi_k^j))P_m^{j,k} = \sum_{m=1}^{2^N-1}\sum_{k=1}^{2^{jN}}(f(E)-f(\xi_k^j))P_m^{j,k}.\end{aligned} \tag{5.17}$$

We note that the ranges of $P_m^{j,k}$ are mutually orthogonal. The ranges of $(f(E) - f(\xi_k^j))P_m^{j,k}$, however, lose the orthogonality for different m with j, k being fixed, while they keep the orthogonality for different k with j, m being fixed. Thus, the orthogonality argument cannot be applied to $\sum_m$ in the middle member of (5.17) but can be applied to $\sum_k$ in the rightmost member. We borrow this clever trick from [21] and estimate B_n as follows.

First, $\sum_k$ is estimated as in the case of A_n with the result

$$\left\|\sum_{k=1}^{2^{jN}}(f(E)-f(\xi_k^j))P_m^{j,k}u\right\|^2 \le d_N^2|f|_w^2 w(2^{-j})^2\|u\|^2.$$

Since $\sum_k$ is of rank at most 2^{jN}, it follows that

$$\left\|\sum_k\right\|_{\mathcal{C}} \le d_N\nu_{\mathcal{C}}(2^{jN})|f|_w w(2^{-j}) \le d_N|f|_w\eta_j.$$

Then we use the triangle inequality to estimate $\mathcal{C}$-norm of the sums $\sum_m$ and $\sum_j$. Since the sum $\sum_m$ consists of at most 2^N-1 non-zero summands, we obtain

$$\|B_n\|_{\mathcal{C}} \le d_N(2^N-1)|f|_w \sum_{j=J_n}^{\infty}\eta_j. \tag{5.18}$$

Finally, it follows from (5.16), (5.18), and (5.10) that

$$\begin{aligned}\|f(E_n)-f(E_{0,n})\|_{\mathcal{C}} &\le (\|A_n\|_{\mathcal{C}} + \|B_n\|_{\mathcal{C}}) \\ &\le 2^N d_N|f|_w \sum_{j=J_n}^{\infty}\eta_j \le 2^N d_N|f|_w c_N 2^{-n}\varepsilon.\end{aligned} \tag{5.19}$$

We now take c_N of (5.10) as $c_N = (2^N d_N)^{-1}$. Then, the right hand side of (5.19) is equal to that of (5.12). This conclude the proof of Proposition 5.2. $\square$

Acknowledgments Part of the present work was done while the author was visiting the Department of Mathematics, Tohoku University, by the support of the global COE program" Weaving Science Web beyond Particle-Matter Hierarchy "of MEXT, Japan. It is our pleasure to thank Professor Hideo Kozono for taking the initiative in inviting us and providing an excellent working condition, and all members of the department for their hospitality. We thank Professor Jürgen Voigt for adequate comments he made upon seeing the first draft of this article and for sending us reference [22]. We also thank Professor Hiroaki Aikawa for a useful conversation which led the author to the use of the modulus of continuity in the formulation of Theorem 4.2.

References

[1] Aikawa, H., Modulus of continuity of the Dirichlet solutions, to appear in *Bull. London Math. Soc.*

[2] Bercovici, H. and Voiculescu, D., The analogue of Kuroda's theorem for n tuples, *Operator Theory: Advances and Applications,* vol. 41, Birkuhäuser, 1989, pp. 57–60.

[3] Berg, I. D., An extension of the Weyl-von Neumann theorem to normal operaotors, *Trans. Amer. Math. Soc.*, **160**(1971), 365–371.

[4] Birman, M. S. and Solomjak, M. Z., Spectral theory of self-adjoint operators in Hilbert space, *Translated from the Russian by S. Khrushchev and V. Peller, Math. Appl. (Soviet Series)*, D. Reidel, Dordrecht, 1987.

[5] Carey, R. W. and Pincus, J. D., Unitary equivalence modulo the trace class self-adjoint operators, *Amer. J. Math.*, **98**(1976), 481–514.

[6] Gohberg , I. C. and Kreĭn, M.G., Introduction to the Theory of Linear Nonselfadjoint Operators, *Translations Amer. Math. Soc.*, **18**, 1969.

[7] Halmos, P., Quasitriangular operators, *Acta Sci. Math.*, **29**(1968), 284–293.

[8] Kato, T., Perturbation of continuous spectra by trace class operators, *Proc. Japan Acad.*, **33**(1957), 260–264.

[9] Kuroda, S. T., On a theorem of Weyl-von Neumann, *Proc. Japan Acad.*, **34**(1958), 11–15.

[10] Kuroda, S. T., Estimates of Kato-Temple type for n-dimensional spectral measures, *Publ. RIMS, Kyoto Univ.,* **43**(2007), 505–520.

[11] Neumann, J. von, Charakterisierung des Spektrums eines Integraloperators, *Actualités Sci. Ind.*, **229**(1935), 33–55.

[12] Okayasu, R., Entropy of subshifts and the Macaev norm, *J. Math. Soc. Japan,* **56**(2004), 177–191.

[13] Rosenblum, M., Perturbation of continuous spectra and unitary equivalence, *Pacific J. Math.*, **7**(1957), 997–1010.

[14] Sikonia, W., The von Neumann converse of Weyl's theorem, *Indiana Univ. Math. J.*, **21**(1971), 121–124.

[15] Voiculescu, D., Some extensions of quasitriangularity, I–II *Rev. Roumaine Math. Pures Appl.* **18**(1973), 1303–1320; *ibid.*, 1440–1456.

[16] Voiculescu, D., Some results on norm-ideal perturbations of Hilbert space operators, I, *J. Operator Theory*, **2**(1979), 3–37.

[17] Voiculescu, D., Some results on norm-ideal perturbations of Hilbert space operators. II, *J. Operator Theory*, **5**(1981), 77–100.

[18] Voiculescu, D., Hilbert space operators modulo norm ideals, *Proc. Intern. Congress of Mathematicians*, Warsaw, 1983, 1041–1047.

[19] Voiculescu, D., On the existence of quasicentral approximate units relative to norm ideals, Part I, *J. Functional Anal.*, **91**(1990), 1-36.

[20] Voiculescu, D., Perturbations of operators, connections with singular integrals, hyperbolicity and entropy, *Harmonic Analysis and Discrete Potential Theory*, Picardello, M. A., ed., *Plenum Press*, 1992, 181–191.

[21] Voigt, J., Perturbation theory for commutative m-tuples of self-adjoint operators, *J. Functional Anal.*, **25**(1977), 317–334.

[22] Weidmann, J., Das Weyl-von Neumann'sche Theorem für normale Operatoren, *unpublished.*

[23] Weyl, H., Über beschränkte quadratische Formen, deren Differenz vollstetig ist, *Rend. Circ. Mat. Palermo*, **27**(1909), 373–392.

[24] Xia, J. Diagonalization modulo norm ideals with Lipschitz estimate, *J. Functional Anal.*, **145**(1997), 491–526.

[25] Xia, J., Diagonalization and unitary equivalence modulo Schatten p-classes, *J. Functional Anal.*, **175**(2000), 279–307.

[26] Xia, J., Singular integral operators and norm ideals satisfying a quantitative variant of Kuroda's condition, *J. Functional Anal.*, **228**(2005), 369–393.

[27] Xia, J., A condition for diagonalization modulo arbitrary norm ideals, *J. Funct. Anal.*, **255**(2008), 1039–1056.

[28] Xia, J., One hundred years of diagonalization,
http://www.acsu.buffalo.edu/~jxia/Preprints.html