The Asymptotic Linearity Theorem for the study of additivity problems of the zero-point vibrational energy of hydrocarbons and the total pi-electron energy of alternant hydrocarbons

Shigeru Arimoto*

Department of Chemistry, University of Saskatchewan, Saskatoon, Saskatchewan, Canada S7N 0W0

and

Mark Spivakovsky

Department of Mathematics, University of Toronto, Toronto, Ontario, Canada M5S 1A4

Received 10 October 1991; in final form 22 July 1992

By using the approach via the aspect of form and general topology, as well as basic notions of abstract algebra, a theoretical framework has been developed which elucidates the mechanism of the additivity relationships between structure and properties in molecules having many identical moieties. The main theorem, the Asymptotic Linearity Theorem (ALT), together with an auxiliary theorem, the α Independence Theorem, implies that the zero-point vibrational energy (or total pi-electron energy for the case of alternant hydrocarbons) E_n of a linearly extended system $B-A_n-B'$ having n repeating identical moieties has the asymptotic expansion $E_n = \alpha n + \beta + o(1)$ as $n \to \infty$, where $\alpha \in \mathbb{R}$ is independent of the choice of the end moieties B and B$'$. The theorem being formulated in a general context, the actual implication of the ALT is much broader than the above two applications would indicate.

1. Introduction

Strong additive correlation between molecular structure and the zero-point vibrational energies in hydrocarbons has been known empirically for decades. Since Cottrell [1] first pointed out that the increment in zero-point vibrational energy per CH_2 in paraffin hydrocarbons was approximately constant, several empirical additive formulae have been proposed.

*On leave from: Institute for Fundamental Chemistry, 34-4 Nishihiraki-cho, Takano, Sakyo-ku, Kyoto 606, Japan.

Pitzer and Catalano [2] proposed an empirical equation with two parameters for paraffins. Shingu and Fujimoto [3–7] gave a more accurate and more general empirical equation with five parameters for paraffins, olefins, and aromatic hydrocarbons. They incorporated their equation as the zero-point energy term into their empirical equation for the atomic heat of formation of hydrocarbons [4,5,7], which also showed excellent agreement with the observed values.

Recently, Schulman and Disch [8] provided a simple but less accurate empirical additive formula for hydrocarbons,

$$\mathrm{ZPE}(n, m) = 3.88n + 7.12m - 6.19 \text{ (kcal/mol)}, \tag{1.1}$$

where n and m denote the numbers of C and H atoms, respectively. This formula is a refinement of what is possibly the most simple formula by Flanigan et al. [9],

$$\mathrm{ZPE}(n, m) = 2n + 7m \text{ (kcal/mol)}. \tag{1.2}$$

However, until recently, no substantial progress had been made to explain theoretically why the zero-point vibrational energies in those molecules should be additive to such a good approximation despite the delocalized and strongly varying terms that contribute to them. To elucidate the mechanism of this additivity is, in our opinion, one of the fundamental and important theoretical problems in structural chemistry.

The purpose of the present paper is to establish a fundamental existence theorem for the above additivity relations using the approach via the aspect of form [10–13] and general topology [10,11,13–16], as well as basic notions of abstract algebra [17,18].

The main theorem, referred to as the Asymptotic Linearity Theorem (ALT), is based on previously established theorems (cf. section 3), the α Existence Theorem [11], the α Independence Theorem [11], and the α Representation Theorem [13], which have also been proved using the approach via the aspect of form and general topology. Thus, it is now possible to grasp and view these four fundamental theorems from a single perspective.

These theorems are all applicable to the additivity problems of the total pi-electron energy [19–29] of alternant hydrocarbons (cf. refs. [11–13] for detailed arguments for the unification of the problems of the ZPVEs and TPEEs).

In section 2, we select a system of linear chains and formulate a concrete additivity problem of the ZPVEs. We then solve this special problem by proving the Asymptotic Linearity Theorem, which ignores the non-essential specific features and conditions involved in the formulated problem and thus makes it possible to unify the solutions of problems from different branches of molecular science.

2. Formulation of the problem and sketch of a solution

Let Ch_N denote a linear chain with free ends consisting of N particles each of mass 1 and separation 1 which can vibrate harmonically under a restoring force due to the first-neighbour interaction 1 [10, 11, 13]. Set $\hbar/2 = 1$ for simplicity. Then the ZPVE E_N of the linear chain Ch_N is expressed as

$$E_N = \sum_{i=1}^{N} \varphi(\lambda_i(K_N)), \tag{2.1}$$

where K_N denotes the $N \times N$ positive semi-definite real-symmetric matrix given by

$$K_N = \begin{pmatrix} 1 & -1 & & & & & \\ -1 & 2 & -1 & & \text{zeros} & & \\ & -1 & 2 & \cdot & & & \\ & & \cdot & \cdot & \cdot & & \\ & & & \cdot & \cdot & \cdot & \\ & & & & \cdot & 2 & -1 \\ & \text{zeros} & & & -1 & 2 & -1 \\ & & & & & -1 & 1 \end{pmatrix}, \tag{2.2}$$

$\lambda_i(K_N)$ denotes the ith eigenvalue $(\lambda_1(K_N) \le \lambda_2(K_N) \le \ldots \le \lambda_i(K_N) \le \ldots \le \lambda_N(K_N))$ of the real-symmetric matrix K_N, and the φ denotes the continuous function $\varphi(t) = |t|^{1/2}$.

Alternatively, E_N may be written concisely as follows by defining the function of the matrix (see refs. [10–13] for details):

$$E_N = \mathrm{Tr}\ \varphi(K_N). \tag{2.3}$$

Inspection of the graph of numerically obtained E_N shows a strong linear correlation between N and E_N. By numerical calculations, one can observe many similar asymptotic linear relationships between N and E'_N in analogous but more complex linear chains Ch'_N.

However, for simplicity we shall here concentrate on a single case and formulate our problem as follows:

PROBLEM 1

To prove the existence of the asymptotic line $\alpha N + \beta$ for the above ZPVEs E_N of linear chains Ch_N.

Although instructive, the following ad hoc solution (cf. section 2 of ref. [10]) is not usable for general cases and we must look for another solution. This solution depends upon the incidental situation in which all the eigenvalues $\lambda_i(K_N)$ are analytically obtainable, and $\sum_{i=1}^{N} \varphi(\lambda_i(K_N))$ can also be analytically expressed,

$$\begin{aligned} E_N &= \sum_{i=1}^{N} \varphi(4\sin^2((i-1)\pi/2N)) \\ &= \cot(\pi/4N) - 1 \\ &= \alpha N + \beta + o(1), \end{aligned} \tag{2.4}$$

where $\alpha = 4/\pi$, $\beta = -1$, and $o(1)$ denotes a function of N such that $o(1) \to 0$ as $N \to \infty$.

In what follows, we shall make a sketch of another solution of problem 1 which is not ad hoc. Before doing so, it is profitable at this point to review the thought process of the solution of a similar but simpler problem [11]:

PROBLEM 2

To prove the existence of $\lim_{N\to\infty} E_N/N \in \mathbb{R}$ for the ZPVEs E_N of linear chains Ch_N, without using the analytic solution of eigenvalue problems of K_N.

Let $C(I)$ denote the real normed space of all real-valued continuous functions defined on a closed interval $I = [a, b]$ ($a, b \in \mathbb{R}$, $a < b$) equipped with the uniform norm $\|\cdot\|_u$ given by

$$\|\varphi\|_u = \sup_{t\in I} |\varphi(t)|. \tag{2.5}$$

Let $\{M_N\} \in X_r(q)$ be a fixed repeat sequence with block-size q.

Remark

See section 3 for the definitions of the repeat space $X_r(q)$ and the repeat sequence. It is easy to see that the matrix K_N defined by (2.2) is the Nth term of a repeat sequence with block-size 1. Throughout this section, if the reader wishes, he may omit the reference to the definitions of the repeat space $X_r(q)$ and the repeat sequence, and assume that $\{M_N\} = \{K_N\}$ whenever the statement $\{M_N\} \in X_r(q)$ appears.

Let I be a fixed closed interval which contains all the eigenvalues of M_N for all $N \in \mathbb{Z}^+$. Such an interval I will be called compatible with $\{M_N\}$ (cf. ref. [10] for the existence of such an interval).

Define a mapping π_α: $C(I) \to \{T, F\}$ by

$$\pi_\alpha(\varphi) = \begin{cases} T & \text{if } \lim\limits_{N\to\infty} [\mathrm{Tr}\, \varphi(M_N)]/N \in \mathbb{R} \text{ exists,} \\ F & \text{if } \lim\limits_{N\to\infty} [\mathrm{Tr}\, \varphi(M_N)]/N \in \mathbb{R} \text{ does not exist.} \end{cases} \tag{2.6}$$

We could solve problem 2 in two steps (cf. ref. [11] for essentially the same formulation of steps (I) and (II)):

(I) $\pi_\alpha(P) = \{T\}$, (2.7)

(II) $\pi_\alpha(\overline{P}) = \{T\}$, (2.8)

where P denotes the subset of $C(I)$ of all polynomial functions with real coefficients and the upper bar denotes the closure operation in the normed space $C(I)$. (Note that $\overline{P} = C(I)$ by the Weierstrass theorem.)

The approaches used in the above steps (I) and (II) were referred to, respectively, as

(I) the approach via the aspect of form,

(II) the approach via general topology.

We shall employ the same approaches in establishing our Asymptotic Linearity Theorem.

In order to make more precise the meaning of the transition from step (I) to step (II), let us define a topology on the set $\{T, F\}$. Henceforth, we let $(\{T, F\}, o_T)$ (or $\{T, F\}$ for short where no confusion arises) denote the topological space with the underlying set $\{T, F\}$ and the system of open sets $o_T = \{\phi, \{F\}, \{T, F\}\}$. We remark that this topology is stronger than the trivial topology $o_t = \{\phi, \{T, F\}\}$ and weaker than the discrete topology $o_d = \{\phi, \{T\}, \{F\}, \{T, F\}\}$, and that the toplogical space $(\{T, F\}, o_T)$ is not Hausdorff.

With this setting, π_α is a mapping from a topological space to a topological space, thus it is meaningful to ask whether or not π_α is a continuous mapping.

In fact, π_α is continuous. To verify this, first define as in ref. [11] two functionals $\overline{\alpha}$, $\underline{\alpha}$: $C(I) \to \mathbb{R}$ by

$$\overline{\alpha}(\varphi) = \limsup_{N\to\infty} [\mathrm{Tr}\, \varphi(M_N)]/N, \tag{2.9}$$

$$\underline{\alpha}(\varphi) = \liminf_{N\to\infty} [\mathrm{Tr}\, \varphi(M_N)]/N. \tag{2.10}$$

Secondly, define a mapping f_α: $C(I) \to \mathbb{R}^2$ by

$$f_\alpha(\varphi) = (\overline{\alpha}(\varphi), \underline{\alpha}(\varphi)). \tag{2.11}$$

Finally, define a mapping f_Δ: $\mathbb{R}^2 \to \{T, F\}$ by

$$f_\Delta((x, y)) = \begin{cases} T & \text{if } (x, y) \in \Delta, \\ F & \text{if } (x, y) \notin \Delta, \end{cases} \tag{2.12}$$

where $\Delta = \{(x, y) \in \mathbb{R}^2,\ x = y\}$ denotes the diagonal set.

Now consider the following diagram:

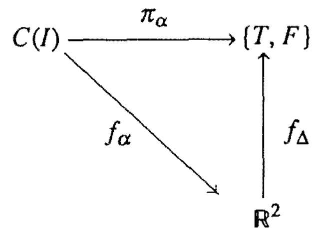


Diagram α.

Recalling the fact that for any bounded real sequence a_n, $\lim_{n\to\infty} a_n$ exists in $\mathbb{R}$ if and only if $\lim\sup_{n\to\infty} a_n = \lim\inf_{n\to\infty} a_n$, we can easily infer that

$$\pi_\alpha = f_\Delta \circ f_\alpha, \tag{2.13}$$

i.e. diagram α is commutative.

Since inverse images of $\{T, F\}$, $\{T\}$, ϕ (all the closed sets of $(\{T, F\}, o_{\mathrm{T}})$) by f_Δ are, respectively, $\mathbb{R}^2$, Δ, ϕ (all closed in $\mathbb{R}^2$), the mapping f_Δ is continuous. On the other hand, we already know that $\overline{\alpha}$ and $\underline{\alpha}$ are both continuous [11]; thus it follows that f_α is also continuous.

Hence, by (2.13), π_α is clearly continuous so that one can deduce eq. (2.8) from eq. (2.7).

The essential feature of the above argument on the transition from step (I) to step (II) can be summarized as follows. Suppose that:

(i) X is a topological space.

(ii) Y is a T_2-space (Hausdorff space), $Y \times Y$ denotes the product space with the box topology, and Δ denotes the diagonal set of $Y \times Y$. (Note that Δ is a closed set since Y is Hausdorff.)

(iii) $Z = (\{T, F\}, o_{\mathrm{T}})$ is the topological space with the system of open sets $o_{\mathrm{T}} = \{\phi, \{F\}, \{T, F\}\}$.

(iv) $f_1, f_2 \colon X \to Y$ are continuous mappings.

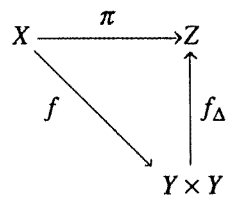
(v) $f \colon X \to Y \times Y$ is a mapping defined by $f(\varphi) = (f_1(\varphi), f_2(\varphi))$. (Note that f is continuous since both f_1 and f_2 are continuous.)

(vi) $f_\Delta \colon Y \times Y \to Z$ is a mapping defined by

$$f_\Delta((u, v)) = \begin{cases} T & \text{if } (u, v) \in \Delta, \\ F & \text{if } (u, v) \notin \Delta, \end{cases} \tag{2.14}$$

(recall the fact that Δ is a closed set and note that f_Δ is continuous).

(vii) $\pi: X \to Z$ is a mapping and the following diagram is commutative:



Then π is continuous. Moreover, for any subset X_0 of X, (I) implies (II):

(I) $\pi(X_0) = \{T\}$, (2.15)

(II) $\pi(\overline{X_0}) = \{T\}$. (2.16)

We shall utilize this thought process for establishing the Asymptotic Linearity Theorem.

Now we can state the outline of the solution of problem 1. Let CBV(I) denote the real normed space of all real-valued continuous functions of bounded variation defined on a closed interval $I = [a, b]$ ($a, b \in \mathbb{R}$, $a < b$) equipped with the norm $\|\cdot\|$ given by

$$\|\varphi\| = \sup_{t \in I} |\varphi(t)| + \sup_{\Delta: a = t_1 \leq t_2 \leq \ldots \leq t_n = b} \sum_{j=1}^{n} |\varphi(t_j) - \varphi(t_{j-1})|. \qquad (2.17)$$

As in the solution of problem 2, let $\{M_N\} \in X_r(q)$ denote a fixed repeat sequence with block-size q. Let I be a fixed closed interval which contains all the eigenvalues of M_N for all $N \in \mathbb{Z}^+$.

Define a mapping π_β: CBV(I) $\to \{T, F\}$ by

$$\pi_\beta(\varphi) = \begin{cases} T & \text{if } \operatorname{Tr} \varphi(M_N) \text{ has an asymptotic line,} \\ F & \text{if } \operatorname{Tr} \varphi(M_N) \text{ does not have an asymptotic line.} \end{cases} \qquad (2.18)$$

We then proceed in three steps to obtain

(I) $\pi_\beta(P) = \{T\}$, (2.19)

(II) $\pi_\beta(\overline{P}) = \{T\}$, (2.20)

(III) $\varphi_\xi \in \overline{P}$ $(\xi > 0)$. (2.21)

Here, P denotes the subset of CBV(I) of all polynomial functions with real coefficients defined on I, the upper bar denotes the closure operation in the normed space CBV(I), and $\varphi_\xi \in$ CBV(I) denotes the function defined by $\varphi_\xi(t) = |t|^\xi$. Note that (II) and (III) solve problem 1.

3. Theory

3.1. PRELIMINARIES

We shall first review necessary terminology and theorems. Throughout, let $\mathbb{Z}^+$, $\mathbb{R}$, $\mathbb{R}_0^+$, $\mathbb{C}$ denote, respectively, the set of all positive integers, real numbers, non-negative real numbers, and complex numbers; and by "for all $N \gg 0$", we mean "for all positive integers N greater than some given positive integer".

Let M denote an $n \times n$ Hermitian matrix and let φ denote a real-valued function defined on a subset $S \subset \mathbb{R}$ such that the subset S contains all the eigenvalues of M. Then we may define the matrix $\varphi(M)$ as

$$\varphi(M) = U \ \mathrm{diag}(\varphi(\lambda_1), \varphi(\lambda_2), \ldots, \varphi(\lambda_n))U^{-1}, \tag{3.1}$$

(cf. refs. [10–13]), where U is an $n \times n$ unitary matrix such that

$$U^{-1}MU = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n). \tag{3.2}$$

Let $I = [a, b]$ ($a, b \in \mathbb{R}$, $a < b$) denote a closed interval and let $P(I)$ denote the set of all polynomial functions with real coefficients defined on I. Suppose that the interval I contains all the eigenvalues of M. Then for any $\varphi \in P(I)$ with $\varphi(t) = c_0t^0 + c_1t^1 + \ldots + c_kt^k$, the matrix $\varphi(M)$, defined as above, may be expressed by

$$\varphi(M) = c_0M^0 + c_1M^1 + \ldots + c_kM^k, \tag{3.3}$$

where M^0 denotes the $n \times n$ unit matrix.

By eq. (3.3), we can rephrase a result which was obtained previously [10] by using the approach via the aspect of form.

THEOREM 1 (PALT)

Let $\{M_N\} \in X_r(q)$ be a fixed repeat sequence, let I be a fixed closed interval compatible with $\{M_N\}$. Then, for any element $\varphi \in P(I)$, there exist $\alpha(\varphi)$, $\beta(\varphi) \in \mathbb{R}$ such that

$$\mathrm{Tr}\ \varphi(M_N\} = \alpha(\varphi)N + \beta(\varphi) \tag{3.4}$$

for all $N \gg 0$.

This theorem is of vital significance for establishing the ALT. We shall call it the Polynomial Asymptotic Linearity Theorem (PALT) in view of the analogous assertion made in the ALT.

Now we shall recall the notion of the repeat space $X_r(q)$ and related terminology [10,11,13].

Fix a $q \in \mathbb{Z}^+$ and let $X(q)$ denote the set of all matrix sequences whose Nth term M_N is an arbitrary $qN \times qN$ real symmetric matrix, $N \in \mathbb{Z}^+$. This set obviously constitutes a linear space over the field $\mathbb{R}$ with term-wise addition and scalar multiplication,

$$\{M_N\} + \{M'_N\} = \{M_N + M'_N\}, \tag{3.5}$$

$$k\{M_N\} = \{kM_N\}, \tag{3.6}$$

$N \in \mathbb{Z}^+$.

We defined three fundamental subspaces $X_r(q)$, $X_\alpha(q)$, and $X_\beta(q)$ of $X(q)$.

Let P_N denote an $N \times N$ real-orthogonal matrix given by

$$P_N = \begin{pmatrix} 0 & 1 & & & & \\ & 0 & 1 & & \text{zeros} & \\ & & 0 & \cdot & & \\ & & & \cdot & \cdot & \\ & & & & \cdot & 1 \\ & \text{zeros} & & & 0 & 1 \\ 1 & & & & & 0 \end{pmatrix}. \tag{3.7}$$

The subspace $X_r(q)$ is defined to be the set of all matrix sequences $\{M_N\} \in X(q)$ such that, for all $N \gg 0$,

$$M_N = A_N + B_N, \tag{3.8}$$

where A_N, B_N are $qN \times qN$ real symmetric matrices having the forms given below:

$$A_N = \sum_{n=-v}^{v} P_N^n \otimes Q_n, \tag{3.9}$$

where v is a non-negative integer, $Q_{-v}, Q_{-v+1}, \ldots, Q_v$ are $q \times q$ real matrices, v and Q_n are constant and independent of N, P_N^n with $n \in \{-2, -3, \ldots\}$ is defined to be $(P_N^{-1})^{-n}$ which equals the transpose of P_N^{-n}, and the symbol $\otimes$ denotes the Kronecker product [30]. (Note that since A_N is symmetric, Q_{-n} must equal the transpose of Q_n for all $n \in \{0, 1, 2, \ldots, v\}$.) A_N is defined for all $N \in \mathbb{Z}^+$ with $N > 2v + 1$.

The matrix B_N has the form

$$B_N = \begin{pmatrix} \left.W_1\right\rfloor & & \left\lfloor W_2\right. \\ & & \\ \left.W_3\right\rceil & & \left\lceil W_4\right. \end{pmatrix}, \tag{3.10}$$

where W_1, W_2, W_3, and W_4 are $qw \times qw$ real matrices, where $w \in \mathbb{Z}^+$, and w and W_j are constant and independent of N. B_N is defined for all $N \in \mathbb{Z}^+$ with $N > 2w$.

Similarly, $X_\alpha(q)$ is defined by setting $M_N = A_N$ in eq. (3.8) and $X_\beta(q)$ by setting $M_N = B_N$.

Note that in the linear space $X(q)$, subspace $X_r(q)$ is the sum of subspaces $X_\alpha(q)$ and $X_\beta(q)$. One can equivalently define $X_r(q)$ to be the sum of these subspaces after defining them first.

We called $X_r(q)$, $X_\alpha(q)$, and $X_\beta(q)$, respectively, the repeat space, alpha space, and beta space with block-size q, and each element of $X_r(q)$, $X_\alpha(q)$, and $X_\beta(q)$, respectively, a repeat sequence, alpha sequence, and beta sequence.

Let $\{M_N\} \in X_r(q)$ be a repeat sequence. A closed interval $I = [a, b]$ ($a, b \in \mathbb{R}$, $a < b$) was said to be compatible with $\{M_N\}$ if all the eigenvalues of M_N are contained in the interval I for all $N \in \mathbb{Z}^+$ (cf. ref. [10] for the existence of such an interval I for any repeat sequence).

The following theorems 2, 3, and 4, together with the above-stated theorem 1, form the basis for proving the ALT.

THEOREM 2 (α EXISTENCE THEOREM)

Let $\{M_N\} \in X_r(q)$ be a fixed repeat sequence, let I be a fixed closed interval compatible with $\{M_N\}$. Then, for any element φ of the normed space $C(I)$, there exists an $\alpha(\varphi) \in \mathbb{R}$ such that

$$[\mathrm{Tr}\ \varphi(M_N)]/N \to \alpha(\varphi) \quad \text{as } N \to \infty. \tag{3.11}$$

THEOREM 3 (α INDEPENDENCE THEOREM)

Let $\{M_N\}$, $\{M'_N\} \in X_r(q)$ be fixed repeat sequences, such that $\{M_N\} - \{M'_N\} \in X_\beta(q)$, let I be a fixed closed interval compatible with both M_N and M'_N. Define two functionals $\alpha, \alpha' : C(I) \to \mathbb{R}$ by

$$\alpha(\varphi) = \lim_{N\to\infty} [\mathrm{Tr}\ \varphi(M_N)]/N, \tag{3.12}$$

$$\alpha'(\varphi) = \lim_{N\to\infty} [\mathrm{Tr}\ \varphi(M'_N)]/N. \tag{3.13}$$

Then,

$$\alpha(\varphi) = \alpha'(\varphi) \tag{3.14}$$

for all $\varphi \in C(I)$; that is, $\alpha = \alpha'$.

THEOREM 4 (α REPRESENTATION THEOREM)

Let $\{M_N\} \in X_r(q)$ be a fixed repeat sequence, and let $F \in H_f(q)$ be the FS map associated with the repeat sequence $\{M_N\}$. Let I be a fixed closed interval compatible with both $\{M_N\}$ and F.

Define two functionals α, α^{int}: $C(I) \to \mathbb{R}$ by

$$\alpha(\varphi) = \lim_{N\to\infty} [\operatorname{Tr} \varphi(M_N)]/N, \tag{3.15}$$

$$\alpha^{\text{int}}(\varphi) = (1/2\pi)\int_{-\pi}^{\pi} \operatorname{Tr} \varphi(F(\theta))\, d\theta, \tag{3.16}$$

$\varphi \in C(I)$. Then,

$$\alpha(\varphi) = \alpha^{\text{int}}(\varphi) \tag{3.17}$$

for all $\varphi \in C(I)$; that is, $\alpha = \alpha^{\text{int}}$.

Remark

For the definition of $H_f(q)$, the FS map, and the compatibility of I with $F \in H_f(q)$, see ref. [13]. In what follows, we shall not use this terminology, but use only hitherto explained notions to minimize the preliminaries.

3.2. ASYMPTOTIC LINEARITY THEOREM

Now we are ready to state our main theorem:

THEOREM 5 (ALT)

Let $\{M_N\} \in X_r(q)$ be a fixed repeat sequence, let I be a fixed closed interval compatible with $\{M_N\}$. Then, for any element $\varphi \in \overline{P}$ in the normed space CBV(I), there exist $\alpha(\varphi)$, $\beta(\varphi) \in \mathbb{R}$ such that

$$\operatorname{Tr} \varphi(M_N) = \alpha(\varphi)N + \beta(\varphi) + o(1) \tag{3.18}$$

as $N \to \infty$.

To prove this theorem, we shall utilize the mapping π_β defined by (2.18). First, observe that statement (I) follows directly from the PALT and that statement (II) is equivalent to the assertion of the ALT:

(I) $\pi_\beta(P) = \{T\} \Leftarrow$ theorem 1 (PALT),

(II) $\pi_\beta(\overline{P}) = \{T\} \Leftrightarrow$ theorem 5 (ALT).

Next, note that if statement (I) is true and π_β is continuous, then statement (II) is true (see the end of this section). Thus, for the proof of the ALT, we see that we have only to prove the continuity of the mapping π_β.

Recall diagram α, and consider the following diagram which involves the π_β, the f_Δ defined in (2.12), and three other mappings f_β, f'_β, and E:

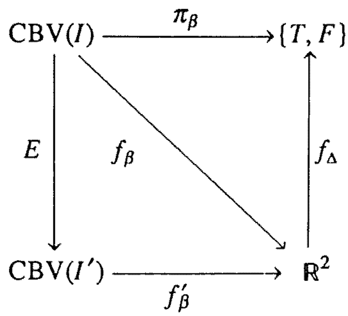


Diagram β.

This diagram possesses, in the upper-right triangular part, an analogous structure to that of diagram α. We shall at first focus attention on this triangular part.

The mapping f_β is defined (recall the analogous definition of f_α) by using two functionals $\overline{\beta}$ and $\underline{\beta}$,

$$f_\beta(\varphi) = (\overline{\beta}(\varphi),\ \underline{\beta}(\varphi)). \tag{3.19}$$

The $\overline{\beta}$, $\underline{\beta}$: $\mathrm{CBV}(I) \to \mathbb{R}$ are defined as follows (that $\overline{\beta}$, $\underline{\beta}$, and f_β are well defined will be clear later):

$$\overline{\beta}(\varphi) = \limsup_{N\to\infty} \beta_N(\varphi), \tag{3.20}$$

$$\underline{\beta}(\varphi) = \liminf_{N\to\infty} \beta_N(\varphi), \tag{3.21}$$

where β_N: $\mathrm{CBV}(I) \to \mathbb{R}$ are linear functionals defined by

$$\beta_N(\varphi) = \mathrm{Tr}\ \varphi(M_N) - \alpha(\varphi)N, \tag{3.22}$$

and where α: $\mathrm{CBV}(I) \to \mathbb{R}$ is the linear functional defined by

$$\alpha(\varphi) = \lim_{N\to\infty} [\mathrm{Tr}\ \varphi(M_N)]/N. \tag{3.23}$$

Note that the limit in (3.23) always exists in $\mathbb{R}$ whenever $\varphi \in \mathrm{CBV}(I) \subset C(I)$ by the α Existence Theorem.

Recall the fact that a real sequence E_N has an asymptotic line, i.e. $E_N = \alpha N + \beta + o(1)$, $N \to \infty$ for some α, $\beta \in \mathbb{R}$, if and only if the limits $\alpha = \lim_{N\to\infty}(E_N/N)$ and $\beta = \lim_{N\to\infty}(E_N - \alpha N)$ exist in $\mathbb{R}$. Thus, bearing in mind that $\pi_\beta(\varphi) = T$ if and only if $\lim_{N\to\infty}\beta_N(\varphi)$ exists in $\mathbb{R}$, one easily verifies that

$$\pi_\beta = f_\Delta \circ f_\beta, \tag{3.24}$$

similar to $\pi_\alpha = f_\Delta \circ f_\alpha$ in diagram α.

Since we know that f_Δ is continuous, to prove that π_β is continuous it is enough to show that f_β is continuous. To prove the continuity of f_β, we now turn our attention to the lower-left triangular part of diagram β.

Fix an $\{M'_N\} \in X_\alpha(q)$ and an $\{M''_N\} \in X_\beta(q)$ such that

$$\{M_N\} = \{M'_N\} + \{M''_N\} \tag{3.25}$$

for all $N \in \mathbb{Z}^+$, and fix a closed interval I' which contains the interval I and is compatible with both $\{M_N\}$ and $\{M'_N\}$.

Define E: $\mathrm{CBV}(I) \to \mathrm{CBV}(I')$ by

$$E(\varphi)(t) = \begin{cases} \varphi(a) & \text{if } t \in [A, a), \\ \varphi(t) & \text{if } t \in [a, b], \\ \varphi(b) & \text{if } t \in (b, B], \end{cases} \tag{3.26}$$

where $I = [a, b]$, $I' = [A, B]$, $(A \le a < b \le B)$. The mapping E is obviously linear, i.e. the relations

$$E(\varphi_1 + \varphi_2) = E(\varphi_1) + E(\varphi_2), \tag{3.27}$$

$$E(k\varphi_1) = kE(\varphi_1), \tag{3.28}$$

hold for all $\varphi_1, \varphi_2 \in \mathrm{CBV}(I)$ and $k \in \mathbb{R}$. Moreover, the linear mapping E is isometric, i.e. the equality

$$\| E(\varphi) \| = \| \varphi \| \tag{3.29}$$

holds for all $\varphi \in \mathrm{CBV}(I)$, which can be easily verified by observing that the CBV norm is expressed as the sum of the uniform norm and the total variation, and that E preserves both of them.

Equations (3.27), (3.28), and (3.29) show that E is a bounded linear operator from the normed space $\mathrm{CBV}(I)$ to the normed space $\mathrm{CBV}(I')$. Thus, we see that E is a continuous mapping.

The definition of f'_β is quite analogous to that of f_β:

$$f'_\beta(\varphi) = (\overline{\beta}^*(\varphi),\ \underline{\beta}^*(\varphi)). \tag{3.30}$$

Here, $\overline{\beta}^*$, $\underline{\beta}^*$: $\mathrm{CBV}(I') \to \mathbb{R}$ are two functionals defined as follows (that they are well defined will be shown below):

$$\overline{\beta}^*(\varphi) = \limsup_{N \to \infty} \beta_N^*(\varphi), \tag{3.31}$$

$$\underline{\beta}^*(\varphi) = \liminf_{N \to \infty} \beta_N^*(\varphi), \tag{3.32}$$

where β_N^*: $\mathrm{CBV}(I') \to \mathbb{R}$ are linear functionals defined by

$$\beta_N^*(\varphi) = \mathrm{Tr}\, \varphi(M_N) - \alpha^*(\varphi) N, \tag{3.33}$$

and where α^*: $\mathrm{CBV}(I') \to \mathbb{R}$ is the linear functional defined by

$$\alpha^*(\varphi) = \lim_{N \to \infty} [\mathrm{Tr}\, \varphi(M_N)]/N. \tag{3.34}$$

Note that the limit in (3.34) always exists in $\mathbb{R}$ whenever $\varphi \in \mathrm{CBV}(I') \subset C(I')$ by the α Existence Theorem.

Now we shall introduce an estimate of $\beta_N^*(\varphi)$ which plays an important role in establishing the ALT.

PROPOSITION ES#

There exists a $c \in \mathbb{R}_0^+$ and an $n \in \mathbb{Z}^+$ such that

$$\text{ES\#:} \quad |\beta_N^*(\varphi)| \leq c \|\varphi\| \tag{3.35}$$

holds for all $\varphi \in \mathrm{CBV}(I')$ and all $N > n$.

In this section, we assume the validity of proposition ES#, whose proof will be given later.

By the definition of $\beta_N^*(\varphi)$ and proposition ES#, $\beta_N^*(\varphi)$ with $\varphi \in \mathrm{CBV}(I')$ is a bounded real sequence so that its limit superior and limit inferior are both real numbers. Thus, we see that both $\overline{\beta}^*$ and $\underline{\beta}^*$ are well defined.

Let $\phi \in \mathrm{CBV}(I)$ be arbitrary, and put $\varphi = E(\phi) \in \mathrm{CBV}(I')$. Then by the definition of E and the function of the matrix, we have $\phi(M_N) = \varphi(M_N)$ for all $N \in \mathbb{Z}^+$, from which is follows that

$$\beta_N(\phi) = \beta_N^*(\varphi) \tag{3.36}$$

for all $N \in \mathbb{Z}^+$. Thus, we can infer that $\overline{\beta}$ and $\underline{\beta}$ in (3.20), (3.21) are also well defined.

Moreover, by (3.36) one easily obtains that

$$f_\beta = f'_\beta \circ E. \tag{3.37}$$

Hence, by (3.24) and (3.37), we obtain:

PROPOSITION 1

Diagram β is commutative.

We wish to have the following:

PROPOSITION 2

All the mappings (E, f'_β, f_β, f_Δ, and π_β) in diagram β are continuous.

Proof

In view of proposition 1, and the continuity of E and f_Δ, which we have already demonstrated, for the proof of proposition 2 we have only to show the continuity of f'_β, or equivalently, the continuity of both $\overline{\beta}^*$ and $\underline{\beta}^*$.

Let $\varphi_1, \varphi_2 \in \mathrm{CBV}(I')$ be arbitrary. Recall the fundamental inequalities

$$|\limsup_{N\to\infty} a_N - \limsup_{N\to\infty} b_N| \le \limsup_{N\to\infty} |a_N - b_N|, \tag{3.38}$$

$$|\liminf_{N\to\infty} a_N - \liminf_{N\to\infty} b_N| \le \limsup_{N\to\infty} |a_N - b_N|, \tag{3.39}$$

which are valid for any bounded real sequences a_N and b_N. Inserting $a_N = \beta_N^*(\varphi_1)$, $b_N = \beta_N^*(\varphi_2)$, we obtain

$$|\overline{\beta}^*(\varphi_1) - \overline{\beta}^*(\varphi_2)| \le \limsup_{N\to\infty} |\beta_N^*(\varphi_1) - \beta_N^*(\varphi_2)|, \tag{3.40}$$

$$|\underline{\beta}^*(\varphi_1) - \underline{\beta}^*(\varphi_2)| \le \limsup_{N\to\infty} |\beta_N^*(\varphi_1) - \beta_N^*(\varphi_2)|. \tag{3.41}$$

On the other hand, by the linearity of β_N^* and proposition ES#, there is a $c \in \mathbb{R}_0^+$ such that the inequality

$$\limsup_{N\to\infty} |\beta_N^*(\varphi_1) - \beta_N^*(\varphi_2)| \le c\,\|\varphi_1 - \varphi_2\| \tag{3.42}$$

holds independently of the choice of $\varphi_1, \varphi_2 \in \mathrm{CBV}(I')$. Therefore, we obtain

$$|\overline{\beta}^*(\varphi_1) - \overline{\beta}^*(\varphi_2)| \le c\,\|\varphi_1 - \varphi_2\|, \tag{3.43}$$

$$|\underline{\beta}^*(\varphi_1) - \underline{\beta}^*(\varphi_2)| \le c\,\|\varphi_1 - \varphi_2\|, \tag{3.44}$$

for all $\varphi_1, \varphi_2 \in \mathrm{CBV}(I')$, which imply that both $\overline{\beta}^*$ and $\underline{\beta}^*$ are Lipschitz continuous, thus continuous. □

Now we can give the

Proof of theorem 5 (ALT)

By theorem 1 (PALT), we have $\pi_\beta(P) = \{T\}$. However, π_β is continuous by proposition 2, so that $\pi_\beta(\overline{P}) \subset \overline{\{T\}} = \{T\}$. On the other hand, clearly $\pi_\beta(\overline{P}) \supset \pi_\beta(P) = \{T\}$. Therefore, $\pi_\beta(\overline{P}) = \{T\}$. □

3.3. ESTIMATE ES#

The proof of proposition ES# can be reduced to proving the following two propositions, ES1 and ES2. Using the α Independence Theorem and the triangle inequality, one easily sees that the assertion of proposition ES# immediately follows from those of propositions ES1 and ES2.

PROPOSITION ES1

There exists a $c \in \mathbb{R}_0^+$ and an $n \in \mathbb{Z}^+$ such that

$$\text{ES1:} \quad |\operatorname{Tr} \varphi(M_N) - \operatorname{Tr} \varphi(M'_N)| \le c \|\varphi\| \tag{3.45}$$

holds for all $\varphi \in \mathrm{CBV}(I')$ and all $N > n$.

PROPOSITION ES2

There exists a $c \in \mathbb{R}_0^+$ and an $n \in \mathbb{Z}^+$ such that

$$\text{ES2:} \quad |\operatorname{Tr} \varphi(M'_N) - \alpha'(\varphi)N| \le c \|\varphi\| \tag{3.46}$$

holds for all $\varphi \in \mathrm{CBV}(I')$ and all $N > n$, where $\alpha'(\varphi)$ is defined by $\alpha'(\varphi) = \lim_{N\to\infty} [\operatorname{Tr} \varphi(M'_N)]/N$.

The following two sections are devoted to establishing the above propositions ES1 and ES2.

3.4. ESTIMATE ES1

To prove proposition ES1, we need the following well-known theorem [31,32] and lemma 1.

THEOREM (STURMIAN SEPARATION THEOREM)

Let H denote an $\infty \times \infty$ Hermitian matrix ($H_{ij} = (H_{ji})^* \in \mathbb{C}$ for all $i, j \in \mathbb{Z}^+$). Consider the sequence H_N of $N \times N$ Hermitian matrices defined by

$$(H_N)_{ij} = (H)_{ij}, \tag{3.47}$$

$i, j \in \{1, 2, \ldots, N\}$, $N \in \mathbb{Z}^+$. Let $\lambda_h(H_N)$ denote the hth eigenvalue of H_N ($\lambda_1(H_N) \leq \lambda_2(H_N) \leq \ldots \leq \lambda_h(H_N) \leq \ldots \leq \lambda_N(H_N)$). Then for each $N \in \mathbb{Z}^+$,

$$\lambda_h(H_{N+1}) \leq \lambda_h(H_N) \leq \lambda_{h+1}(H_{N+1}) \tag{3.48}$$

holds for all $h \in \{1, 2, \ldots, N\}$.

Before stating lemma 1, we fix some notation. Let $I = [a, b]$ be a closed interval. By BV(I), we shall denote the set of all real-valued functions of bounded variation defined on I. By $V_I(\varphi)$, with $\varphi \in \mathrm{BV}(I)$, we shall denote the total variation of function φ on interval I.

LEMMA 1

Let $n \in \mathbb{Z}^+$ with $n \geq 2$, let $K = \{k_1, k_2, \ldots, k_r\}$ be a subset of $\{1, 2, \ldots, n\}$ consisting of r distinct elements ($1 \leq r < n$), and let $L = \{1, 2, \ldots, n\} \backslash K$. Let M and M' be $n \times n$ Hermitian matrices such that the ijth entries of M and M' coincide for all $(i, j) \in L \times L$, i.e. such that

$$(M - M')_{ij} = 0 \tag{3.49}$$

for all $(i, j) \in L \times L$. Let $I = [a, b]$ be a closed interval which contains all the eigenvalues of both M and M'. Then we have

$$|\mathrm{Tr}\, \varphi(M) - \mathrm{Tr}\, \varphi(M')| \leq r V_I(\varphi) \tag{3.50}$$

for all $\varphi \in \mathrm{BV}(I)$.

Proof

(1) In the case $K = \{k_1\}$, $k_1 \in \{1, 2, \ldots, n\}$: If $k_1 \neq n$, apply the similarity transformation to M and M' which exchanges the k_1th column (row) with the nth column (row) of the matrices. Under the similarity transformation, the eigenvalues of matrices are invariant; thus, we may assume $k_1 = n$ without loss of generality.

Let M_0 denote the $(n-1) \times (n-1)$ Hermitian matrix defined by

$$(M_0)_{ij} = (M)_{ij}, \tag{3.51}$$

$i, j \in \{1, 2, \ldots, n-1\}$. Put $\lambda_0 = a$, $\lambda_j = \lambda_j(M_0)$ for $j \in \{1, 2, \ldots, n-1\}$, and $\lambda_n = b$. Then, by the Sturmian Separation Theorem, we have

$$\lambda_{h-1} \leq \lambda_h(M),\ \ \lambda_h(M') \leq \lambda_h \tag{3.52}$$

for each $h \in \{1, 2, \dots, n\}$. It follows that

$$|\operatorname{Tr} \varphi(M) - \operatorname{Tr} \varphi(M')| = \left| \sum_{h=1}^{n} \varphi(\lambda_h(M)) - \sum_{h=1}^{n} \varphi(\lambda_h(M')) \right|$$

$$\leq \sum_{h=1}^{n} |\varphi(\lambda_h(M)) - \varphi(\lambda_h(M'))|$$

$$\leq V_I(\varphi). \tag{3.53}$$

(2) $K = \{k_1, k_2, \dots, k_r\}$: Recalling the assumed relation between M and M', M can obviously be expressed as M' + (modification on k_jth column and row, $j \in \{1, 2, \dots, r\}$);

$$M = \delta_0 + (\delta_1 + \delta_2 + \dots + \delta_r), \tag{3.54}$$

where $\delta_0 = M'$, δ_t with $t \in \{1, 2, \dots, r\}$ are $n \times n$ Hermitian matrices whose components are all zero except for the k_tth column and k_tth row. By considering an extension of φ as in (3.26), we may assume that I contains all the eigenvalues of $\sum_{t=0}^{u} \delta_t$.

Thus, applying (1) r times, we have

$$|\operatorname{Tr} \varphi(M) - \operatorname{Tr} \varphi(M')| = \left| \operatorname{Tr} \varphi\left(\sum_{t=0}^{r} \delta_t \right) - \operatorname{Tr} \varphi\left(\sum_{t=0}^{0} \delta_t \right) \right|$$

$$\leq \sum_{s=1}^{r} \left| \operatorname{Tr} \varphi\left(\sum_{t=0}^{s} \delta_t \right) - \operatorname{Tr} \varphi\left(\sum_{t=0}^{s-1} \delta_t \right) \right| \leq r V_I(\varphi). \tag{3.55}$$

□

Now we are ready to give the

Proof of proposition ES1

By the definition of the beta sequence with block-size q, for all integers N greater than some positive integer n, we have

$$M_N - M'_N = M''_N = \begin{pmatrix} \boxed{W_1} & & \boxed{W_2} \\ & \text{zeros} & \\ \boxed{W_3} & & \boxed{W_4} \end{pmatrix}, \tag{3.56}$$

where W_1, W_2, W_3, and W_4 are $qw \times qw$ real matrices, and where $w \in \mathbb{Z}^+$, w and W_j are constant and independent of N. Note that both M_N and M'_N are Hermitian matrices.

Thus, we may apply lemma 1 and obtain

$$|\mathrm{Tr}\,\varphi(M_N) - \mathrm{Tr}\,\varphi(M'_N)| \le 2qwV_{I'}(\varphi) \le 2qw\,\|\varphi\| \tag{3.57}$$

for all $\varphi \in \mathrm{CBV}(I')$ and all integers N greater than the positive integer n. □

3.5. ESTIMATE ES2

Recall first the notions of the alpha space and alpha sequence with block-size q. By the definition of the alpha sequence, M'_N can be expressed by

$$M'_N = \sum_{n=-\upsilon}^{\upsilon} P_N^n \otimes Q_n \tag{3.58}$$

for all $N \gg 0$, where the symbols on the right-hand side are as in (3.9).

To prove proposition ES2, we shall utilize the Hermitian matrix

$$F(\theta) = \sum_{n=-\upsilon}^{\upsilon} (\exp(\mathrm{i}n\theta))Q_n, \tag{3.59}$$

$\theta \in \mathbb{R}$, and the integral representation α^{int} of α given in theorem 4. Namely, we use the equalities

$$\mathrm{Tr}\,\varphi(M'_N) = \sum_{r=1}^{N} \sum_{h=1}^{q} \varphi(\lambda_h(F(2\pi r/N))), \tag{3.60}$$

$$\alpha'(\varphi) = (1/2\pi)\int_0^{2\pi} \sum_{h=1}^{q} \varphi(\lambda_h(F(\theta)))\ \mathrm{d}\theta, \tag{3.61}$$

which are valid for all $\varphi \in \mathrm{CBV}(I') \subset C(I')$ and all $N \gg 0$.

Remark ※ (See ref. [13] for details)

(i) We let $\lambda_h(F(\theta))$ denote the hth eigenvalue of the Hermitian matrix $F(\theta)$ as in the case of real-symmetric matrices.

(ii) For all $N \gg 0$, the set of all the eigenvalues of M'_N is given by $\{\lambda_h(F(2\pi r/N)) \colon h \in \{1, 2, \dots, q\},\ r \in \{1, 2, \dots, N\}\}$.

(iii) The function $f_h\colon \theta \mapsto \lambda_h(F(\theta))$ defined on $\mathbb{R}$ is real-valued, continuous, and periodic with period 2π, for each $h \in \{1, 2, \dots, q\}$. The range of each function f_h is contained in the interval I', which was assumed to contain all the eigenvalues of M'_N for all $N \in \mathbb{Z}^+$. (In other words, the mapping F is compatible with the interval I'.)

Let $I'' = [0, 2\pi]$, and consider the following diagram:

$$\begin{array}{ccc} \mathrm{CBV}(I') & & \\ {\scriptstyle G}\downarrow & \searrow^{\beta'_N} & \\ \mathrm{CBV}(I'') & \xrightarrow[\beta''_N]{} & \mathbb{R} \end{array}$$

Diagram ·※·

Here, mappings β'_N, G, and β''_N are defined by

$$\beta'_N(\varphi) = \mathrm{Tr}\,\varphi(M'_N) - \alpha'(\varphi)N, \tag{3.62}$$

$$G(\varphi)(\theta) = \sum_{h=1}^{q} \varphi(\lambda_h(F(\theta))), \tag{3.63}$$

$$\beta''_N(\varphi) = \sum_{r=1}^{N} \varphi(2\pi r/N) - (1/2\pi)\left(\int_0^{2\pi} \varphi(\theta)\,\mathrm{d}\theta\right)N. \tag{3.64}$$

We shall demonstrate the validity of proposition ES2 in four steps by proving

PROPOSITION ES2′

Let β'_N, G, β''_N be as above. Then the following statements are all true:

(i) There exists a $c \in \mathbb{R}_0^+$ such that

$$\| G(\varphi) \| \leq c \| \varphi \| \tag{3.65}$$

holds for all $\varphi \in \mathrm{CBV}(I')$.

(ii) For all $g \in \mathrm{CBV}(I'')$ and all $N \in \mathbb{Z}^+$,

$$| \beta''_N(g) | \leq \| g \| \tag{3.66}$$

holds.

(iii) Diagram ·※· is commutative for all $N \gg 0$, i.e.

$$\beta'_N = \beta''_N \circ G \tag{3.67}$$

for all $N \gg 0$.

(iv) Proposition ES2 holds.

In what follows, we assume that the following theorem holds:

THEOREM 6

Let υ be a non-negative integer, let $Q_{-\upsilon}, Q_{-\upsilon+1},\ldots,Q_{\upsilon}$ be $q \times q$ real matrices such that $Q_{-n} = Q_n^{\mathrm{T}}$ for all $n \in \{0, 1, 2, \ldots, \upsilon\}$. Let $F(\theta)$ denote the $q \times q$ Hermitian matrix: $F(\theta) = \sum_{n=-\upsilon}^{\upsilon} \mathrm{e}^{in\theta} Q_n$, $\theta \in [0, 2\pi]$. Let $f_h(\theta)$ denote the hth eigenvalue of $F(\theta)$. Then function f_h is piecewise monotone for each $h \in \{1, 2, \ldots, q\}$. (By "piecewise monotone", we mean that there is a finite partition $0 = x_0 < x_1 < \ldots < x_m = 2\pi$ of the interval $[0, 2\pi]$ such that f_h is monotone on each subinterval $[x_{j-1}, x_j]$, $j \in \{1, 2, \ldots, m\}$.)

The proof of this theorem will be given later.

Proof of theorem ES2

(i) Let φ be any element of CBV(I'). Consider any partition Δ: $0 = \theta_0 \le \theta_1 \le \theta_2 \le \ldots \le \theta_n = 2\pi$ of the interval $I'' = [0, 2\pi]$. Then we obtain

$$\sum_{i=1}^{n} |G(\varphi)(\theta_i) - G(\varphi)(\theta_{i-1})| \le \sum_{h=1}^{q} \sum_{i=1}^{n} |\varphi(f_h(\theta_i)) - \varphi(f_h(\theta_{i-1}))|$$

$$\le \left(\sum_{h=1}^{q} m(h) \right) V_{I'}(\varphi), \tag{3.68}$$

where $m(h)$ denotes the number of intervals of monotonicity of f_h on $[0, 2\pi]$. Here we have used theorem 6, which states that $m(h) < \infty$ for all $h \in \{1, 2, \ldots, q\}$. Straight from the definition of the total variation, we obtain

$$V_{[0,2\pi]}(G(\varphi)) \le \left(\sum_{h=1}^{q} m(h) \right) V_{I'}(\varphi). \tag{3.69}$$

By (3.68) and the easily verifiable relation

$$\sup_{\theta \in [0, 2\pi]} |G(\varphi)(\theta)| \le q \sup_{t \in I'} |\varphi(t)|, \tag{3.70}$$

one immediately obtains

$$\|G(\varphi)\| \leq q \sup_{t \in I'} |\varphi(t)| + \left(\sum_{h=1}^{q} m(h) \right) V_{I'}(\varphi)$$

$$\leq \max\left(q, \sum_{h=1}^{q} m(h) \right) \|\varphi\| \tag{3.71}$$

for all $\varphi \in \mathrm{CBV}(I')$.

(ii) The validity of part (ii) follows from the following

THEOREM 7

Let $[a, b]$ denote a fixed closed interval on $\mathbb{R}$. Let $\beta_N : \mathrm{CBV}[a, b] \to \mathbb{R}$ denote linear functionals defined by

$$\beta_N(g) = \sum_{r=1}^{N} g(x(N, r)) - (1/(b-a))\left(\int_a^b g(\theta)\,\mathrm{d}\theta \right) N, \tag{3.72}$$

$N \in \mathbb{Z}^+$, where $x(N, r)$ denotes the real number:

$$x(N, r) = a + (b - a)r/N, \tag{3.73}$$

i.e. the coordinate of the rth point of N-equipartition of $[a, b]$. Then we have

$$(1) \qquad |\beta_N(g)| \leq V_{[a,b]}(g), \tag{3.74}$$

$$(2) \qquad |\beta_N(g)| \leq \|g\| \tag{3.75}$$

for all $N \in \mathbb{Z}^+$ and $g \in \mathrm{CBV}[a, b]$, where $V_{[a,b]}(g)$ denotes the total variation of g on $[a, b]$, $\|\cdot\|$ denotes the norm of $\mathrm{CBV}[a, b]$.

Proof of theorem 7

(1) Let g be any element of $\mathrm{CBV}[a, b]$, let N be any positive integer. Since g is continuous on $[a, b]$, by the Mean Value Theorem there exists a real number $U_r \in (x(N, r-1), x(N, r))$ such that

$$\int_{x(N,r-1)}^{x(N,r)} g(\theta)\,\mathrm{d}\theta = g(U_r)((b-a)/N). \tag{3.76}$$

Therefore, we have

$$(1/(b-a))\left(\int_a^b g(\theta)\,\mathrm{d}\theta\right)N = \sum_{r=1}^{N} g(U_r). \tag{3.77}$$

From this, it follows that

$$\begin{aligned}
|\beta_N(g)| &= \left|\sum_{r=1}^{N} g(x(N,r)) - \sum_{r=1}^{N} g(U_r)\right| \\
&\le \sum_{r=1}^{N} \{|g(x(N,r)) - g(U_r)| + |g(U_r) - g(x(N,r-1))|\} \\
&\le \sum_{r=1}^{N} V_{[x(N,r-1),\,x(N,r)]}(g) \\
&= V_{[a,b]}(g).
\end{aligned} \tag{3.78}$$

(2) By the definition of the norm of CBV$[a, b]$,

$$\|g\| = \sup_{\theta\in[a,b]} |g(\theta)| + V_{[a,b]}(g), \tag{3.79}$$

the conclusion immediately follows from (1).

(iii) This is an immediate consequence of (3.60), (3.61) and the definitions of β'_N, G, and β''_N.

(iv) By (i), (ii), and (iii), there exist a $c \in \mathbb{R}_0^+$ and an $n \in \mathbb{Z}^+$ such that

$$|\beta'_N(\varphi)| \le c\,\|\varphi\| \tag{3.80}$$

holds for all $\varphi \in \mathrm{CBV}(I')$ and all $N > n$, so that the assertion of proposition ES2 is true. □

It remains to prove theorem 6. We need three lemmas:

LEMMA 2

Let D be a unique factorization domain. Let f, g be two polynomials given by

$$f = a_0x^n + a_1x^{n-1} + \ldots + a_n \in D[x] \quad (a_0 \ne 0), \tag{3.81}$$

$$g = b_0x^m + b_1x^{m-1} + \ldots + b_m \in D[x]. \tag{3.82}$$

Let $R(f, g)$ stand for the resultant of the polynomials f, g. Then f and g have a common non-constant factor if and only if

$$R(f, g) = 0. \tag{3.83}$$

Remark

The proof of lemma 2 is analogous to that in ref. [17] for the case $a_0 b_0 \neq 0$. See also ref. [18] for the proof of the case in which D is a field.

LEMMA 3

Let D be a unique factorization domain of characteristic zero. Let f be a polynomial given by

$$f = a_0 x^n + a_1 x^{n-1} + \ldots + a_n \in D[x] \quad (a_0 \neq 0). \tag{3.84}$$

Then a factorization of the polynomial f into irreducible factors has multiple non-constant factors if and only if the discriminant of f

$$\delta(f) = R(f, f') = 0. \tag{3.85}$$

(cf. ref. [17] for the proof).

Notation

$C^{\omega}(I)$ with $I = [a, b]$ denotes the unique factorization domain of all real analytic functions on the closed interval I. A real-valued function on a subset $S \subset \mathbb{R}$ is called real analytic on S if it is the restriction to S of a function which is real analytic on some open set $0 \supset S$.

LEMMA 4

Let $f \in (C^{\omega}[a, b])[\lambda]$ be a monic polynomial of degree $q \in \mathbb{Z}^+$ over the unique factorization domain $C^{\omega}[a, b]$ given by

$$f = \lambda^q + a_1 \lambda^{q-1} + \ldots + a_q. \tag{3.86}$$

Suppose that for any fixed $\theta \in [a, b]$, the polynomial

$$f_\theta = \lambda^q + a_1(\theta) \lambda^{q-1} + \ldots + a_q(\theta) \tag{3.87}$$

over the field $\mathbb{R}$ has q real roots, which we denote by $\lambda_{1\theta} \leq \lambda_{2\theta} \leq \ldots \leq \lambda_{q\theta}$. Then all the $\lambda_{i\theta}$'s are piecewise monotone on $[a, b]$.

Proof

Let ω_1 denote the discriminant of f.

(I) If $\omega_1(\theta) \equiv 0$ on $[a, b]$, a factorization of f into irreducible factors has multiple non-constant factors. Therefore, the proof of the lemma is reduced to the case:

(II) $\omega_1(\theta) \not\equiv 0$ on $[a, b]$. Then the number of zeros of ω_1 on the compact set $[a, b]$ is finite since $\omega_1 \in C^{\omega}[a, b]$.

Suppose one of the $\lambda_{i\theta}$'s, say $\lambda_{1\theta}$, were not piecewise monotone on $[a, b]$, then there would exist an open interval $(a', b') \subset [a, b] \backslash \{\theta \in [a, b] : \omega_1(\theta) = 0\}$ with the following properties:

(1) $\lambda_{1\theta}, \lambda_{2\theta}, \ldots, \lambda_{q\theta}$ are real analytic on (a', b') (by the Implicit Function Theorem).

(2) For $\theta \in (a', b')$ and $i \neq j$, we have $\lambda_{i\theta} \neq \lambda_{j\theta}$.

(3) $\mathrm{d}\lambda_{1\theta}/\mathrm{d}\theta$ has infinitely many zeros on (a', b').

Let the function $\omega_2 \in C^{\omega}[a, b]$ be defined by

$$\omega_2 = R(f, \partial f/\partial\theta), \tag{3.88}$$

where $\partial f/\partial\theta = a_1'\lambda^{q-1} + a_2'\lambda^{q-2} + \ldots + a_q' \in (C^{\omega}[a, b])[\lambda]$. By the fundamental property of the resultant, ω_2 can be written as

$$\omega_2 = Af + B(\partial f/\partial\theta), \tag{3.89}$$

where A and B are some polynomials in $(C^{\omega}[a, b])[\lambda]$.

All the zeros of $\mathrm{d}\lambda_{1\theta}/\mathrm{d}\theta$ on (a', b') are zeros of ω_2, by (3.89) and the equalities

$$f(\lambda_{1\theta}) = 0, \tag{3.90}$$

$$(\partial f/\partial\theta)(\lambda_{1\theta}) = -(\partial f/\partial\lambda)(\lambda_{1\theta})\, \mathrm{d}\lambda_{1\theta}/\mathrm{d}\theta, \tag{3.91}$$

valid for all $\theta \in (a', b')$. Thus, ω_2 has infinitely many zeros on $[a, b]$ so that $\omega_2(\theta) \equiv 0$ on $[a, b]$. Using lemma 2, one concludes that f and $\partial f/\partial\theta$ have a common non-constant factor.

Without loss of generality, we may assume that f is irreducible by the following easily verifiable fact: Suppose we are given a factorization of $f : f = \prod_{j=1}^{m} f_j$, where all the f_j are irreducible monic polynomials. Then there exists a $j \in \{1, 2, \ldots, m\}$ such that for any $\theta \in (a', b')$, $f_j(\lambda_{1\theta}) = 0$.

Now recalling the fact that f is a monic polynomial, clearly we have

$$\deg(\partial f/\partial\theta) < \deg(f). \tag{3.92}$$

Bearing in mind that f is irreducible and that f and $\partial f/\partial\theta$ have a common non-constant factor, we infer that $\partial f/\partial\theta = 0$. Hence, $\lambda_{1\theta}$ is piecewise monotone (in fact, constant) on $[a, b]$. A contradiction. Therefore, all the $\lambda_{1\theta}$'s are piecewise monotone on $[a, b]$. □

Finally, we can give

Proof of theorem 6

Consider the characteristic polynomial

$$|\lambda I_q - F(\theta)| = 0. \tag{3.93}$$

The left-hand side can be written as a monic polynomial of degree q,

$$\lambda^q + a_1(\theta)\lambda^{q-1} + \ldots + a_q(\theta). \tag{3.94}$$

In view of the analyticity of each entry of $F(\theta)$ and the definition of the determinant, the a_j: $\theta \mapsto a_j(\theta)$ are obviously all analytic functions on $\mathbb{C}$. They are real-valued on $\mathbb{R}$ since for each $\theta \in \mathbb{R}$, all the roots of eq. (3.93) are real. In fact, the left-hand side of eq. (3.93) can be expressed by

$$|\lambda I_q - F(\theta)| = \prod_{h=1}^{q} (\lambda - \lambda_h(F(\theta))), \tag{3.95}$$

where $\lambda_h(F(\theta))$ denotes the hth eigenvalue of the Hermitian matrix $F(\theta)$, which is clearly real.

Hence, we may apply lemma 4, and the conclusion of theorem 6 follows immediately. □

3.6. STATEMENT (III) AND SOLUTION OF PROBLEM 1

To solve problem 1, it remains to prove statement (III) (statement (2.21)). We divide the proof of statement (III) into two parts.

Proof of statement (III)

Let C^1 denote the subset of CBV(I) of all the C^1 functions. Recall the definition of φ_ξ and put

$$U = \{\varphi_\xi : \xi > 0\}. \tag{3.96}$$

First, we shall show that C^1 is dense in the set U, i.e. $U \subset \overline{C^1}$. Secondly, we shall prove that $\overline{C^1} = \overline{P}$.

(1) Put $I = [a, b]$, where $a < b$. If $a > 0$ or $b < 0$, then clearly we have $U \subset C^1 \subset \overline{C^1}$. Here, we shall assume that $\{0\} \in (a, b)$. The argument for the case in which $a = 0$ or $b = 0$ is analogous and we omit it.

If $\xi > 1$, then φ_ξ is continuously differentiable on $I = [a, b]$ so that $\varphi_\xi \in \overline{C^1}$ for all $\xi > 1$.

For any fixed $\xi \in (0, 1]$, consider a C^1 function ϕ_s with $0 < s \le |a|, |b|$ defined by

$$\phi_s(t) = \begin{cases} (\xi s^{\xi-2}/2)t^2 + (s^\xi - \xi s^\xi/2) & \text{if } t \in [-s, s], \\ |t|^\xi & \text{if } t \in I \setminus [-s, s]. \end{cases} \tag{3.97}$$

It is easy to verify that $\varphi_\xi - \phi_s$ is an even function on $[-s, s]$, monotone decreasing on $[0, s]$, and vanishing on $I\setminus[-s, s]$. Thus, we have

$$\sup_{t\in I} |\varphi_\xi(t) - \phi_s(t)| = \phi_s(0), \tag{3.98}$$

$$V_I(\varphi_\xi - \phi_s) = 2\phi_s(0). \tag{3.99}$$

Hence,

$$\|\varphi_\xi - \phi_s\| = 3\phi_s(0) \to 0, \tag{3.100}$$

$s \to 0$. Therefore, for any $\xi \in (0, 1]$, φ_ξ is also an element of $\overline{C^1}$. Then it immediately follows that $U \subset \overline{C^1}$.

(2) We first restrict our attention to the set C^1 and consider two normed spaces $X_1 = (C^1, \|\cdot\|_1)$ and $X_2 = (C^1, \|\cdot\|)$, where $\|\cdot\|_1$ is given by

$$\|\varphi\|_1 = \sup_{t\in I} |\varphi(t)| + \sup_{t\in I} |\varphi'(t)|. \tag{3.101}$$

We wish to verify that P is dense in X_2. To do this, we recall the well-known fact that P is dense in X_1 [33]. Then to prove that P is dense in X_2, it is obviously enough to show that the $\|\cdot\|_1$ norm is stronger than the $\|\cdot\|$ norm; in other words, there exists a real constant K such that

$$\|\varphi\| \le K\|\varphi\|_1 \tag{3.102}$$

for all $\varphi \in C^1$.

Let φ be any element of C^1 and let Δ: $a = t_0 \le t_1 \le t_2 \le \ldots \le t_n = b$ denote any partition of interval $[a, b]$. By the Mean Value Theorem, it follows that

$$\sum_{i=1}^n |\varphi(t_i) - \varphi(t_{i-1})| \le \sum_{i=1}^n \sup_{t\in I} |\varphi'(t)| \; |t_i - t_{i-1}|$$

$$= (b-a) \sup_{t\in I} |\varphi'(t)|. \tag{3.103}$$

Recalling the definition of the total variation,

$$V_I(\varphi) \le (b-a) \sup_{t \in I} |\varphi'(t)|. \tag{3.104}$$

Therefore, we obtain

$$\| \varphi \| \le \max(1, (b-a)) \| \varphi \|_1, \tag{3.105}$$

proving that P is dense in X_2.

Return to the original space CBV(I); by what has been proved above, we know that

$$C^1 \subset \overline{P}. \tag{3.106}$$

From this, we infer that $\overline{C^1} \subset \overline{P}$. On the other hand, clearly $P \subset C^1$, whence $\overline{P} \subset \overline{C^1}$. Therefore, we conclude that

$$\overline{C^1} = \overline{P}. \tag{3.107}$$

By (1) and (2), for each $\xi > 0$, we have

$$\varphi_\xi \in \overline{P}, \tag{3.108}$$

as desired. □

Thus, the validity of statement (II) (or the assertion of the ALT) and statement (III) has been proved. Setting $q = 1$, $\xi = 1/2$, one obtains a solution of problem 1 immediately.

We note that the validity of hypotheses I and II in ref. [10] follows directly from the continuity of π_β and the validity of statement (III), respectively.

For a better understanding of the ALT and its utility, we shall recall the following problem which was previously formulated in [10]:

PROBLEM 3

For a fixed matrix sequence $\{M_N\} \in X_r(q)$, obtain a large enough set F of functions φ such that for all $\varphi \in F$, the real sequence Tr $\varphi(M_N)$ has an asymptotic line.

As can be easily seen by the argument of the present section, a solution of problem 3 also gives a solution of problem 1.

In fact, the ALT provides a direct solution of problem 3, and it gives, in a broader context, an indirect solution of problem 1. The unifying power, utility, and novelty of the ALT stem from the new viewpoint to the additivity problems, which is associated with the formulation of problem 3 (cf. ref. [10] for a detailed explanation of the thought process leading to the formulation of this "reversed" problem (see especially p. 131 of ref. [10])).

Now that the ALT has been established, it is interesting to ask the following question:

Let $\{M_N\} \in X_r(q)$ be a fixed repeat sequence, let $\overline{P}$ be as in the ALT, and let E_N: $\overline{P} \to \mathbb{R}$ denote the sequence of the functionals defined by $E_N(\varphi) = \mathrm{Tr}\ \varphi(M_N)$. Fix a $j \in \mathbb{Z}^+$ and a $\varphi \in \overline{P}$. Then, do $a_1(\varphi), a_2(\varphi), \ldots, a_j(\varphi) \in \mathbb{R}$ exist such that

$$E_N(\varphi) = a_1(\varphi)N + a_2(\varphi)N^0 + \ldots + a_j(\varphi)N^{2-j} + o(N^{2-j}) \tag{3.109}$$

as N tends to infinity?

If $j = 1$ or 2, then the answer to this question is obviously affirmative by the ALT. If $\varphi \in P$, the answer is also affirmative by theorem 1 (PALT). However, for general $j \in \mathbb{Z}^+$ and $\varphi \in \overline{P}$, the problem is yet to be investigated.

Suppose, on the other hand, that we are explicitly given a special matrix sequence $\{M_N\} \in X_r(q)$ which can be (block) diagonalized, together with a special function $\varphi \in \overline{P}$. Then the answer to the above question could possibly be easily obtainable.

For example, let us assume that the $\{M_N\}$ equals $\{K_N\}$ given by (2.2) and $\varphi(t) = |t|^{1/2}$ as in (2.1). Then, by using (2.4), for any $j \in \mathbb{Z}^+$, the corresponding real number sequence E_N can be expanded by

$$E_N = a_1N + a_2N^0 + \ldots + a_jN^{2-j} + o(N^{2-j}) \tag{3.110}$$

as N tends to infinity, where $a_1, a_2, \ldots, a_j$ are real constants. ($a_1 = 4/\pi$, $a_2 = -1$; one can express $a_3, \ldots, a_j$ in terms of Bernoulli's number B_n by using the well-known power expansion $(x/2)\cot(x/2) = 1 - \sum_{n=1}^{\infty}(B_nx^{2n}/(2n)!)$.)

In the present paper, however, we are not concerned with those special cases in which the matrix sequence $\{M_N\}$ and the function φ are incidentally such that $E_N(\varphi) = \mathrm{Tr}\ \varphi(M_N)$ can be explicitly expressed and directly expanded as in (3.110). Nor are we concerned with similarly special cases in which $E_N(\varphi) = \mathrm{Tr}\ \varphi(M_N)$ can be semi-explicitly expressed and the asymptotic behavior of $E_N(\varphi)$ is fairly easily tractable (cf. ref. [13], section 5, for such cases where the M_N are block-diagonalizable and $E_N(\varphi)$ can be expressed in terms of Riemann sums.)

Although heuristic and instructive, the conventional approach that could, under special circumstances, lead to an equation of the type (3.110) is too ad hoc to be effective for establishing the ALT and related theorems which are formulated in a new and broader context.

Using the ALT and related theorems, one can also deal with the additivity problems of small or intermediate size molecules which can be represented by $\{M_N\} \in X_r(q)$ with $N = 1, 2, 3, \ldots$. The details of this application of the ALT to smaller molecules will be published elsewhere.

It should also be mentioned that the ALT, by virtue of the extensiveness of $\overline{P} \subset \mathrm{CBV}(I)$ in the theorem, can be applied as well to the addivitity problems of thermodynamic quantities of molecules [34], which have been extensively investigated by empirical chemists (cf. e.g. refs. [4–7]). In this connection, the reader is referred to ref. [35] for a prototypical study of the size effect of vibrational thermodynamic quantities of solids, which employs the approximation of elastic continuum and the techniques essentially analogous to those used in deriving the asymptotic expansion of the real number sequence E_N given by (3.110).

It is also interesting and instructive to refer to the modern theory of bulk matter and thermodynamic limit. One can see, for example in a review article [36] and references therein, various techniques with which to obtain the asymptotic relations between the number of particles in physical systems and the magnitude of physical quantities. The reference to the modern theory of bulk matter might be profitable, especially to those readers who are familiar with the asymptotic relationships between structure and properties in hydrocarbons, although the theory of bulk matter can neither handle nor predict the properties of hydrocarbon molecules that have been experimentally studied in chemistry.

4. Concluding remarks

By statement (III), $\overline{P}$ in CBV(I) contains the absolute value function φ_1, defined by $\varphi_1(t) = |t|$, $t \in I$. Thus, the Asymptotic Linearity Theorem can be applied to the additivity problems of the TPEEs of alternant hydrocarbons. We remark that P is not dense in CBV(I), i.e. $\overline{P} \neq$ CBV(I). It can be shown that the Cantor function [37] defined on I, for instance, is an element of CBV$(I)\setminus\overline{P}$. A concise characterization of $\overline{P}$ is possible. This will make application of the ALT easier. Details will be published later, together with applications of the ALT to the additivity problems of thermodynamic quantities of molecules [34].

Acknowledgements

It is a great pleasure to express our sincere thanks to Professor K. Saito and Professor I. Naruki of the Research Institute for Mathematical Sciences, Kyoto University, who provided valuable suggestions and help for the proof of lemma 4. One of the authors (SA) wishes to express his deep-felt gratitude to the late Professor H. Shingu, Professors K. Okugawa, K. Fukui, G.G. Hall, P.G. Mezey and K. Taylor for their help and encouragement.

References

[1] T.L. Cottrell, J. Chem. Soc. (1948)1448.
[2] K.S. Pitzer and E. Catalano, J. Am. Chem. Soc. 78(1956)4844.
[3] H. Shingu and T. Fujimoto, J. Chem. Phys. 31(1959)556.
[4] H. Shingu and T. Fujimoto, Bull. Japan Petroleum Inst. 1(1959)11.
[5] T. Fujimoto and H. Shingu, Nippon Kagaku Zasshi 82(1961)789, 794, 945, 948.
[6] T. Fujimoto and H. Shingu, Nippon Kagaku Zasshi 83(1962)19.

[7] T. Fujimoto and H. Shingu, Nippon Kagaku Zasshi 83(1962)23, 359, 364.
[8] J.M. Schulman and R.L. Disch, Chem. Phys. Lett. 133(1985)291.
[9] M.C. Flanigan, A. Kormornicki and J.W. McIver, Jr., in: *Electronic Structure Calculation*, ed. G.A. Segel (Plenum Press, New York, 1977).
[10] S. Arimoto, Phys. Lett. A113(1985)126.
[11] S. Arimoto, Phys. Lett. A124(1987)131.
[12] S. Arimoto, Phys. Lett. A124(1987)275.
[13] S. Arimoto and G.G. Hall, Int. J. Quant. Chem. 41(1992)613.
[14] J.K. Kelly, *General Topology* (Van Nostrand, Princeton, 1955).
[15] S. Willard, *General Topology* (Addison–Wesley, Reading, MA, 1970).
[16] J. Dieudonné, *Foundations of Modern Analysis* (Academic Press, New York, 1960).
[17] R.J. Walker, *Algebraic Curves* (Dover, New York, 1962).
[18] L. Rédel, *Algebra* (Pergamon, Oxford, 1967).
[19] G.G. Hall, Proc. Roy. Soc. A229(1955)251.
[20] G.G. Hall, Int. J. Math. Educ. Sci. Technol. 4(1973)233.
[21] G.G. Hall, Bull. Inst. Math. Appl. 17(1981)70.
[22] J. Aihara, J. Am. Chem. Soc. 98(1976)2750.
[23] A. Graovac, I. Gutman and N. Trinajstić, *Topological Approach to the Chemistry of Conjugated Molecules* (Springer, Berlin, 1977).
[24] I. Gutman and N. Trinajstić, Chem. Phys. Lett. 20(1973)257.
[25] A.T. Balaban, *Chemical Applications of Graph Theory* (Academic Press, New York, 1976).
[26] A. Tang, Y. Kiang, G. Yan and S. Tai, *Graph Theoretical Molecular Orbitals* (Science Press, Beijing, 1986).
[27] Y. Jiang, A. Tang and R. Hoffmann, Theor. Chim. Acta 66(1984)183.
[28] A. Motoyama and H. Hosoya, J. Math. Phys. 18(1977)1485.
[29] H. Hosoya and A. Motoyama, J. Math. Phys. 26(1985)157.
[30] A. Graham, *Kronecker Products and Matrix Calculus with Applications* (Ellis Horwood, Chichester, 1981).
[31] A.S. Householder, *The Theory of Matrices in Numerical Analysis* (Dover, New York, 1975).
[32] G.E. Shilov, *Linear Algebra* (Dover, New York, 1977).
[33] R. Courant and D. Hilbert, *Methoden der mathematischen Physik* (Springer, Berlin, 1933).
[34] S. Arimoto and K. Taylor, J. Math. Chem. 13(1993)249, 265.
[35] E.W. Motroll, J. Chem. Phys. 18(1950)183.
[36] E.H. Lieb, Rev. Mod. Phys. 48(1976)553.
[37] G.E. Shilov and B.L. Gurevich, *Integral, Measure and Derivative: A Unified Approach* (Dover, New York, 1977).