数理解析研究所講究録 1458 巻 2005 年 1-9

Extensions of the BMV-conjecture

Frank Hansen

November 16, 2005

Abstract

The Bessis-Moussa-Villani conjecture asserts that for any $n \times n$ matrices A and B such that A is Hermitian and B is positive semi-definite, the function $t \to \operatorname{Tr}\exp(A - tB)$ is the Laplace transform of a positive measure. We say that a function f, defined on the positive half-line, has the BMV-property if for arbitrary $n \times n$ matrices A and B such that A is positive definite and B is positive semi-definite, the function $t \to \operatorname{Tr} f(A + tB)$ is the Laplace transform of a positive measure. The BMV-conjecture is thus equivalent to the assertion that the function $t \to \exp(-t)$ has the BMV-property.

We prove that any non-negative and operator monotone decreasing function defined on the positive half-line has the BMV-property.

Key words: Trace functions, BMV-conjecture.

1 Introduction

Studying perturbations of exactly solvable Hamiltonian systems in statistical mechanics Bessis, Moussa and Villani [2] noted that the Padé approximant to the partition function $Z(\beta) = \operatorname{Tr}\exp(-\beta(H_0 + \lambda H_1))$ may be efficiently calculated, if the function

$$\lambda \to \operatorname{Tr}\exp(-\beta(H_0 + \lambda H_1))$$

is the Laplace transform of a positive measure. The authors then noted that this is indeed true for a system of spinless particles with local interactions bounded from below. The statement also holds if H_0 and H_1 are commuting operators, or if they are just 2×2 matrices. These observations led to the formulation of the following conjecture:

Conjecture (BMV). *Let A and B be $n\times n$ matrices for some natural number n, and suppose that A is self-adjoint and B is positive semi-definite. Then there is a positive measure μ with support in the closed positive half-axis such that*

$$\operatorname{Tr}\exp(A-tB)=\int_0^\infty e^{-ts}\,d\mu(s)$$

for every $t\geq 0$.

The Bessis-Moussa-Villani (BMV) conjecture may be reformulated as an infinite series of inequalities.

Theorem (Bernstein). *Let f be a real C^∞-function defined on the positive half-axis. If f is completely monotone, that is*

$$(-1)^n f^{(n)}(t)\geq 0 \qquad t>0,\, n=0,1,2,\ldots,$$

then there exists a positive measure μ on the positive half-axis such that

$$f(t)=\int_0^\infty e^{-st}\,d\mu(s)$$

for every $t>0$.

The BMV-conjecture is thus equivalent to saying that the function

$$f(t)=\operatorname{Tr}\exp(A-tB) \qquad t>0$$

is completely monotone. A proof of Bernstein's theorem can be found in [4].

Assuming the BMV-conjecture one may derive a similar statement for free semicircularly distributed elements in a type II_1 von Neumann algebra with a faithful trace. This consequence of the conjecture has been proved by Fannes and Petz [6]. A hypergeometric approach by Drmota, Schachermayer and Teichmann [5] gives a proof of the BMV-conjecture for some types of 3×3 matrices. This paper is a review article based on [10].

1.1 Equivalent formulations

The BMV-conjecture can be stated in several equivalent forms.

Theorem 1.1. *The following conditions are equivalent:*

(i). For arbitrary $n\times n$ matrices A and B such that A is self-adjoint and B is positive semi-definite the function $f(t)=\operatorname{Tr}\exp(A-tB)$, defined on the positive half-axis, is the Laplace transform of a positive measure supported in $[0,\infty)$.

(ii). For arbitrary $n \times n$ matrices A and B such that A is self-adjoint and B is positive semi-definite the function $g(t) = \mathrm{Tr}\exp(A + itB)$, defined on the positive half-axis, is of positive type.

(iii). For arbitrary positive definite $n \times n$ matrices A and B the polynomial $P(t) = \mathrm{Tr}(A + tB)^p$ has non-negative coefficients for any $p = 1, 2, \ldots$.

(iv). For arbitrary positive definite $n \times n$ matrices A and B the function $\varphi(t) = \mathrm{Tr}\exp(A + tB)$ is m-positive on some open interval of the form $(-\alpha, \alpha)$.

The first statement is the BMV-conjecture, and it readily implies the second statement by analytic continuation. The sufficiency of the second statement is essentially Bochner's theorem. The implication $(iii) \Rightarrow (i)$ is obtained by applying Bernstein's theorem and approximation of the exponential function by its Taylor expansion. The implication $(i) \Rightarrow (iii)$ was proved by Lieb and Seiringer [16]. A function $\varphi : (-\alpha, \alpha) \to \mathbf{R}$ is said to be m-positive, if for arbitrary self-adjoint $k \times k$ matrices X with non-negative entries and spectra contained in $(-\alpha, \alpha)$ the matrix $\varphi(X)$ has non-negative entries. The implication $(iii) \Rightarrow (iv)$ follows by approximation, while the implication $(iv) \Rightarrow (i)$ follows by Bernstein's theorem and [8, Theorem 3.3] which states that an m-positive function is real analytic with non-negative derivatives in zero.

In a recent paper [13] Hillar studied the coefficients of the above polynomial $P(t) = \mathrm{Tr}(A + tB)^p$. The coefficient of t^k in $P(t)$ is the trace of the so called kth Hurwitz product $S_{p,k}(A, B)$ of A and B, which is the sum of all words of lenght p in A and B in which B appears k times. This polynomial has real coefficients, and in [15] it is proved that each constituent word in $S_{p,k}(A, B)$ has positive trace for $p < 6$ and all n. The first case in which the conjecture is in doubt is thus for $n = 3$ and $p = 6$. Even in this case all coefficients except $\mathrm{Tr}\, S_{6,3}(A, B)$ were known to be positive. The question is very subtle since some of the words in the Hurwitz product may have negative trace. It was shown in [15] that the word $ABABBA$ may have negative trace for some positive definite 3×3 matrices A and B. Finally it was proved in [14], using heavy computation, that the polynomial $P(t)$ has positive coefficients[1] also in the case $n = 3$ and $p = 6$.

[1]This means that the non-zero coefficients of the polynomial are positive.

2 Preliminaries and main result

Let f be a real function of one variable defined on a real interval I. We consider for each natural number n the associated matrix function $x \to f(x)$ defined on the set of self-adjoint matrices of order n with spectra in I. The matrix function is defined by setting

$$f(x) = \sum_{i=1}^{p} f(\lambda_i)P_i \qquad \text{where} \quad x = \sum_{i=1}^{p} \lambda_i P_i$$

is the spectral resolution of x. The matrix function $x \to f(x)$ is Fréchet differentiable [7] if I is open and f is continuously differentiable [3].

2.1 The BMV-property

Definition 2.1. *A function $f\colon \mathbf{R}_+ \to \mathbf{R}$ is said to have the BMV-property, if to each $n = 1, 2, \ldots$ and each pair of $n \times n$ matrices A and B, such that A is positive definite and B is positive semi-definite, there is a positive measure μ with support in $[0, \infty)$ such that*

$$\operatorname{Tr} f(A + tB) = \int_0^\infty e^{-st}\, d\mu(s)$$

for every $t > 0$.

The BMV-conjecture is thus equivalent to the statement that the function $t \to \exp(-t)$ has the BMV-property.

Main Theorem. *Every non-negative operator monotone decreasing function defined on the open positive half-line has the BMV-property.*

3 Differential analysis

An simple proof of the following result can be found in [11, Proposition 1.3].

Proposition 3.1. *The Fréchet differential of the exponential operator function $x \to \exp(x)$ is given by*

$$d\exp(x)h = \int_0^1 \exp(sx)h\exp((1-s)x)\, ds = \int_0^1 A(s)\exp(x)\, ds$$

where $A(s) = \exp(sx)h\exp(-sx)$ for $s \in \mathbf{R}$.

This is only a small part of the Dyson formula which contains formalisme developed earlier by Tomonaga, Schwinger and Feynman. The subject was given a rigorous mathematical treatment by Araki in terms of expansionals in Banach algebras. In particular [1, Theorem 3], the expansional

$$E_r(h;x) = \sum_{n=0}^{\infty} \int_0^1 \int_0^{s_1} \cdots \int_0^{s_{n-1}} A(s_n)A(s_{n-1})\cdots A(s_1)\, ds_n\, ds_{n-1}\cdots ds_1$$

is absolutely convergent in the norm topology with limit

$$E_r(h;x) = \exp(x+h)\exp(-x).$$

We therefore obtain the pth Fréchet differential of the exponential operator function by the expression

$$\begin{aligned} &d^p \exp(x)h^p \\ &= p! \int_0^1 \int_0^{s_1} \cdots \int_0^{s_{p-1}} A(s_p)A(s_{p-1})\cdots A(s_1)\exp(x)\, ds_p\, ds_{p-1}\cdots ds_1. \end{aligned}$$

3.1 Divided differences

The following representation of divided differences is due to Hermite [12].

Proposition 3.2. *Divided differences can be written in the following form*

$$\begin{aligned} [x_0,x_1]_f &= \int_0^1 f'\Big((1-t_1)x_0 + t_1 x_1\Big)\, dt \\ [x_0,x_1,x_2]_f &= \int_0^1 \int_0^{t_1} f''\Big((1-t_1)x_0 + (t_1-t_2)x_1 + t_2 x_2\Big)\, dt_2\, dt_1 \\ &\ \ \vdots \\ [x_0,x_1,\cdots,x_n]_f &= \int_0^1 \int_0^{t_1} \cdots \int_0^{t_{n-1}} f^{(n)}\Big((1-t_1)x_0 + (t_1-t_2)x_1 + \cdots \\ &\qquad + (t_{n-1}-t_n)x_{n-1} + t_n x_n\Big)\, dt_n \cdots dt_2\, dt_1 \end{aligned}$$

where f is an n-times continuously differential function defined on an open interval I, and $x_0, x_1, \ldots, x_n$ are (not necessarily distinct) points in I.

3.2 Main technical tools

Taking the trace of the pth Fréchet differential of the exponential operator function [10, Theorem 3.4] one derive:

Theorem 3.3. *Let x and h be operators on a Hilbert space of finite dimension n written on the form*

$$x = \sum_{i=1}^{n} \lambda_i e_{ii} \quad \textit{and} \quad h = \sum_{i,j=1}^{n} h_{ij} e_{ij}$$

where $\{e_{ij}\}_{i,j=1}^{n}$ is a system of matrix units, and $\lambda_1, \ldots, \lambda_n$ and $h_{i,j}$ for $i, j = 1, \ldots, n$ are complex numbers. Then the pth derivative

$$\begin{aligned}&\frac{d^p}{dt^p} \operatorname{Tr} \exp(x + th)\Big|_{t=0} \\ &= p! \sum_{i_1=1}^{n} \cdots \sum_{i_p=1}^{n} h_{i_p i_{p-1}} \cdots h_{i_2 i_1} h_{i_1 i_p} [\lambda_{i_1}, \lambda_{i_2}, \cdots, \lambda_{i_p}, \lambda_{i_p}]_{\exp},\end{aligned}$$

where $[\lambda_{i_1}, \lambda_{i_2}, \cdots, \lambda_{i_p}, \lambda_{i_p}]_{\exp}$ are divided differences of order $p+1$ of the exponential function.

Making use of the linearity of the function $f \to [x_0, x_1, \ldots, x_n]_f$ one obtains [10, Lemma 3.5 and Corollary 3.6] the following:

Corollary 3.4. *Let $f : I \to \mathbf{R}$ be a C^∞-function defined on an open and bounded interval I, and let x and h be self-adjoint operators on a Hilbert space of finite dimension n written on the form*

$$x = \sum_{i=1}^{n} \lambda_i e_{ii} \quad \textit{and} \quad h = \sum_{i,j=1}^{n} h_{ij} e_{ij}$$

where $\{e_{ij}\}_{i,j=1}^{n}$ is a system of matrix units, and $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of x counted with multiplicity. If the spectrum of x is in I, then the trace function $t \to \operatorname{Tr} f(x+th)$ is infinitely differentiable in a neighborhood of zero and the pth derivative

$$\begin{aligned}&\frac{d^p}{dt^p} \operatorname{Tr} f(x + th)\Big|_{t=0} \\ &= p! \sum_{i_1=1}^{n} \cdots \sum_{i_p=1}^{n} h_{i_1 i_2} h_{i_2 i_3} \cdots h_{i_{p-1} i_p} h_{i_p i_1} [\lambda_{i_1}, \lambda_{i_2}, \cdots, \lambda_{i_p}, \lambda_{i_1}]_f,\end{aligned}$$

where $[\lambda_{i_1}, \lambda_{i_2}, \cdots, \lambda_{i_p}, \lambda_{i_1}]_f$ are divided differences of order $p+1$ of the function f.

4 Proof of the main theorem

Proposition 4.1. *Consider for a constant $c \geq 0$ the function*

$$g(t) = \frac{1}{c+t} \qquad t > 0.$$

For arbitrary $n \times n$ matrices x and h such that x is positive definite and h is positive semi-definite we have

$$(-1)^p \frac{d^p}{dt^p} \operatorname{Tr} g(x+th)\Big|_{t=0} \geq 0$$

for $p = 1, 2, \ldots$.

Proof. Note that the divided differences of g are of the form

(1) $\quad [\lambda_1, \lambda_2, \ldots, \lambda_p]_g = (-1)^{p-1} g(\lambda_1) g(\lambda_2) \cdots g(\lambda_p) \qquad p = 1, 2, \ldots.$

In the statement of Corollary 3.4 we set $\xi_i = g(\lambda_i) a_i$ and $b_i = g(\lambda_i)^{1/2} a_i$ where a_i is the ith row in a matrix a such that $h = aa^*$, and consequently $h_{ij} = (a_i \mid a_j)$. By calculation we then obtain:

$$\begin{aligned} &\frac{(-1)^p}{p!} \frac{d^p}{dt^p} \operatorname{Tr} g(x+th)\Big|_{t=0} \\ &= \sum_{i_1=1}^{n} \cdots \sum_{i_p=1}^{n} (\xi_{i_1} \mid b_{i_2})(b_{i_2} \mid b_{i_3}) \cdots (b_{i_{p-1}} \mid b_{i_p})(b_{i_p} \mid \xi_{i_1}), \end{aligned}$$

and it is not difficult to prove that such a sum is non-negative. **QED**

Proof of the main theorem. Consider again the function

$$g(t) = \frac{1}{c+t} \qquad t > 0$$

for $c \geq 0$ and arbitrary $n \times n$ matrices x and h such that x is positive definite and h is positive semi-definite. We first note that

$$\frac{d^p}{dt^p} \operatorname{Tr} g(x+th)\Big|_{t=t_0} = \frac{d^p}{d\varepsilon^p} \operatorname{Tr} g(x+t_0 h+\varepsilon h)\Big|_{\varepsilon=0}$$

for $p = 1, 2, \ldots$ and $t_0 \geq 0$. The function $t \to \operatorname{Tr} g(x+th)$ is therefore completely monotone. Let now $f\colon \mathbf{R}_+ \to \mathbf{R}$ be a non-negative operator monotone decreasing function. One may show [10] that f allows the representation

$$f(t) = \beta + \int_0^\infty \frac{1}{c+t} \, d\nu(c)$$

for a positive measure ν. The function $t \to \operatorname{Tr} f(x+th)$ is hence completely monotone and thus by Bernstein's theorem the Laplace transform of a positive measure with support in $[0, \infty)$. **QED**

4.1 Further analysis

One may try to use the Hermite expression in Proposition 3.2 to obtain a proof of the BMV-conjecture. Applying Theorem 3.3 and calculating the third derivative of the trace function we obtain

$$\frac{-1}{3!}\frac{d^3}{dt^3}\operatorname{Tr}\exp(x-th)\Big|_{t=0} = \sum_{p,i,j=1}^{n}(a_p \mid a_i)(a_i \mid a_j)(a_j \mid a_p)[\lambda_p\lambda_i\lambda_j\lambda_p]_{\exp}$$

$$= \int_0^1\int_0^{t_1}\int_0^{t_2}\sum_{p,i,j=1}^{n}(a_p \mid a_i)(a_i \mid a_j)(a_j \mid a_p)\exp\big((1-(t_1-t_3))\lambda_p + (t_1-t_2)\lambda_i + (t_2-t_3)\lambda_j\big)\,dt_3\,dt_2\,dt_1$$

where $h = aa^*$ and a_i is the ith row in a. Assuming the BMV-conjecture this integral should be non-negative, and this would obviously be the case if the integrand is a non-negative function. However, there are examples [10, Example 4.2] where the integrand takes negative values.

Another way forward would be to examine the value of loops of the form

$$(a_1 \mid a_2)(a_2 \mid a_3)\cdots(a_{p-1} \mid a_p)(a_p \mid a_1)$$

since they, apart from an alternating sign, are the only possible negative factors in the expression of the derivatives of the trace functions. By applying a variational principle the lower bound

$$-\cos^p\left(\frac{\pi}{p}\right) \le (a_1 \mid a_2)(a_2 \mid a_3)\cdots(a_{p-1} \mid a_p)(a_p \mid a_1)$$

was established in [9]. The lower bound converges very slowly to -1 as p tends to infinity, and it is attained essentially only when all the vectors form a "*fan*" in a two-dimensional subspace.

Remark 4.2. *If we only consider one-dimensional perturbations, that is if $h = cP$ for a constant $c > 0$ and a one-dimensional projection P, then h is of the form $h = (\xi_i\bar{\xi}_j)_{i,j=1,\dots,n}$ for some vector $\xi = (\xi_1,\dots,\xi_n)$ and each loop*

$$h_{i_1i_2}h_{i_2i_3}\cdots h_{i_{p-1}i_p}h_{i_pi_1} = \|\xi_{i_1}\|^2\cdots\|\xi_{i_p}\|^2$$

is manifestly real and non-negative. This implies that the trace function

$$t \to \operatorname{Tr}\exp(-(x+th)),$$

for any self-adjoint $n\times n$ matrix x, is the Laplace transform of a positive measure with support in $[0,\infty)$.

References

[1] H. Araki. Expansional in Banach algebras. *Ann. scient. Éc. Norm. Sup.*, 6:67–84, 1973.

[2] D. Bessis, P. Moussa, and M. Villani. Monotonic converging variational approximations to the functional integrals in quantum statistical mechanics. *J. Math. Phys.*, 16:2318–2325, 1975.

[3] A.L. Brown and H.L. Vasudeva. *The calculus of operator functions and operator convexity.* Dissertationes Mathematicae. Polska Akademia Nauk, Instytut Matematyczny, 2000.

[4] W. Donoghue. *Monotone matrix functions and analytic continuation.* Springer, Berlin, Heidelberg, New York, 1974.

[5] M. Drmota, W. Schachermayer, and J. Teichmann. A hyper-geometric approach to the BMV-conjecture. *Preprint*, 2004.

[6] M. Fannes and D. Petz. On the function $\operatorname{Tr} e^{H+itK}$. *Int. J. Math. Math. Sci.*, 29:389–393, 2002.

[7] T.M. Flett. *Differential Analysis.* Cambridge University Press, Cambridge, 1980.

[8] F. Hansen. Functions of matrices with nonnegative entries. *Linear Algebra Appl.*, 166:29–43, 1992.

[9] F. Hansen. Lower bounds on products of correlation coefficients. *J. Inequal. Pure and Appl. Math.*, 5(1):Article 16, 2004.

[10] F. Hansen. Trace functions as Laplace transforms. *arXiv:math.OA/0507018v3*, pages 1–16, 2005.

[11] F. Hansen and G.K. Pedersen. Perturbation formulas for traces on C^*-algebras. *Publ. RIMS, Kyoto Univ.*, 31:169–178, 1995.

[12] Ch. Hermite. Sur l'interpolation. *C.R. Acad. sc. Paris*, 48:62–67, 1859.

[13] C.F. Hillar. Advances on the Bessis-Moussa-Villani trace conjecture. *arXiv:math.OA/0507166v1*, pages 1–13, 2005.

[14] C.F. Hillar and C.R. Johnson. On the positivity of the coefficients of a certain polynomial defined by two positive definite matrices. *J. Stat. Phys.*, 118:781–789, 2005.

[15] C.R. Johnson and C.J. Hillar. Eigenvalues of words in two positive definite letters. *SIAM J. Matrix Anal. Appl.*, 23:916–928, 2002.

[16] E. Lieb and R. Seiringer. Equivalent forms of the Bessis-Moussa-Villani conjecture. *J. Stat. Phys.*, 115:185–190, 2004.

Frank Hansen: Institute of Economics, University of Copenhagen, Studiestraede 6, DK-1455 Copenhagen K, Denmark.