Complex Anal. Oper. Theory (2017) 11:843–855
DOI 10.1007/s11785-015-0513-4

Complex Analysis and Operator Theory

On the BMV Conjecture for 2 × 2 Matrices and the Exponential Convexity of the Function $\cosh(\sqrt{at^2+b})$

Victor Katsnelson[1]

Received: 24 May 2015 / Accepted: 5 November 2015 / Published online: 11 December 2015

Abstract The BMV conjecture states that for $n \times n$ Hermitian matrices A and B the function $f_{A,B}(t) = \operatorname{trace} e^{tA+B}$ is exponentially convex. Recently the BMV conjecture was proved by Herbert Stahl. The proof of Herbert Stahl is based on ingenious considerations related to Riemann surfaces of algebraic functions. In the present paper we give a purely "matrix" proof of the BMV conjecture for 2×2 matrices. This proof is based on the Lie product formula for the exponential of the sum of two matrices. The proof also uses the commutation relations for the Pauli matrices and does not use anything else.

Keywords BMV conjecture · Absolutely monotonic functions · Exponentially convex functions · Positive definite functions

Mathematics Subject Classification Primary 15A15 · 15A16; Secondary 30F10 · 44A10

1 Herbert Stahl's Theorem

In the paper [4] a conjecture was formulated which now is commonly known as the BMV conjecture:

The BMV Conjecture Let A and B be Hermitian matrices of size $n \times n$. Then the function

$$f_{A,B}(t) = \operatorname{trace}\{\exp[tA + B]\} \tag{1.1}$$

Communicated by Dmitry Kaliuzhnyi-Verbovetskyi.

✉ Victor Katsnelson
victor.katsnelson@weizmann.ac.il; victorkatsnelson@gmail.com

1 Department of Mathematics, The Weizmann Institute, 76100 Rehovot, Israel

of the variable t is representable as a bilateral Laplace transform of a non-negative measure $d\sigma_{A,B}(\lambda)$ compactly supported on the real axis:

$$f_{A,B}(t) = \int\limits_{\lambda \in (-\infty,\infty)} \exp(t\lambda)\, d\sigma_{A,B}(\lambda), \quad \forall\, t \in (-\infty, \infty). \tag{1.2}$$

Definition 1.1 Let A, B be a pair of square matrices of the same size $n \times n$. The function $f_{A,B}(t)$ of the variable $t \in \mathbb{R}$ defined by (1.1) is said to be *the trace-exponential function generated by the pair* A, B.

Let us note that the function $f_{A,B}(t)$, considered for $t \in \mathbb{C}$, is an entire function of exponential type. The indicator diagram of the function $f_{A,B}$ is the closed interval $[\lambda_{\min}, \lambda_{\max}]$, where $\lambda_{\min}$ and $\lambda_{\max}$ are the least and the greatest eigenvalues of the matrix A respectively. Thus if the function $f_{A,B}(t)$ is representable in the form (1.2) with a non-negative measure $d\sigma_{A,B}(\lambda)$, then $d\sigma_{A,B}(\lambda)$ is actually supported on the interval $[\lambda_{\min}, \lambda_{\max}]$ and the representation

$$f_{A,B}(t) = \int\limits_{\lambda \in [\lambda_{\min},\lambda_{\max}]} \exp(t\lambda)\; d\sigma_{A,B}(\lambda), \qquad \forall\, t \in \mathbb{C}, \tag{1.3}$$

holds for every $t \in \mathbb{C}$.

The representability of the function $f_{A,B}(t)$, (1.1), in the form (1.3) with a non-negative $d\sigma_{A,B}$ is evident if the matrices A and B commute. In this case $d\sigma(\lambda)$ is an atomic measure supported on the spectrum of the matrix A. In general case, if the matrices A and B do not commute, the BMV conjecture remained an open question for longer than 35 years. In 2011, Herbert Stahl gave an affirmative answer to the BMV conjecture.

Theorem (H. Stahl) *Let A and B be $n \times n$ hermitian matrices. Then the function $f_{A,B}(t)$ defined by* (1.1) *is representable as the bilateral Laplace transform* (1.3) *of a non-negative measure $d\sigma_{A,B}(\lambda)$ supported on the closed interval* $[\lambda_{\min}, \lambda_{\max}]$.

The first arXiv version of H.Stahl's Theorem appeared in [10], the latest arXiv version—in [11], the journal publication—in [12]. The proof of Herbert Stahl is based on ingenious considerations related to Riemann surfaces of algebraic functions. In [5,6] a simplified version of the Herbert Stahl proof is presented.

In the present paper we focus on the BMV conjecture for 2×2 matrices. In this special case the BMV conjecture was proved in [9, section 2] using a perturbation series. We give a purely "matrix" proof of the BMV conjecture for 2×2 matrices.

2 Exponentially Convex Functions

Definition 2.1 A function f on $\mathbb{R}$, $f : \mathbb{R} \to [0, \infty)$, is said to be *exponentially convex* if

1. For every positive integer N, for every choice of real numbers $t_1, t_2, \ldots, t_N$, and complex numbers $\xi_1, \xi_2, \ldots, \xi_N$, the inequality holds
$$\sum_{r,s=1}^{N} f(t_r + t_s)\xi_r\overline{\xi_s} \geq 0; \tag{2.1}$$
2. The function f is continuous on $\mathbb{R}$.

The class of exponentially convex functions was introduced by S.N.Bernstein, [2], see Sect. 15 there. The Russian translation of the paper [2] can be found in [3], pp. 370–425.

From (2.1) it follows that the inequality $f(t_1 + t_2) \leq \sqrt{f(2t_1)f(2t_2)}$ holds for every $t_1 \in \mathbb{R}$, $t_2 \in \mathbb{R}$. Thus the alternative takes place:
If f is an exponentially convex function, then either $f(t) \equiv 0$, or $f(t) > 0$ for every $t \in \mathbb{R}$.

2.1 Properties of the Class of Exponentially Convex Functions

- P1. If $f(t)$ is an exponentially convex function and $c \geq 0$ is a nonnegative constant, then the function $cf(t)$ is exponentially convex.
- P2. If $f_1(t)$ and $f_2(t)$ are exponentially convex functions, then their sum $f_1(t)+f_2(t)$ is exponentially convex.
- P3. If $f_1(t)$ and $f_2(t)$ are exponentially convex functions, then their product $f_1(t) \cdot f_2(t)$ is exponentially convex.
- P4. If $f(t)$ is an exponentially convex function and a, b are real numbers, then the function $f(at + b)$ is exponentially convex.
- P5. Let $\{f_n(t)\}_{1\leq n<\infty}$ be a sequence of exponentially convex functions. We assume that for each $t \in \mathbb{R}$ there exists the limit $f(t) = \lim_{n\to\infty} f_n(t)$, and that $f(t) < \infty\ \forall t \in \mathbb{R}$. Then the limiting function $f(t)$ is exponentially convex.

From the functional equation for the exponential function it follows that for each real number μ, for every choice of real numbers $t_1, t_2, \ldots, t_N$ and complex numbers $\xi_1, \xi_2, \ldots, \xi_N$, the equality holds
$$\sum_{r,s=1}^{N} e^{(t_r+t_s)\mu}\xi_r\overline{\xi_s} = \Big|\sum_{p=1}^{N} e^{t_p\mu}\xi_p\Big|^2 \geq 0. \tag{2.2}$$

The relation (2.2) can be formulated as

Lemma 2.2 *For each real number μ, the function $e^{t\mu}$ of the variable t is exponentially convex.*

For $z \in \mathbb{C}$, the function $\cosh z$, which is called *the hyperbolic cosine of* z, is defined as
$$\cosh z = \frac{1}{2}(e^z + e^{-z}). \tag{2.3}$$
From Lemma 2.2 and property P2 we obtain

Lemma 2.3 *For each real μ, the function* $\cosh(t\,\mu)$ *of the variable t is exponentially convex.*

The following result is well known.

Theorem 2.4 (The representation theorem)

1. *Let $\sigma(d\mu)$ be a nonnegative measure on the real axis, and let the function $f(t)$ be defined as the two-sided Laplace transform of the measure $\sigma(d\mu)$:*

$$f(t) = \int\limits_{\mu\in\mathbb{R}} e^{t\mu}\,\sigma(d\mu), \tag{2.4}$$

   where the integral in the right hand side of (2.4) *is finite for any $t \in \mathbb{R}$. Then the function f is exponentially convex.*
2. *Let $f(t)$ be an exponentially convex function. Then this function f can be represented on $\mathbb{R}$ as a two-sided Laplace transform* (2.4) *of a nonnegative measure $\sigma(d\mu)$.* (*In particular, the integral in the right hand side of* (2.4) *is finite for any $t \in \mathbb{R}$.*) *The representing measure $\sigma(d\mu)$ is unique.*

The assertion 1 of the representation theorem is an evident consequence of Lemma 2.2, of the properties P1, P2, P5, and of the definition of the integration operation.

The proof of the assertion 2 can be found in [1, Theorem 5.5.4], and in [13, Theorem 21].

Of course, Lemma 2.3 is a special case of the representation theorem which corresponds to the representing measure

$$\sigma(d\nu) = 1/2(\delta(\nu-\mu) + \delta(\nu+\mu))\,d\nu,$$

where $\delta(\nu \mp \mu)$ are Dirak's δ-functions supported at the points $\pm\mu$.

Thus the Herbert Stahl theorem can be reformulated as follows: *Let A and B be Hermitian $n\times n$ matrices. Let the function $f_{A,B}(t)$ is defined by* (1.1) *for $t \in (-\infty, \infty)$. Then the function $f_{A,B}(t)$, considered as a function of the variable t, is exponentially convex.*

3 Reduction the BMV Conjecture for General 2 × 2 Hermitian Matrices *A* and *B* to the Case of Special *A* and *B*

Lemma 3.1 *Let A and B be an arbitrary pair of 2×2 Hermitian matrices. Then there exists a pair A_0, B_0 of Hermitian 2×2 matrices possessing the properties:*

1. *The conditions*

$$(\mathrm{a}).\ \operatorname{trace} A_0 = 0, \quad (\mathrm{b}).\ \operatorname{trace} B_0 = 0, \quad (\mathrm{c}).\ \operatorname{trace} A_0B_0 = 0. \tag{3.1}$$

   are satisfied.

2. *The trace-exponential functions $f_{A,B}$ and f_{A_0,B_0} generated by these pairs are related by the equality*

$$f_{A,B}(t) = ce^{t\lambda} f_{A_0,B_0}(t + t_0), \tag{3.2}$$

where λ and t_0 are some real numbers, c is a positive number.

Remark 3.2 From Lemma 3.1 it follows that in order to prove the BMV conjecture for arbitrary pair A, B of Hermitian 2×2 matrices, it is sufficient to prove this conjecture only for pairs A_0, B_0 satisfying the conditions (3.1).

Proof of Lemma 3.1 Let A and B be Hermitian matrices of size 2×2 and I be the identity matrix of size 2×2. Let us define

$$A_0 = A - \frac{\operatorname{trace} A}{2} I. \tag{3.3}$$

Without loss of generality we can assume that

$$A_0 \neq 0. \tag{3.4}$$

Otherwise

$$f_{A,B}(t) = ce^{t\lambda}, \quad \text{where } \lambda = \frac{\operatorname{trace} A}{2}, \quad c = \operatorname{trace} e^B > 0,$$

and (3.2) holds with $A_0 = 0$, $B_0 = 0$. Since the matrix A_0 is Hermitian, from (3.4) it follows that $A_0^2 \geq 0$, $A_0^2 \neq 0$. Thus

$$\operatorname{trace} A_0^2 > 0. \tag{3.5}$$

Let us define

$$t_0 = \frac{\operatorname{trace} A_0 B}{\operatorname{trace} A_0^2}, \tag{3.6}$$

$$B_0 = B - \frac{\operatorname{trace} B}{2} I - t_0 A_0. \tag{3.7}$$

Since $\operatorname{trace} I = 2$ and $\operatorname{trace} X$ depends on 2×2 matrix X linearly, the conditions $\operatorname{trace} A_0 = 0$, $\operatorname{trace} B_0 = 0$ are fulfilled. According to (3.6), the condition $\operatorname{trace} A_0 B_0 = 0$ is fulfilled as well. Since

$$A = A_0 + \lambda I, \quad B = B_0 + \mu I + t_0 A_0, \quad \text{where } \lambda = \frac{\operatorname{trace} A}{2}, \quad \mu = \frac{\operatorname{trace} B}{2},$$

the linear matrix pencils $tA + B$ and $tA_0 + B_0$ are related by the equality

$$At + B = (t\lambda + \mu)I + ((t + t_0)A_0 + B_0).$$

Therefore $e^{tA+B} = e^{t\lambda+\mu} e^{(t+t_0)A_0+B_0}$, that is the equality (3.2) holds with $c = e^{\mu}$. □

Lemma 3.3 *Let A_0, B_0 be Hermitian matrices of size 2×2 satisfying the condition (3.1), $A_0 \neq 0$. Then there exists an unitary matrix U which reduces the matrices A_0, B_0 to the form*

$$UA_0U^* = \alpha\sigma, \quad UB_0U^* = \beta\tau, \tag{3.8}$$

where $\alpha > 0$, $\beta \geq 0$ are numbers and σ, τ are the Pauli matrices:

$$\sigma = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \qquad \tau = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \tag{3.9}$$

Proof Let U be an unitary matrix which reduces the Hermitian matrix A_0 to the diagonal form: $UA_0U^* = \left[\begin{smallmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{smallmatrix}\right]$. Since trace $A_0 = 0$, the equality $\lambda_1 = -\lambda_2$ holds. Since $A_0 \neq 0$, also $\lambda_1, \lambda_2 \neq 0$. Thus for some unitary matrix U, the first of the equalities (3.8) holds with some number $\alpha > 0$. We fix this matrix U and *define* the matrix $\left[\begin{smallmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{smallmatrix}\right] = UB_0U^*$. Since trace $B_0 = 0$ and the matrix trace of is an unitarily invariant, the equality $b_{11} + b_{22} = 0$ holds. Since $UA_0B_0U^* = UA_0U^* \cdot UB_0U^* = \left[\begin{smallmatrix} \alpha & 0 \\ 0 & -\alpha \end{smallmatrix}\right] \cdot \left[\begin{smallmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{smallmatrix}\right] = \left[\begin{smallmatrix} \alpha b_{11} & \alpha b_{12} \\ -\alpha b_{21} & -\alpha b_{22} \end{smallmatrix}\right]$ and trace $A_0B_0 = 0$, also trace$\left[\begin{smallmatrix} \alpha b_{11} & \alpha b_{12} \\ -\alpha b_{21} & -\alpha b_{22} \end{smallmatrix}\right] = 0$, that is $\alpha(b_{11} - b_{22}) = 0$. Since $\alpha \neq 0$, $b_{11} - b_{22} = 0$. Finally, $b_{11} = b_{22} = 0$. Since the matrix $\left[\begin{smallmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{smallmatrix}\right]$ is Hermitian, its entries b_{12} and b_{21} are conjugate complex numbers: $b_{12} = \overline{b_{21}}$. The additional unitary equivalence transformation $X \to \left[\begin{smallmatrix} e^{i\vartheta} & 0 \\ 0 & 1 \end{smallmatrix}\right] X \left[\begin{smallmatrix} e^{-i\vartheta} & 0 \\ 0 & 1 \end{smallmatrix}\right]$ does not change the matrix σ, but allows to reduce the matrix $\left[\begin{smallmatrix} 0 & b_{12} \\ \overline{b_{12}} & 0 \end{smallmatrix}\right]$ to the form $\beta\tau$. □

Lemma 3.4 *Let A_0 and B_0 be 2×2 Hermitian matrices satisfying the conditions (3.1), (3.4), and U be the unitary matrix which reduces the pair A_0, B_0 to the pair $\alpha\sigma$, $\beta\tau$ according to (3.8), (3.9). Then the trace-exponential functions generated by the pairs A_0, B_0 and $\alpha\sigma$, $\beta\tau$ coincide:*

$$f_{A_0,B_0}(t) = f_{\alpha\sigma,\,\beta\tau}(t). \tag{3.10}$$

Proof

$$\begin{aligned} f_{A_0,B_0}(t) &= \text{trace}\, e^{tA_0+B_0} = \text{trace}\, Ue^{tA_0+B_0}U^* \\ &= e^{U(tA_0+B_0)U^*} = e^{t\alpha\sigma+\beta\tau} = f_{\alpha\sigma,\beta\tau}(t). \end{aligned}$$

□

Remark 3.5 From Lemmas 3.1, 3.3 and 3.4 it follows that in order to prove the BMV conjecture for arbitrary pair A, B of Hermitian 2×2 matrices, it is enough to prove this conjecture for any pair of the form $A = \alpha\sigma$, $B = \beta\tau$ with $\alpha > 0$, $\beta \geq 0$.

4 The Formulation of the Main Theorem

Theorem 4.1 (The main theorem) *Let α, β be arbitrary non-negative numbers and σ, τ be the Pauli matrices defined by (3.9).*

Then the trace-exponential function $f_{\alpha\sigma,\beta\tau}(t)$ *generated by the pair of matrices* $\alpha\sigma$, $\beta\tau$ *is exponentially convex.*

The trace-exponential function $f_{\alpha\sigma,\beta\tau}(t)$ can be easily found explicitly:

$$f_{\alpha\sigma,\beta\tau}(t) = 2\cosh\sqrt{\alpha^2t^2+\beta^2}, \tag{4.1}$$

where $\cosh\zeta$ is the hyperbolic cosine function. However the exponential convexity of the function $\cosh\sqrt{\alpha^2t^2+\beta^2}$ is not evident.

There are different ways to prove the exponential convexity of the function $f_{\alpha\sigma,\beta\tau}(t)$. One can forget the "matrix" origin of the function $f_{\alpha\sigma,\beta\tau}(t)$ and work with its analytic expression $\cosh\sqrt{\alpha^2t^2+\beta^2}$ only. The function $\cosh\sqrt{\alpha^2t^2+\beta^2}$ can be presented as a bilateral Laplace transform of some measure. The density of this measure can be expressed in terms of the modified Bessel function I_1. From this expression it is evident that the representing measure is non-negative. However the calculation of the representing measure is not so transparent.

In the present paper we give a purely "matrix" proof of the BMV conjecture for 2×2 matrices. This proof is based on the Lie product formula for the exponential of the sum of two matrices. The proof also uses the commutation relations for the Pauli matrices and does not use anything else.

5 The Proof of Theorem 4.1

Since the trace-exponential function $f_{\alpha\sigma,\beta\tau}(t)$ is even in β, the equality

$$f_{\alpha\sigma,\beta\tau}(t) = f_{\alpha\sigma,-\beta\tau}(t)$$

holds for any numbers α, β. Therefore,

$$f_{\alpha\sigma,\beta\tau}(t) = \operatorname{trace}\mathscr{E}(t;\alpha,\beta), \tag{5.1}$$

where $\mathscr{E}(t;\alpha,\beta)$ is the 2×2 matrix-function:

$$\mathscr{E}(t;\alpha,\beta) = \frac{1}{2}[e^{t\alpha\sigma+\beta\tau} + e^{t\alpha\sigma-\beta\tau}]. \tag{5.2}$$

Lemma 5.1 (A version of the Lie product formula) *Let X and Y be square matrices of the same size, say $n\times n$. Then*

$$e^{X+Y} = \lim_{N\to\infty}\Big(e^{\frac{X}{N}}\big(I+\tfrac{Y}{N}\big)\Big)^N. \tag{5.3}$$

Proof Proof of the equality (5.3) can be modified from the proof which is presented in [7, Theorem 2.10]. □

Proof of Theorem 4.1 We apply the equality (5.3) in the cases $X = t\alpha\sigma$ and Y is one in two matrices $Y = \beta\tau$, $Y = -\beta\tau$.

The equality

$$\tau^2 = I \tag{5.4}$$

and the commutation relation

$$\tau\sigma\tau = -\sigma \tag{5.5}$$

play crucial role in the proof of Theorem 4.1.

For every number λ, the matrix exponential $e^{\lambda\sigma}$ is a diagonal 2×2 matrix:

$$e^{\lambda\sigma} = \begin{bmatrix} e^{\lambda} & 0 \\ 0 & e^{-\lambda} \end{bmatrix}. \tag{5.6}$$

From (5.4) and (5.5) the commutation relation for the matrix exponentials $e^{\lambda\sigma}$, follows:

$$\tau e^{\lambda\sigma}\tau = e^{-\lambda\sigma}, \quad \forall \lambda \in \mathbb{R}. \tag{5.7}$$

According to (5.2) and Lemma 5.1,

$$\mathscr{E}(t;\alpha,\beta) = \lim_{N\to\infty} \mathscr{E}_N(t;\alpha,\beta), \tag{5.8}$$

where

$$\mathscr{E}_N(t;\alpha,\beta) = \frac{1}{2}[\mathscr{E}_N^+(t;\alpha,\beta) + \mathscr{E}_N^-(t;\alpha,\beta)], \tag{5.9}$$

$$\mathscr{E}_N^+(t;\alpha,\beta) = \Big(e^{\frac{t\alpha\sigma}{N}}\big(I + \tfrac{\beta\tau}{N}\big)\Big)^N, \quad \mathscr{E}_N^-(t;\alpha,\beta) = \Big(e^{\frac{t\alpha\sigma}{N}}\big(I - \tfrac{\beta\tau}{N}\big)\Big)^N. \tag{5.10}$$

From (5.10) it follows that

$$\mathscr{E}_N^+(t;\alpha,\beta) = \sum_{\varepsilon_1,\varepsilon_2\dots\varepsilon_N} e^{\frac{t\alpha\sigma}{N}} M_{\varepsilon_1}^+ e^{\frac{t\alpha\sigma}{N}} M_{\varepsilon_2}^+ \dots e^{\frac{t\alpha\sigma}{N}} M_{\varepsilon_N}^+, \tag{5.11a}$$

$$\mathscr{E}_N^-(t;\alpha,\beta) = \sum_{\varepsilon_1,\varepsilon_2\dots\varepsilon_N} e^{\frac{t\alpha\sigma}{N}} M_{\varepsilon_1}^- e^{\frac{t\alpha\sigma}{N}} M_{\varepsilon_2}^- \dots e^{\frac{t\alpha\sigma}{N}} M_{\varepsilon_N}^-, \tag{5.11b}$$

where each of ε_j, $j = 1, 2, \dots N$, takes value either 0, or 1, and the factors $M_\varepsilon^\pm$ are:

$$M_0^+ = I, \quad M_0^- = I, \quad M_1^+ = \tfrac{\beta\tau}{N}, \quad M_1^- = -\tfrac{\beta\tau}{N}. \tag{5.12}$$

The sums in (5.11) run over all possible combinations $\varepsilon_1, \varepsilon_2 \dots \varepsilon_N$ with either $\varepsilon_j = 0$ or $\varepsilon_j = 1$. (There are 2^N such combinations.) Grouping terms, we present the sums (5.11) as iterated sums, where the summation index m in the external sum runs over the set $0, 1, 2, \dots, N$. Each term in the internal sum is a product[1] which

[1] We omit the "trivial" factors $M_0^+ = I$, $M_0^- = I$.

contains N factors of the form $e^{t\alpha\frac{1}{N}\sigma}$ and m factors of the form $\pm\frac{\beta\boldsymbol{\tau}}{N}$. These factors in general do not commute. So the generic term of the internal sum is the "word" $W = F_1 \cdot F_2 \cdot \cdots \cdot F_k \cdot \cdots \cdot F_{N+m}$, consisting of two letters only: either $F_k = e^{t\alpha\frac{1}{N}\sigma}$ or $F_k = \pm\frac{\beta\boldsymbol{\tau}}{N}$. In the word W, the letters $F_k = \pm\frac{\beta\boldsymbol{\tau}}{N}$ occupy m, $0 \leq m \leq N$, positions enumerated by $k = p_1, k = p_2, \ldots, k = p_m$. Since each two neighbouring letters of the form $\pm\frac{\beta\boldsymbol{\tau}}{N}$ must be separated by at least one letter of the form $e^{t\alpha\frac{1}{N}\sigma}$, the subscripts p_j, $1 \leq j \leq m$ enumerating positions of letters of the form $\pm\frac{\boldsymbol{\tau}}{N}$ must satisfy the conditions

$$1 < p_1,\ p_1 + 1 < p_2,\ p_2 + 1 < p_3, \ldots, p_{m-1} + 1 < p_m,\ p_m \leq N + m. \quad (5.13)$$

The letters $F_k = e^{t\alpha\frac{1}{N}\sigma}$ occupy the remaining N position.

Thus

$$\mathscr{E}_N^{+}(t;\alpha,\beta) = \sum_{0\leq m\leq N}\left(\frac{1}{N^m}\sum_{p_1,p_2,\ldots\, p_m} W^{+}_{p_1,p_2,\ldots,p_m}\right), \quad (5.14a)$$

$$\mathscr{E}_N^{-}(t;\alpha,\beta) = \sum_{0\leq m\leq N}\left(\frac{1}{N^m}\sum_{p_1,p_2,\ldots\, p_m} W^{-}_{p_1,p_2,\ldots,p_m}\right), \quad (5.14b)$$

where

$$\begin{aligned} W^{+}_{p_1,p_2,\ldots,p_m} &= \beta^m \cdot e^{t\alpha\frac{p_1-1}{N}\sigma} \cdot \boldsymbol{\tau} \cdot e^{t\alpha\frac{p_2-p_1-1}{N}\sigma} \cdot \boldsymbol{\tau} \cdot e^{t\alpha\frac{p_3-p_2-1}{N}\sigma} \cdot \boldsymbol{\tau} \\ &\cdot e^{t\alpha\frac{p_4-p_3-1}{N}\sigma} \cdot \boldsymbol{\tau} \cdot \cdots \cdot e^{t\alpha\frac{p_m-p_{m-1}-1}{N}\sigma} \cdot \boldsymbol{\tau} \cdot e^{t\alpha(1-\frac{p_m-m}{N})\sigma}, \quad (5.15a) \\ W^{-}_{p_1,p_2,\ldots,p_m} &= (-\beta)^m \cdot e^{t\alpha\frac{p_1-1}{N}\sigma} \cdot \boldsymbol{\tau} \cdot e^{t\alpha\frac{p_2-p_1-1}{N}\sigma} \cdot \boldsymbol{\tau} \cdot e^{t\alpha\frac{p_3-p_2-1}{N}\sigma} \cdot \boldsymbol{\tau} \\ &\cdot e^{t\alpha\frac{p_4-p_3-1}{N}\sigma} \cdot \boldsymbol{\tau} \cdot \cdots \cdot e^{t\alpha\frac{p_m-p_{m-1}-1}{N}\sigma} \cdot \boldsymbol{\tau} \cdot e^{t\alpha(1-\frac{p_m-m}{N})\sigma}, \quad (5.15b) \end{aligned}$$

and the inner sums in (5.14) runs over all sets of m integers $p_1, p_2, \ldots p_m$ satisfying the conditions (5.13). There are $\binom{N}{m} = \frac{N!}{m!(N-m)!}$ such sets of m integers.

By definition, the terms of the sums (5.14) corresponding to $m = 0$ are equal to $e^{t\alpha\sigma}$.

In the expressions (5.14), we should consider separately terms with even and odd indices m.

If m is odd, then in the expressions (5.15) for the words $W^{+}_{p_1,p_2,\ldots,p_m}$ and $W^{-}_{p_1,p_2,\ldots,p_m}$, the factors β^m and $(-\beta)^m$ are of opposed signs. All other factors in these expressions coincide term by term. Therefore

$$\begin{aligned} &W^{+}_{p_1,p_2,\ldots,p_m} + W^{-}_{p_1,p_2,\ldots,p_m} = 0 \quad \text{for each odd } m, \quad \text{for each} \\ &\quad \text{set of subscripts } p_1,\ p_2, \ldots, p_m \text{ satisfying the conditions (5.13).} \quad (5.16) \end{aligned}$$

If m is even, then the factors β^m and $(-\beta)^m$ in the expressions (5.15) for the words $W^{+}_{p_1,p_2,\ldots,p_m}$ and $W^{-}_{p_1,p_2,\ldots,p_m}$ coincide. All other factors in these expressions coincide term by term as well. Therefore

$$W^{+}_{p_1,p_2,\dots,p_m} = W^{-}_{p_1,p_2,\dots,p_m} \quad \text{for each even } m, \quad \text{for each}$$
$$\text{set of subscripts } p_1,\, p_2,\, \dots,\, p_m \text{ satisfying the conditions (5.13).} \tag{5.17}$$

For even m, say $m = 2l$, the expression (5.15) for the word $W^{+}_{p_1,p_2,\dots,p_{2l}} = W^{-}_{p_1,p_2,\dots,p_{2l}}$ can be simplified. Let us choose and fix the set $p_1,\, p_2,\, \dots,\, p_{2l}$ of subscripts satisfying the conditions (5.13). The factors τ-s in the expression in (5.14a) and (5.14b) corresponding to this set of subscripts can be grouped by pairs of adjacent factors:

$$\begin{aligned} W^{\pm}_{p_1,p_2,\dots p_{2l}} &= \beta^{2l}\cdot e^{t\alpha\frac{p_1-1}{N}\sigma}\cdot(\tau\, e^{t\alpha\frac{p_2-p_1-1}{N}\sigma}\tau)\cdot e^{t\alpha\frac{p_3-p_2-1}{N}\sigma} \\ &\cdot\ \cdots\ \cdot e^{t\alpha\frac{p_{2l-1}-p_{2l-2}-1}{N}\sigma}\cdot(\tau\, e^{t\alpha\frac{p_{2l}-p_{2l-1}-1}{N}\sigma}\tau)\cdot e^{t\alpha(1-\frac{p_{2l}-2l}{N})\sigma}, \end{aligned} \tag{5.18}$$

Using (5.7), we obtain that

$$\tau\, e^{t\alpha\frac{p_{2j}-p_{2j-1}-1}{N}\sigma}\tau = e^{-t\alpha\frac{p_{2j}-p_{2j-1}-1}{N}\sigma}, \quad 1\le j\le l. \tag{5.19}$$

Hence

$$W^{+}_{p_1,p_2,\dots,p_{2l}} = W^{-}_{p_1,p_2,\dots,p_{2l}} = \beta^{2l}e^{t\alpha\mu_{p_1,\dots p_{2l};N}\sigma}, \tag{5.20}$$

where

$$\mu_{p_1,\dots p_{2l};N} = \tfrac{1}{N}[2p_1-2p_2+2p_3-\cdots+2p_{2l-1}-2p_{2l}+N+2l]. \tag{5.21}$$

The numbers $\mu_{p_1,\dots p_{2l};N}$ satisfy the inequalities

$$-(1-2/N)\le\mu_{p_1,\dots p_{2l};N}\le 1. \tag{5.22}$$

From (5.9), (5.14), (5.16), and (5.20) it follows that

$$\mathscr{E}_N(t;\alpha,\beta) = e^{t\alpha\sigma} + \sum_{l:1\le l\le N/2}\left(\frac{\beta^{2l}}{N^{2l}}\sum_{p_1,p_2,\dots,p_{2l}} e^{t\alpha\mu_{p_1,p_2,\dots,p_{2l};N}\sigma}\right), \tag{5.23}$$

where $p_1,\, p_2,\, \dots,\, p_{2l}$ run over the set of integers satisfying the conditions (5.13), the numbers $\mu_{p_1,\dots p_{2l};N}$ are defined in (5.21).

The equality (5.23) expresses the matrix function $\mathscr{E}_N(t;\alpha,\beta)$ as a linear combination of the matrix functions $e^{t\alpha\mu\sigma}$ with *non-negative* coefficients, which depend on β:

$$\mathscr{E}_N(t;\alpha,\beta) = \int\limits_{\mu\in[-1,1]} e^{t\alpha\mu\sigma}\,\rho_N(d\mu), \tag{5.24}$$

where

$$\rho_N(d\mu) = \sum_{0 \le l \le N/2} \rho_{N,l}(d\mu), \tag{5.25a}$$

$$\begin{aligned} \rho_{N,0}(d\mu) &= \delta(\mu - 1)\, d\mu, \quad \rho_{N,l}(d\mu) \\ &= \tfrac{\beta^{2l}}{N^{2l}} \sum_{p_1, p_2, \dots, p_{2l}} \delta(\mu - \mu_{p_1,p_2,\dots,p_{2l};N})\, d\mu, \end{aligned} \tag{5.25b}$$

$\delta(\mu)$ is the Dirac δ-function, the summation in (5.25b) runs over all sets of integers $p_1, p_2, \dots, p_{2l}$ satisfying the conditions (5.13) with $m = 2l$, the numbers $\mu_{p_1,p_2,\dots,p_{2l};N}$ are the same that in (5.21).

In view of (5.6), the matrix-function $\mathscr{E}_N(t; \alpha, \beta)$ is diagonal:

$$\mathscr{E}_N(t; \alpha, \beta) = \begin{bmatrix} e_{1,N}(t; \alpha, \beta) & 0 \\ 0 & e_{2,N}(t; \alpha, \beta) \end{bmatrix}. \tag{5.26}$$

The diagonal entries $e_{1,N}(t; \alpha, \beta),\ e_{2,N}(t; \alpha, \beta)$ are representable as

$$e_{1,N}(t; \alpha, \beta) = \int\limits_{\mu \in [-1,1]} e^{t\alpha\mu} \rho_N(d\mu), \quad e_{2,N}(t; \alpha, \beta) = \int\limits_{\mu \in [-1,1]} e^{-t\alpha\mu} \rho_N(d\mu). \tag{5.27}$$

According to Theorem 2.4, each of the functions $e_{1,N}(t; \alpha, \beta)$, $e_{2,N}(t; \alpha, \beta)$ is exponentially convex. Their sum, which is the trace of the matrix $\mathscr{E}_N(t; \alpha, \beta)$, is exponentially convex. In view of (5.8), the function trace $\mathscr{E}(t; \alpha, \beta)$ is exponentially convex. The reference to (5.1) completes the proof of Theorem 4.1. □

Remark 5.2 For each $\beta \ge 0$, the family of the measures $\{\rho_N(d\mu)\}_N$ is uniformly bounded with respect to N:

$$\int\limits_{\mu \in [-1,1]} \rho_N(d\mu) \le e^{\beta}. \tag{5.28}$$

Indeed, for each N, the cardinality of the set of integers $p_1, p_2, \dots, p_m$ satisfying the conditions (5.13) is equal to $\binom{N}{m} = \frac{N!}{m!(N-m)!}$. According to (5.25b),

$$\int\limits_{\mu \in [-1,1]} \rho_{N,l}(d\mu) = \binom{N}{2l} \frac{\beta^{2l}}{N^{2l}}, \quad \forall l : 0 \le 2l \le N.$$

Taking into account (5.25a), we obtain

$$\int\limits_{\mu \in [-1,1]} \rho_N(d\mu) = \sum_{l: 0 \le 2l \le N} \binom{N}{2l} \tfrac{\beta^{2l}}{N^{2l}} < \sum_{0 \le k \le N} \binom{N}{k} \left(\frac{\beta}{N}\right)^k = \left(1 + \tfrac{\beta}{N}\right)^N < e^{\beta}.$$

6 A Theorem on the Integral Representation of a 2 × 2 Matrix Function

Theorem 6.1 *Let β be a non-negative, σ and τ be the Pauli matrices which were defined in (3.9). For each $\beta \geq 0$, let $\mathscr{E}(t;\beta)$ be the matrix function of the variable $t \in \mathbb{R}$ which is defined by the equality*

$$\mathscr{E}(t;\beta) = \frac{e^{t\sigma+\beta\tau} + e^{t\sigma-\beta\tau}}{2}. \tag{6.1}$$

(*The value β is considered as a parameter.*)

Then there exists a non-negative scalar *measure $\rho(d\mu)$ supported on the interval* $[-1, 1]$ *such that the integral representation*

$$\mathscr{E}(t;\beta) = \int\limits_{\mu\in[-1,1]} e^{t\mu\sigma}\rho(d\mu), \quad \forall t \in \mathbb{R}. \tag{6.2}$$

holds. The measure ρ admits the estimate

$$\int\limits_{\mu\in[-1,1]} \rho(d\mu) \leq e^{\beta}. \tag{6.3}$$

Proof We start from the integral representation (5.24), where we can set $\alpha = 1$. The inequality (5.28) means that for each β, the family of measures $\{\rho_N\}$ is bounded with respect to N. Therefore the family of measures $\{\rho_N\}$ is weakly compact. From (5.24) and (5.8) it follows that representation (6.2) holds with every measure ρ which is a weak limiting point of the family $\{\rho_N\}$. Actually such ρ is unique. □

Remark 6.2 The measure $\rho(d\mu)$ which appears in the integral representation (6.2) can be presented explicitly. The matrix-function $\mathscr{E}(t;\beta)$ is diagonal:

$$\mathscr{E}(t;\beta) = \begin{bmatrix} e_1(t;\beta) & 0 \\ 0 & e_2(t;\beta) \end{bmatrix}. \tag{6.4}$$

From (6.1) we find that

$$e_1(t;\beta) = \cosh\sqrt{t^2+\beta^2} + t\cdot\frac{\sinh\sqrt{t^2+\beta^2}}{\sqrt{t^2+\beta^2}}, \tag{6.5a}$$

$$e_2(t;\beta) = \cosh\sqrt{t^2+\beta^2} - t\cdot\frac{\sinh\sqrt{t^2+\beta^2}}{\sqrt{t^2+\beta^2}}. \tag{6.5b}$$

The function $\cosh\sqrt{t^2+\beta^2}$ admits the integral representation

$$\cosh\sqrt{t^2+\beta^2} = \cosh t + \int\limits_{\mu\in[-1,1]} \widehat{d}(\mu,\beta)e^{\mu t}\,d\mu, \tag{6.6}$$

where

$$\widehat{d}(\mu,\beta) = \frac{\beta}{2\sqrt{1-\mu^2}} I_1(\beta\sqrt{1-\mu^2}), \quad -1 \le \mu \le 1. \tag{6.7}$$

$I_1(\,.\,)$ is the modified Bessel function. The appropriate calculation can be found in [8, Section 3], in particular Lemma 3.2 there. From (6.5), (6.6) and (6.7) we obtain the following expression for the measure $d\rho(\mu)$ from (6.2):

$$\rho(d\mu) = \delta(\mu-1)\,d\mu + (1+\mu)\widehat{d}(\mu,\beta)\,d\mu. \tag{6.8}$$

References

1. Н.И. Ахиезер. *Классическая проблема моментов.* Физматгиз, Москва, (1965) (in Russian). English translation: Akhiezer, N.I.: The Clasical Moment Problem. Oliver and Boyd, Edinburgh (1965)
2. Bernstein, S.N.: Sur les functions absolument monotones. Acta Math. **52** (1928), 1–66. (In French)
3. С.Н. Бернштейн. *Собрание Сочинений. Том 1.* Издательство АН СССР, 1952
4. Bessis, D., Moussa, P., Villani, M.: Monotonic converging variational approximations to the functional integrals in quantum statistical mechanics. J. Math. Phys. **16**(11), 2318–2325 (1975)
5. Eremenko, A.: Herbert Stahl's proof of the BMV conjecture. arXiv:1312.6003
6. А.Э.Еременко. *Доказательство Герберта Шталя гипотезы БМВ.* Матем. Сб., **206**(1), 97–102 (2015) (in Russian). English translation: Eremenko, A.: Herbert Stahl's proof of the BMV conjecture. Sb. Math. **206**(1), 87–92 (2015)
7. Hall, B.C.: Lie Groups, Lie Algebras and Representations. Springer, New York (2003)
8. Katsnelson, V.: The function cosh $\sqrt{at^2+b}$ is exponentially convex. arXiv:1502.04201
9. Mehta, M.L., Kumar, K.: On an integral representation of the function Tr $[e^{A-\lambda B}]$. J. Phys. A Math. Gen. **9**(2), 197–206 (1976)
10. Stahl H.: Proof of the BMV conjecture, pp. 1–56. arXiv:1107.4875v1. 25 July 2011 (2011)
11. Stahl H.: Proof of the BMV conjecture, pp. 1–25. arXiv:1107.4875v3. 17 August 2012 (2012)
12. Stahl, H.: Proof of the BMV conjecture. Acta Math. **211**, 255–290 (2013)
13. Widder, D.V.: Laplace Transform. Princeton University Press, Princeton (1946)