Publ. RIMS, Kyoto Univ.
33 (1997), 443–463

Operator Convex Functions of Several Variables

By

Frank HANSEN*

Abstract

The functional calculus for functions of several variables associates to each tuple $x=(x_1,\cdots,x_k)$ of selfadjoint operators on Hilbert spaces $H_1,\cdots,H_k$ an operator $f(x)$ in the tensor product $B(H_1)\otimes\cdots\otimes B(H_k)$. We introduce the notion of generalized Hessian matrices associated with f. Those matrices are used as the building blocks of a structure theorem for the second Fréchet differential of the map $x\to f(x)$. As an application we derive that functions with positive semi-definite generalized Hessian matrices of arbitrary order are operator convex. The result generalizes a theorem of Kraus [15] for functions of one variable.

§1. Introduction

Let $f:I_1\times\cdots\times I_k\to\mathbf{R}$ be a real function of k variables defined on the product of k intervals, and let $x=(x_1,\cdots,x_k)$ be a tuple of selfadjoint matrices of order $n_1,\cdots,n_k$ such that the eigenvalues of x_i are contained in I_i for each $i=1,\cdots,k$. We say that such a tuple is in the domain of f and define $f(x)=f(x_1,\cdots,x_k)$ to be the matrix of order $n_1\cdots n_k$ constructed in the following way. For each $i=1,\cdots,k$ we consider the possibly degenerate spectral resolution

$$x_i=\sum_{m_i=1}^{n_i}\lambda_{m_i}(i)e^i_{m_im_i}$$

where $\{e^i_{s_iu_i}\}^{n_i}_{s_i,u_i=1}$ is the corresponding system of matrix units and let the formula

$$f(x_1,\cdots,x_n)=\sum_{m_1=1}^{n_1}\cdots\sum_{m_k=1}^{n_k}f(\lambda_{m_1}(1),\cdots,\lambda_{m_k}(k))e^1_{m_1m_1}\otimes\cdots\otimes e^k_{m_km_k}$$

Communicated by Y. Ihara, December 16, 1996.
1991 Mathematics Subject Classification(s): 26B25, 15A45, 46, 47 (as secondary)
* Institute of Economics, University of Copenhagen, Studiestraede 6, DK-1455 Copenhagen K, Denmark.

define the functional calculus. If f can be written as a product of k functions $f=f_1\cdots f_k$ where f_i is a function only of the ith coordinate, then $f(x_1,\cdots,x_k)=f_1(x_1)\otimes\cdots\otimes f_k(x_k)$. The given definition is readily extended to bounded normal operators on a Hilbert space, cf. [14].

The above function f of k real variables is said to be matrix convex of order $(n_1,\cdots,n_k)$, if

$$(*)\qquad f(\lambda x_1+(1-\lambda)y_1,\cdots,\lambda x_k+(1-\lambda)y_k)\leq\lambda f(x_1,\cdots,x_k)+(1-\lambda)f(y_1,\cdots,y_k)$$

for every $\lambda\in[0,1]$ and all tuples of selfadjoint matrices $(x_1,\cdots,x_k)$ and $(y_1,\cdots,y_k)$ such that the orders of x_i and y_i are n_i and their eigenvalues are contained in I_i for $i=1,\cdots,k$. The definition is meaningful since also the spectrum of $\lambda x_i+(1-\lambda)y_i$ is contained in the interval I_i for each $i=1,\cdots,k$. It is clear that the pointwise limit of a sequence of matrix convex functions of order $(n_1,\cdots,n_k)$ is again matrix convex of order $(n_1,\cdots,n_k)$. If f is matrix convex of order $(n_1,\cdots,n_k)$, then it is also matrix convex of any order $(n'_1,\cdots,n'_k)$ such that $n'_i\leq n_i$ for $i=1,\cdots,k$. If f is matrix convex of all orders, then we say that f is operator convex. If $I_1,\cdots,I_k$ are open intervals, then it is enough to assume that f is mid-point matrix convex of arbitrary order. This follows because such a function is real analytic and hence continuous, cf. the discussion in the introduction of [10]. It is the aim of this article to develop tools that make it possible to investigate the notion of operator convexity for functions of several variables, thus generalizing the theorem of Kraus [15] for functions of one variable.

§2. The Fréchet Differential

Let X and Y be Banach spaces. We say that a function $f:A\to Y$ defined on a subset A of X is Fréchet differentiable at an inner point $x_0\in A$, if there exists a bounded linear operator $df(x_0)\in B(X,Y)$ such that

$$\lim_{h\to 0}\|h\|^{-1}(f(x_0+h)-f(x_0)-df(x_0)h)=0.$$

Likewise f is said to be Fréchet differentiable in an open set A, if f is Fréchet differentiable at every point $x_0\in A$. We say that f is continuously Fréchet differentiable, if the differential mapping $A\ni x\to df(x)\in B(X,Y)$ is continuous. This notion of differentiability has been used to study perturbation formulas associated with the functional calculus in C^*-algebras, cf. [12]. The present notation and various results from the theory of Fréchet differentiable functions between Banach spaces are taken from [8]. The first result is quoted from [12].

Proposition 2.1. *If $\mathscr{A}$ is a Banach algebra, then the exponential function $A \to \exp(A)$ is continuously Fréchet differentiable, and*

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

for all x and h in $\mathscr{A}$.

The Fréchet differential df of a Fréchet differentiable function $f: A \to Y$ defined on an open subset $A \subseteq X$ is a function from A into the Banach space $B(X, Y)$ of bounded linear functions from X to Y. If df is Fréchet differentiable, then we define the second Fréchet differential of f, denoted by d^2f, to be the Fréchet differential of df. The second order Fréchet differential can be considered as a function $d^2f: A \to B_2(X, Y)$ from A into the Banach space of bounded bilinear functions from X to Y. We notice that $d(df(x)h)k = d^2f(x)(h,k)$ for $h,k \in X$, and that $d^2f(x)$ is symmetric in the sense that $d^2f(x)(h,k) = d^2f(x)(k,h)$, cf. the standard reference [8]. The following proposition is the starting point in our investigation of operator convex functions.

Proposition 2.2. *If A is an open convex subset of a real Banach space X and $B(H)_{sa}$ is the space of bounded selfadjoint operators on a Hilbert space H, then a twice Fréchet differentiable function $f: A \to B(H)_{sa}$ is convex, if and only if $d^2f(x)(h,h) \geq 0$ for each $x \in A$ and $h \in X$.*

The result follows by adapting the reasoning of classical analysis to the present situation and can be found in [8, Exercises 3.1.8 and 3.6.4]. The following elementary result is stated without proof.

Lemma 2.3. *Let X be a Banach space and Y a Banach algebra, and let $F, G: A \to Y$ be mappings which are Fréchet differentiable at an interior point $x_0 \in A \subseteq X$. Then the mapping $(FG)(x) = F(x)G(x)$ is Fréchet differentiable at x_0, and the Fréchet differential is*

$$d(FG)(x_0)h = (dF(x_0)h)G(x_0) + F(x_0)dG(x_0)h$$

for each $h \in X$.

Applying the above lemma to the Fréchet differential of the exponential mapping, we obtain

Proposition 2.4. *If* $\mathscr{A}$ *is a complex Banach algebra and* $t \in \mathbf{R}$, *then the function* $x \to \exp(itx)$ *is twice continuously Fréchet differentiable, and*

$$d\exp(itx)h = it\int_0^1 \exp(itsx)h\exp(it(1-s)x)ds$$

$$d^2\exp(itx)(h,h') = -t^2\int_0^1\int_0^1 [s\exp(itusx)h'\exp(it(1-u)sx)h\exp(it(1-s)x)$$
$$+(1-s)\exp(itsx)h\exp(itu(1-s)x)h'\exp(it(1-u)(1-s)x)]du\,ds$$

for all x,h,h' *in* $\mathscr{A}$.

Let $x=(x_1,\cdots,x_k)$ be a tuple of bounded operators on Hilbert spaces $H_1,\cdots,H_k$. The exponential function

$$\exp(it\cdot x)=\exp(it_1x_1)\otimes\cdots\otimes\exp(it_kx_k)$$

is everywhere defined in the product space $B(H_1)\times\cdots\times B(H_k)$ and maps it into the tensor product $B(H_1)\otimes\cdots\otimes B(H_k)$. The definition is consistent with the functional calculus of selfadjoint or normal operators as given in the introduction. The following result is a direct application of [8, Theorem 3.3.1 and Theorem 3.3.2].

Proposition 2.5. *The exponential function* $\exp(it\cdot x)=\exp(it_1x_1)\otimes\cdots\otimes\exp(it_kx_k)$ *is continuously Fréchet differentiable, and*

$$d\exp(it\cdot x)a=\sum_{i=1}^{k}\exp(it_1x_1)\otimes\cdots\otimes d\exp(it_ix_i)a_i\otimes\cdots\otimes\exp(it_kx_k)$$

for each $t=(t_1,\cdots,t_k)\in\mathbf{R}^k$ *and each tuple of operators* $a=(a_1,\cdots,a_k)$. *It is understood that the differential operator is applied only in position number i of the tensor product in each term.*

Corollary 2.6. *The exponential function is infinitely many times Fréchet differentiable, and the second derivative is given by*

$$d^2 \exp(it\cdot x)(a,b) = d(d\exp(it\cdot x)a)b$$

$$= \sum_{i,j=1;i\neq j}^{k} \exp(it_1x_1)\otimes\cdots\otimes d\exp(it_ix_i)a_i\otimes\cdots\otimes d\exp(it_jx_j)b_j\otimes\cdots\otimes\exp(it_kx_k)$$

$$+\sum_{i=1}^{k}\exp(it_1x_1)\otimes\cdots\otimes d^2\exp(it_ix_i)(a_i,b_i)\otimes\cdots\otimes\exp(it_kx_k).$$

We consider the set $C_0^p(\mathbf{R}^k)$ of real functions of k variables with continuous partial derivatives of order p and compact support.

Lemma 2.7. *Let $f\in C_0^p(\mathbf{R}^k)$ and let*

$$\tilde{f}(s) = \frac{1}{(2\pi)^k}\int_{\mathbf{R}^k} e^{it\cdot s} f(t)\, d^k t$$

denote the Fourier transform. Then $|s|\tilde{f}(s)$ is an integrable function for $p>1+k/2$ and $|s|^2\tilde{f}(s)$ is integrable for $p>2+k/2$.

Proof. The function f and its partial derivatives up to order p are continuous functions with compact support thus square integrable. It follows from Plancherel's theorem that the function $g(s)=(1+|s|)^p\tilde{f}(s)$ is square integrable. Since

$$|s|\tilde{f}(s) = g(s)\frac{|s|}{(1+|s|)^p}$$

we conclude that $|s|\tilde{f}(s)$ is integrable, if

$$\int_{\mathbf{R}^k}\frac{|s|^2}{(1+|s|)^{2p}}\, d^k s = \Omega_k\int_0^\infty \frac{r^2}{(1+r)^{2p}}\, r^{k-1}\, dr < \infty$$

where Ω_k is the volume of the surface of the unit sphere in $\mathbf{R}^k$. But this integral is finite, if and only if $2p-k-1>1$. Similarly

$$|s|^2\tilde{f}(s) = g(s)\frac{|s|^2}{(1+|s|)^p}$$

and the function $|s|^2(1+|s|)^{-p}$ is square integrable, if and only if $2p-k-3>1$.
QED

The following result is similar to [12, Theorem 1.5].

Theorem 2.8. *Let $f\in C_0^2(\mathbf{R})$. The function $x\to f(x)$ defined on the selfadjoint operators $B(H)_{sa}$ on a Hilbert space H is continuously Fréchet differentiable and*

$$\begin{aligned} df(x)h &= \int_{-\infty}^{\infty}(-is)\int_0^1 \exp(-istx)h\exp(-is(1-t)x)dt\tilde{f}(s)ds \\ &= \int_{-\infty}^{\infty} d\exp(-isx)h\tilde{f}(s)ds \end{aligned}$$

for all x,h in $B(H)_{sa}$.

Proof. The proposed expression of the Fréchet differential is bounded because $\exp(-istx)$ and $\exp(-is(1-t)x)$ are unitary operators, and the function $s\to s\tilde{f}(s)$ is integrable according to Lemma 2.7. The linear form depends continuously on x. We obtain

$$\begin{aligned} &f(x+h)-f(x)-\int_{-\infty}^{\infty}(-is)\int_0^1 \exp(-istx)h\exp(-is(1-t)x)dt\tilde{f}(s)ds \\ &=\int_{-\infty}^{\infty}\left(\exp(-is(x+h))-\exp(-isx)+is\int_0^1 \exp(-istx)h\exp(-is(1-t)x)dt\right)\tilde{f}(s)ds \\ &=\int_{-\infty}^{\infty}(-is)\int_0^1\Big(\exp(-ist(x+h))-\exp(-istx)\Big)h\exp(-is(1-t)x)dt\tilde{f}(s)ds, \end{aligned}$$

where we used the Dyson formula, cf. [12]. The norm of this expression is bounded by

$$\|h\|\int_{-\infty}^{\infty}|s|\int_0^1\|\exp(-ist(x+h))-\exp(-istx)\|dt|\tilde{f}(s)|ds$$

and even after division by $\|h\|$ this does tend to zero as $h\to 0$ by Lebesgue's theorem of dominated convergence. QED

Corollary 2.9. *Let $f \in C^2(I)$ where I is an open interval. The function $x \to f(x)$ defined on operators $x \in B(H)_{sa}$ with spectra contained in I is continuously Fréchet differentiable.*

Proof. Let $x \in B(H)_{sa}$ have spectrum $Sp(x) \subset I$. Since $Sp(x)$ is compact, we can find an open and bounded interval J such that

$$Sp(x) \subset J \subset \bar{J} \subset I.$$

The function f is bounded on the closure $\bar{J}$, so we can extend the restriction of f to J to a function in $C_0^2(\mathbf{R})$. Since continuous Fréchet differentiability of the mapping $x \to f(x)$ in a point x only depends on f in a neighborhood of the spectrum of x, the assertion follows. QED

Let d_i denote the partial Fréchet differential operator associated with a function defined on a product space, cf. [8, Section 3.3], and let P_i denote the insertion mapping which inserts h_i into the ith coordinate of the zero vector in $B(H) \times \cdots \times B(H)$. The partial Fréchet differential $d_i f(x)$ is the differential of the function $h_i \to f(x + P_i h_i)$ at $h_i = 0$.

Corollary 2.10. *Let $f \in C^p(I)$ where $I = I_1 \times \cdots \times I_k$ is a product of open intervals and $p > 1 + k/2$. The function $x \to f(x)$ defined on tuples of selfadjoint operators $x = (x_1, \cdots, x_k)$ in $B(H)$ contained in the domain of f is continuously Fréchet differentiable and*

$$df(x)h = \sum_{i=1}^{k} d_i f(x)h_i$$

for every tuple $h = (h_1, \cdots, h_k)$ with $h_i \in B(H)_{sa}$ for $i = 1, \cdots, k$. If f has compact support, then the Fréchet differential can be written as

$$df(x)h = \int_{\mathbf{R}^k} d\exp(-is \cdot x)h\tilde{f}(s)d^k s$$

where the Fréchet differential under the integral is taken with respect to $x = (x_1, \cdots, x_k)$.

Proof. We may assume that f has compact support. Since

$$f(x + P_i h_i) - f(x) - \int_{\mathbf{R}^k} d_i \exp(-is \cdot x)h_i \tilde{f}(s)d^k s$$

$$= \int_{\mathbf{R}^k} \left(\exp(-is\cdot(x+P_i h_i)) - \exp(-is\cdot x) - d_i \exp(-is\cdot x)h_i \right) \tilde{f}(s) d^k s$$

and this expression is bounded in norm by

$$\|h_i\| \int_{\mathbf{R}^k} |s_i| \int_0^1 \|\exp(-is_i t(x_i+h_i)) - \exp(-is_i t x_i)\| dt \, |\tilde{f}(s)| d^k s$$

we obtain that the function $x \to f(x)$ has partial Fréchet differentials given by

$$d_i f(x) h_i = \int_{\mathbf{R}^k} d_i \ \exp(-is\cdot x) h_i \tilde{f}(s) d^k s.$$

This entails that $x \to f(x)$ is Fréchet differentiable with Fréchet differential

$$df(x)h = \sum_{i=1}^k d_i f(x) h_i = \sum_{i=1}^k \int_{\mathbf{R}^k} d_i \ \exp(-is\cdot x) h_i \tilde{f}(s) d^k s$$
$$= \int_{\mathbf{R}^k} d \exp(-is\cdot x) h \tilde{f}(s) d^k s$$

according to [8, Theorem 3.3.2] and Proposition 2.5. QED

We then consider twice Fréchet differentiable functions. Since the reasoning is very similar to the above arguments, the exposition is brief.

Proposition 2.11. *Let $f \in C_0^3(\mathbf{R})$. The function $x \to f(x)$ defined on the selfadjoint operators $B(H)_{sa}$ on a Hilbert space H is twice continuously Fréchet differentiable and*

$$d^2 f(x)(a,b) = \int_{\mathbf{R}} d^2 \exp(-isx)(a,b) \tilde{f}(s) ds$$

for all $a, b \in B(H)_{sa}$ where the Fréchet differential under the integral is taken with respect to x.

Proof. The function $s \to |s|^2 \tilde{f}(s)$ is integrable according to Lemma 2.7 and the result now follows as in Theorem 2.8. QED

Corollary 2.12. *Let $f \in C^p(I)$ where $I = I_1 \times \cdots \times I_k$ is a product of open intervals and $p > 2 + k/2$. The function $x \to f(x)$ defined on tuples of selfadjoint operators $x = (x_1, \cdots, x_k)$ in $B(H)$ contained in the domain of f is twice continuously Fréchet differentiable and*

$$d^2 f(x)(a,b) = \sum_{i,j=1}^{k} d_i d_j f(x)(a_i, b_j)$$

for all tuples $a = (a_1, \cdots, a_k)$ and $b = (b_1, \cdots, b_k)$ with $a_i, b_i \in B(H)_{sa}$ for $i = 1, \cdots, k$. If f has compact support, then the second Fréchet differential can be written as

$$d^2 f(x)(a,b) = \int_{\mathbf{R}^k} d^2 \exp(-is \cdot x)(a,b) \tilde{f}(s) d^k s$$

where the Fréchet differential under the integral is taken with respect to $x = (x_1, \cdots, x_k)$.

§3. Generalized Hessian Matrices

Let f be a twice continuously differentiable real function defined on an open interval $I \subseteq \mathbf{R}$. The divided difference $[\lambda\mu]$ of f taken in the points λ, $\mu \in I$ is defined as

$$[\lambda\mu] = \begin{cases} \dfrac{f(\lambda) - f(\mu)}{\lambda - \mu} & \text{for } \lambda \neq \mu \\[2ex] f'(\lambda) & \text{for } \lambda = \mu \end{cases}$$

and it is a symmetric function of the two arguments with partial derivatives in each of the two variables. The second divided difference $[\lambda\mu\zeta]$ taken in the points $\lambda, \mu, \zeta \in I$ is defined as

$$[\lambda\mu\zeta] = \begin{cases} \dfrac{[\lambda\mu] - [\mu\zeta]}{\lambda - \zeta} & \text{for } \lambda \neq \zeta \\[2ex] \dfrac{\partial}{\partial\lambda} [\lambda\mu] & \text{for } \lambda = \zeta \end{cases}$$

and it is a symmetric function of the three arguments, cf. [6] for a more systematic introduction to divided differences for functions of one variable.

If f is a real function defined on the product $I_1 \times I_2$ of two open intervals with continuous partial derivatives up to the second order, then we can consider the divided differences $[\lambda\mu|\xi]$ and $[\lambda\mu\zeta|\xi]$ which are just the previously defined divided differences for the function of one variable obtained by fixing the second variable to ξ. We define the divided differences $[\xi|\lambda\mu]$ and $[\xi|\lambda\mu\zeta]$ similarly. There are, however, also mixed second derivatives defined as

$$[\lambda\mu|\zeta\xi] = \begin{cases} \dfrac{[\lambda|\zeta\xi]-[\mu|\zeta\xi]}{\lambda-\mu} & \text{for } \lambda\neq\mu \\ \dfrac{\partial}{\partial\lambda}[\lambda|\zeta\xi] & \text{for } \lambda=\mu. \end{cases}$$

We could have defined the mixed derivatives by dividing to the right instead of dividing to the left, but this gives the same result. Finally, if f is a real function defined on the product $I_1 \times \cdots \times I_k$ of k open intervals with continuous partial derivatives up to the second order, then we consider the second divided differences that appear by fixing all but one or two of the k coordinates of f. They are labeled as

$$[\lambda_1 | \cdots |\mu_1\mu_2\mu_3| \cdots | \lambda_k]^i$$

where the superscript i indicates that the partial divided difference of the second order is taken at the ith coordinate and all other coordinates are fixed at the values $\lambda_1, \cdots, \lambda_{i-1}$ and $\lambda_{i+1}, \cdots, \lambda_k$ or as

$$[\lambda_1 | \cdots |\mu_1\mu_2| \cdots |\xi_1\xi_2| \cdots |\lambda_k]^{ij}$$

where the superscripts ij indicate that the mixed partial divided difference of the second order is taken at the distinctly different coordinates i and j and all other coordinates are fixed at the values $\lambda_1, \cdots, \lambda_{i-1}, \lambda_{i+1}, \cdots, \lambda_{j-1}$ and $\lambda_{j+1}, \cdots, \lambda_k$. The notation does not imply any particular order of the coordinates which can be chosen from the full range $1, \cdots, k$.

Definition 3.1. *Let $f: I_1 \times \cdots \times I_k \to \mathbf{R}$ be a real function of k variables defined on the product of k open intervals with continuous partial derivatives up to the second order. We define a data set Λ of order $(n_1, \cdots, n_k)$ for f to be an element $\Lambda \in I_1^{n_1} \times \cdots \times I_k^{n_k}$, and we usually write it the form*

$$(*) \qquad \Lambda = \{\lambda_{m_i}(i)\}_{m_i=1,\cdots,n_i} \qquad i=1,\cdots,k.$$

To a given data set Λ *we associate so-called generalized Hessian matrices. First we define to each tuple of natural numbers* $(m_1, \cdots, m_k) \le (n_1, \cdots, n_k)$ *and to any* $s, u = 1, \cdots, k$ *a matrix denoted* $H_{us}(m_1, \cdots, m_k)$ *of order* $n_u \times n_s$ *in the following way*:

1. *If* $s \ne u$, *then we set*

$$H_{us}(m_1, \cdots, m_k) =$$
$$([\lambda_{m_1}(1)| \cdots |\lambda_{m_s}(s)\lambda_j(s)| \cdots |\lambda_p(u)\lambda_{m_u}(u)| \cdots |\lambda_{m_k}(k)]^{su})_{p=1,\cdots,n_u; j=1,\cdots,n_s}$$

2. *If* $s = u$, *then we set*

$$H_{ss}(m_1, \cdots, m_k) = 2([\lambda_{m_1}(1)| \cdots |\lambda_{m_s}(s)\lambda_p(s)\lambda_j(s)| \cdots |\lambda_{m_k}(k)]^s)_{p,j=1,\cdots,n_s}.$$

We then define the generalized Hessian matrix as the block matrix

$$H(m_1, \cdots, m_k) = (H_{us}(m_1, \cdots, m_k))_{u,s=1,\cdots,k}$$

which is quadratic and symmetric and of order $n_1 + \cdots + n_k$.

If $n_i = 1$ for $i = 1, \cdots, k$ then the data set (*) reduces to k numbers $\lambda(1), \cdots, \lambda(k)$ and there is only one (generalized) Hessian matrix H. The submatrix H_{us} is a 1×1 matrix with the partial derivative $f''_{us}(\lambda(1), \cdots, \lambda(k))$ as matrix element for $s, u = 1, \cdots, k$. Therefore H can be identified with the usual Hessian matrix associated with a function of k variables. The notion of generalized Hessian matrices can be extended from real valued functions to complex valued functions of k real variables. We shall do this for the exponential functions $s \to e^{it \cdot s}$ without further remarks.

The generalized Hessian matrices are used in the structure theorem for the second Fréchet differential of the mapping associated with the functional calculus for f, and they are useful to investigate the notion of matrix convexity.

§4. The Structure of the Second Fréchet Differential

Lemma 4.1. *Let* x *be a bounded selfadjoint operator on a Hilbert space* H *of finite dimension* n, *and let* $\{e_{ij}\}_{i,j=1}^n$ *be a system of matrix units in* $B(H)$ *such that* $x = \sum_{i=1}^n \lambda_i e_{ii}$. *Then*

$$d\exp(itx)e_{ij} = [\lambda_i \lambda_j]e_{ij}$$
$$d^2\exp(itx)(e_{ij}, e_{pq}) = \delta_{iq}[\lambda_i \lambda_p \lambda_j]e_{pj} + \delta_{jp}[\lambda_j \lambda_i \lambda_q]e_{iq}$$

for $i,j,p,q=1,\cdots,n$; *where the divided differences are taken with respect to the function* $s \to e^{its}$.

Proof. Applying Proposition 2.4 we obtain

$$d\exp(itx)e_{ij}=it\int_0^1 \exp(itsx)e_{ij}\exp(it(1-s)x)ds=it\int_0^1 e^{its\lambda_i}e_{ij}e^{it(1-s)\lambda_j}ds$$

which is evaluated to $[\lambda_i\lambda_j]e_{ij}$ and similarly

$$d^2\exp(itx)(e_{ij},e_{pq})=-t^2\int_0^1\int_0^1 [s\exp(itusx)e_{pq}\exp(it(1-u)sx)e_{ij}\exp(it(1-s)x)$$

$$+(1-s)\exp(itsx)e_{ij}\exp(itu(1-s)x)e_{pq}\exp(it(1-u)(1-s)x)]du\,ds$$

$$=-t^2\int_0^1\int_0^1 [\delta_{qi}se^{itus\lambda_p}e^{it(1-u)s\lambda_i}e^{it(1-s)\lambda_j}e_{pj}$$

$$+\delta_{jp}(1-s)e^{its\lambda_i}e^{itu(1-s)\lambda_j}e^{it(1-u)(1-s)\lambda_q}e_{iq}]du\,ds$$

which is evaluated to $\delta_{iq}[\lambda_i\lambda_p\lambda_j]e_{pj}+\delta_{jp}[\lambda_j\lambda_i\lambda_q]e_{iq}$. The cases where indices or eigenvalues coincide are considered separately. QED

Lemma 4.2. *Let* $x=(x_1,\cdots,x_k)$ *be selfadjoint matrices acting on finite dimensional Hilbert spaces* $H_1,\cdots,H_k$ *of dimensions* $n_1,\cdots,n_k$ *and consider for each* $i=1,\cdots,k$ *a possibly degenerate spectral resolution*

$$x_i=\sum_{m_i=1}^{n_i}\lambda_{m_i}(i)e^i_{m_im_i}$$

where $\{e^i_{s_iu_i}\}^{n_i}_{s_i,u_i=1}$ *is the corresponding system of matrix units. We shall for later use adopt the notation* $\{e^i_{m_i}\}^{n_i}_{m_i=1}$ *for a system of unital eigenvectors of the one-dimensional projections* $e^i_{m_im_i}$.

The second Fréchet differential of the exponential function $x\to\exp(it\cdot x)$ *satisfy*

$$d^2\exp(it\cdot x)\Big((\alpha_1e^1_{ij},\cdots,\alpha_ke^k_{ij}),(\beta_1e^1_{pq},\cdots,\beta_ke^k_{pq})\Big)$$

$$=\sum_{s,u=1;s\neq u}^{k}\ \sum_{m_1=1}^{n_1}\cdots\sum_{m_{s-1}=1}^{n_{s-1}}\ \sum_{m_{s+1}=1}^{n_{s+1}}\cdots\sum_{m_{u-1}=1}^{n_{u-1}}\ \sum_{m_{u+1}=1}^{n_{u+1}}\cdots\sum_{m_k=1}^{n_k}$$

$$\alpha_s\beta_u[\lambda_{m_1}(1)|\cdots|\lambda_i(s)\lambda_j(s)|\cdots|\lambda_p(u)\lambda_q(u)|\cdots|\lambda_{m_k}(k)]^{su}$$

$$e^1_{m_1m_1}\otimes\cdots\otimes e^s_{ij}\otimes\cdots\otimes e^u_{pq}\otimes\cdots\otimes e^k_{m_km_k}$$

$$+\sum_{s=1}^{k}\sum_{m_1=1}^{n_1}\cdots\sum_{m_{s-1}=1}^{n_{s-1}}\sum_{m_{s+1}=1}^{n_{s+1}}\cdots\sum_{m_k=1}^{n_k}\alpha_s\beta_s$$

$$(\delta_{iq}[\lambda_{m_1}(1)|\cdots|\lambda_i(s)\lambda_p(s)\lambda_j(s)|\cdots|\lambda_{m_k}(k)]^s e^1_{m_1m_1}\otimes\cdots\otimes e^s_{pj}\otimes\cdots\otimes e^k_{m_km_k}$$

$$+\delta_{jp}[\lambda_{m_1}(1)|\cdots|\lambda_j(s)\lambda_i(s)\lambda_q(s)|\cdots|\lambda_{m_k}(k)]^s e^1_{m_1m_1}\otimes\cdots\otimes e^s_{iq}\otimes\cdots\otimes e^k_{m_km_k})$$

for all complex sequences $\alpha_1,\cdots,\alpha_k$ *and* $\beta_1,\cdots,\beta_k$. *We use the convention that* $e^s_{ij}=\underline{0}$ *if* $\max\{i,j\}>n_s$ *for* $s=1,\cdots,k$. *The partial divided differences are taken with respect to the function* $s\to e^{it\cdot s}$.

Proof. By Corollary 2.6 we obtain

$$d^2\exp(it\cdot x)\left((\alpha_1e^1_{ij},\cdots,\alpha_ke^k_{ij}),(\beta_1e^1_{pq},\cdots,\beta_ke^k_{pq})\right)=$$

$$\sum_{s,u=1,s\neq u}^{k}\alpha_s\beta_u\exp(it_1x_1)\otimes\cdots\otimes d\exp(it_sx_s)e^s_{ij}\otimes\cdots\otimes d\exp(it_ux_u)e^u_{pq}\otimes\cdots\otimes\exp(it_kx_k)$$

$$+\sum_{s=1}^{k}\alpha_s\beta_s\exp(it_1x_1)\otimes\cdots\otimes d^2\exp(it_sx_s)(e^s_{ij},e^s_{pq})\otimes\cdots\otimes\exp(it_kx_k)$$

which combined with Lemma 4.1 and the spectral theorem applied in each of the variables not subject to differentiation give the desired formula. QED

Theorem 4.3. *Let* $f\in C^p(I_1\times\cdots\times I_k)$ *with* $p>2+k/2$ *where* $I_1,\cdots,I_k$ *are open intervals, and let* $x=(x_1,\cdots,x_k)$ *be a tuple of selfadjoint matrices of order* $(n_1,\cdots,n_k)$ *in the domain of* f. *We consider the data set*

$$\Lambda=\{\lambda_{m_i}(i)\}_{m_i=1,\cdots,n_i}\qquad i=1,\cdots,k$$

consisting of the (possibly degenerate) eigenvalues of $(x_1,\cdots,x_k)$, *and the ensemble of generalized Hessian matrices associated with* f *and* Λ. *The second Fréchet differential is then given by*

$$d^2f(x)(h,h)=\sum_{m_1=1}^{n_1}\cdots\sum_{m_k=1}^{n_k}\sum_{s,u=1}^{k}\sum_{j=1}^{n_s}\sum_{p=1}^{n_u}h^s_{m_sj}h^u_{pm_u}(H_{us}(m_1,\cdots,m_k))_{pj}$$

$$(e^1_{m_1m_1}\otimes\cdots\otimes e^u_{pm_u}\otimes\cdots\otimes e^k_{m_km_k})(e^1_{m_1m_1}\otimes\cdots\otimes e^s_{m_sj}\otimes\cdots\otimes e^k_{m_km_k})$$

for any tuple $h=(h^1,\cdots,h^k)$ *of selfadjoint matrices of order* $(n_1,\cdots,n_k)$.

Proof. We first prove the statement for the exponential functions $s \to e^{it\cdot s}$ where $t=(t_1,\cdots,t_k)$ is a fixed parameter in $\mathbf{R}^k$. We set $n=\max\{n_1,\cdots,n_k\}$ and expand $h=(h^1,\cdots,h^k)$ with respect to the corresponding matrix units

$$h=(h^1,\cdots,h^k)=\sum_{i,j=1}^{n}(h_{ij}^1e_{ij}^1,\cdots,h_{ij}^ke_{ij}^k)$$

where we set $h_{ij}^s=0$ and $e_{ij}^s=\underline{0}$ if $\max\{i,j\}>n_s$ for $s=1,\cdots,k$. We obtain

$$d^2\exp(it\cdot x)(h,h)=d^2\exp(it\cdot x)\Big((h^1,\cdots,h^k),(h^1,\cdots,h^k)\Big)$$

$$=d^2\exp(it\cdot x)\Big(\sum_{i,j=1}^{n}(h_{ij}^1e_{ij}^1,\cdots,h_{ij}^ke_{ij}^k),\sum_{p,q=1}^{n}(h_{pq}^1e_{pq}^1,\cdots,h_{pq}^ke_{pq}^k)\Big)$$

$$=\sum_{i,j,p,q=1}^{n}d^2\exp(it\cdot x)\Big((h_{ij}^1e_{ij}^1,\cdots,h_{ij}^ke_{ij}^k),(h_{pq}^1e_{pq}^1,\cdots,h_{pq}^ke_{pq}^k)\Big).$$

Applying Lemma 4.2 one gets

$$d^2\exp(it\cdot x)(h,h)$$

$$=\sum_{i,j,p,q=1}^{n}\Bigg[\sum_{s,u=1;s\neq u}^{k}\sum_{m_1=1}^{n_1}\cdots\sum_{m_{s-1}=1}^{n_{s-1}}\sum_{m_{s+1}=1}^{n_{s+1}}\cdots\sum_{m_{u-1}=1}^{n_{u-1}}\sum_{m_{u+1}=1}^{n_{u+1}}\cdots\sum_{m_k=1}^{n_k}$$

$$h_{ij}^sh_{pq}^u[\lambda_{m_1}(1)|\cdots|\lambda_i(s)\lambda_j(s)|\cdots|\lambda_p(u)\lambda_q(u)|\cdots|\lambda_{m_k}(k)]^{su}$$

$$e_{m_1m_1}^1\otimes\cdots\otimes e_{ij}^s\otimes\cdots\otimes e_{pq}^u\otimes\cdots\otimes e_{m_km_k}^k$$

$$+\sum_{s=1}^{k}\sum_{m_1=1}^{n_1}\cdots\sum_{m_{s-1}=1}^{n_{s-1}}\sum_{m_{s+1}=1}^{n_{s+1}}\cdots\sum_{m_k=1}^{n_k}h_{ij}^sh_{pq}^s$$

$$(\delta_{iq}[\lambda_{m_1}(1)|\cdots|\lambda_i(s)\lambda_p(s)\lambda_j(s)|\cdots|\lambda_{m_k}(k)]^se_{m_1m_1}^1\otimes\cdots\otimes e_{pj}^s\otimes\cdots\otimes e_{m_km_k}^k$$

$$+\delta_{jp}[\lambda_{m_1}(1)|\cdots|\lambda_j(s)\lambda_i(s)\lambda_q(s)|\cdots|\lambda_{m_k}(k)]^se_{m_1m_1}^1\otimes\cdots\otimes e_{iq}^s\otimes\cdots\otimes e_{m_km_k}^k)\Bigg].$$

We evaluate the terms containing the Kronecker symbols, and after applying the transformation $(i,q,j)\to(p,j,i)$ in the last term of the second sum, one obtains

$$d^2\exp(it\cdot x)(h,h)$$

$$=\sum_{s=1}^{k}\Bigg[\sum_{i,j,p,q=1}^{n}\sum_{u=1;u\neq s}^{k}\sum_{m_1=1}^{n_1}\cdots\sum_{m_{s-1}=1}^{n_{s-1}}\sum_{m_{s+1}=1}^{n_{s+1}}\cdots\sum_{m_{u-1}=1}^{n_{u-1}}\sum_{m_{u+1}=1}^{n_{u+1}}\cdots\sum_{m_k=1}^{n_k}$$

$$h^s_{ij}h^u_{pq}[\lambda_{m_1}(1)|\cdots|\lambda_i(s)\lambda_j(s)|\cdots|\lambda_p(u)\lambda_q(u)|\cdots|\lambda_{m_k}(k)]^{su}$$

$$e^1_{m_1m_1}\otimes\cdots\otimes e^s_{ij}\otimes\cdots\otimes e^u_{pq}\otimes\cdots\otimes e^k_{m_km_k}$$

$$+\sum_{i,j,p=1}^{n}\sum_{m_1=1}^{n_1}\cdots\sum_{m_{s-1}=1}^{n_{s-1}}\sum_{m_{s+1}=1}^{n_{s+1}}\cdots\sum_{m_k=1}^{n_k}$$

$$2h^s_{ij}h^s_{pi}[\lambda_{m_1}(1)|\cdots|\lambda_i(s)\lambda_p(s)\lambda_j(s)|\cdots|\lambda_{m_k}(k)]^s e^1_{m_1m_1}\otimes\cdots\otimes e^s_{pj}\otimes\cdots\otimes e^k_{m_km_k}\Bigg].$$

Since every term with $i>n_s$ is zero and every term with $q>n_u$ is zero, we can change the variables $(i,q)\to(m_s,m_u)$ and obtain

$$d^2\exp(it\cdot x)(h,h)=\sum_{s=1}^{k}\Bigg[\sum_{j,p=1}^{n}\sum_{u=1;u\neq s}^{k}\sum_{m_1=1}^{n_1}\cdots\sum_{m_k=1}^{n_k}$$

$$h^s_{m_sj}h^u_{pm_u}[\lambda_{m_1}(1)|\cdots|\lambda_{m_s}(s)\lambda_j(s)|\cdots|\lambda_p(u)\lambda_{m_u}(u)|\cdots|\lambda_{m_k}(k)]^{su}$$

$$e^1_{m_1m_1}\otimes\cdots\otimes e^s_{m_sj}\otimes\cdots\otimes e^u_{pm_u}\otimes\cdots\otimes e^k_{m_km_k}$$

$$+\sum_{j,p=1}^{n}\sum_{m_1=1}^{n_1}\cdots\sum_{m_k=1}^{n_k}2h^s_{m_sj}h^s_{pm_s}[\lambda_{m_1}(1)|\cdots|\lambda_{m_s}(s)\lambda_p(s)\lambda_j(s)|\cdots|\lambda_{m_k}(k)]^s$$

$$e^1_{m_1m_1}\otimes\cdots\otimes e^s_{pj}\otimes\cdots\otimes e^k_{m_km_k}\Bigg].$$

Rearranging the sums and splitting the tensor products then give

$$d^2\exp(it\cdot x)(h,h)=\sum_{m_1=1}^{n_1}\cdots\sum_{m_k=1}^{n_k}\sum_{s=1}^{k}\Bigg[\sum_{u=1;u\neq s}^{k}\sum_{j=1}^{n_s}\sum_{p=1}^{n_u}$$

$$h^s_{m_sj}h^u_{pm_u}[\lambda_{m_1}(1)|\cdots|\lambda_{m_s}(s)\lambda_j(s)|\cdots|\lambda_p(u)\lambda_{m_u}(u)|\cdots|\lambda_{m_k}(k)]^{su}$$

$$(e^1_{m_1m_1}\otimes\cdots\otimes e^u_{pm_u}\otimes\cdots\otimes e^k_{m_km_k})(e^1_{m_1m_1}\otimes\cdots\otimes e^s_{m_sj}\otimes\cdots\otimes e^k_{m_km_k})$$

$$+\sum_{j,p=1}^{n_s}2h^s_{m_sj}h^s_{pm_s}[\lambda_{m_1}(1)|\cdots|\lambda_{m_s}(s)\lambda_p(s)\lambda_j(s)|\cdots|\lambda_{m_k}(k)]^s$$

$$(e^1_{m_1m_1}\otimes\cdots\otimes e^s_{pm_s}\otimes\cdots\otimes e^k_{m_km_k})(e^1_{m_1m_1}\otimes\cdots\otimes e^s_{m_sj}\otimes\cdots\otimes e^k_{m_km_k})\Bigg].$$

Inserting the elements of the generalized Hessian matrices $H_{us}(m_1,\cdots,m_k)$ associated with the function $s\to e^{it\cdot s}$ and the data set Λ we obtain

$$d^2\exp(it\cdot x)(h,h)=\sum_{m_1=1}^{n_1}\cdots\sum_{m_k=1}^{n_k}\sum_{s,u=1}^{k}\sum_{j=1}^{n_s}\sum_{p=1}^{n_u}h^s_{m_sj}h^u_{pm_u}(H_{us}(m_1,\cdots,m_k))_{pj}$$

$$(e^1_{m_1m_1}\otimes\cdots\otimes e^u_{pm_u}\otimes\cdots\otimes e^k_{m_km_k})(e^1_{m_1m_1}\otimes\cdots\otimes e^s_{m_sj}\otimes\cdots\otimes e^k_{m_km_k})$$

and the statement is proved for the exponential functions. If $f\in C_0^p(\mathbb{R}^k)$ for $p>2+k/2$, then we apply Corollary 2.12. The evaluation of the second Fréchet differential can be carried out under the integral by Lebesgue's theorem of dominated convergence since the function $|s|^2\tilde{f}(s)$ is integrable. The statement now follows from the linearity in f of the generalized Hessian matrices. If f does not have compact support, then we consider the restriction of f to $J_1\times\cdots\times J_k$ where $J_1,\cdots,J_k$ are bounded open intervals such that the spectrum of x_i is contained in J_i and the closure $\bar{J}_i\subset I_i$ for $i=1,\cdots,k$. The restriction of f is then extended to a function in $C_0^p(\mathbb{R}^k)$ and the statement follows because Fréchet differentiability of the mapping $x\to f(x)$ in a point x only depends on f in a neighborhood of the spectra of $x=(x_1,\cdots,x_k)$. QED

Corollary 4.4. *Let in the setting of Lemma* 4.2 *and Theorem* 4.3

$$\varphi=\sum_{m_1=1}^{n_1}\cdots\sum_{m_k=1}^{n_k}\varphi(m_1,\cdots,m_k)e^1_{m_1}\otimes\cdots\otimes e^k_{m_k}$$

be an arbitrary vector in the tensor product $H_1\otimes\cdots\otimes H_k$. *There exists a hermitian, sesquilinear form* $\omega_\varphi(a,b)$ *defined on the complex vector space of tuples of matrices of order* $(n_1,\cdots,n_k)$ *such that the expectation value of the second Fréchet differential*

$$(d^2f(x)(a,b)\varphi\mid\varphi)=\operatorname{Re}\ \omega_\varphi(a,b)$$

for all tuples a and b of selfadjoint matrices of order $(n_1,\cdots,n_k)$. *It is given by*

$$\omega_\varphi(a,b)=\sum_{m_1=1}^{n_1}\cdots\sum_{m_k=1}^{n_k}\Big(H(m_1,\cdots,m_k)\,\Phi^a(m_1,\cdots,m_k)\mid\Phi^b(m_1,\cdots,m_k)\Big)$$

where $H(m_1,\cdots,m_k)$ *are the generalized Hessian matrices. The vectors*

$$\Phi^a(m_1,\cdots,m_k)=\left(\begin{array}{c}\Phi^a_1(m_1,\cdots,m_k)\\ \vdots\\ \Phi^a_k(m_1,\cdots,m_k)\end{array}\qquad m_i=1,\cdots,n_i\quad\textit{for}\quad i=1,\cdots,k\right)$$

are given by

$$\Phi^a_s(m_1,\cdots,m_k)_{j_s}=a^s_{m_sj_s}\varphi(m_1,\cdots,m_{s-1},j_s,m_{s+1},\cdots,m_k)$$

for $j_s=1,\cdots,n_s$ *and* $s=1,\cdots,k$.

Proof. Since the generalized Hessian matrices $H(m_1,\cdots,m_k)$ are real and symmetric thus selfadjoint, it is clear that $\omega_\varphi(a,b)$ as defined is a hermitian sesquilinear form on the complex vector space of tuples of matrices of order $(n_1,\cdots,n_k)$. Its real part is a real symmetric form, and it is therefore sufficient (by real polarization) to prove that $(d^2f(x)(h,h)\varphi\,|\,\varphi)=\omega_\varphi(h,h)$ for all tuples $h=(h^1,\cdots,h^k)$ of selfadjoint matrices of order $(n_1,\cdots,n_k)$.

We take the expectation value of the second Fréchet differential as given by Theorem 4.3 by the vector φ. Then we insert the vectors Φ_s^h and obtain

$$(d^2f(x)(h,h)\varphi\,|\,\varphi)=\sum_{m_1=1}^{n_1}\cdots\sum_{m_k=1}^{n_k}\sum_{s,u=1}^{k}\sum_{j=1}^{n_s}\sum_{p=1}^{n_u}h^s_{m_sj}h^u_{pm_u}(H_{us}(m_1,\cdots,m_k))_{pj}$$

$$((e^1_{m_1m_1}\otimes\cdots\otimes e^s_{m_sj}\otimes\cdots\otimes e^k_{m_km_k})\varphi\,|\,(e^1_{m_1m_1}\otimes\cdots\otimes e^u_{m_up}\otimes\cdots\otimes e^k_{m_km_k})\varphi)$$

$$=\sum_{m_1=1}^{n_1}\cdots\sum_{m_k=1}^{n_k}\sum_{s,u=1}^{k}\sum_{j=1}^{n_s}\sum_{p=1}^{n_u}h^s_{m_sj}h^u_{pm_u}(H_{us}(m_1,\cdots,m_k))_{pj}$$

$$\varphi(m_1,\cdots,m_{s-1},j,m_{s+1},\cdots,m_k)\bar{\varphi}(m_1,\cdots,m_{u-1},p,m_{u+1},\cdots,m_k)$$

$$=\sum_{m_1=1}^{n_1}\cdots\sum_{m_k=1}^{n_k}\sum_{s,u=1}^{k}\sum_{j=1}^{n_s}\sum_{p=1}^{n_u}(H_{us}(m_1,\cdots,m_k))_{pj}\Phi_s^h(m_1,\cdots,m_k)_j\bar{\Phi}_u^h(m_1,\cdots,m_k)_p$$

$$=\sum_{m_1=1}^{n_1}\cdots\sum_{m_k=1}^{n_k}\sum_{s,u=1}^{k}(H_{us}(m_1,\cdots,m_k)\Phi_s^h(m_1,\cdots,m_k)\,|\,\Phi_u^h(m_1,\cdots,m_k))$$

$$=\sum_{m_1=1}^{n_1}\cdots\sum_{m_k=1}^{n_k}(H(m_1,\cdots,m_k)\Phi^h(m_1,\cdots,m_k)\,|\,\Phi^h(m_1,\cdots,m_k)).$$

QED

The structure theorem for the second Fréchet differential is useful because the generalized Hessian matrices $H(m_1,\cdots,m_k)$ only depend on f and the eigenvalues of the matrices $x_1,\cdots,x_k$. The vectors $\Phi^h(m_1,\cdots,m_k)$ depend on $h=(h^1,\cdots,h^k)$ and φ and indirectly on the systems of matrix units diagonalizing $x_1,\cdots,x_k$ through the coordinates of h and φ.

§5. The Matrix and Operator Convex Functions

Proposition 5.1. *Let* $f\in C^p(I_1\times\cdots\times I_k)$ *where* $I_1,\cdots,I_k$ *are open intervals* and p$>2+k/2$. *If for a tuple* $(n_1,\cdots,n_k)$ *all of the generalized Hessian matrices associated with* f *and any data set* $\Lambda\in I_1^{n_1}\times\cdots\times I_k^{n_k}$ *are positive semi-definite, then* f *is matrix convex of order* $(n_1,\cdots,n_k)$.

Proof. Let $x=(x_1,\cdots,x_k)$ be a tuple of selfadjoint matrices of order $(n_1,\cdots,n_k)$ in the domain of f. The assumption ensures that the second Fréchet differential $d^2f(x)(h,h)$ is positive semi-definite for any tuple $h=(h^1,\cdots,h^k)$ of selfadjoint matrices of order $(n_1,\cdots,n_k)$. But then f is matrix convex of order $(n_1,\cdots,n_k)$ according to Proposition 2.2. QED

If the function f is everywhere defined in $\mathbf{R}^k$ and does not grow too fast at infinity, then we may relax the differentiability condition in the above proposition and only require f to be in $C^2(\mathbf{R}^k)$. This can be derived by first considering the convolution of f by an appropriate positive approximating unit and then make use of the fact that the matrix convex functions of a particular order is a closed set.

One may ask whether the condition in Proposition 5.1 is also necessary. This is indeed so for functions of one variable, and the result is due to Kraus [15]. It is easily recovered in this setting by localization of the vectors $\Phi(m)$. Following Kraus one choose

$$\Phi(m)_j=\xi(m)\bar{\xi}(j)\varphi(j) \qquad j=1,\cdots,n$$

with the condition that $\bar{\xi}(j)\varphi(j)=\eta_j$ for some fixed but arbitrary vector $\eta\in\mathbf{C}^n$. This can be done for any positive ξ which then can be chosen arbitrarily close to the indicator function of some fixed m. It thus follows from Corollary 4.4 that each of the generalized Hessian matrices is positive semi-definite. However, it can be shown that such localization of the vectors $\Phi(m_1,\cdots,m_k)$ is in general not possible for functions of more than one variable. The functions

$$(*) \qquad f(t_1,t_2)=\frac{1}{(1-\mu_1t_1)(1-\mu_2t_2)} \qquad t_1,t_2\in\,]-1,1[$$

are known to be operator convex for any $\mu_1,\mu_2\in[-1,1]$. The result follows from a theorem of Ando, and it is noticed in [5].

Theorem 5.2. *Let* $\mu_1,\cdots,\mu_k\in[-1,1]$ *and consider the functions*

$$f_i(t_i) = \frac{1}{1-\mu_i t_i} \qquad t_i \in]-1,1[$$

for $i=1,\cdots,k$. *The function*

$$f(t_1,\cdots,t_k) = f_1(t_1)\cdots f_k(t_k)$$
$$= \prod_{i=1}^{k} \frac{1}{1-\mu_i t_i} \qquad t_1,\cdots,t_k \in]-1,1[$$

of k variables is operator convex, and the ensemble of generalized Hessian matrices associated with f and any set of data

$$\Lambda \in]-1,1[^{n_1} \times \cdots \times]-1,1[^{n_k} \qquad n_1,\cdots,n_k \in \mathbf{N}$$

consists of mutually proportional and positive semi-definite matrices.

Proof. Take $(n_1,\cdots,n_k) \in \mathbf{N}^k$ and any data set $\Lambda \in]-1,1[^{n_1} \times \cdots \times]-1,1[^{n_k}$ which we write on the form

$$\Lambda = \{\lambda_{m_i}(i)\}_{m_i=1,\cdots,n_i} \qquad i=1,\cdots,k.$$

We define the vectors

$$a(i) = \mu_i \Big(f_i(\lambda_1(i)),\cdots,f_i(\lambda_{n_i}(i)) \Big) \in \mathbf{R}^{n_i}$$

for $i=1,\cdots,k$. The divided differences associated with the functions f_i are of the form

$$[x_1 x_2] = \mu_i f_i(x_1) f_i(x_2)$$
$$[x_1 x_2 x_3] = \mu_i^2 f_i(x_1) f_i(x_2) f_i(x_3)$$

for $x_1, x_2, x_3 \in]-1,1[$. It follows that the partial divided differences associated with f are given by

$$[\lambda_{m_1}(1)|\cdots|\lambda_{m_s}(s)\lambda_j(s)|\cdots|\lambda_p(u)\lambda_{m_u}(u)|\cdots|\lambda_{m_k}(k)]^{su}$$
$$= \mu_s \mu_u f(\lambda_{m_1}(1),\cdots,\lambda_{m_k}(k)) f_s(\lambda_j(s)) f_u(\lambda_p(u))$$

for $s \neq u$ while

$$[\lambda_{m_1}(1)|\cdots|\lambda_{m_s}(s)\lambda_p(s)\lambda_j(s)|\cdots|\lambda_{m_k}(k)]^s$$
$$= \mu_s^2 f(\lambda_{m_1}(1),\cdots,\lambda_{m_k}(k)) f_s(\lambda_j(s)) f_s(\lambda_p(s))$$

for $s=u$. Notice that the expression has the same form for $s\neq u$ and for $s=u$. This is only by coincidence for the very special function f. It then follows that

$$H_{us}(m_1,\cdots,m_k)=f(\lambda_{m_1}(1),\cdots,\lambda_{m_k}(k))a(u)^t a(s)$$

for $s\neq u$ and

$$H_{ss}(m_1,\cdots,m_k)=2f(\lambda_{m_1}(1),\cdots,\lambda_{m_k}(k))a(s)^t a(s)$$

for $s=u$. The generalized Hessian matrices are consequently of the form

$$H(m_1,\cdots,m_k)=f(\lambda_{m_1}(1),\cdots,\lambda_{m_k}(k))\begin{pmatrix} 2a(1)^t a(1) & a(1)^t a(2) & \cdots & a(1)^t a(k) \\ a(2)^t a(1) & 2a(2)^t a(2) & \cdots & a(2)^t a(k) \\ \vdots & \vdots & \ddots & \vdots \\ a(k)^t a(1) & a(k)^t a(2) & \cdots & 2a(k)^t a(k) \end{pmatrix}$$

and they are bounded from below by

$$f(\lambda_{m_1}(1),\cdots,\lambda_{m_k}(k))\begin{pmatrix} a(1)^t a(1) & \cdots & a(1)^t a(k) \\ \vdots & \ddots & \vdots \\ a(k)^t a(1) & \cdots & a(k)^t a(k) \end{pmatrix}$$

$$=f(\lambda_{m_1}(1),\cdots,\lambda_{m_k}(k))\Big(a(1)\cdots a(k)\Big)^t\Big(a(1)\cdots a(k)\Big)$$

which are positive semi-definite matrices. QED

We notice that all the generalized Hessian matrices above associated with an arbitrary data set are proportional.

Corollary 5.3. *Let ν be a non-negative Borel measure on the cube $[-1,1]^k$ for $k\in\mathbf{N}$ and let $a_1,\cdots,a_k$ be real numbers. The function*

$$f(t_1,\cdots,t_k)=a_1t_1+\cdots+a_kt_k+\int_{-1}^{1}\cdots\int_{-1}^{1}\prod_{i=1}^{k}\frac{1}{1-\mu_i t_i}\,d\nu(\mu_1,\cdots,\mu_k)$$

is operator convex on the open cube $]-1,1[^k$.

Acknowledgement

The author wishes to thank Professor I. Ojima of RIMS and the former director of the institute, Professor H. Araki, for kind hospitality. The author is endebted to Mombusho and JSPS for financial support.

References

[1] Ando, T., Concavity of certain maps of positive definite matrices and applications to Hadamard products, *Linear Algebra Appl.*, **26** (1979), 203–241.

[2] Araki, H., Expansional in Banach algebras, *Ann. Sci. Éc. Norm. Sup.*, **6** (1973), 67–84.

[3] Aujla, J. S., Matrix convexity of functions of two variables, *Linear Algebra Appl.*, **194** (1993), 149–160.

[4] Bendat, J. and Sherman, S., Monotone and convex operator functions, *Trans. Am. Math. Soc.*, **79** (1955), 58–71.

[5] Brown, A. L. and Vasudeva, H. L., The calculus of operator functions, operator monotonocity and operator convexity, Submitted to *Linear Algebra Appl.*, 1996.

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

[7] Wigner, E. P. and Neumann, J. v., Significance of Loewner's theorem in the quantum theory of collisions, *Ann. of Math.*, **59** (1954), 418–433.

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

[9] Hansen, F., An operator inequality, *Math. Ann.*, **246** (1980), 249–250.

[10] ———, Jensen's operator inequality for functions of two variables, *Proc. Amer. Math. Soc*, 1996.

[11] Hansen, F. and Pedersen, G. K., Jensen's inequality for operators and Löwner's theorem, *Proc. Amer. Math. Soc.*, **125** (1997), 2093–2102.

[12] ———, Perturbation formulas for traces on C^*-algebras, *Publ. RIMS, Kyoto Univ.*, **31** (1995), 169–178.

[13] Kadison, R. V. and Ringrose, J. R., *Fundamentals of the Theory of Operator Algebras, I-II.*, Academic Press, 1982 & 1986.

[14] Korányi, A., On some classes of analytic functions of several variables, *Trans Amer. Math. Soc.*, **101** (1961), 520–554.

[15] Kraus, F., Über konvekse Matrixfunktionen, *Math. Z.*, **41** (1936), 18–42.

[16] Löwner, K., Über monotone Matrixfunktionen, *Math. Z.*, **38** (1934), 177–216.