Journal of Convex Analysis
Volume 9 (2002), No. 2, 401–414

On Subgradients of Spectral Functions

Marc Ciligot-Travain
Department of mathematics, University of Avignon,
33 rue Pasteur, 84000 Avignon, France
marc.ciligot@univ-avignon.fr

Sado Traore
U.P.B./ E.S.I., 01 BP 1091,
Bobo-Dioulasso 01, Burkina–Faso
sado@univ-ovaga.bf

Received December 1, 2000
Revised manuscript received January 25, 2002

Let $F : \mathbf{S}(m) \to \overline{\mathbb{R}}$ be a *spectral function* (i.e. $\mathbf{S}(m)$ is the space of $m \times m$ real symmetric matrices, $\forall O \in \mathbf{O}(m), \forall X \in \mathbf{S}(m),\ F(OX^tO) = F(X)$, where $\mathbf{O}(m)$ is the orthogonal group and ${}^tO$ is the transpose of O). We associate to it the symmetric function $s_F : \mathbb{R}^m \to \overline{\mathbb{R}}$ by restricting it to the subspace of diagonal matrices. In this work, on the one hand, we give a new, natural proof of the formula which binds the Fréchet subgradients of a spectral function F and the Fréchet subgradients of the function s_F (identical formulas follow for the subgradients and the horizon subgradients); on the other hand we deduce from the previous results and from convexity arguments that, in the general case, a similar formula holds for the Clarke subgradients.

Keywords: Spectral function, eigenvalues, eigenvalue optimization, perturbation theory, Clarke subgradient, subgradient, nonsmooth analysis

1991 Mathematics Subject Classification: Primary: 90C31, 15A18, Secondary: 49K40, 26B05

1. Introduction

This work is about some variational properties of functions defined on the space of (real) symmetric matrices and which depend only of the eigenvalues of the matrix. They are called here "spectral functions". More precisely, a function F is spectral if

$$F(OX^tO) = F(X) \quad \text{for all } X \text{ symmetric and } O \text{ orthogonal.} \tag{1}$$

Our interest for these functions arises from their use in Optimization. One can consult the survey [15]. The problem we are interested in can be subsumed as follow: using the invariance property (1) satisfied by F, relate the variational properties of F with those of its restriction to the subspace of diagonal matrices.

We complete the existing results in two ways. First, we use a new approach, based on a certain "projection map", to recover some results of A. Lewis (see [14]) about the subgradients of spectral functions. Note that this method can be and will be applied by us to other problems about variational properties of spectral functions in future papers.

Secondly, we establish a new result about the Clarke subgradients of a spectral map in the general case, which means that the function considered is not necessarily locally Lipschitz

ISSN 0944-6532 / \$ 2.50

(the result is known in the locally Lipschitz case, see [14]). Our method is based on some convexity arguments.

Let us introduce some definitions and notations in order to make precise our problem.

Let $m \in \mathbb{N}^*$ and $\mathbf{M}(m)$ denotes the space of $m \times m$ real matrices. If $M \in \mathbf{M}(m)$, let ${}^tM$ be the transpose of M. We denote by $\mathbf{S}(m)$ the subspace of $m \times m$ real symmetric matrices (i.e. $\mathbf{S}(m) = \{M \in \mathbf{M}(m) | {}^tM = M\}$) and by $\mathbf{O}(m)$ the group of orthogonal $m \times m$ matrices (i.e. $\mathbf{O}(m) = \{O \in \mathbf{M}(m) | {}^tOO = \mathrm{I}_m\}$, where I_m is the $m \times m$ identity matrix).

We define a (left) action $\star$ of $\mathbf{O}(m)$ on $\mathbf{S}(m)$ by

$$(O, X) \in \mathbf{O}(m) \times \mathbf{S}(m) \longrightarrow O \star X = OX{}^tO.$$

Definition 1.1 (Spectral function). Let S be a set . A function $F : \mathbf{S}(m) \to S$ is a *spectral function* if F is invariant under the action $\star$, i.e.

$$\forall O \in \mathbf{O}(m), \forall X \in \mathbf{S}(m), \quad F(O \star X) = F(X). \tag{2}$$

In other words, if we denote by Orb_X the *orbit* of $X \in \mathbf{S}(m)$ relatively to the action $\star$, i.e. $\mathrm{Orb}_X = \{O \star X, O \in \mathbf{O}(m)\}$, spectral functions are functions which are constant on each orbit (associated with the action $\star$). These functions are sometimes called *orthogonally invariant* or *eigenvalue functions*.

Now let $X \in \mathbf{S}(m)$, μ_i, $i = 1, ..., p$, be the eigenvalues of X with $\mu_1 > \mu_2 > ... > \mu_p$, $p \in \mathbb{N}^*$, and $m_i \in \mathbb{N}^*$, $i = 1, ..., p$, be the multiplicity of μ_i. We set

$$\lambda(X) = (\underbrace{\mu_1, ..., \mu_1}_{m_1 \text{ times}}, \underbrace{\mu_2, ..., \mu_2}_{m_2 \text{ times}}, ..., \underbrace{\mu_p, ..., \mu_p}_{m_p \text{ times}}).$$

This defines a spectral map $\lambda : \mathbf{S}(m) \longrightarrow \mathbb{R}^m$. If $f : \mathbb{R}^m \to \overline{\mathbb{R}}$ is a function, then

$$f \circ \lambda : \mathbf{S}(m) \to \overline{\mathbb{R}} \tag{3}$$

is clearly spectral. In fact all spectral functions $F : \mathbf{S}(m) \to \overline{\mathbb{R}}$ have the form (3). Denote by diag : $\mathbb{R}^m \to \mathbf{S}(m)$ the map which associates to $(x_1, ..., x_m)$ the diagonal matrix $[x_i \delta_{ij}]_{i,j=1,...,m}$ (where $\delta_{ij} = 1$ if $i = j$ and 0 if $i \neq j$). Set

$$s_F = F \circ \mathrm{diag} : \mathbb{R}^m \to \overline{\mathbb{R}},$$

then one has:

$$F = s_F \circ \lambda. \tag{4}$$

The function s_F is symmetric (i.e. invariant under coordinates permutations) and is the unique symmetric function $f : \mathbb{R}^m \to \overline{\mathbb{R}}$ satisfying $F = f \circ \lambda$. Moreover, note that s_F is (almost) the restriction of F to the subspace of diagonal matrices.

The main objective of our work is to relate the first order variational properties of F with those of s_F. More precisely, we compute several kinds of subgradients of F from the corresponding ones of s_F.

Let us make a brief tour of the existing results. Some variational properties (at the first, the second or higher order) of the function λ (or its components) and related functions can be considered as classical theory. This covers the case $t \to \lambda(X(t))$, $t \in I$, I interval of $\mathbb{R}$, with X analytic or differentiable or C^1. This includes perturbations in fixed directions ($t \to \lambda(Z + tV)$). One can consult the books of Rellich [19], Kato [7] or [6], and Baumegärtel [1]. One can find alternative approaches and more results in the works of Hiriart–Urruty and Ye [10], [11], Overton and Womersley [18], Hiriart–Urruty and Lewis [8], Hiriart–Urruty and Torki [9] and Torki [23]. A study of the first and second order epi-differentiability of the components of λ can be found in Torki's paper [24].

Concerning spectral functions, the convex case has been investigated in the works of Friedland [5], Martínez–Legaz [17], Lewis [12], Seeger [22] and Torki [25] . Differentiability properties have been studied in the works of Lewis [13], Lewis and Sendov [16]; analycity in a paper of Tsing N.-K., Fan M.K.H. and Verriest E.I. [26]; nonsmooth variational properties in papers of Lewis [13] and [14].

2. Preliminaries and notations

In all the sequel, an Euclidean space is a finite dimensional real Hilbert space. Let $(\mathcal{E}, \langle ., . \rangle)$ be an Euclidean space and let $\mathbf{G}$ be a subgroup of the group $\mathbf{O}(\mathcal{E})$ of linear isometries of $\mathcal{E}$. Let S be a set and $h : \mathcal{E} \to S$. We say that h is $\mathbf{G}$–invariant if

$$\forall g \in \mathbf{G}, \quad h \circ g = h.$$

In other words, h is $\mathbf{G}$–invariant if h is invariant under the action $(g, x) \to g(x)$ of $\mathbf{G}$ on $\mathcal{E}$. For example, if $\mathbf{P}(m)$ denotes the group of permutation matrices, a symmetric function on $\mathbb{R}^m$ is exactly a $\mathbf{P}(m)$–invariant function.

We endow $\mathbf{M}(m)$ with the following usual scalar product:

$$\langle A, B \rangle = \mathrm{Tr}(A^t B) = \sum_{i,j=1,\dots,m} A_{ij} B_{ij},$$

and $\mathbb{R}^m$ with its natural scalar product. We denote all the norms by $\|.\|$. If $X \in \mathbf{S}(m)$ then $\|X\| = \|\lambda(X)\|$.

The space $\mathbf{S}(m)$, endowed with the scalar product just defined, is an Euclidean space. Let $O \in \mathbf{O}(m)$. Set $\mathrm{Int}(O) : \mathbf{S}(m) \to \mathbf{S}(m)$ defined by: $\forall X \in \mathbf{S}(m)$, $\mathrm{Int}(O)(X) = O \star X$. $\mathrm{Int} : \mathbf{O}(m) \to \mathbf{O}(\mathbf{S}(m))$ is a group homomorphism, and a spectral function is a $\mathrm{Int}(\mathbf{O}(m))$–invariant function. If $\mathbf{G}$ is a subgroup of $\mathbf{O}(m)$, we say "$\mathbf{G}$–invariant" instead of "$\mathrm{Int}(\mathbf{G})$–invariant" (for example, a spectral function is an $\mathbf{O}(m)$–invariant function).

The function λ is Lipschitzian (see, for example, [2] III.6.15),

$$\forall A, B \in \mathbf{S}(m), \quad \|\lambda(A) - \lambda(B)\| \leq \|A - B\|. \tag{5}$$

This last property is equivalent to the following inequality ([2] III.6.14) which is of importance for the sequel (see, for example, [14] Theorem 2 for an elegant geometric proof):

$$\forall A, B \in \mathbf{S}(m), \quad \mathrm{Tr}(AB) \leq \langle \lambda(A), \lambda(B) \rangle. \tag{6}$$

Let $\delta = (m_1, ..., m_p)$ with $p \in \mathbb{N}^*$, $m_i \in \mathbb{N}^*$ for all $i \in \{1, ..., p\}$, $m = m_1 + \cdots + m_p$. Given

$M_i \in \mathbf{M}(m_i)$, $i = 1, ..., p$, we denote $M_1 \oplus M_2 \oplus \cdots \oplus M_p$ the block diagonal matrix

$$M_1 \oplus M_2 \oplus \cdots \oplus M_p = \begin{bmatrix} M_1 & 0 & \dots & 0 \\ 0 & M_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & M_p \end{bmatrix}.$$

We set $\mathbf{S}(\delta) = \mathbf{S}(m_1) \oplus \mathbf{S}(m_2) \oplus \cdots \oplus \mathbf{S}(m_p)$, $\mathbf{O}(\delta) = \mathbf{O}(m_1) \oplus \mathbf{O}(m_2) \oplus \cdots \oplus \mathbf{O}(m_p)$. The group $\mathbf{O}(\delta)$ acts on $\mathbf{S}(\delta)$ (by $\star$) and one can generalize what has just been said.

We define the map $\lambda_\delta : \mathbf{S}(\delta) \to \mathbb{R}^m$ by

$$\lambda_\delta(S_1 \oplus \cdots \oplus S_p) = (\lambda(S_1), ..., \lambda(S_p)).$$

The function λ_δ is $\mathbf{O}(\delta)$–invariant, for all $X \in \mathbf{S}(\delta)$, $\|X\| = \|\lambda_\delta(X)\|$ and from (5) and (6), one deduces for all $A, B \in \mathbf{S}(\delta)$:

$$\|\lambda_\delta(A) - \lambda_\delta(B)\| \leq \|A - B\|, \tag{7}$$

$$\mathrm{Tr}(AB) \leq \langle \lambda_\delta(A), \lambda_\delta(B) \rangle. \tag{8}$$

Recall that $\mathbf{P}(m)$ is the subgroup (of $\mathbf{O}(m)$) of all permutation matrices. We set $\mathbf{P}(\delta) = \mathbf{P}(m_1) \oplus \mathbf{P}(m_2) \oplus \cdots \oplus \mathbf{P}(m_p)$. If $\mathcal{A}$ is a $\mathbf{P}(\delta)$–invariant subset of $\mathbb{R}^m$ (i.e. $\forall P \in \mathbf{P}(\delta)$, $P\mathcal{A} \subset \mathcal{A}$), one has

$$\lambda_\delta^{-1}(\mathcal{A}) = \mathbf{O}(\delta) \star \mathrm{diag}\, \mathcal{A}. \tag{9}$$

In fact, if $X \in \lambda_\delta^{-1}(\mathcal{A})$, there exists $O \in \mathbf{O}(\delta)$ such that $X = O \star \mathrm{diag}\, \lambda_\delta(X)$ and $\lambda_\delta(X) \in \mathcal{A}$ so $X \in \mathbf{O}(\delta) \star \mathrm{diag}\, \mathcal{A}$. Now if $X = O \star \mathrm{diag}\, x$, $O \in \mathbf{O}(\delta)$, $x \in \mathcal{A}$, then there exists $P \in \mathbf{P}(\delta)$ such that $\lambda_\delta(X) = Px \in \mathcal{A}$ ($x \in \mathcal{A}$ and $\mathcal{A}$ is $\mathbf{P}(\delta)$–invariant). Finally, we denote by $\mathbf{D}^{\downarrow}(m)$ the set of $m \times m$ diagonal matrices with entries arranged in the decreasing order.

3. A "projection" map

Due the invariance property (2) satisfied by a spectral function, knowing that each orbit is a submanifold, it is very natural, in a first step, to relate the variational properties of a spectral function F at a point with the variational properties of its restriction to the "normal" space to the orbit (of this point) at this point.

Knowing some elements about the geometry of the orbits, we make the following construction.

We denote by $\mathbf{A}(m)$ the subspace of skew-symmetric matrices of $\mathbf{M}(m)$ ($\mathbf{A}(m) = \{A \in \mathbf{M}(m) | {}^tA = -A\}$). The Lie bracket of $X, Y \in \mathbf{M}(m)$ is $[X, Y] = XY - YX$. We set $\mathrm{ad}(X) : Y \to [X, Y]$.

Now fix $Z \in \mathbf{S}(m)$. We consider the map $\mathrm{ad}(Z)$ as a map from $\mathbf{A}(m)$ to $\mathbf{S}(m)$. We set $\mathcal{T}_Z = \mathrm{Im}\ \mathrm{ad}(Z) = \{[Z, A], A \in \mathbf{A}(m)\}$ and define $\mathcal{N}_Z = \mathcal{T}_Z^{\perp}$ (in $\mathbf{S}(m)$). One has (see, for example, [14] Theorem 1 for a proof):

$$\mathcal{N}_Z = \{X \in \mathbf{S}(m) | XZ = ZX\}. \tag{10}$$

Note that $Z \in \mathcal{N}_Z$. We set

$$\operatorname{ad}(Z)^\dagger = \left(\operatorname{ad}(Z)\Big|_{[\operatorname{Ker}\operatorname{ad}(Z)]^\perp}^{\operatorname{Im}\ \operatorname{ad}(Z)} \right)^{-1} \circ Q_Z,$$

where Q_Z is the orthogonal projection on $\mathcal{T}_Z$. Now set

$$\psi_Z : \left\{ \begin{array}{lcl} \mathbf{S}(m) & \longrightarrow & \mathbf{S}(m) \\ X & \longrightarrow & e^{-\operatorname{ad}(Z)^\dagger X} \star P_Z X \end{array} \right. ,$$

where P_Z is the orthogonal projection on $\mathcal{N}_Z$.

The map ψ_Z is smooth (C^∞) and $\psi_Z(Z) = Z$. Using, for example, the fact that, for all $A, B \in \mathbf{M}(m)$, $e^A B e^{-A} = e^{\operatorname{ad}(A)}(B)$, we also obtain:

$$\psi_Z(Z+V) = \sum_{n=0}^{+\infty} \frac{1}{n!} \operatorname{ad}(-\operatorname{ad}(Z)^\dagger V)^n (Z + P_Z V),$$

because $\operatorname{ad}(Z)^\dagger Z = 0$ ($Z \in \mathcal{N}_Z$). So

$$\psi_Z(Z+V) = Z + P_Z V + [Z, \operatorname{ad}(Z)^\dagger V] + \mathrm{o}(V).$$

But $[Z, \operatorname{ad}(Z)^\dagger V] = \operatorname{ad}(Z) \circ \operatorname{ad}(Z)^\dagger(V) = Q_Z V$, so

$$\psi_Z(Z+V) = Z + V + \mathrm{o}(V). \tag{11}$$

According to (11), we deduce from the local inversion Theorem that there exist an open neighbourhood $\mathcal{U}$ of the origin in $\mathcal{T}_Z$ and an open neighbourhood $\mathcal{V}$ of Z in $\mathcal{N}_Z$ such that $\mathcal{W} = \psi_Z(\mathcal{U} + \mathcal{V})$ is an open neighbourhood of Z in $\mathbf{S}(m)$ and $\psi_Z : \mathcal{U} + \mathcal{V} \to \mathcal{W}$ is a smooth diffeomorphism.

We now define $\pi_Z : \mathcal{W} \to \mathcal{V}$, our "projection map", by

$$\pi_Z = P_Z \circ \psi_Z^{-1}.$$

The map π_Z is smooth, $\pi_Z(Z) = Z$ and $\pi'_Z(Z) = P_Z$. Moreover,

$$\forall X \in \mathcal{W}, \quad \pi_Z(X) \in \operatorname{Orb}_X . \tag{12}$$

In fact, we can define $\mathrm{O}_Z : \mathcal{W} \to \mathbf{O}(m)$ by

$$\mathrm{O}_Z(X) = e^{-\operatorname{ad}(Z)^\dagger \psi_Z^{-1}(X)},$$

and we obtain, from the definition of ψ_Z,

$$\forall X \in \mathcal{W}, \quad X = \mathrm{O}_Z(X) \star \pi_Z(X),$$

which demonstrates (12). It is property (12) which makes us say that π_Z "respects the orbits".

We subsume some properties of π_Z:

Proposition 3.1. *The map $\pi_Z : \mathcal{W} \to \mathcal{V}$ is smooth, surjective and $\pi_Z(Z) = Z$, $\pi'_Z(Z) = P_Z$. Moreover it satisfies*

$$\forall X \in \mathcal{W}, \quad X = \mathrm{O}_Z(X) \star \pi_Z(X). \tag{13}$$

Therefore, for every spectral function $F : \mathbf{S}(m) \to \overline{\mathbb{R}}$,

$$F|_{\mathcal{W}} = F_Z \circ \pi_Z, \tag{14}$$

where F_Z denotes the restriction of the spectral function F to $\mathcal{N}_Z$ and $F|_{\mathcal{W}}$ the restriction of F to $\mathcal{W}$.

4. Fréchet subgradients, subgradients and horizon subgradients of a spectral function

In this section, we recover, by a new and natural way, some results due to A.S. Lewis (see [14]) and which will be needed in the sequel. Note that we don't use any result about the generalized differentiability of λ. We compute the Fréchet subgradients of F at Z from the Fréchet subgradients of s_F at $\lambda(Z)$. Similar formulas follow for the subgradients and the horizon subgradients. We follow, essentially, the definitions and notations of [21].

Definition 4.1 (Fréchet subgradients). Let $(\mathcal{E}, \langle ., . \rangle)$ be an Euclidean space, $h : \mathcal{O} \to \overline{\mathbb{R}}$ be a function, where $\mathcal{O}$ is an open subset of $\mathcal{E}$. Let $z \in \mathcal{O}$ with $h(z) \in \mathbb{R}$. An element x^* of $\mathcal{E}$ is called a *Fréchet subgradient* of h at z, if

$$h(x) \geq h(z) + \langle x^*, x - z \rangle + \mathrm{o}(x - z). \tag{15}$$

One can consult [21] 8.B for more details. We denote by $\widehat{\partial}\, h(z)$ the set of all Fréchet subgradients of h at z.

Definition 4.2 (Subgradients and horizon subgradients). An element x^* of $\mathcal{E}$ is called a *subgradient* of h at z if there are sequences (x_n^*) and (x_n) such that $x_n^* \to x^*$, $x_n \to z$, $h(x_n) \to h(z)$ and $\forall n \in \mathbb{N}, x_n^* \in \widehat{\partial} h(x_n)$. We denote by $\partial h(z)$ the set of all subgradients of h at z.

An element x^* of $\mathcal{E}$ is called an *horizon subgradient* of h at z if there are sequences (x_n^*), (x_n) and (ϵ_n) such that $\epsilon_n \downarrow 0$, $\epsilon_n x_n^* \to x^*$, $x_n \to z$, $h(x_n) \to h(z)$ and $\forall n \in \mathbb{N}, x_n^* \in \widehat{\partial} h(x_n)$. We denote by $\partial^\infty h(z)$ the set of all subgradients of h at z.

In the last definition, $\varepsilon_n \downarrow 0$ means that $\forall n \in \mathbb{N}, \varepsilon_n \in]0, +\infty[$ and $\varepsilon_n \to 0$ When two functions are equal in a neighbourhood of a point then their sets of Fréchet subgradients are equal. The same is true for the subgradients and the horizon subgradients. We will need the following property:

Proposition 4.3 ([21], 10.7). *Let $H : \mathcal{O} \to \mathcal{F}$ be a C^1-map, where $\mathcal{F}$ is an Euclidean space and $\mathcal{O}$ is an open subset of the Euclidean space $\mathcal{E}$. Let $z \in \mathcal{O}$ and $h : \mathcal{F} \to \overline{\mathbb{R}}$ be a function with $h(H(z)) \in \mathbb{R}$. Then*

$$H'(z)^* \,\widehat{\partial}\, h(H(z)) \subset \widehat{\partial}(h \circ H)(z), \tag{16}$$

where $H'(z)^*$ is the transpose of $H'(z)$. Moreover, if $H'(z)$ is surjective, then equality holds in (16).

Due to Proposition 4.3, if H is a linear surjective from $\mathcal{E}$ to $\mathcal{F}$ then (16) becomes:

$$\widehat{\partial}(h \circ H)(z) = H^* \widehat{\partial} h(H(z)). \tag{17}$$

Now suppose that for a certain subgroup $\mathbf{G}$ of linear isometries of $\mathcal{E}$, $h : \mathcal{E} \to \overline{\mathbb{R}}$ is $\mathbf{G}$-invariant. Then, using (17), we deduce that:

$$\forall G \in \mathbf{G}, \quad \widehat{\partial} h(G(z)) = G \widehat{\partial} h(z). \tag{18}$$

In particular, if $\forall G \in \mathbf{G}, G(z) = z$, then

$$\forall G \in \mathbf{G}, \quad \widehat{\partial} h(z) = G \widehat{\partial} h(z). \tag{19}$$

Let us return now to spectral functions. Let $F : \mathbf{S}(m) \to \overline{\mathbb{R}}$ be a spectral map, $Z \in \mathbf{S}(m)$ with $F(Z) \in \mathbb{R}$. Due to the invariance property satisfied by the spectral map F, it is very natural to "reduce the problem" to $\mathcal{N}_Z$. Recall that F_Z is the restriction of F to the subspace $\mathcal{N}_Z$.

From Proposition 4.3, from (14), from the fact that $\pi'_Z(Z) = P_Z$ is surjective and P_Z^* is the canonical injection from $\mathcal{N}_Z$ to $\mathbf{S}(m)$, one obtains:

Lemma 4.4.

$$\widehat{\partial} F(Z) = \widehat{\partial} F_Z(Z).$$

Now, we have to compute $\widehat{\partial} F_Z(Z)$ from $\widehat{\partial} s_F(\lambda(Z))$. From (18), and the fact that $F \circ \operatorname{Int}(O) = F$, we deduce, for $X \in \mathbf{S}(m)$,

$$\widehat{\partial} F(O \star X) = O \star \widehat{\partial} F(X). \tag{20}$$

This property allows us to suppose that $Z \in \mathbf{D}^{\downarrow}(m)$, i.e. $Z = \operatorname{diag} \lambda(Z)$, which simplifies the study. We denote by $\delta(Z)$ the multi-index $(m_1, m_2, ..., m_p) \in \mathbb{N}^p$ such that

$$\lambda(Z) = (\underbrace{\mu_1, ..., \mu_1}_{m_1 \text{ times}}, \underbrace{\mu_2, ..., \mu_2}_{m_2 \text{ times}}, ..., \underbrace{\mu_p, ..., \mu_p}_{m_p \text{ times}}).$$

and $\mu_1 > \mu_2 > \cdots > \mu_p$. We have $\mathcal{N}_Z = \mathbf{S}(\delta(Z))$ and for all $O \in \mathbf{O}(\delta(Z))$, $O \star Z = Z$.

Lemma 4.5. *For $Z \in \mathbf{D}^{\downarrow}(m)$, one has:*

$$\widehat{\partial} F_Z(Z) = \mathbf{O}(\delta(Z)) \star \operatorname{diag} \widehat{\partial} s_F(\lambda(Z)) = \lambda_{\delta(Z)}^{-1}(\widehat{\partial} s_F(\lambda(Z))). \tag{21}$$

Proof. The second equality is due to (9) using the fact that $\widehat{\partial} s_F(\lambda(Z))$ is a $\mathbf{P}(\delta(Z))$–invariant subset (this is due to (19), knowing that s_F is $\mathbf{P}(m)$–invariant and $\forall P \in \mathbf{P}(\delta(Z)), P\lambda(Z) = \lambda(Z)$).

Let $X^* \in \widehat{\partial} F_Z(Z)$. There exists $O \in \mathbf{O}(\delta(Z))$ such that $O \star X^* = \operatorname{diag} \lambda_{\delta(Z)}(X^*)$. Due to the fact that F_Z is $\mathbf{O}(\delta(Z))$-invariant, $O \star Z = Z$ and the property (19), $\operatorname{diag} \lambda_{\delta(Z)}(X^*) \in \widehat{\partial} F_Z(Z)$. Then, due to (16), seeing diag as a map from $\mathbb{R}^m$ to $\mathbf{S}(\delta(Z))$,

$$\operatorname{diag}^*(\operatorname{diag} \lambda_{\delta(Z)}(X^*)) \in \widehat{\partial}(F_Z \circ \operatorname{diag})(\lambda(Z)),$$

but $\mathrm{diag}^* \circ \mathrm{diag}$ is the identity of $\mathbb{R}^m$ and $F_Z \circ \mathrm{diag} = s_F$ so $\lambda_{\delta(Z)}(X^*) \in \widehat{\partial}\, s_F(\lambda(Z))$.

This shows that $\widehat{\partial}\, F_Z(Z) \subset \lambda_{\delta(Z)}^{-1}(\widehat{\partial}\, s_F(\lambda(Z)))$.

Let $X^* \in \mathbf{S}(\delta(Z))$ such that $\lambda_{\delta(Z)}(X^*) \in \widehat{\partial}\, s_F(\lambda(Z))$. There exists $\varepsilon : \mathbb{R}^m \to \overline{\mathbb{R}}$ such that $\lim_{x\to 0} \varepsilon(x) = 0$ and

$$\forall x \in \mathbb{R}^m, \quad s_F(\lambda(Z) + x) \geq s_F(\lambda(Z)) + \langle \lambda_{\delta(Z)}(X^*), x\rangle + \|x\|\varepsilon(x).$$

Let $X \in \mathbf{S}(\delta(Z))$. There exists $O \in \mathbf{O}(\delta(Z))$ such that $X = O \star \mathrm{diag}\, \lambda_{\delta(Z)}(X)$. We have

$$F(Z + \mathrm{diag}\, \lambda_{\delta(Z)}(X)) \geq F(Z) + \langle \lambda_{\delta(Z)}(X^*), \lambda_{\delta(Z)}(X)\rangle + \|\lambda_{\delta(Z)}(X)\|\varepsilon(\lambda_{\delta(Z)}(X)),$$

but $\|\lambda_{\delta(Z)}(X)\| = \|X\|$, $\lim_{X\to 0} \lambda_{\delta(Z)}(X) = 0$ and

$$F(Z + \mathrm{diag}\, \lambda_{\delta(Z)}(X)) = F(O \star (Z + \mathrm{diag}\, \lambda_{\delta(Z)}(X))) = F(Z + X),$$

so, due to the property (7),

$$F(Z + X) \geq F(Z) + \langle X^*, X\rangle + \mathrm{o}(X),$$

and $X^* \in \widehat{\partial}\, F_Z(Z)$. Finally $\lambda_{\delta(Z)}^{-1}(\widehat{\partial}\, s_F(\lambda(Z))) \subset \widehat{\partial}\, F_Z(Z)$. □

We can now deduce the main theorem of this section, which is Theorem 6 of [14]:

Theorem 4.6. *Let $F : \mathbf{S}(m) \to \overline{\mathbb{R}}$ be a spectral function and $Z \in \mathbf{S}(m)$, with $F(Z) \in \mathbb{R}$. Then*

$$\widehat{\partial} F(Z) = \mathrm{Trans}_Z \star \mathrm{diag}\, \widehat{\partial} s_F(\lambda(Z)), \tag{22}$$

where $\mathrm{Trans}_Z = \{O \in \mathbf{O}(m) | Z = O \star \mathrm{diag}\, \lambda(Z)\}$. Similar formulas hold for the subgradients and the horizon subgradients.

Proof. If $Z \in \mathbf{D}^{\downarrow}(m)$, then the result follows immediately from Lemmas 4.4 and 4.5, knowing that, in this case, $\mathrm{Trans}_Z = \mathbf{O}(\delta(Z))$.

In the general case, let $O \in \mathrm{Trans}_Z$. Then, using property (20), one has

$$\widehat{\partial}\, F(Z) = O \star \widehat{\partial}\, F(\mathrm{diag}\, \lambda(Z)) = O \star [\mathbf{O}(\delta(Z)) \star \mathrm{diag}\, \widehat{\partial}\, s_F(\lambda(Z))].$$

We conclude, using the fact that $O\mathbf{O}(\delta(Z)) = \mathrm{Trans}_Z$.

For the case of subgradients and horizon subgradients, one can follow the method of Lewis [14]. □

5. The Clarke subgradients of a spectral function

Now, we are going to establish a formula similar to (22) for the Clarke subgradients. The case when F is locally Lipschitz has been considered in [13] and [14]. Let us first recall some definitions about Clarke subgradients (see [4] 2.4 and [21] for more details).

Let $C \subset \mathcal{E}$, where $(\mathcal{E}, \langle .,. \rangle)$ is an Euclidean space, and $z \in C$.

The *Clarke tangent cone* (or *regular tangent cone*) to C at z, $\widehat{T}_C(z)$, is the closed convex cone defined by

$$\widehat{T}_C(z) = \{v \in \mathcal{E} \mid \forall x_n \xrightarrow{C} z, \forall t_n \downarrow 0, \exists v_n \to v, \forall n \in \mathbb{N}, x_n + t_n v_n \in C\}.$$

where $x_n \xrightarrow{C} z$ means that $x_n \to z$ and $\forall n \in \mathbb{N}, x_n \in C$, and $t_n \downarrow 0$ means that $\forall n \in \mathbb{N}, t_n \in]0, +\infty[$ and $t_n \to 0$. The *Clarke normal cone* of C at z is

$$\overline{N}_C(z) = \widehat{T}_C(z)^* = \{x^* \in \mathcal{E} \mid \forall v \in \widehat{T}_C(z), \langle x^*, v\rangle \leq 0\}.$$

It is a closed convex cone. Using this notion of normal cone, one can define the Clarke subgradients of a function. Recall that the epigraph of $h : \mathcal{E} \to \overline{\mathbb{R}}$ is the set

$$\operatorname{epi} h = \{(x, r) \in \mathcal{E} \times \mathbb{R} \mid h(x) \leq r\}.$$

Definition 5.1 (Clarke subgradients). Let $h : \mathcal{E} \to \overline{\mathbb{R}}$, $z \in \mathcal{E}$ such that $h(z) \in \mathbb{R}$. The set of *Clarke subgradients* of h at z is defined as

$$\overline{\partial}\, h(z) = \{x^* \in \mathcal{E} \mid (x^*, -1) \in \overline{N}_{\operatorname{epi} h}(z, h(z))\}.$$

When two functions are equal in a neighbourhood of a point then their sets of Clarke subgradients at this point are equal.

We will use in the sequel that properties similar to (18) and (19) can be established, for example from 10.7 [21], for the subgradients, the horizon subgradients and the Clarke subgradients.

We will say that $h : \mathcal{E} \to \overline{\mathbb{R}}$ is *lower semicontinuous* (lsc in the sequel) if $\operatorname{epi} h$ is closed.

We are going to deduce our result about Clarke subgradients of spectral functions from a relation between the subgradients, horizon subgradients and the Clarke subgradients.

Theorem 5.2 ([20]). *Let $h : \mathcal{E} \to \overline{\mathbb{R}}$ be a lsc function and z with $h(z) \in \mathbb{R}$ then:*

$$\overline{\partial}\, h(z) = \overline{\text{co}}\left[\partial h(z) + \partial^\infty h(z)\right], \tag{23}$$

where $\overline{\text{co}}$ denotes the operation of taking the closed convex hull of a set.

The formula (23) can also be deduced from [21] 6(19), 8(32) and 8.9.

In all the sequel, $F : \mathbf{S}(m) \to \overline{\mathbb{R}}$ is a lsc spectral function and $Z \in \mathbf{S}(m)$ with $F(Z) \in \mathbb{R}$. Due to the relations $F = s_F \circ \lambda$, $s_F = F \circ \operatorname{diag}$ and the continuity of diag and λ, one has that F is lsc if and only if s_F is lsc.

Due to the invariance property satisfied by F, one has, if $X \in \mathbf{S}(m)$,

$$\forall O \in \mathbf{O}(m), \quad \overline{\partial}\, F(O \star X) = O \star \overline{\partial}\, F(X).$$

So, if $O \in \operatorname{Trans}_Z$,

$$\overline{\partial}\, F(Z) = O \star \overline{\partial}\, F(\operatorname{diag} \lambda(Z)). \tag{24}$$

This allows us to reduce to the case $Z \in \mathbf{D}^{\downarrow}(m)$.

We will use a particular preorder on $\mathbb{R}^m$. If $(x_1, ..., x_m) \in \mathbb{R}^m$, we denote by $x^{\downarrow}$ the vector obtained by rearranging the coordinates of x in the decreasing order. If $x, y \in \mathbb{R}^m$, we say that x is *majorised* by y, in symbols $x \prec y$, if

$$\sum_{j=1}^{k} x_j^{\downarrow} \leq \sum_{j=1}^{k} y_j^{\downarrow}, \quad \forall k \in \{1, ...m\},$$

and

$$\sum_{j=1}^{m} x_j^{\downarrow} = \sum_{j=1}^{m} y_j^{\downarrow}.$$

The relation $\prec$ is a preorder on $\mathbb{R}^m$, in the sense that it is a reflexive, transitive relation. One can consult [2], chapter II for more details. One has the following characterization of this preorder, due to the Birkhoff's Theorem (see [2] II.2.3 and II.1.10). Let us recall that $\mathbf{P}(m)$ is the finite group of permutation matrices.

Proposition 5.3. *Let $x, y \in \mathbb{R}^m$. Then $x \prec y$ if and only if there exists a family $(\alpha_P)_{P\in\mathbf{P}(m)}$ of non-negative reals $(\forall P \in \mathbf{P}(m), \alpha_P \geq 0)$ such that $\sum_{P\in\mathbf{P}(m)} \alpha_P = 1$ and such that*

$$x = \sum_{P\in\mathbf{P}(m)} \alpha_P P y.$$

Our interest in this preorder arises from the property:

Theorem 5.4 ([2] (III.13)). *Let $A, B \in \mathbf{S}(m)$, then*

$$\lambda(A + B) \prec \lambda(A) + \lambda(B).$$

Let $\delta = (m_1, ..., m_p) \in (\mathbb{N}^*)^p$ such that $m_1 + ... + m_p = m$. We see the space $\mathbb{R}^m$ as $\prod_{i=1,...,p} \mathbb{R}^{m_i}$ and define the preorder $\prec_\delta$ as $\prec^{\times p}$ so that

$$(x_1, ..., x_p) \prec_\delta (y_1, ..., y_p) \Leftrightarrow x_i \prec y_i, \quad \forall i = 1, ..., p.$$

We will use the following lemma which is a consequence of Proposition 5.3:

Lemma 5.5. *Let $x, y \in \mathbb{R}^m$. Then $x \prec_\delta y$ if and only if there exists a family of non–negative reals $(\alpha_P)_{P\in\mathbf{P}(\delta)}$ such that $\sum_{P\in\mathbf{P}(\delta)} \alpha_P = 1$ and*

$$x = \sum_{P\in\mathbf{P}(\delta)} \alpha_P P y.$$

Proof. Let $(x_1, ..., x_p) \prec_\delta (y_1, ..., y_p)$ then, from proposition 5.3, one deduces that there exists $(\alpha^i_P)_{P\in\mathbf{P}(m_i)}$, $i = 1, ..., p$, p families of non–negative reals such that

$$\forall i \in \{1, ..., p\}, \qquad \sum_{P\in\mathbf{P}(m_i)} \alpha^i_P = 1, \quad x_i = \sum_{P\in\mathbf{P}(m_i)} \alpha^i_P P y_i.$$

So

$$(x_1, ..., x_p) = \sum_{P_1 \in \mathbf{P}(m_1)} \alpha^1_{P_1} \big(P_1 \oplus \mathrm{I}_{m_2} \oplus \cdots \oplus \mathrm{I}_{m_p}\big)(y_1, x_2 ..., x_p)$$

$$(y_1, x_2 ..., x_p) = \sum_{P_2 \in \mathbf{P}(m_2)} \alpha^2_{P_2} \big(\mathrm{I}_{m_1} \oplus P_2 \oplus \mathrm{I}_{m_3} \oplus \cdots \oplus \mathrm{I}_{m_p}\big)(y_1, y_2, x_3, ..., x_p)$$

$$\vdots$$

$$(y_1, ..., y_{p-1}, x_p) = \sum_{P_p \in \mathbf{P}(m_p)} \alpha^p_{P_p} \big(\mathrm{I}_{m_1} \oplus \cdots \oplus P_p\big)(y_1, ..., y_p).$$

It follows that:

$$(x_1, ..., x_p) = \sum_{P_1 \oplus \cdots \oplus P_p \in \mathbf{P}(\delta)} \alpha^1_{P_1} \alpha^2_{P_2} ... \alpha^p_{P_p} P_1 \oplus \cdots \oplus P_p(y_1, ..., y_p),$$

with

$$\alpha^1_{P_1} \alpha^2_{P_2} ... \alpha^p_{P_p} \geq 0, \qquad \sum_{P_1 \oplus \cdots \oplus P_p \in \mathbf{P}(\delta)} \alpha^1_{P_1} \alpha^2_{P_2} ... \alpha^p_{P_p} = 1.$$

The other implication is direct. □

From Theorem 5.4, one deduces that for all $A, B \in \mathbf{S}(\delta)$,

$$\lambda_\delta(A + B) \prec_\delta \lambda_\delta(A) + \lambda_\delta(B). \tag{25}$$

The following properties will be useful in the sequel. Let $\mathcal{A}$ be a $\mathbf{P}(\delta)$-invariant subset of $\mathbb{R}^m$. Then, due to Lemma 5.5,

$$x \prec_\delta y \text{ and } y \in \mathcal{A} \Rightarrow x \in \operatorname{co} \mathcal{A}, \tag{26}$$

where co denotes the operation of taking the convex hull of a set. Note also that if $\mathcal{A}$ is a $\mathbf{P}(\delta)$–invariant subset then $\operatorname{co} \mathcal{A}$ and $\overline{\operatorname{co}}\, \mathcal{A}$ are $\mathbf{P}(\delta)$–invariant subsets. If $\mathcal{A}$ is a $\mathbf{P}(\delta)$–invariant subset of $\mathbb{R}^m$,

$$\lambda_\delta^{-1}(\operatorname{cl}(\mathcal{A})) = \operatorname{cl}[\lambda_\delta^{-1}(\mathcal{A})], \tag{27}$$

where $\operatorname{cl}(\mathcal{A})$, for example, is the closure of $\mathcal{A}$.

The main result is based essentially on the following two lemmas.

Lemma 5.6. *Let $\mathcal{C}$ be a $\mathbf{P}(\delta)$-invariant, convex subset of $\mathbb{R}^m$. Then $\lambda_\delta^{-1}(\mathcal{C})$ is a convex set. Let $\mathcal{A}$ be a $\mathbf{P}(\delta)$-invariant subset of $\mathbb{R}^m$ then*

$$\lambda_\delta^{-1}(\operatorname{co} \mathcal{A}) = \operatorname{co}(\lambda_\delta^{-1}(\mathcal{A})), \quad \lambda_\delta^{-1}(\overline{\operatorname{co}}\, \mathcal{A}) = \overline{\operatorname{co}}(\lambda_\delta^{-1}(\mathcal{A})). \tag{28}$$

Proof. Let $A, B \in \lambda_\delta^{-1}(\mathcal{C})$ and $\alpha, \beta \geq 0$, $\alpha + \beta = 1$. From (25), one deduces

$$\lambda_\delta(\alpha A + \beta B) \prec_\delta \alpha \lambda_\delta(A) + \beta \lambda_\delta(B).$$

But $\lambda_\delta(A), \lambda_\delta(B) \in \mathcal{C}$ which is convex, so $\alpha \lambda_\delta(A) + \beta \lambda_\delta(B) \in \mathcal{C}$ and we conclude using (26).

Now, if $\mathcal{A}$ is $\mathbf{P}(\delta)$-invariant then $\operatorname{co}\mathcal{A}$ is $\mathbf{P}(\delta)$-invariant too and convex. It follows that $\lambda_\delta^{-1}(\operatorname{co}\mathcal{A})$ is convex and contains $\lambda_\delta^{-1}(\mathcal{A})$, so $\operatorname{co}\lambda_\delta^{-1}(\mathcal{A}) \subset \lambda_\delta^{-1}(\operatorname{co}\mathcal{A})$. On the other hand,

$$\operatorname{diag}\operatorname{co}\mathcal{A} = \operatorname{co}\operatorname{diag}\mathcal{A} \subset \operatorname{co}(\mathbf{O}(\delta)\star\operatorname{diag}\mathcal{A})$$

and $\operatorname{co}(\mathbf{O}(\delta)\star\operatorname{diag}\mathcal{A})$ is $\mathbf{O}(\delta)$-invariant, so, using (9),

$$\lambda_\delta^{-1}(\operatorname{co}\mathcal{A}) = \mathbf{O}(\delta)\star\operatorname{diag}\operatorname{co}\mathcal{A} \subset \operatorname{co}(\mathbf{O}(\delta)\star\operatorname{diag}\mathcal{A}) = \operatorname{co}(\lambda_\delta^{-1}(\mathcal{A})).$$

This demonstrates the first equality in (28). For the second, using (27),

$$\lambda_\delta^{-1}(\overline{\operatorname{co}}\,\mathcal{A}) = \lambda_\delta^{-1}(\operatorname{cl}(\operatorname{co}\mathcal{A})) = \operatorname{cl}(\lambda_\delta^{-1}(\operatorname{co}\mathcal{A})) = \operatorname{cl}(\operatorname{co}(\lambda_\delta^{-1}(\mathcal{A}))) = \overline{\operatorname{co}}(\lambda_\delta^{-1}(\mathcal{A})).$$

$\square$

Lemma 5.7. *Let $\mathcal{A}$ and $\mathcal{B}$ be two $\mathbf{P}(\delta)$-invariant subsets of $\mathbb{R}^m$. Then*

$$\overline{\operatorname{co}}\left[\lambda_\delta^{-1}(\mathcal{A}) + \lambda_\delta^{-1}(\mathcal{B})\right] = \lambda_\delta^{-1}(\overline{\operatorname{co}}\,[\mathcal{A}+\mathcal{B}]). \tag{29}$$

Proof. $\mathcal{A}+\mathcal{B}$ is $\mathbf{P}(\delta)$-invariant so

$$\lambda_\delta^{-1}(\mathcal{A}+\mathcal{B}) = \mathbf{O}(\delta)\star\operatorname{diag}(\mathcal{A}+\mathcal{B}) \subset \mathbf{O}(\delta)\star\operatorname{diag}\mathcal{A} + \mathbf{O}(\delta)\star\operatorname{diag}\mathcal{B} = \lambda_\delta^{-1}(\mathcal{A}) + \lambda_\delta^{-1}(\mathcal{B}),$$

and one inclusion is true, using (28).

On the other hand, let $X \in \lambda_\delta^{-1}(\mathcal{A})$ and $Y \in \lambda_\delta^{-1}(\mathcal{B})$. One has

$$\lambda_\delta(X+Y) \prec_\delta \lambda_\delta(X) + \lambda_\delta(Y),$$

so $\lambda_\delta(X+Y)$ is δ–majorised by an element of $\mathcal{A}+\mathcal{B}$, $\mathbf{P}(\delta)$-invariant, which implies that $\lambda_\delta(X+Y) \in \overline{\operatorname{co}}\,[\mathcal{A}+\mathcal{B}]$ and the second inclusion, using the fact that, due to lemma 5.6, $\lambda_\delta^{-1}(\overline{\operatorname{co}}(\mathcal{A}+\mathcal{B}))$ is closed convex and $\mathbf{P}(\delta)$–invariant. $\square$

Theorem 5.8. *Let $F : \mathbf{S}(m) \to \overline{\mathbb{R}}$ be a lsc spectral function and $Z \in \mathbf{S}(m)$ with $F(Z) \in \mathbb{R}$. Then*

$$\overline{\partial}\,F(Z) = \operatorname{Trans}_Z\star\operatorname{diag}\overline{\partial}\,s_F(\lambda(Z)) \tag{30}$$

where $\operatorname{Trans}_Z = \{O \in \mathbf{O}(m) | Z = O\star\operatorname{diag}\lambda(Z)\}$.

Proof. First, let $O \in \operatorname{Trans}_Z$, then (see (24))

$$\overline{\partial}\,F(Z) = O\star\overline{\partial}\,F(\operatorname{diag}\lambda(Z)). \tag{31}$$

Now, using (23) and Theorem 4.6,

$$\begin{aligned}\overline{\partial}\,F(\operatorname{diag}\lambda(Z)) &= \overline{\operatorname{co}}\left[\partial F(\operatorname{diag}\lambda(Z)) + \partial^\infty F(\operatorname{diag}\lambda(Z))\right]\\ &= \overline{\operatorname{co}}\left[\lambda_{\delta(Z)}^{-1}(\partial s_F(\lambda(Z))) + \lambda_{\delta(Z)}^{-1}(\partial^\infty s_F(\lambda(Z)))\right].\end{aligned}$$

The sets $\partial s_F(\lambda(Z))$, $\partial^\infty s_F(\lambda(Z))$ and $\overline{\partial}\,s_F(\lambda(Z))$ are $\mathbf{P}(\delta(Z))$–invariant, so, using the Lemma 5.7, (23) and (9),

$$\begin{aligned}\overline{\partial}\,F(\operatorname{diag}\lambda(Z)) &= \lambda_{\delta(Z)}^{-1}\big(\overline{\operatorname{co}}\left[\partial s_F(\lambda(Z)) + \partial^\infty s_F(\lambda(Z))\right]\big)\\ &= \mathbf{O}(\delta(Z))\star\operatorname{diag}\overline{\partial}\,s_F(\lambda(Z)).\end{aligned} \tag{32}$$

Finally, one has $O\mathbf{O}(\delta(Z)) = \operatorname{Trans}_Z$, and using (31), (32), one deduces:

$$\overline{\partial}\,F(Z) = \operatorname{Trans}_Z\star\operatorname{diag}\overline{\partial}\,s_F(\lambda(Z)).$$

$\square$

Acknowledgements. The authors are grateful to Professor L. Thibault for his hints, and to the referees for their careful reading and their useful remarks.

References

[1] H. Baumgärtel: Analytic Perturbation Theory for Matrices and Operators, Birkhäuser (1985).

[2] R. Bhatia: Matrix Analysis, Graduate Texts in Mathematics 169, Springer, Berlin (1996).

[3] J. M. Borwein, A. S. Lewis: Convex Analysis and Nonlinear Optimization, Theory and Examples, CMS Books in Mathematics, Springer, Berlin (2000).

[4] F. H. Clarke: Optimization and Nonsmooth Analysis, Wiley, New York (1983).

[5] S. Friedland: Convex spectral functions, Linear and Multilinear Algebra 9 (1981) 299–316.

[6] T. Kato: Perturbation Theory for Linear Operators, Springer, Berlin (1966).

[7] T. Kato: A Short Introduction to Perturbation Theory for Linear Operators, Springer, Berlin (1982).

[8] J.-B. Hiriart–Urruty, A. S. Lewis: The Clarke and Michel–Penot subdifferentials of the eigenvalues of a symmetric matrix, Comput. Optim. Appl. 13(1–3) (1999) 13–23.

[9] J.-B. Hiriart–Urruty, M. Torki: Analyse de la sensibilité du second ordre de toutes les valeurs propres d'une matrice symétrique, C. R. Acad. Sci. Paris, Sér. I, Math. 324 (1997) 1071–1074.

[10] J.-B. Hiriart-Urruty, D. Ye: Analyse de la sensibilité du premier ordre de toutes les valeurs propres d'une matrice symétrique, C. R. Acad. Sci. Paris, Sér. I, Math. 318 (1994) 177–182.

[11] J.-B. Hiriart-Urruty, D. Ye: Sensitivity analysis of all eigenvalues of a symmetric matrix, Numer. Math. 70 (1995) 45–72.

[12] A. S. Lewis: Convex analysis on the Hermitian matrices, SIAM J. Optim. 6 (1996) 164–177.

[13] A. S. Lewis: Derivatives of spectral functions, Math. Oper. Res. 21 (1996) 576–588.

[14] A. S. Lewis: Nonsmooth analysis of eigenvalues, Math. Programming 84 (1999) 1–24.

[15] A. S. Lewis, M. L. Overton: Eigenvalue optimization, Acta Numer. 5 (1996) 149–190.

[16] A. S. Lewis, H. S. Sendov: Twice differentiable spectral functions, To appear.

[17] J. E. Martínez-Legaz: On convex and quasiconvex spectral functions, in: Proceedings of the 2nd Catalan Days on Applied Mathematics, Presses Universitaires de Perpignan (1995).

[18] M. L. Overton, R. S. Womersley: Optimality conditions and duality theory for minimizing sums of the largest eigenvalues of symmetric matrices, Math. Programming 62 (1993) 321–357.

[19] F. Rellich: Perturbation Theory of Eigenvalue Problems, Gordon and Breach Science Publishers, New York (1969).

[20] R. T. Rockafellar: Proximal subgradients, marginal values, and augmented Lagrangians in nonconvex optimization, Math. Oper. Res. 6 (1981) 424–436.

[21] R. T. Rockafellar, R. J.-B. Wets: Variational Analysis, Springer, Berlin (1998).

[22] A. Seeger: Convex analysis of spectrally defined matrix functions, SIAM J. Optim. 7 (1997) 679–696.

[23] M. Torki: Second-order directional derivatives of all eigenvalues of a symmetric matrix, Nonlinear Anal. 46 (2001) 1133–1150.

[24] M. Torki: First- and second-order epi-differentiability in eigenvalue optimization, J. Math. Anal. Appl. 234 (1999) 391–416.

[25] M. Torki: Second-order epi-differentiability and subdifferentiability of spectral convex functions, to appear.

[26] N.-K. Tsing, M. K. H. Fan, E. I. Verriest: On analycity of functions involving eigenvalues, Linear Algebra Appl. 207 (1994) 159–180.