Bouligand and Clarke subdifferentials of spectral mappings in Euclidean Jordan algebras

Michel Baes*

September 23, 2008

Abstract

We study in this note the Bouligand subdifferential and the Clarke subdifferential of gradients of spectral functions in Euclidean Jordan algebras, that is, function that only depend on the eigenvalues of their argument. More generally, we study these subdifferentials for spectral mappings, relating in a close form formula the Bouligand subdifferential of a symmetric mapping to Bouligand subdifferential of the spectral mapping it generates. We also obtain explicit upper and lower approximations for the Clarke subdifferential of spectral mappings.

Keywords: Euclidean Jordan algebras, spectral mappings, subdifferentiability.

2000 AMS Classification: 90C25, 52A41, 47A55

1 Introduction

Dealing with non-differentiabilities has always been a challenging task in optimization. Differentials offer a clear and compact characterization for optimality, they follow simple combination rules and so on. When the objective function is convex but not necessarily differentiable, the natural substitute for the differential is the *subdifferential* $\partial f(x)$, that is, the set of normals to hyperplanes through $(x., f(x))$ that support the epigraph of f. We still have a simple optimality characterization, and some calculus rules still hold (see e.g. Theorem 3.5.5 in [BL06]).

In 1975, F. Clarke extended the notion of subdifferential for non-convex functions so that these two desirable properties hold, at least partially (see [Cla75] of Chapter 6 in [BL06] for details). For locally Lipschitz continuous functions, the so-called *Clarke subdifferential* can be defined as follows. Let $f : \mathbb{R}^N \to \mathbb{R}$ be a locally Lipschitz continuous function, and x be a point of its domain. In view of Rademacher's Theorem ([Rad19], or, for a modern exposition, Section 9.1 in [BL06]), the function f is differentiable on a dense subset D of its domain. Now, we consider all the sequences $\{x_k : k \geq 0\}$ of D that converge to x and for which $\{f'(x_k) : k \geq 0\}$ converges as well. The *Bouligand subdifferential* is the set of all the limit points for these sequences $\{f'(x_k) : k \geq 0\}$. Finally, the Clarke subdifferential is the convex hull of this set. Many equivalent definitions of Clarke subdifferential exists. The advantage of this one is that it can be immediately extended to vector-valued functions.

The interest for considering Bouligand and Clarke subdifferentials of vector-valued functions originates in the study and the implementation of *non-smooth* Newton methods, where the

*M. Baes is with the Department of Electrical Engineering (ESAT), Research Group SCD-SISTA, Katholieke Universiteit Leuven, Kasteelpark Arenberg 10, B-3001 Heverlee, Belgium. E-mail: Michel.Baes@esat.kuleuven.be.

function to minimize is not necessarily differentiable. These schemes have been studied in [QS93] and further developed in [SS03]. The idea of these schemes is quite simple. The usual Newton algorithm for twice continuously differentiable functions f generates a sequence $\{x_k : k \geq 0\}$ from an initial point x_0:

$$x_{k+1} = x_k - f''(x_k)^{-1} f'(x_k).$$

Now, if the function f has a locally Lipschitz continuous gradient, the Hessian at x_k is not guaranteed to exists. We can therefore replace it by an element of its Bouligand or Clarke subdifferential. Conditions for Q-quadratic convergence of the resulting scheme have been investigated in [SS03] and in [Qi93], namely the *strong semismoothness* of the gradient of the objective function.

Nevertheless, evaluating Clarke and Bouligand subdifferentials presents in general heavy numerical difficulties; for instance, determining whether a Bouligand subdifferential contains one single element is numerically unstable, as a tiny perturbation of the data can completely change the output. In the direct continuation of our work on inherited properties of spectral functions on Jordan algebras [Bae07], we address in this note the Bouligand subdifferentiability of spectral mappings in Euclidean Jordan algebras by determining an explicit formula for it. This formula shows the link between the Bouligand subdifferential of a symmetric function g that is locally Lipschitz continuous and of the function $g \circ \lambda$, where λ is the eigenvalue mapping in the corresponding Euclidean Jordan algebra. We can further use this exact formula to obtain lower and upper approximations of the Clarke subdifferential of such functions.

2 Preliminaries on Euclidian Jordan algebras

Let $\mathcal{J}$ be a N-dimensional Euclidean Jordan algebra of dimension r. For definitions of the necessary Jordan algebraic concepts, we refer the reader to standard textbooks such as [BK66], [FK94], or [Koe99].

For every $u \in \mathcal{J}$, we denote by $\lambda(u) = (\lambda_1(u), \cdots, \lambda_r(u))^T$ the vector of eigenvalues of u ordered decreasingly. The multiplication symbol in $\mathcal{J}$ is denoted by $\cdot$, although it is often omitted for the sake of notational concision. The multiplication operator by u is written $L(u)$ and the quadratic operator by Q_u, so that $Q_u = 2L(u)^2 - L(u^2)$. For any $u, v \in \mathcal{J}$, its polarization $L(u)L(v) + L(v)L(u) - L(uv)$ is denoted by $Q_{u,v}$. The trace function, which maps its argument to the sum of its eigenvalues, is represented by tr; this function is linear, as shown in Proposition II.2.1 in [FK94].

Let $\mathcal{P}$ be the set of all $r \times r$ permutation matrices. We denote by $Q \subset \mathbb{R}^r$ a Lebesgue measurable set with nonempty interior. We also assume that S is *symmetric*, i.e. that $PS = S$ for every permutation P.

We consider a symmetric function $g : Q \to \mathbb{R}^r$, that is, a function satisfying $g(P\lambda) = Pg(\lambda)$ for every $\lambda \in Q$ and $P \in \mathcal{P}$. We assume that g is locally Lipschitz continuous. In view of Rademacher's Theorem, this function is therefore differentiable on a subset D_g of Q whose complement in Q is Lebesgue-negligible.

The Bouligand subdifferential of g in $\lambda \in Q$, denoted as $\partial_B g(\lambda)$, is the set of accumulation points of all the sequences $\{g'(\lambda_m) : m \geq 0\}$ for which $\{\lambda_m : m \geq 0\}_m \subseteq D_g$ converges to λ. The Clarke subdifferential $\partial_C g(\lambda)$ is defined as the convex hull of $\partial_B g(\lambda)$.

Let $Q_{\mathcal{J}} : \{u \in \mathcal{J} : \lambda(u) \in Q\}$. We proceed to define the *spectral mapping* $G : Q_{\mathcal{J}} \to \mathcal{J}$ generated by g. Considering a complete spectral decomposition $u = \sum_{i=1}^r \lambda_i(u) c_i$ of $u \in Q_{\mathcal{J}}$, we define $G(u) := \sum_{i=1}^r g_i(\lambda(u)) c_i$. In view of the symmetry of g, it is not difficult to prove that this definition does not depend on the specific spectral decomposition of u we have taken.

Following Proposition 4.5.1 in [Bae06], we can easily show that G is locally Lipschitz continuous. Moreover, From Theorem 5.5.1 in [Bae06], we know that the subset $D_G\ of Q_{\mathcal{J}}$ where G is differentiable is precisely $\{u : \lambda(u) \in D_g\}$. Therefore its Bouligand and its Clarke subdifferentials are well defined. In order to compute them, we need to use the formula for the Jacobian of spectral mapping in Jordan algebras, which has been established in [].

Let $u \in D_G$. We define the integers $s, k_1, \ldots, k_s$ for which:

$$\lambda_1(u) = \cdots = \lambda_{k_1}(u) > \lambda_{k_1+1}(u) = \cdots = \lambda_{k_2}(u) > \cdots \lambda_{k_s}(u),$$

and the sets of numbers $M_j := \{k_{j-1}+1, \ldots, k_j\}$, with $k_0 = 0$. Let $u = \sum_{j=1}^{s} \lambda_{k_j}(u) e_j$ be the spectral decomposition of u.

By symmetry of g, one can show that $g'(\lambda(u))$ can be rewritten as $B(\lambda(u)) + \mathrm{diag}(b(\lambda(u)))$, where $B(\lambda(u))$ is a block-matrix, where all the coefficients in a block $[B(\lambda(u))_{i,j}]_{i \in M_\alpha, j \in M_\beta}$ have the same value. The matrix $b(\lambda(u))$ is diagonal, and all the coefficients $[b(\lambda(u))_i]_{i \in M_\alpha}$ are the same.

Theorem 2.1 (Theorem 51 in [Bae07]) *Let $u \in D_G$ and $h \in \mathcal{J}$. Abbreviating $\lambda(u)$ to λ, we have:*

$$\begin{aligned} G'(u)h = & \sum_{\alpha,\beta=1}^{s} g'_{k_\alpha k_\beta}(\lambda)\mathrm{tr}(e_\beta h)e_\alpha + \sum_{\alpha=1}^{s} b_{k_\alpha}(\lambda)(Q_{e_\alpha}h - \mathrm{tr}(e_\alpha h)e_\alpha) \\ & + \sum_{\alpha \neq \beta} \frac{g_{k_\beta}(\lambda) - g_{k_\alpha}(\lambda)}{\lambda_{k_\beta} - \lambda_{k_\alpha}} Q_{e_\alpha, e_\beta} h. \end{aligned} \tag{1}$$

The following corollary gives an equivalent expression for the Jacobian of G.

Corollary 2.1 *Let $u \in D_G$ and $h \in \mathcal{J}$, and let $u = \sum_{i=1}^{r} \lambda_i(u) c_i$ be a complete spectral decomposition of u. Then, writing λ for $\lambda(u)$, we have:*

$$G'(u)h = \sum_{i,j=1}^{r} g'_{ij}(\lambda)\mathrm{tr}(c_j h)c_i + \sum_{i \neq j} \lim_{t \downarrow \lambda_j - \lambda_i} \frac{g_i(\lambda + t(\hat{v}_i - \hat{v}_j)) - g_i(\lambda)}{t} Q_{c_i, c_j} h,$$

where $\hat{v}_i$ is the vector $(\delta_{1,i}, \ldots, \delta_{n,1})^T$.

Proof

The proof is essentially a careful rewriting of (1). The first term is, given the block structure of $g'(\lambda)$:

$$\begin{aligned} \sum_{\alpha,\beta=1}^{s} g'_{k_\alpha k_\beta}(\lambda)\mathrm{tr}(e_\beta h)e_\alpha = & \sum_{\alpha \neq \beta} \sum_{i \in M_\alpha} \sum_{j \in M_\beta} g'_{ij}(\lambda)\mathrm{tr}(c_j h)c_i \\ & + \sum_{\alpha=1}^{s} \sum_{\substack{i,j \in M_\alpha \\ i \neq j}} g'_{ii}(\lambda)\mathrm{tr}(c_j h)c_i + \sum_{\alpha=1}^{s} \sum_{i \in M_\alpha} g'_{ii}(\lambda)\mathrm{tr}(c_i h)c_i. \end{aligned}$$

According to the definition of b, the second term can be expanded as:

$$\sum_{\alpha=1}^{s} b_{k_\alpha}(\lambda)(Q_{e_\alpha}h - \mathrm{tr}(e_\alpha h)e_\alpha) = \sum_{\alpha=1}^{s} \sum_{\substack{i,j \in M_\alpha \\ i \neq j}} (g'_{ii}(\lambda) - g'_{ij}(\lambda))(Q_{c_i,c_j}h - \mathrm{tr}(c_j h)c_i).$$

Note that $\mathrm{tr}(c_k h)c_k = Q_{c_k} h$, so that the terms with $i = j$ are null. Therefore, the two first terms of (1) sum up to :

$$\sum_{\alpha\neq\beta}\sum_{i\in M_\alpha}\sum_{j\in M_\beta} g'_{ij}(\lambda)\mathrm{tr}(c_j h)c_i + \sum_{\alpha=1}^{s}\sum_{\substack{i,j\in M_\alpha\\ i\neq j}} g'_{ij}(\lambda)\mathrm{tr}(c_j h)c_i + \sum_{\alpha=1}^{s}\sum_{i\in M_\alpha} g'_{ii}(\lambda)\mathrm{tr}(c_i h)c_i$$

$$+ \sum_{\alpha=1}^{s}\sum_{\substack{i,j\in M_\alpha\\ i\neq j}} g'_{ii}(\lambda)\mathrm{tr}(c_j h)c_i - \sum_{\alpha=1}^{s}\sum_{\substack{i,j\in M_\alpha\\ i\neq j}} g'_{ii}(\lambda)\mathrm{tr}(c_j h)c_i$$

$$+ \sum_{\alpha=1}^{s}\sum_{\substack{i,j\in M_\alpha\\ i\neq j}} (g'_{ii}(\lambda) - g'_{ij}(\lambda))Q_{c_i,c_j} h$$

$$= \sum_{i,j=1}^{r} g'_{ij}(\lambda)\mathrm{tr}(c_j h)c_i + 0 + \sum_{\alpha=1}^{s}\sum_{\substack{i,j\in M_\alpha\\ i\neq j}} (g'_{ii}(\lambda) - g'_{ij}(\lambda))Q_{c_i,c_j} h.$$

It remains to observe that

$$g'_{ii}(\lambda) - g'_{ij}(\lambda) = \lim_{t\downarrow 0} \frac{g_i(\lambda + t(\hat{v}_i - \hat{v}_j)) - g_i(\lambda)}{t}.$$

■

Finally, we need the following technical result, reproduced from [Bae06], Proposition 4.3.2.

Proposition 2.2 *Suppose that we have an element $u \in \mathcal{J}$ and a sequence $(u_m)_{m\geq 0}$ of $\mathcal{J}$ that converges to u. We denote the complete spectral decompositions of these elements by $u_m = \sum_{i=1}^{r} \lambda_i(u_m)c_{i,m}$ and $u = \sum_{i=1}^{r} \lambda_i(u)c_i$ respectively. We define the numbers $s, k_1, \ldots, k_s$ so that $k_s := r$ and:*

$$\lambda_1(u) = \cdots = \lambda_{k_1}(u) > \lambda_{k_1+1}(u) = \cdots = \lambda_{k_2}(u) > \cdots \lambda_{k_s}(u).$$

We set $M_j := \{k_{j-1} + 1, \ldots, k_j\}$, $e_j := \sum_{i\in M_j} c_i$, and $e_{j,m} := \sum_{i\in M_j} c_{i,m}$.
For every $1 \leq j \leq s$, the sequence $(e_{j,m})_{m\geq 0}$ converges to e_j as m goes to infinity.

3 Bouligand subdifferential of spectral mappings

Let us fix an element $u \in \mathcal{J}$. In this section, we aim at describing the set $\partial_B G(u)$, relating it to the set $\partial_B g(\lambda(u))$.

We define the integers $s, k_1, \ldots, k_s$ for which:

$$\lambda_1(u) = \cdots = \lambda_{k_1}(u) > \lambda_{k_1+1}(u) = \cdots = \lambda_{k_2}(u) > \cdots \lambda_{k_s}(u),$$

and the sets of numbers $M_\alpha := \{k_{\alpha-1} + 1, \ldots, k_\alpha\}$, with $k_0 = 0$. Let $u = \sum_{\alpha=1}^{s} \lambda_{k_\alpha}(u)e_\alpha$ be the spectral decomposition of u.

The symmetry of g, and the subsequent block + diagonal structure of $g'(u)$, induces some symmetry property on the set $\partial_B g(u)$. Indeed, if the sequence $\lambda_m \in D_g$ converges to $\lambda(u)$ while $g'(\lambda_m)$ converges to A, then the sequence $P\lambda_m$ is also contained in D_g, and converges also to $\lambda(u)$. As $g'(P\lambda_m) = P^T g'(\lambda_m)P$, we deduce that $P^T AP \in \partial_B g(\lambda(u))$.

However, as the function g is not differentiable anymore, we cannot conclude from this short reasoning that the elements of $\partial_B g(\lambda(u))$ share the same structure as Jacobians of g. Indeed, even if all the coefficients of a same block are identical for Jacobians, this is not

necessarily the case for Bouligand subdifferentials. As illustration, consider the symmetric function $(x,y)^T \mapsto g(x,y) = (|x|,|y|)^T$. We have

$$\partial_B g(0,0) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} :$$

the diagonal elements are not all identical.

Theorem 3.1 *A linear application $V : \mathcal{J} \to \mathcal{J}$ belongs to $\partial_B G(u)$ if and only if there exists a Jordan frame $\{c_1, \ldots, c_r\}$ for which $u = \sum_{i=1}^r \lambda_i(u) c_i$ and a matrix $A = [A_{ij}]_{i,j=1}^r \in \partial_B g(\lambda(u))$ such that, for every $h \in \mathcal{J}$:*

$$Vh = \sum_{i,j=1}^{r} A_{ij}\mathrm{tr}(c_j h)c_i + \sum_{\alpha=1}^{s} \sum_{\substack{i,j \in M_\alpha \\ i \neq j}} (A_{ii} - A_{ij}) Q_{c_i,c_j} h + \sum_{\alpha=1}^{s} \sum_{\beta \neq \alpha} \frac{g_{k\alpha}(\lambda(u)) - g_{k_\beta}(\lambda(u))}{\lambda_{k\alpha}(u) - \lambda_{k_\beta}(u)} Q_{e_\alpha,e_\beta} h. \tag{2}$$

Proof

Let V be a linear application of the form prescribed in the statement of the theorem. We construct here a sequence $\{u_m : m \geq 0\}$ of D_G that converges to u, and for which $G'(u_m)$ goes to V. Let $A \in \partial_B g(\lambda(u))$ be the application used in the description of V, and let $\{\lambda_m : m \geq 0\} \subseteq D_g$ be a sequence converging to λ such that $g'(\lambda_m) \to A$. Let $\{c_1, \ldots, c_r\}$ be a Jordan frame defining V, so that $u = \sum_{i=1}^r \lambda_i(u) c_i$ is a complete spectral decomposition of u. We let $u_m := \sum_{i=1}^r \lambda_{m,i} c_i$. We aim to prove that $G'(u_m) \to V$.

Let us fix an arbitrary $h \in \mathcal{J}$. Using Corollary 2.1, we have, with $\lambda := \lambda(u)$:

$$\begin{aligned} G'(u_m)h &= \sum_{i,j=1}^{r} g'_{ij}(\lambda_m)\mathrm{tr}(c_j h)c_i + \sum_{i \neq j} \lim_{t \downarrow \lambda_{m,j} - \lambda_{m,i}} \frac{g_i(\lambda_m + t(\hat{v}_i - \hat{v}_j)) - g_i(\lambda_m)}{t} Q_{c_i,c_j} h \\ &\to \sum_{i,j=1}^{r} A_{ij}\mathrm{tr}(c_j h)c_i + \sum_{\substack{i \neq j \\ \lambda_i \neq \lambda_j}} \frac{g_j(\lambda) - g_i(\lambda_m)}{\lambda_j - \lambda_i} Q_{c_i,c_j} h + \sum_{\substack{i \neq j \\ \lambda_i = \lambda_j}} (A_{ii} - A_{ij}) Q_{c_i,c_j} h. \end{aligned}$$

Reciprocally, let $\{u_m : m \geq 0\}$ be a sequence of D_G that converges to u and such that $G'(u_m)$ converges to a linear application V. We define the numbers $s_m, k_{m,1}, \ldots, k_{m,s_m}$ such that:

$$\lambda_1(u_m) = \cdots = \lambda_{k_{m,1}}(u_m) > \lambda_{k_{m,1}+1}(u) = \cdots = \lambda_{k_{m,2}}(u) > \cdots \lambda_{k_{m,s}}(u).$$

By possibly considering a subsequence of u_m, we can assume without loss of generality that these numbers do not depend on m. To simplify the notation, we set $\hat{s} := s_m$ and $\hat{k}_j = k_{m,j}$. Now, let $\hat{M}_j := \{\hat{k}_{j-1} + 1, \ldots, \hat{k}_j\}$, with $\hat{k}_0 = 0$. As a trivial consequence of Proposition 2.2, each set $\hat{M}_j$ is entirely contained in some set M_α.

Let $u_m = \sum_{j=1}^{\hat{s}} \lambda_{\hat{k}_j}(u) d_{m,j}$ be the spectral decomposition of u_m. Without loss of generality, we can further assume that the sequences of idempotents $\{d_{m,j} : m \geq 0\}$ all converge. We denote by d_j their limit. As another consequence of Proposition 2.2, we can write $u = \sum_{j=1}^{\hat{s}} \lambda_{\hat{k}_j}(u) d_j = \sum_{j=1}^{s} \lambda_{k_\alpha}(u) e_\alpha$.

We prove first that $g'(\lambda(u_m))$ converges to some linear application. In view of Theorem 2.1, we know that for every $h \in \mathcal{J}$:

$$\begin{aligned} G'(u_m)h &= \sum_{i,j=1}^{\hat{s}} g'_{\hat{k}_i \hat{k}_j}(\lambda(u_m))\mathrm{tr}(d_{m,j} h) d_{m,i} + \sum_{i=1}^{\hat{s}} b_{\hat{k}_i}(\lambda(u_m))(Q_{d_{m,i}} h - \mathrm{tr}(d_{m,i} h) d_{m,i}) \\ &\quad + \sum_{i \neq j} \frac{g_{\hat{k}_j}(\lambda(u_m)) - g_{\hat{k}_i}(\lambda(u_m))}{\lambda_{\hat{k}_j}(u_m) - \lambda_{\hat{k}_i}(u_m)} Q_{d_{m,i},d_{m,j}} h. \end{aligned}$$

According to the formula above, we have for $i \neq j$:

$$\begin{aligned} Q_{d_i} V d_j &= \lim_{m\to\infty} Q_{d_{m,i}} G'(u_m) d_{m,j} \\ &= \lim_{m\to\infty} g'_{\hat{k}_i\hat{k}_j}(\lambda(u_m))\mathrm{tr}(d_{m,j})d_{m,i} = \lim_{m\to\infty} g'_{\hat{k}_i\hat{k}_j}(\lambda(u_m))\mathrm{tr}(d_j)d_i. \end{aligned}$$

This indicates that $g'_{\hat{k}_i\hat{k}_j}(\lambda(u_m))$ converges when $i \neq j$. Also:

$$\begin{aligned} Q_{d_i} V d_i &= \lim_{m\to\infty} Q_{d_{m,i}} G'(u_m) d_{m,i} \\ &= \lim_{m\to\infty} g'_{\hat{k}_i\hat{k}_i}(\lambda(u_m))\mathrm{tr}(d_{m,i})d_{m,i} + b_{\hat{k}_i}(\lambda(u_m))(d_{m,i} - \mathrm{tr}(d_{m,i})d_{m,i}). \end{aligned}$$

If $|\hat{M}_i| = 1$, this limit reduces to $\lim_{m\to\infty} g'_{\hat{k}_i\hat{k}_i}(\lambda(u_m))d_i$, and $g'_{\hat{k}_i\hat{k}_i}(\lambda(u_m))$ converges. Now, if d_i is not a minimal idempotent, we can select a number j in $\hat{M}_i$ different from $\hat{k}_i$ and write:

$$\begin{aligned} Q_{d_i} V d_i &= \lim_{m\to\infty} Q_{d_{m,i}} G'(u_m) d_{m,i} \\ &= \lim_{m\to\infty} g'_{\hat{k}_i\hat{k}_i}(\lambda(u_m))\mathrm{tr}(d_{m,i})d_{m,i} + b_{\hat{k}_i}(\lambda(u_m))(1 - \mathrm{tr}(d_{m,i}))d_{m,i} \\ &= \lim_{m\to\infty} g'_{\hat{k}_i\hat{k}_i}(\lambda(u_m))\mathrm{tr}(d_i)d_i + (g'_{\hat{k}_i\hat{k}_i}(\lambda(u_m)) - g'_{\hat{k}_i j}(\lambda(u_m)))(1 - \mathrm{tr}(d_i))d_i. \end{aligned}$$

Since $g'_{\hat{k}_i j}(\lambda(u_m))$ converges, we conclude that $g'_{\hat{k}_i\hat{k}_i}(\lambda(u_m))$ converges as well. In summary, and given the block structure of $g'(\lambda(u_m))$, we have proved that $g'(\lambda(u_m))$ converges to a linear application A of $\partial_B g(\lambda)$.

Now, let us study the limit:

$$\begin{aligned} &\lim_{m\to\infty} G'(u_m)h \\ &= \lim_{m\to\infty} \sum_{i,j=1}^{\hat{s}} g'_{\hat{k}_i\hat{k}_j}(\lambda(u_m))\mathrm{tr}(d_{m,j}h)d_{m,i} + \sum_{i=1}^{\hat{s}} b_{\hat{k}_i}(\lambda(u_m))(Q_{d_{m,i}}h - \mathrm{tr}(d_{m,i}h)d_{m,i}) \\ &\quad + \sum_{i\neq j} \frac{g_{\hat{k}_j}(\lambda(u_m)) - g_{\hat{k}_i}(\lambda(u_m))}{\lambda_{\hat{k}_j}(u_m) - \lambda_{\hat{k}_i}(u_m)} Q_{d_{m,i},d_{m,j}}h \end{aligned}$$

Now, let $\{c_1, \ldots, c_r\}$ be a Jordan frame compatible with $\{d_1, \ldots, d_{\hat{s}}\}$, that is, a Jordan frame for which $u_m = \sum_{i=1}^r \lambda_i(u_m)c_i$. Obviously, we also have $u = \sum_{i=1}^r \lambda_i(u)c_i$. We rewrite the first term of the above limit as follows.

$$\begin{aligned} &\lim_{m\to\infty} \sum_{i,j=1}^{\hat{s}} g'_{\hat{k}_i\hat{k}_j}(\lambda(u_m))\mathrm{tr}(d_{m,j}h)d_{m,i} = \sum_{i\neq j} A_{\hat{k}_i\hat{k}_j}\mathrm{tr}(d_j h)d_i + \sum_{i=1}^{\hat{s}} A_{\hat{k}_i\hat{k}_i}\mathrm{tr}(d_i h)d_i \\ &= \sum_{i\neq j}\sum_{k\in M_i'}\sum_{l\in M_j'} A_{kl}\mathrm{tr}(c_l h)c_k + \sum_{i=1}^{\hat{s}}\sum_{\substack{l,k\in M_i' \\ l\neq k}} A_{kk}\mathrm{tr}(c_l h)c_k + \sum_{i=1}^{\hat{s}}\sum_{k\in M_i'} A_{kk}\mathrm{tr}(c_k h)c_k. \end{aligned} \tag{3}$$

We work the second term as follows.

$$\begin{aligned} &\lim_{m\to\infty} \sum_{i=1}^{\hat{s}} b_{\hat{k}_i}(\lambda(u_m))(Q_{d_{m,i}}h - \mathrm{tr}(d_{m,i}h)d_{m,i}) = \sum_{i=1}^{\hat{s}} \sum_{\substack{l,k\in M_i' \\ l\neq k}} (A_{kk} - A_{kl})(Q_{c_k,c_l}h - \mathrm{tr}(c_l h)c_k) \\ &= \sum_{i=1}^{\hat{s}} \sum_{\substack{l,k\in M_i' \\ l\neq k}} A_{kk}Q_{c_k,c_l}h - A_{kl}Q_{c_k,c_l}h - A_{kk}\mathrm{tr}(c_l h)c_k + A_{kl}\mathrm{tr}(c_l h)c_k. \end{aligned} \tag{4}$$

Finally, the third term is:

$$\lim_{m\to\infty}\sum_{i\neq j}\frac{g_{\hat{k}_j}(\lambda(u_m))-g_{\hat{k}_i}(\lambda(u_m))}{\lambda_{\hat{k}_j}(u_m)-\lambda_{\hat{k}_i}(u_m)}Q_{d_{m,i},d_{m,j}}h$$

$$=\sum_{\alpha=1}^{s}\sum_{\beta\neq\alpha}\frac{g_{k_\beta}(\lambda)-g_{k_\alpha}(\lambda)}{\lambda_{k_\beta}-\lambda_{k_\alpha}}Q_{e_\alpha,e_\beta}h+\sum_{\alpha=1}^{s}\sum_{\substack{M_i',M_j'\subseteq M_\alpha\\ i\neq j}}A_{\hat{k}_i\hat{k}_i}Q_{d_i,d_j}h-A_{\hat{k}_i\hat{k}_j}Q_{d_i,d_j}hQ_{d_i,d_j}h \quad (5)$$

Now, we sum up (3), (4), and (5). The second term of (3) and the third term of (4) cancel. The first and the third term of (3), and the fourth of (4) sum up to

$$\sum_{i,j=1}^{r}A_{ij}\mathrm{tr}(c_jh)c_i.$$

The two first of (4) and the two last of (5) give:

$$\sum_{k\neq l}(A_{kk}-A_{kl})Q_{c_k,c_l}h.$$

It remains to add the first term of (5) to get the desired formula. ■

Example 3.1 *Our theorem provides us in particular with a close form formula for the Bouligand subdifferential of spectral functions of symmetric matrices. We consider here the Euclidian Jordan algebra $\mathcal{J}:=\mathbb{S}^n$ of $n\times n$ symmetric matrices with the symmetrized matrix product. Let $g:\mathbb{R}^n\to\mathbb{R}^n$ be a symmetric mapping, and G be the spectral mapping generated by g. In other words, for every $A=U\mathrm{diag}(\lambda(A))U^T$, with U orthogonal, we have:*

$$G(A):=U\mathrm{diag}(g(\lambda(A)))U^T,$$

where $\mathrm{diag}(\lambda)$ *is the diagonal matrix that has the vector λ on its diagonal. As there is no possible confusion between the two writings, we also use the notation* $\mathrm{diag}(X)$ *when X is a matrix to designate the vector build by the diagonal elements of X.*

Following the notation of Sendov and Lewis [], we introduce for every $\in\partial_B g(\lambda(A))$ the operator: $\mathcal{A}^V:\mathbb{R}^r\to\mathbb{S}^n$, with, for every $\lambda\in\mathbb{R}^n$:

$$\mathcal{A}^V_{ij}(\lambda)=\begin{cases}0, & \text{if } i=j;\\ V_{ii}-V_{ij}, & \text{if } i,j\in M_\alpha,\ i\neq j\\ \dfrac{g_i(\lambda)-g_j(\lambda)}{\lambda_i-\lambda_j}, & \text{otherwise.}\end{cases}$$

We use $X\circ Y$ to denote the Hadamard (i.e. componentwise) product of two matrices X and Y, $\mathrm{Diag}(X)$ *is the diagonal matrix filled with the diagonal elements of the matrix X.*

The Bouligand subdifferential $\partial_B G(A)$ can be described as follows:

$$\partial_B G(A)=\Big\{\ H\mapsto U[\mathrm{Diag}(V\mathrm{diag}(U^THU))+\mathcal{A}^V(\lambda(A))\circ(U^THU)]U^T,$$
$$\text{such that } V\in\partial_B g(\lambda(A)), U^TU=I, A=U\mathrm{diag}(\lambda(A))U^T\Big\}.$$

In order to relate this formula to (2), we consider a $V\in\partial_B g(\lambda(A))$ and an orthogonal matrix U as above. We denote by u_i the i-th column of U and by C_i the idempotent matrix $u_iu_i^T$. Now, $\sum_{i,j=1}^r V_{ij}\mathrm{Tr}(C_jH)C_i=\sum_{i=1}^r[V\mathrm{diag}(U^THU)]_iC_i=U\mathrm{Diag}(V\mathrm{diag}(UHU^T))U^T$. Similarly, the two last sums of (2) can easily be seen to match the term $U[\mathcal{A}^V(\lambda(A))\circ(UHU^T)]U^T$, because $Q_{C_i,C_j}H=(C_iHC_j+C_jHC_i)/2=(u_i[U^THU]_{ij}u_j^T+u_j[U^THU]_{ij}u_i^T)/2$. ■

4 Towards a formula for Clarke subdifferentials

In order to get the Clarke subdifferential, it "suffices" to take the convex hull of the Bouligand subdifferential we described in the previous section. This task appears quite delicate in general, but we can still describe not to difficultly a lower and an upper bound on these sets.

The following proposition is immediate.

Proposition 4.1 *Let G be a spectral mapping on $\mathcal{J}$ generated by a symmetric mapping g, and let u be a point of its domain. Suppose that a linear application V can be described as follows for every $h \in \mathcal{J}$:*

$$Vh = \sum_{i,j=1}^{r} A_{ij}\mathrm{tr}(c_j h)c_i + \sum_{\alpha=1}^{s} \sum_{\substack{i,j \in M_\alpha \\ i \neq j}} (A_{ii} - A_{ij}) Q_{c_i,c_j} h + \sum_{\alpha=1}^{s} \sum_{\beta \neq \alpha} \frac{g_{k\alpha}(\lambda(u)) - g_{k_\beta}(\lambda(u))}{\lambda_{k\alpha}(u) - \lambda_{k_\beta}(u)} Q_{e_\alpha,e_\beta} h,$$

where $A = [A_{ij}]_{i,j=1}^{r} \in \partial_C g(\lambda(u))$, and $\{c_1, \ldots, c_r\}$ is a Jordan frame for which u can be written as $u = \sum_{i=1}^{r} \lambda_i(u) c_i$.

An upper bound on this set can also be obtained with a little bit more effort. We need the following technical lemma first.

Lemma 4.2 *Let E and F be two real vector spaces. Let $Q \subseteq E$ and $S \subseteq F$ be two convex sets, and let $B : E \times F \to \mathbb{R}$ be a bilinear function. Given $a_1, a_2 \in Q$, $b_1, b_2 \in S$, and $\lambda \in [0,1]$, there exists $a_3 \in Q$ and $b_3 \in Q$ such that:*

$$B(a_3, b_3) = \lambda B(a_1, b_1) + (1 - \lambda) B(a_2, b_2). \tag{6}$$

Proof

The assertion is trivial when $\lambda \in \{0, 1\}$. We can assume that $0 < \lambda < 1$. We are looking for points a_3, b_3 of the form:

$$a_3 = \alpha a_1 + (1 - \alpha) a_2, \quad b_3 = \alpha b_1 + (1 - \alpha) b_2,$$

where $\alpha \in [0, 1]$. The relation (6) takes the form:

$$\alpha^2 B(a_1 - a_2, b_1 - b_2) + \alpha B(a_1 - a_2, b_2) + \alpha B(a_2, b_1 - b_2) = \lambda(B(a_1, b_1) - B(a_2, b_2)). \tag{7}$$

Let $u_1 := B(a_1 - a_2, b_1 - b_2)$, $u_2 := B(a_1 - a_2, b_1)$, and $u_3 := B(a_2, b_1 - b_2)$. We also define $s := u_1 + u_2 + u_3$. Observe that the right-hand side of (7) is λs. The case $s = 0$ is trivial, so we assume that $s \neq 0$. Let

$$g(\alpha) := \alpha^2 \frac{u_1}{s} + \alpha \frac{u_2}{s} + \alpha \frac{u_3}{s}.$$

If $u_1/s > 0$, we have $g(\lambda) = \lambda - \lambda(1 - \lambda)u_1/s < \lambda$ and $g(1) = 1 > \lambda$. By continuity, there exists an $\alpha \in]\lambda, 1[$ for which $g(\alpha) = \lambda$. Similarly, if $u_1/s < 0$, we have $g(\lambda) > \lambda$ and $g(0) = 0 < \lambda$ so that there exists an $\alpha \in]0, \lambda[$ for which $g(\alpha) = \lambda$. Finally, when u_1, one can take $\alpha = \lambda$. ■

For every $\gamma \in \mathbb{R}^r$, we denote by $\mathcal{SC}(\gamma)$ the convex hull of $\{P\gamma : P \in \mathcal{P}\}$. By Birkhoff's Theorem, the set $\mathcal{SC}(\gamma)$ can also be described by $\{B\gamma : B \text{ is doubly stochastic }\}$. Also, for every $\gamma \in \mathbb{R}^r$, and every partition $M_1, \ldots, M_s$ of $\{1, \ldots, r\}$, we write γ_{M_α} for the vector of $\mathbb{R}^{|M_\alpha|}$ constituted by the components of γ that are in M_α, ordered by their indices. Finally, for every $u \in \mathcal{J}$ that belongs also to a subalgebra $\mathcal{J}'$ of $\mathcal{J}$, we denote by $\lambda(u; \mathcal{J}')$ the vector of eigenvalues of u in the subalgebra $\mathcal{J}'$. This notation was used in [Bae07], where the reader can find some comments on these vectors.

Theorem 4.1 *Every linear application $V \in \partial_C G(u)$ can be written in the following form:*

$$h \mapsto Vh = \sum_{i,j=1}^{r} A_{ij}\gamma_i c_i + \sum_{\alpha=1}^{s} \sum_{\substack{i,j \in M_\alpha \\ i \neq j}} (A_{ii} - A_{ij}) h_{ij} + \sum_{\alpha=1}^{s} \sum_{\beta \neq \alpha} \frac{g_{k\alpha}(\lambda(u)) - g_{k\beta}(\lambda(u))}{\lambda_{k\alpha}(u) - \lambda_{k\beta}(u)} Q_{e_\alpha, e_\beta} h, \quad (8)$$

where $A = [A_{ij}]_{i,j=1}^{r} \in \partial_C g(\lambda(u))$, and $\{c_1, \ldots, c_r\}$ is a Jordan frame for which $\sum_{i \in M_\alpha} c_i = e_\alpha$ for all α. The vector $\gamma \in \mathbb{R}^r$ satisfies $\gamma_{M_\alpha} \in SC(\lambda(Q_{e_\alpha} h; \mathcal{J}_1(e_\alpha)))$, and the element h_{ij} belongs to $\mathcal{J}_1(e_\alpha)$, with $\lambda(h_{ij}; \mathcal{J}_1(e_\alpha)) \in SC((\|Q_{e_\alpha} h)\|/\sqrt{2}, 0, \cdots, 0, -\|Q_{e_\alpha} h)\|/\sqrt{2})$ for all $i \neq j$ in M_α.

Proof

Let $V \in \partial_C G(u)$. According to Caratheodory's Theorem, the application V can be represented as the convex combination of only a finite number of elements $A^{(1)}, \cdots, A^{(K)}$ of $\partial_B G(u)$. Let us fix $h \in \mathcal{J}$. From Theorem 3.1, we know that each $A^{(k)}h$ can be decomposed into three parts. The first part looks like

$$\sum_{\alpha,\beta=1}^{s} \sum_{i \in M_\alpha} \sum_{j \in M_\beta} A_{ij}^{(k)} \mathrm{tr}(c_j^{(k)} h) c_i^{(k)},$$

where $\sum_{i \in M_\alpha} c_i^{(k)} = e_\alpha$ for each k. Now, every convex combination of elements of the type:

$$\sum_{i \in M_\alpha} \sum_{j \in M_\beta} A_{ij}^{(k)} \mathrm{tr}(c_j^{(k)} h) c_i^{(k)} =: \sum_{i \in M_\alpha} b_i^{(k)} c_i^{(k)}$$

belongs to $\mathcal{J}_1(e_\alpha)$, and has eigenvalues in $\mathcal{J}_1(e_\alpha)$ that are in the convex hull of $\{b_{M_\alpha}^{(k)} : 1 \leq k \leq K\}$ (see Theorem 27 in [Bae07], and especially the inclusion (6)). We can now apply Lemma 4.2 with χ being the standard dot product in $\mathbb{R}^{|M_\alpha|}$ to get the first term of characterization (8).

We move now to the second term:

$$\sum_{\alpha=1}^{s} \sum_{\substack{i,j \in M_\alpha \\ i \neq j}} (A_{ii}^{(k)} - A_{ij}^{(k)}) Q_{c_i^{(k)}, c_j^{(k)}} h$$

It can be proved (See Proposition IV.1.4 in [FK94]) that the only possibly nonzero eigenvalues of $Q_{c_i^{(k)}, c_j^{(k)}} h$ are $\pm \|h\|_2/\sqrt{2}$, each of them with multiplicity one. By a similar reasoning as for the first term, we obtain the characterization of this second term given in (8).

Finally, the third term remains the same for every k, and is invariant under every convex combination of these. ■

This theorem is only a partial result, but it can nevertheless be used to prove that $0 \in \partial_C G(u)$ iff $0 \in \partial_C g(\lambda(u))$.

Acknowledgments

This research was supported by: ◇ Research Council KUL: GOA AMBioRICS, CoE EF/05/006 Optimization in Engineering(OPTEC), IOF-SCORES4CHEM; ◇ Belgian Federal Science Policy Office: IUAP P6/04 (DYSCO, Dynamical systems, control and optimization, 2007-2011).

References

[Bae06] M. Baes, *Spectral Functions and Smoothing Techniques in Jordan Algebras*, Ph.D. thesis, Université Catholique de Louvain, Louvain-la-Neuve, Belgium, September 2006.

[Bae07] ———, *Convexity and differentiability properties of spectral functions and spectral mappings on Euclidean Jordan algebras*, Linear Algebra and its Applications **422** (2007), no. 2-3, 664–700.

[BK66] H. Braun and M. Koecher, *Jordan-Algebren*, Grundlehren der mathematischen Wissenschaften, no. 128, Springer, 1966.

[BL06] J. Borwein and A. Lewis, *Convex analysis and nonlinear optimization*, second ed., CMS Books in Mathematics,, Springer, 2006, Theory and examples.

[Cla75] F. Clarke, *Generalized gradients and applications*, Transactions of the American Mathematical Society **205** (1975), 247–262.

[FK94] J. Faraut and A. Koranyi, *Analysis on Symmetric Cones*, Oxford University Press, 1994.

[Koe99] M. Koecher, *The Minnesota Notes on Jordan Algebras and Their Applications*, Lecture Notes in Mathematics, no. 1710, Springer-Verlag, 1999.

[Qi93] L. Qi, *Convergence analysis of some algorithms for solving nonsmooth equations*, Mathematics of Operations Research **18** (1993), 227–244.

[QS93] L. Qi and J. Sun, *A nonsmooth version of Newton's method*, Mathematical Programming **58** (1993), 353–367.

[Rad19] H. Rademacher, *Über partielle und totale Differenzierbarkeit von Functionen meherer Variabeln und über die Transformation der Doppelintegrale*, Mathematische Annalen **79** (1919), 340–359.

[SS03] D. Sun and J. Sun, *Strong Semismoothness of Eigenvalues of Symmetric Matrices and Its Application to Inverse Eigenvalue Problems*, SIAM Journal on Numerical Analysis **40** (2003), no. 6, 2352–2367.