to appear in Journal of Mathematical Analysis and Applications, 2011

A merit function method for infinite-dimensional SOCCPs

Yungyen Chiang [1]
Department of Applied Mathematics
National Sun Yat-sen University
Kaohsiung 80424, Taiwan

Shaohua Pan [2]
School of Mathematical Sciences
South China University of Technology
Guangzhou 510640, China

Jein-Shan Chen [3]
Department of Mathematics
National Taiwan Normal University
Taipei 11677, Taiwan

July 25, 2009
(revised December 20, 2010)

Abstract. We introduce the Jordan product associated with the second-order cone $\mathbb{K}$ into the real Hilbert space $\mathcal{H}$, and then define a one-parametric class of complementarity functions Φ_t on $\mathcal{H} \times \mathcal{H}$ with the parameter $t \in [0, 2)$. We show that the squared norm of Φ_t with $t \in (0, 2)$ is a continuously F(réchet)-differentiable merit function. By this, the second-order cone complementarity problem (SOCCP) in $\mathcal{H}$ can be converted into an unconstrained smooth minimization problem involving this class of merit functions, and furthermore, under the monotonicity assumption, every stationary point of this minimization problem is shown to be a solution of the SOCCP.

Key words: Hilbert space, complementarity, second-order cone, merit functions.

---

[1]The author's work is partially supported by grants from the National Science Council of the Republic of China. E-mail : chiangyy@math.nsysu.edu.tw.

[2]The author's work is supported by Guangdong Natural Science Foundation (No. 9251802902000001) and the Fundamental Research Funds for the Central Universities (SCUT). E-mail: shhpan@scut.edu.cn

[3]Member of Mathematics Division, National Center for Theoretical Sciences, Taipei Office. The author's work is partially supported by National Science Council of Taiwan. E-mail: jschen@math.ntnu.edu.tw.

1 Introduction

Let $\mathcal{H}$ be a real Hilbert space endowed with an inner product $\langle \cdot, \cdot \rangle$. The complementarity problem $\mathrm{CP}(K,T)$ in $\mathcal{H}$ is, for any given closed convex cone $K \subseteq \mathcal{H}$ and a continuously F(réchet)-differentiable mapping $T\colon \mathcal{H} \to \mathcal{H}$, to find a vector $x \in \mathcal{H}$ such that

$$x \in K, \quad T(x) \in K^* \ \text{ and } \ \langle x, T(x) \rangle = 0 \tag{1}$$

where $K^* := \{x \in \mathcal{H} \mid \langle x, y \rangle \geq 0 \ \ \forall y \in K\}$ is the dual cone of K. A closed convex cone K in $\mathcal{H}$ is called *self-dual* if K coincides with its dual cone K^*; for example, the non-negative orthant cone $\mathbb{R}^n_+ := \{(x_1, \dots x_n) \in \mathbb{R}^n \mid x_j \geq 0,\ j = 1, 2, \dots, n\}$ and the second-order cone (also called Lorentz cone) $\mathbb{K}^n := \{(r,\, x') \in \mathbb{R} \times \mathbb{R}^{n-1} \mid r \geq \|x'\|\}$. This paper is concerned with the complementarity problem associated with the infinite-dimensional second-order cone $\mathbb{K}$ in $\mathcal{H}$ which is closed, convex and self-dual (see Section 2 for its definition). The problem, denoted by $\mathrm{CP}(\mathbb{K}, T)$, is to find an $x \in \mathbb{K}$ such that

$$x \in \mathbb{K}, \quad T(x) \in \mathbb{K} \ \text{ and } \ \langle x, T(x) \rangle = 0. \tag{2}$$

This class of problems arises directly from the optimality conditions of certain types of infinite-dimensional optimization problems such as the one in [10], which is the reformulation of a min-max optimization problem with linear constraints in a Hilbert space.

Recently, nonlinear symmetric cone optimization and complementarity problems in finite-dimensional spaces such as semidefinite cone optimization and complementarity problems, second-order cone (SOC) optimization and complementarity problems, and general symmetric cone optimization and complementarity problems, become an active research field of mathematical programming. Taking SOC optimization and complementarity problems for example, there have proposed many effective solution methods, including the interior point methods [2, 18, 21, 22], the smoothing Newton methods [5, 9, 11], the semismooth Newton methods [16, 19], and the merit function method [6, 3]. However, to our best knowledge, there are few works about nonlinear symmetric cone optimization and complementarity problems in infinite-dimensional spaces except [10], in which with the JB algebras of finite rank primal-dual interior-point methods are presented for some special type of infinite-dimensional cone optimization problems.

In this paper, we consider a merit function method for solving the problem $\mathrm{CP}(\mathbb{K}, T)$. The method aims to seek a smooth merit function $\Psi : \mathcal{H} \times \mathcal{H} \to \mathbb{R}_+$ satisfying

$$\Psi(x, y) = 0 \iff x \in \mathbb{K}, \quad y \in \mathbb{K}, \quad \langle x, y \rangle = 0, \tag{3}$$

and reformulates the problem $\mathrm{CP}(\mathbb{K}, T)$ as a smooth minimization problem

$$\min_{x \in \mathcal{H}} \Psi(x, T(x)) \tag{4}$$

in the sense that x^* is a solution of $\mathrm{CP}(\mathbb{K},T)$ if and only if x^* solves (4) with zero optimal value. We call such Ψ a merit function associated with $\mathbb{K}$. Like handling complementarity problems in finite-dimensional spaces, we seek a merit function associated with $\mathbb{K}$ with a complementarity function (C-function for short) associated with $\mathbb{K}$. Specifically, a mapping $\Phi\colon \mathcal{H}\times\mathcal{H}\to\mathcal{H}$ is called a *C-function* associated with $\mathbb{K}$ if for any $x,y\in\mathcal{H}$,

$$\Phi(x,y)=0 \iff x\in\mathbb{K},\quad y\in\mathbb{K}\quad\text{and}\quad \langle x,y\rangle=0.$$

Clearly, the squared norm of Φ induces a merit function associated with $\mathbb{K}$.

When $\mathcal{H}$ is the Euclidean space $\mathbb{R}^n$, the Fischer-Burmeister (FB) and natural residual (NR) C-functions associated with the SOC $\mathbb{K}^n$ [9] are respectively defined as

$$\Phi_{\text{FB}}(x,y) := (x^2+y^2)^{1/2}-(x+y)\quad\ \forall x,y\in\mathbb{R}^n \tag{5}$$

and

$$\Phi_{\text{NR}}(x,y) := x-(x-y)_+\quad\ \forall x,y\in\mathbb{R}^n, \tag{6}$$

where $x^2=x\bullet x$ with "$\bullet$" means the Jordan product in $\mathbb{R}^n$, $x^{1/2}$ with $x\in\mathbb{K}^n$ is a vector such that $x^{1/2}\bullet x^{1/2}=x$, and $(x)_+$ denotes the projection onto $\mathbb{K}^n$. The function Φ_{FB} was well-studied in [6, 20], and particularly its squared norm was shown to be a smooth merit function in [6]. Since the squared norm of Φ_{NR} is not differentiable, it is often involved in the smoothing methods for the SOCCPs [5, 11]. The above two C-functions are subsumed in Kanzow and Kleinmichel C-function associated with $\mathbb{K}^n$:

$$\Phi_t(x,y) := \left[(x-y)^2+2tx\bullet y\right]^{1/2}-(x+y)\quad\ \forall x,y\in\mathbb{R}^n \tag{7}$$

where t is an arbitrary but fixed real number from $[0,2)$. This function was studied in [4] and its squared norm with $t\in(0,2)$ was shown to be continuously differentiable. Note that, when $n=1$, Φ_{FB}, Φ_{NR} and Φ_t reduce to the FB NCP-function [8], the minimum function [14], and the Kanzow and Kleinmichel NCP-function [12], respectively.

To define these C-functions in the Hilbert space $\mathcal{H}$, we introduce the Jordan product associated with the cone $\mathbb{K}$, and extend the Kanzow and Kleinmichel C-function defined in (7) to $\mathcal{H}$ and show that it satisfies the property (3) for each $t\in[0,2)$. In Section 4, we prove that the squared norm of this class of C-functions with $t\in(0,2)$ are continuously F-differentiable in $\mathcal{H}\times\mathcal{H}$. Note that the corresponding results in [4, 6] were proved by the spectral factorization of vectors, but here we shall not formally use this concept. In Section 5, under the monotonicity assumption, we establish that every stationary point of the unconstrained minimization problem involving this class of merit functions is a solution of $\mathrm{CP}(\mathbb{K},T)$, which generalizes the results of [4, Prop.4.1] and [6, Prop.3].

Throughout this paper, $\|\cdot\|$ denote the norm induced by the inner product $\langle\cdot,\cdot\rangle$ in $\mathcal{H}$. For any given Banach spaces $\mathcal{X}$ and $\mathcal{Y}$, let $\mathcal{L}(\mathcal{X},\mathcal{Y})$ denote the Banach space of all continuous linear mappings from $\mathcal{X}$ into $\mathcal{Y}$. We simply write $\mathcal{L}(\mathcal{X}, \mathcal{X}) = \mathcal{L}(\mathcal{X})$ and denote $\mathrm{GL}(\mathcal{X})$ by the set of all invertible mappings in $\mathcal{L}(\mathcal{X})$. The norm of any $l \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ is defined by $\|l\| := \sup\{\|l(x)\| \mid x \in \mathcal{X} \text{ and } \|x\| = 1\}$. In addition, for any self-adjoint linear operator l from $\mathcal{X} \to \mathcal{X}$, we write $l \succ 0$ (respectively, $l \succeq 0$) to mean that l is positive definite (respectively, positive semidefinite).

2 Lorentz cone and Jordan product

This section is devoted to introducing the Lorentz cone $\mathbb{K}$ mentioned above which is the unique self-dual cone in a family of pointed closed convex cones K in $\mathcal{H}$. Every cone in K is the image of $\mathbb{K}$ under some mapping in $\mathrm{GL}(\mathcal{H})$. Associated with the self-dual closed convex cone, the Jordan product is introduced into the Hilbert space $\mathcal{H}$.

For every integer $n > 1$, the Lorentz cone $\mathbb{K}^n$ given in Section 1 can be written as

$$\mathbb{K}^n := \left\{ x \in \mathbb{R}^n \mid \langle x, e \rangle \geq \frac{1}{\sqrt{2}} \|x\| \right\} \quad \text{with} \quad e = (1, 0) \in \mathbb{R} \times \mathbb{R}^{n-1}.$$

This motivates us to consider the following closed convex cone in the Hilbert space $\mathcal{H}$:

$$K(e, r) \;:=\; \Big\{ x \in \mathcal{H} \mid \langle x, e \rangle \geq r\|x\| \Big\}$$

where $e \in \mathcal{H}$ with $\|e\| = 1$ and $r \in \mathbb{R}$ with $0 < r < 1$. Observe that $K(e, r)$ is pointed, i.e., $K(e, r) \cap (-K(e, r)) = \{0\}$. Let $\langle e \rangle^{\perp} := \big\{ x \in \mathcal{H} \mid \langle x, e \rangle = 0 \big\}$. Then any $x \in \mathcal{H}$ can be written as $x = x' + \lambda e$ with $x' \in \langle e \rangle^{\perp}$ and $\lambda \in \mathbb{R}$. By noting that

$$\langle x, e \rangle \geq r\|x\| \iff \lambda \geq r(\|x'\|^2 + \lambda^2)^{1/2} \iff \lambda \geq \frac{r}{\sqrt{1 - r^2}} \|x'\|,$$

the closed convex cone $K(e, r)$ can be expressed as

$$K(e, r) = \left\{ x' + \lambda e \in \mathcal{H} \mid \; x' \in \langle e \rangle^{\perp} \text{ and } \lambda \geq \frac{r}{\sqrt{1 - r^2}} \, \|x'\| \right\}.$$

Proposition 2.1 *For any unit vector $e \in \mathcal{H}$ and $0 < r < 1$, the dual cone of $K(e, r)$ is $K(e, \sqrt{1 - r^2})$. Hence, the cone $K\left(e, \frac{1}{\sqrt{2}}\right) = \big\{ x' + \lambda e \in \mathcal{H} \mid \lambda \geq \|x'\| \big\}$ is self-dual.*

Proof. Let $x = x' + \lambda e \in K(e, \sqrt{1 - r^2}\,)$ and $y = y' + \mu e \in K(e, r)$ be arbitrary. Since $\lambda\mu \geq \|x'\| \cdot \|y'\|$, we have $\langle x,\, y \rangle \geq \langle x',\, y' \rangle + \|x'\| \cdot \|y'\| \geq 0$. This proves that

$$K(e, \sqrt{1 - r^2}\,) \subset K^*(e, r).$$

Conversely, let $x = x' + \lambda e \in K^*(e,r)$ be arbitrary, and we will prove $x \in K(e, \sqrt{1-r^2})$, i.e., $\lambda \geq r^{-1}\sqrt{1-r^2}\,\|x'\|$. This is trivial when $x' = 0$. When $x' \neq 0$, by considering the element $v = -r^{-1}\sqrt{1-r^2}\,x' + \|x'\|\,e$ of $K(e,r)$, we have

$$0 \leq \langle x, v\rangle = \left(\lambda - r^{-1}\sqrt{1-r^2}\|x'\|\right)\|x'\|,$$

which implies the result. The proof is complete. $\square$

Note that the unit vector $e \in \mathcal{H}$ is not unique. Every unit vector e determines a Lorentz cone $K(e, \frac{1}{\sqrt{2}})$. In this work, we consider a fixed unit vector e and write

$$\mathbb{K} = K\left(e, \frac{1}{\sqrt{2}}\right) = \left\{x' + \lambda e \in \mathcal{H} \mid \lambda \geq \|x'\|\right\}.$$

Unless stated otherwise, we shall alternatively write any $x \in \mathcal{H}$ as $x = x' + \lambda e$ with $x' \in \langle e\rangle^\perp$ and $\lambda = \langle x, e\rangle$. This expression is needed for stating many results and simplifying the computation in the subsequent analysis. In addition, for any $x, y \in \mathcal{H}$, we shall write $x \succ_{\mathbb{K}} y$ (respectively, $x \succeq_{\mathbb{K}} y$) if $x - y \in \mathrm{int}\mathbb{K}$ (respectively, $x - y \in \mathbb{K}$).

Next we show that the solution sets of complementarity problems associated with any $K(e,r)$ are related to those associated with $\mathbb{K}$ via the mappings in $\mathrm{GL}(\mathcal{H})$.

Lemma 2.1 *For any given $0 < r, s < 1$, let $\Lambda_{(r,s)}\colon \mathcal{H} \to \mathcal{H}$ be the mapping defined by*

$$\Lambda_{(r,s)}(x' + \lambda e) \ := \ \frac{\sqrt{1-s^2}}{\sqrt{1-r^2}}\,x' + \frac{s\lambda}{r}\,e \quad \forall x' + \lambda e \in \mathcal{H}.$$

Then, the following statements hold.

(a) *$\Lambda_{(r,s)} \in \mathrm{GL}(\mathcal{H})$ with $\Lambda_{(r,s)}^{-1} = \Lambda_{(s,r)}$, and $\Lambda_{(r,s)}$ maps $K(e,r)$ onto $K(e,s)$.*

(b) *Let $\Lambda_r := \Lambda_{(r,\frac{1}{\sqrt{2}})}$. If $r^2 + s^2 = 1$, then $\langle \Lambda_r(x),\, \Lambda_s(y)\rangle = \frac{1}{2rs}\langle x, y\rangle$ for all $x, y \in \mathcal{H}$.*

Proof. (a) It is clear that $\Lambda_{(r,s)}$ is linear and $\Lambda_{(r,s)}^{-1} = \Lambda_{(s,r)}$. For $x' \in \langle e\rangle^\perp$ and $\lambda \in \mathbb{R}$,

$$\|\Lambda_{(r,s)}(x' + \lambda e)\|^2 = \frac{1-s^2}{1-r^2}\,\|x'\|^2 + \frac{s^2}{r^2}\,\lambda^2 \leq \max\left\{\frac{1-s^2}{1-r^2},\, \frac{s^2}{r^2}\right\}\|x' + \lambda e\|^2.$$

This proves the continuity of $\Lambda_{(r,s)}$. Also, $\Lambda_{(r,s)}$ maps $K(e,r)$ onto $K(e,s)$ by noting that

$$\begin{aligned} x' + \lambda e \in \Lambda_{(r,s)}(K(e,r)) &\iff \Lambda_{(s,r)}(x' + \lambda e) \in K(e,r) \\ &\iff \frac{r\lambda}{s} \geq \frac{r}{\sqrt{1-r^2}} \cdot \frac{\sqrt{1-r^2}}{\sqrt{1-s^2}}\,\|x'\| \\ &\iff \lambda \geq \frac{s}{\sqrt{1-s^2}}\,\|x'\|. \end{aligned}$$

(b) We write $x = x' + \lambda e$ and $y = y' + \mu e$. Then,

$$\begin{aligned}\Lambda_r(x' + \lambda e) &= \frac{1}{\sqrt{2(1-r^2)}}\, x' + \frac{\lambda}{\sqrt{2}r}\, e = \frac{1}{\sqrt{2}s}\, x' + \frac{\lambda}{\sqrt{2}r}\, e\,;\\ \Lambda_s(y' + \mu e) &= \frac{1}{\sqrt{2}r}\, y' + \frac{\mu}{\sqrt{2(1-r^2)}}\, e = \frac{1}{\sqrt{2}r}\, y' + \frac{\mu}{\sqrt{2}s}\, e\,.\end{aligned}$$

Now, the assertion follows immediately by a direct computation. $\square$

From Lemma 2.1, we immediately obtain the following proposition.

Proposition 2.2 *Let $0 < r, s < 1$ be such that $r^2 + s^2 = 1$, and $T\colon \mathcal{H} \to \mathcal{H}$ be given.*

(a) *A point $x \in \mathcal{H}$ solves the problem $\mathrm{CP}(K(e,r),T)$ if and only if $\Lambda_r(x)$ solves the problem $\mathrm{CP}(\mathbb{K}, \Lambda_s \circ T \circ \Lambda_r^{-1})$.*

(b) *If $\Phi : \mathcal{H} \times \mathcal{H} \to \mathcal{H}$ is a C-function associated with $\mathbb{K}$, then the mapping $\Phi_r(x,y) := \Phi(\Lambda_r(x), \Lambda_s(y))$ is a C-function associated with $K(e,r)$.*

Next we introduce the Jordan product associated with the Lorentz cone $\mathbb{K}$. For any $x = x' + \lambda e \in \mathcal{H}$ and $y = y' + \mu e \in \mathcal{H}$, we define the Jordan product of x and y by

$$x \bullet y \;:=\; (\mu x' + \lambda y') + \langle x, y\rangle e, \tag{8}$$

and write $x^2 = x \bullet x$. Clearly, when $\mathcal{H} = \mathbb{R}^n$ and $e = (1,0) \in \mathbb{R} \times \mathbb{R}^{n-1}$, this definition is same as the one given by [7, Chapter II]. By the definition in (8) and a direct computation, it is easy to verify that the following properties hold.

Property 2.1 **(i)** *$x \bullet y = y \bullet x$ and $x \bullet e = x$ for all x, $y \in \mathcal{H}$.*

(ii) *$(x + y) \bullet z = x \bullet z + y \bullet z$ for all x, y, $z \in \mathcal{H}$.*

(iii) *$\langle x,\, y \bullet z\rangle = \langle y,\, x \bullet z\rangle = \langle z,\, x \bullet y\rangle$ for all x, y, $z \in \mathcal{H}$.*

(iv) *For any $x = x' + \lambda e \in \mathcal{H}$, $x^2 = x \bullet x = 2\lambda x' + \|x\|^2 e \in \mathbb{K}$ and $\langle x^2, e\rangle = \|x\|^2$.*

(v) *If $x = x' + \lambda e \in \mathbb{K}$, then there is a unique $x^{1/2} \in \mathbb{K}$ such that $(x^{1/2})^2 = x$, where*

$$x^{1/2} = \begin{cases} 0 & \text{if } x = 0;\\ x'/(2\tau) + \tau e & \text{othewise}\end{cases} \quad \text{with} \quad \tau = \sqrt{\frac{\lambda + \sqrt{\lambda^2 - \|x'\|^2}}{2}}. \tag{9}$$

(vi) *Every $x = x' + \lambda e \in \mathcal{H}$ with $\lambda^2 - \|x'\|^2 \neq 0$ is invertible w.r.t. the Jordan product, i.e., there is a unique point $x^{-1} \in \mathcal{H}$ such that $x \bullet x^{-1} = e$, where*

$$x^{-1} = \frac{-x' + \lambda e}{\lambda^2 - \|x'\|^2}. \tag{10}$$

Moreover, $x \in \mathrm{int}\mathbb{K}$ if and only if $x^{-1} \in \mathrm{int}\mathbb{K}$.

Associated with every $x \in \mathcal{H}$, we define a linear mapping L_x from $\mathcal{H}$ to $\mathcal{H}$ by

$$L_x y := x \bullet y \quad \text{for any } y \in \mathcal{H}. \tag{11}$$

Clearly, $L_x \in \mathcal{L}(\mathcal{H})$. Also, the mapping possesses the following favorable properties.

Lemma 2.2 *For any $x \in \mathcal{H}$, let $L_x \in \mathcal{L}(\mathcal{H})$ be defined as above. Then, we have*

(a) $x \succ_{\mathbb{K}} 0 \Longleftrightarrow L_x \succ 0$ and $x \succeq_{\mathbb{K}} 0 \Longleftrightarrow L_x \succeq 0$.

(b) *If $x = x' + \lambda e$ with $\lambda \neq 0$ and $|\lambda| \neq \|x'\|$, then $L_x \in \mathrm{GL}(\mathcal{H})$ with the inverse given by*

$$L_x^{-1} y = \lambda^{-1}\left(y' - \langle x^{-1}, y\rangle x'\right) + \langle x^{-1}, y\rangle e \quad \text{for any } y = y' + \mu e \in \mathcal{H}. \tag{12}$$

Proof. (a) Fix any $x = x' + \lambda e \in \mathcal{H}$. It suffices to prove the first equivalence, and the second equivalence follow from the first equivalence and the closedness of $\mathbb{K}$. Note that $L_x \succ 0$ if and only if $\langle h, L_x h\rangle > 0$ for any $h = h' + \xi e \in \mathcal{H}\backslash\{0\}$, whereas

$$\begin{aligned} \langle h, L_x h\rangle > 0 \quad &\Longleftrightarrow \quad \lambda\|h'\|^2 + 2\xi\langle x', h'\rangle + \lambda\xi^2 > 0 \\ &\Longleftrightarrow \quad \lambda > 0 \ \text{ and } \ 4\langle x', h'\rangle^2 - 4\lambda^2\|h'\|^2 < 0 \\ &\Longleftrightarrow \quad \lambda > 0 \ \text{ and } \ \|x'\| < \lambda. \end{aligned}$$

(b) To prove $L_x \in \mathrm{GL}(\mathcal{H})$, it suffices to prove that $L_x y = 0$ for some $y = y' + \mu e \in \mathcal{H}$ implies $y = 0$. Indeed, since $L_x y = 0$ implies $\|x \bullet y\|^2 = 0$, which is equivalent to

$$\lambda y' + \mu y' = 0 \ \text{ and } \ \langle x', y'\rangle + \lambda\mu = 0.$$

Since $\lambda \neq 0$, from the first equality we have $y' = -\lambda^{-1}\mu x'$. Substituting it into the second equality yields $\mu = 0$, and so $y' = 0$. A direct computation verifies (12). $\square$

3 Kanzow-Kleinmichel merit function

In this section, we will extend Kanzow-Kleinmichel C-function in (7) to the real Hilbert space $\mathcal{H}$, and present some technical lemmas that will be used in the subsequent analysis. Let t be an arbitrary real number in $[0, 2)$. Define the mapping $\Phi_t \colon \mathcal{H} \times \mathcal{H} \to \mathcal{H}$ by

$$\Phi_t(x, y) \ := \ \left[(x - y)^2 + 2t(x \bullet y)\right]^{1/2} - (x + y). \tag{13}$$

Note that, for any $t \in [0, 2)$ and any $x, y \in \mathcal{H}$,

$$(x - y)^2 + 2t(x \bullet y) = (x + (t - 1)y)^2 + t(2 - t)y^2 \in \mathbb{K}. \tag{14}$$

Hence, the function Φ_t is well-defined. It is easy to see that when $t = 1$ and $t = 0$, Φ_t reduces to the FB and the NR C-function associated with $\mathbb{K}$, respectively.

To show that each Φ_t is a C-function associated with $\mathbb{K}$, we need the following result which is an infinitely dimensional version of [9, Prop.2.1]. The proof given in [9] was based on the geometry of vectors in Euclidean spaces, that is, the notion of an angle between vectors. We here give another proof without using this notion.

Lemma 3.1 *For any $x,\, y \in \mathcal{H}$, the following statements are equivalent:*

(a) $x \in \mathbb{K}$, $y \in \mathbb{K}$ *and* $\langle x,\, y\rangle = 0$*;*

(b) $x \in \mathbb{K}$, $y \in \mathbb{K}$ *and* $x \bullet y = 0$*;*

(c) $x + y \in \mathbb{K}$ *and* $x \bullet y = 0$.

(d) *It holds that (i)* $x = 0$, $y \in \mathbb{K}$*; or (ii)* $x \in \mathbb{K}$, $y = 0$*; or (iii)* $x \in \partial\mathbb{K}$, $y \in \partial\mathbb{K}$ *and* $\langle x, y\rangle = 0$, *where* $\partial\mathbb{K} := \{x' + \lambda e \in \mathcal{H} \mid \ \lambda = \|x'\|\}$ *denotes the boundary of* $\mathbb{K}$.

Proof. Clearly, $(b) \Rightarrow (c)$ and $(d) \Rightarrow (a)$. We need to prove $(a) \Rightarrow (b)$ and $(c) \Rightarrow (d)$.

(a) $\Rightarrow$ (b). Write $x = x' + \lambda e$ and $y = y' + \mu e$. By (8) and $\langle x, y\rangle = 0$, we have $x \bullet y = (\mu x' + \lambda y')$. Since $\lambda \geq \|x'\|$ and $\mu \geq \|y'\|$ by $x, y \in \mathbb{K}$, it follows that

$$\|\mu x' + \lambda y'\|^2 = \mu^2\|x'\|^2 - 2\lambda^2\mu^2 + \lambda^2\|y'\|^2 \leq 0,$$

and $\mu x' + \lambda y' = 0$ follows. Thus, we obtain $x \bullet y = 0$, and hence (a) implies (b).

(c) $\Rightarrow$ (d). Since $x \bullet y = 0$ implies $\|(\mu x' + \lambda y') + \langle x, y\rangle e\|^2 = \|\mu x' + \lambda y'\|^2 + \langle x, y\rangle^2 = 0$, we have $\langle x, y\rangle = 0$ and $\mu x' + \lambda y' = 0$. If $\lambda = 0, \mu \neq 0$, then from $\mu x' + \lambda y' = 0$ and $\langle x, y\rangle = 0$, we get $x' = 0$, and then $x = 0$. Together with $x + y \in \mathbb{K}$, we obtain $y \in \mathbb{K}$, and so Case (i) holds. If $\lambda \neq 0, \mu = 0$, a similar argument yields that Case (ii) holds. If $\lambda = \mu = 0$, then from $x + y \in \mathbb{K}$ it follows that $\|x' + y'\| = 0$. This along with $\langle x, y\rangle = 0$ and $\lambda = 0, \mu = 0$ yields that $x' = 0$ and $y' = 0$, and consequently, $x = y = 0$. Hence, Cases (i), (ii) and (iii) hold. Now, assume that $\lambda\mu \neq 0$. From $\mu x' + \lambda y' = 0$ and $\langle x,\, y\rangle = 0$, we obtain $\lambda^2 = \|x'\|^2$ and $\mu^2 = \|y'\|^2$. This, together with $x + y \in \mathbb{K}$, i.e. $(\lambda + \mu)^2 \geq \|x' + y'\|^2$, implies $\lambda\mu \geq \langle x',\, y'\rangle = -\lambda\mu$, and hence $\lambda\mu > 0$. Since $\lambda + \mu \geq \|x' + y'\|$, we get $\lambda > 0$ and $\mu > 0$. Thus, $\lambda = \|x'\|$ and $\mu = \|y'\|$, which implies that $x, y \in \mathbb{K}$. That is, Case (iii) follows. $\square$

Let $\Psi_t \colon \mathcal{H} \times \mathcal{H} \to \mathbb{R}_+$ denote the squared norm of the function Φ_t, that is,

$$\Psi_t(x, y) := \|\Phi_t(x,\, y)\|^2 \quad \forall x, y \in \mathcal{H}. \tag{15}$$

From the expression of Φ_t and Lemma 3.1, it follows that

$$\begin{aligned}\Psi_t(x,y)=0 \Leftrightarrow \Phi_t(x,\,y)=0 \quad &\Leftrightarrow \quad x+y\in \mathbb{K} \text{ and } (x-y)^2+2t(x\bullet y)=(x+y)^2\\ &\Leftrightarrow \quad x+y\in \mathbb{K} \text{ and } x\bullet y=0\\ &\Leftrightarrow \quad x\in \mathbb{K},\ y\in\mathbb{K},\ \langle x,y\rangle=0.\end{aligned}$$

These equivalence immediately implies the following result.

Proposition 3.1 *The functions Φ_t and Ψ_t are respectively a C-function and a merit function associated with $\mathbb{K}$.*

In what follows, we provide some necessary technical lemmas that will be used later.

Lemma 3.2 *For any given $0<t<2$, $x=x'+\lambda e\in\mathcal{H}$ and $y=y'+\mu e\in\mathcal{H}$, we have*

$$\begin{aligned}(x-y)^2+2t(x\bullet y)\in\partial\mathbb{K} \quad &\Longleftrightarrow \quad x^2+y^2\in\partial\mathbb{K}\\ &\Longleftrightarrow \quad |\lambda|=\|x'\|,\ |\mu|=\|y'\|,\ \lambda\mu=\langle x',y'\rangle\\ &\Longrightarrow \quad \lambda y'=\mu x'\end{aligned}\tag{16}$$

Proof. Using $|\lambda|=\|x'\|, |\mu|=\|y'\|$ and $\lambda\mu=\langle x',y'\rangle$, it is easy to verify $\|\lambda y'-\mu x'\|^2=0$. So, the implication in (16) holds. Now we prove the second equivalence. Noting that

$$\begin{aligned}x^2+y^2 &= 2(\lambda x'+\mu y')+(\|x\|^2+\|y\|^2)e,\\ 2\|\lambda x'+\mu y'\| &\le 2\|\lambda x'\|+2\|\mu y'\|\le\|x\|^2+\|y\|^2,\end{aligned}$$

we have $x^2+y^2\in\partial\mathbb{K}$ if and only if $\|x\|^2+\|y\|^2=2\|\lambda x'\|+2\|\mu y'\|=2\|\lambda x'+\mu y'\|$, i.e., $(|\lambda|-\|x'\|)^2+(|\mu|-\|y'\|)^2=0$ and $\|\lambda x'\|+\|\mu y'\|=\|\lambda x'+\mu y'\|$. Thus, we have

$$x^2+y^2\in\partial\mathbb{K}\Longleftrightarrow|\lambda|=\|x'\|,\ |\mu|=\|y'\|,\ \lambda\mu\langle x',y'\rangle=|\lambda\mu|\cdot\|x'\|\cdot\|y'\|.$$

We may argue that, when $|\lambda|=\|x'\|$ and $|\mu|=\|y'\|$, there holds that

$$\lambda\mu\langle x',y'\rangle=|\lambda\mu|\cdot\|x'\|\cdot\|y'\| \iff \lambda\mu=\langle x',y'\rangle.$$

Indeed, if the equality on the right hand side holds, then $\lambda\mu\langle x',y'\rangle=\lambda^2\mu^2=|\lambda\mu|\cdot\|x'\|\cdot\|y'\|$, which implies the equality of the left hand side. Assume that the equality of the left hand side holds. If $\lambda\mu=0$ or $\|x'\|\cdot\|y'\|=0$, then $x=0$ or $y=0$, and thus $\lambda\mu=0=\langle x',y'\rangle$; and if $\lambda\mu\neq0$ and $\|x'\|\cdot\|y'\|\neq0$, using $\lambda\mu\langle x',y'\rangle=|\lambda\mu|\cdot\|x'\|\cdot\|y'\|>0$ then yields that $|\langle x',y'\rangle|=\|x'\|\cdot\|y'\|=|\lambda\mu|$ and $\langle x',y'\rangle=\lambda\mu$. This proves that the equality of the right hand side holds, and the second equivalence in (16) follows.

To establish the first equivalence in (16), it suffices to prove that

$$(x-y)^2+2t(x\bullet y)\in\partial\mathbb{K} \iff |\lambda|=\|x'\|,\ |\mu|=\|y'\|,\ \lambda\mu=\langle x',y'\rangle.\tag{17}$$

Recall that $(x-y)^2+2t(x\bullet y)=(x+(t-1)y)^2+(\sqrt{t(2-t)}\ y)^2$. By the result above,

$$\begin{aligned}(x-y)^2+2t(x\bullet y)\in\partial\mathbb{K}\iff\ & |\lambda+(t-1)\mu|=\|x'+(t-1)y'\|,|\mu|=\|y'\|,\\ & \mu(\lambda+(t-1)\mu)=\langle x'+(t-1)y',y'\rangle.\end{aligned}$$

Taking into account that $|\mu|=\|y'\|$ implies the following equivalences

$$\begin{aligned}|\lambda+(t-1)\mu|=\|x'+(t-1)y'\|&\iff\lambda^2+2(t-1)\lambda\mu=\|x'\|^2+2(t-1)\langle x',y'\rangle,\\ \mu(\lambda+(t-1)\mu)=\langle x'+(t-1)y',y'\rangle&\iff\lambda\mu=\langle x',y'\rangle,\end{aligned}$$

we immediately obtain (17). Thus, the proof is complete. $\square$

The following lemma is essentially proved in [6, Lemma 3]. We give a simpler proof.

Lemma 3.3 *For $j=1,2$, let $x_j=x'_j+\lambda_je\in\mathcal{H}$. If $\lambda_1x'_1+\lambda_2x'_2\neq0$, then for $j=1,\,2$,*

$$\begin{aligned}\left(\lambda_j+(-1)^j\left\langle\frac{\lambda_1x'_1+\lambda_2x'_2}{\|\lambda_1x'_1+\lambda_2x'_2\|},x'_j\right\rangle\right)^2&\le\left\|x'_j+(-1)^j\lambda_j\frac{\lambda_1x'_1+\lambda_2x'_2}{\|\lambda_1x'_1+\lambda_2x'_2\|}\right\|^2\\ &\le\|x_1\|^2+\|x_2\|^2+2(-1)^j\|\lambda_1x'_1+\lambda_2x'_2\|.\end{aligned}$$

Proof. It suffices to prove the inequalities for $j=1$. The first inequality holds trivially since $|\langle v,w\rangle|\le\|v\|\cdot\|w\|$ for all $v,w\in\mathcal{H}$. The second inequality is proved as follows.

$$\begin{aligned}\left\|x'_1-\lambda_1\frac{\lambda_1x'_1+\lambda_2x'_2}{\|\lambda_1x'_1+\lambda_2x'_2\|}\right\|^2&=\|x'_1\|^2-\frac{2}{\|\lambda_1x'_1+\lambda_2x'_2\|}\langle\lambda_1x'_1,\lambda_1x'_1+\lambda_2x'_2\rangle+\lambda_1^2\\ &=\|x_1\|^2-2\|\lambda_1x'_1+\lambda_2x'_2\|+\frac{2\langle\lambda_2x'_2,\lambda_1x'_1+\lambda_2x'_2\rangle}{\|\lambda_1x'_1+\lambda_2x'_2\|}\\ &\le\|x_1\|^2-2\|\lambda_1x'_1+\lambda_2x'_2\|+2|\lambda_2|\|x'_2\|\\ &\le\|x_1\|^2-2\|\lambda_1x'_1+\lambda_2x'_2\|+\|x_2\|^2,\end{aligned}$$

where the last inequality is using $\|x_2\|^2=\lambda_2^2+\|x'_2\|^2$. Thus, the proof is complete. $\square$

To end the contents of this section, we recall the concept of F(réchet)-differentiability and present some continuously F-differentiable mappings for later use. For given Banach spaces $\mathcal{X}$ and $\mathcal{Y}$, a mapping f from a nonempty open subset X of $\mathcal{X}$ into $\mathcal{Y}$ is said to be *F-differentiable at $x\in X$* if there exists $l_x\in\mathcal{L}(\mathcal{X},\mathcal{Y})$ such that

$$\lim_{h\to0}\frac{f(x+h)-f(x)-l_xh}{\|h\|}=0,$$

and l_x is called *the F-differential* of f at x, written by $f'(x)$. When f is F-differentiable at every point of X, we say that f is F-differentiable on X. If f is F-differentiable on a neighborhood $U \subset X$ of a point $x_0 \in X$, and if, as a mapping from U into the Banach space $\mathcal{L}(\mathcal{X}, \mathcal{Y})$, the mapping $x \longmapsto f'(x)$ is continuous at x_0, then f is said to be *continuously F-differentiable at* x_0. The mapping f is called continuously F-differentiable on X if it is continuously F-differentiable at every point of X. Note that if $f \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$, then f is continuously F-differentiable on $\mathcal{X}$ with $f'(x) = f$ for every $x \in \mathcal{X}$, i.e., $f'(x)v = f(v)$ for all $v \in \mathcal{X}$. By the definition, it is easy to verify the continuous F-differentiability of the mappings given below.

Example 3.1 **(i)** *$f(x) = \langle x, e\rangle$ for any $x \in \mathcal{H}$ with $f'(x)v = \langle v, e\rangle$ for all $v \in \mathcal{H}$.*

(ii) *$f(x) = x - \langle x, e\rangle e$ for any $x \in \mathcal{H}$ with $f'(x)v = v - \langle v, e\rangle e$ for all $v \in \mathcal{H}$.*

(iii) *$f(x) = x^2 = x \bullet x$ for any $x \in \mathcal{H}$ with $f'(x)v = 2x \bullet v$ for all $v \in \mathcal{H}$.*

(iv) *$f(x) = \|x\|^2$ for any $x \in \mathcal{H}$ with $f'(x)v = 2\langle x, v\rangle$ for all $x, v \in \mathcal{H}$.*

(v) *$f(x) = \|x\| = \langle x, x\rangle^{1/2}$ for any $x \in \mathcal{H}$. Such f is continuously F-differentiable only on $\mathcal{H} \setminus \{0\}$ with $f'(x)v = \frac{1}{\|x\|}\langle x, v\rangle$ for all $v \in \mathcal{H}$.*

4 Smoothness of merit function

This section is devoted to establishing the continuous F-differentiability (smoothness) of Ψ_t. For this purpose, we first investigate the F-differentiability of two special mappings defined as in the following two lemmas, respectively.

Lemma 4.1 *Let $\sigma(x) := x^{1/2}$ for any $x \in \mathbb{K}$. Then, the following statements hold.*

(a) σ is continuously F-differentiable on int$\mathbb{K}$, *and for all $v \in \mathcal{H}$,*

$$\sigma'(x)v = \frac{\sqrt{\lambda^2 - \|x'\|^2}}{2\tau}\langle x^{-1/2}, v\rangle x^{-1/2} + \frac{v - \langle v, e\rangle e}{2\tau}$$

where τ is given as in (9).

(b) For every $x \in \operatorname{int}\mathbb{K}$, $2\sigma'(x)v = L^{-1}_{\sigma(x)}v$ for all $v \in \mathcal{H}$.

(c) For every $x \in \operatorname{int}\mathbb{K}$, the F-differential $\sigma'(x)$ is a self-adjoint operator in $\mathcal{L}(\mathcal{H})$, i.e., $\langle \sigma'(x)v, w\rangle = \langle v, \sigma'(x)w\rangle$ for all $v, w \in \mathcal{H}$.

Proof. (a) Recall that $\sigma(x) = \frac{x'}{2\tau} + \tau e$ for $x = x' + \lambda e \in \mathbb{K} \setminus \{0\}$. Since τ as a mapping of $x \in \mathbb{K} \setminus \{0\}$ is F-differentiable on int$\mathbb{K}$, the function σ is F-differentiable on int$\mathbb{K}$.

The differential of σ is computed as follows. Taking into account $2\tau^2 = \lambda + \sqrt{\lambda^2 - \|x'\|^2}$, by Example 3.1 it is not hard to calculate that for all $v \in \mathcal{H}$,

$$4\tau\tau'(x)v \;=\; \langle v, e\rangle + \frac{\lambda\langle v, e\rangle - \langle v, x'\rangle}{\sqrt{\lambda^2 - \|x'\|^2}} = \frac{\langle v, 2\tau^2 e - x'\rangle}{\sqrt{\lambda^2 - \|x'\|^2}},$$

and consequently,

$$\tau'(x)v \;=\; \frac{1}{2\sqrt{\lambda^2 - \|x'\|^2}}\left\langle \tau e - \frac{x'}{2\tau}, v\right\rangle = \frac{1}{2}\left\langle x^{-1/2}, v\right\rangle.$$

Together with the expression $\sigma(x) = \frac{x'}{2\tau} + \tau e$, we obtain that

$$\begin{aligned}
\sigma'(x)v &= \frac{-\tau'(x)v}{2\tau^2}x' + \frac{1}{2\tau}\left(v - \langle v, e\rangle e\right) + \left(\tau'(x)v\right)e \\
&= \frac{\left\langle x^{-1/2}, v\right\rangle}{2\tau}\left(\frac{-x'}{2\tau} + \tau e\right) + \frac{1}{2\tau}\left(v - \langle v, e\rangle e\right) \qquad (18)\\
&= \frac{\left\langle x^{-1/2}, v\right\rangle}{2\tau}\cdot\sqrt{\lambda^2 - \|x'\|^2}x^{-1/2} + \frac{1}{2\tau}\left(v - \langle v, e\rangle e\right).
\end{aligned}$$

We next prove that the F-differential σ' is continuous at any given point $a = a' + \alpha e \in \mathrm{int}\mathbb{K}$. For any $x = x' + \lambda e \in \mathrm{int}\mathbb{K}$, we write

$$\tau(x) = \tau = \sqrt{\frac{\lambda + \sqrt{\lambda^2 - \|x'\|^2}}{2}} \quad \text{and} \quad p(x) = \frac{\sqrt{\lambda^2 - \|x'\|^2}}{2\tau(x)}.$$

Then, from the last equality in (18), it follows that for all $v \in \mathcal{H}$,

$$\begin{aligned}
&\|\sigma'(x)v - \sigma'(a)v\| \\
\le\;& \left\|p(x)\langle x^{-1/2}, v\rangle x^{-1/2} - p(a)\langle a^{-1/2}, v\rangle a^{-1/2}\right\| + \left|\frac{1}{2\tau(x)} - \frac{1}{2\tau(a)}\right|\cdot\|v - \langle v,\, e\rangle e\| \\
\le\;& |p(x) - p(a)|\cdot\left|\langle x^{-1/2}, v\rangle\right|\cdot\left\|x^{-1/2}\right\| + p(a)\cdot\left|\langle x^{-1/2} - a^{-1/2}, v\rangle\right|\cdot\left\|x^{-1/2}\right\| \\
&+p(a)\cdot\left|\langle a^{-1/2}, v\rangle\right|\cdot\left\|x^{-1/2} - a^{-1/2}\right\| + \left|\frac{1}{2\tau(x)} - \frac{1}{2\tau(a)}\right|\cdot\|v\| \\
\le\;& |p(x) - p(a)|\cdot\|x^{-1/2}\|^2\cdot\|v\| + \left|\frac{1}{2\tau(x)} - \frac{1}{2\tau(a)}\right|\cdot\|v\| \\
&+p(a)\cdot\left\|x^{-1/2} - a^{-1/2}\right\|\left(\left\|x^{1/2}\right\| + \left\|a^{-1/2}\right\|\right)\cdot\|v\|.
\end{aligned}$$

This implies that

$$\begin{aligned}
\|\sigma'(x) - \sigma'(a)\| \;\le\;& |p(x) - p(a)|\cdot\left\|x^{-1/2}\right\|^2 + \left|\frac{1}{2\tau(x)} - \frac{1}{2\tau(a)}\right| \\
&+p(a)\cdot\left\|x^{-1/2} - a^{-1/2}\right\|\left(\|x^{1/2}\| + \|a^{-1/2}\|\right),
\end{aligned}$$

and consequently $\|\sigma'(x)-\sigma'(a)\| \to 0$ as $x \to a$.

(b) From the second equality in (18) and equation (12), we obtain for any $v = v'+\theta e \in \mathcal{H}$,

$$\begin{aligned} 2\sigma'(x)v &= \frac{1}{\tau}\langle \sigma(x)^{-1}, v\rangle \cdot \left(\frac{-x'}{2\tau} + \tau e\right) + \frac{v'}{\tau} \\ &= \frac{1}{\tau}\left(v' - \langle \sigma(x)^{-1}, v\rangle \cdot \frac{x'}{2\tau}\right) + \langle \sigma(x)^{-1}, v\rangle e = L^{-1}_{\sigma(x)} v. \end{aligned}$$

(c) For any given $v, w \in \mathcal{H}$, we write $\sigma'(x)v = v_1$ and $\sigma'(x)w = w_1$. Then, by part (b), we have $v = 2\sigma(x) \bullet v_1$ and $w = 2\sigma(x) \bullet w_1$, and consequently

$$\langle \sigma'(x)v, w\rangle = 2\langle v_1, \sigma(x) \bullet w_1\rangle = 2\langle \sigma(x) \bullet v_1, w_1\rangle = \langle v, \sigma'(x)w\rangle.$$

This shows that $\sigma'(x)$ is a self-adjoint operator. The proof is completed. $\square$

Lemma 4.2 *For any $x, y \in \mathcal{H}$ and $r \in \mathbb{R}$, let $\psi_r(x,y) := 2\langle (x^2+y^2)^{1/2}, x+ry\rangle$. Then,*

(a) *ψ_r is F-differentiable at every point $(a,b) \in \mathcal{H} \times \mathcal{H}$ with $a^2+b^2 \in \partial\mathbb{K}$.*

(b) *For any given $x = x'+\lambda e \in \mathcal{H}$ and $y = y'+\mu e \in \mathcal{H}$ with $x^2+y^2 \in \partial\mathbb{K} \setminus \{0\}$,*

$$\psi_r'(x,y)(v,w) = \frac{2(\lambda + r\mu)}{\sqrt{\lambda^2+\mu^2}} \cdot (\langle v, x\rangle + \langle w, y\rangle) + 2\left\langle \left(x^2+y^2\right)^{1/2}, v+rw\right\rangle$$

for all $v, w \in \mathcal{H}$. Furthermore, $\|\psi_r'(x,y)\| \le 4(1+|r|)\sqrt{\|x\|^2+\|y\|^2}$.

Proof. (a) For any $(x,y) \neq (0,0)$, it can be seen that ψ_r is F-differentiable at $(0,0)$ since

$$|\psi_r(x,y) - \psi_r(0,0)| = 2\left|\left\langle (x^2+y^2)^{1/2}, x+ry\right\rangle\right| \le 2\sqrt{\|x\|^2+\|y\|^2} \cdot \|x+ry\|.$$

Next, we consider the case where $(a,b) \neq (0,0)$. Write $a = a'+\alpha e$ and $b = b'+\beta e$. Since $a^2+b^2 \in \partial\mathbb{K}$, we have $2\|\alpha a' + \beta b'\| = \|a\|^2+\|b\|^2 > 0$. So, there exist a convex and bounded open neighborhood U of (a,b) in $\mathcal{H} \times \mathcal{H}$ and a constant $\rho > 0$ such that $\|\lambda x' + \beta y'\| \ge \rho$ for any $(x,y) \in U$ with $x = x'+\lambda e$ and $y = y'+\mu e$. Notice that

$$(x^2+y^2)^{1/2} = \frac{\lambda x' + \mu y'}{\tau(x,y)} + \tau(x,y)e$$

where

$$\tau(x,y) = \sqrt{\frac{\|x\|^2+\|y\|^2+\sqrt{(\|x\|^2+\|y\|^2)^2 - 4\|\lambda x' + \mu y'\|^2}}{2}}.$$

Write

$$\tau_j = \tau_j(x,\, y) := \|x\|^2 + \|y\|^2 + 2(-1)^j\, \|\lambda x' + \mu y'\| \quad \text{for } j = 1, 2.$$

It is not difficult to verify that

$$\tau(x, y) = \frac{\sqrt{\tau_1} + \sqrt{\tau_2}}{2} \quad \text{and} \quad \frac{1}{\tau(x, y)} = \frac{\sqrt{\tau_2} - \sqrt{\tau_1}}{2\|\lambda x' + \mu y'\|}. \tag{19}$$

Consequently,

$$\begin{aligned}
\psi_r(x, y) &= 2\left\langle \frac{\lambda x' + \mu y'}{\tau(x, y)}, x' + r y' \right\rangle + 2\tau(x, y)(\lambda + r\mu) \\
&= \left(\sqrt{\tau_2} - \sqrt{\tau_1}\right) \left\langle \frac{\lambda x' + \mu y'}{\|\lambda x' + \mu y'\|}, x' + r y' \right\rangle + \left(\sqrt{\tau_1} + \sqrt{\tau_2}\right)(\lambda + r\mu) \\
&:= \varphi_1(x, y) + \varphi_2(x, y)
\end{aligned}$$

where

$$\varphi_j(x, y) := \sqrt{\tau_j(x, y)} \left(\lambda + r\mu + (-1)^j \left\langle \frac{\lambda x' + \mu y'}{\|\lambda x' + \mu y'\|}, x' + r y' \right\rangle \right) \quad \text{for } j = 1, 2.$$

Since $\lambda x' + \mu y' \neq 0$ for any $(x, y) \in U$, the mappings

$$(x, y) \longmapsto \|\lambda x' + \mu y'\| \quad \text{and} \quad (x, y) \longmapsto \|\lambda x' + \mu y'\|^{-1}$$

are continuously F-differentiable on U, and then $\sqrt{\tau_2(x, y)}$ is continuously F-differentiable on U since $\tau_2(x, y) > 0$ for $(x, y) \neq (0, 0)$. Hence, φ_2 is continuously Fréchet differentiable on U. To prove that φ_1 is F-differentiable at (a, b), we let

$$\begin{gathered}
f(x, y) := \lambda + r\mu, \quad g(x, y) := \lambda x' + \mu y', \quad p(x, y) := \frac{g(x, y)}{\|g(x, y)\|}, \\
h(x, y) := x' + r y', \quad \varphi_3(x, y) := f(x, y) - \langle p(x, y), h(x, y) \rangle
\end{gathered}$$

for any $(x, y) \in U$ with $x = x' + \lambda e$ and $y = y' + \mu e$. Then,

$$\tau_1(x, y) = \|x\|^2 + \|y\|^2 - 2\|g(x, y)\| \quad \text{and} \quad \varphi_1(x, y) = \sqrt{\tau_1(x, y)}\; \varphi_3(x, y). \tag{20}$$

By Example 3.1, it is not hard to calculate that for any $(v, w) \in \mathcal{H} \times \mathcal{H}$,

$$\begin{aligned}
f'(x, y)(v, w) &= \langle v, e \rangle + r\langle w, e \rangle; \\
g'(x, y)(v, w) &= \lambda v + \langle v, e \rangle (x' - \lambda e) + \mu w + \langle w, e \rangle (y' - \mu e); \\
p'(x, y)(v, w) &= \frac{g'(x, y)(v, w)}{\|g(x, y)\|} - \frac{\langle g'(x, y)(v, w), g(x, y) \rangle}{\|g(x, y)\|^3} g(x, y), \\
h'(x, y)(v, w) &= v - \langle v, e \rangle e + r w - r\langle w, e \rangle e.
\end{aligned}$$

Note that $\|g(x, y)\| \geq \rho$ for all $(x, y) \in U$. By the boundedness of U, there is a constant $c > 0$ such that $\|x\| + \|y\| \leq c$ for all $(x, y) \in U$. Thus, for any $(x, y) \in U$ and any $(v, w) \in \mathcal{H} \times \mathcal{H}$, from the last four equalities it follows that

$$\begin{aligned} &\|f'(x,y)(v,w)\| \leq (|r|+1)(\|v\|+\|w\|), \quad \|h'(x,y)(v,w)\| \leq (|r|+1)(\|v\|+\|w\|), \\ &\|g'(x,y)(v,w)\| \leq 2c(\|v\|+\|w\|), \quad \|p'(x,y)(v,w)\| \leq \frac{4c(\|v\|+\|w\|)}{\rho}. \end{aligned}$$

Consequently,

$$\begin{aligned} \|\tau_1'(x,y)(v,w)\| &= \left\| 2\langle x, v\rangle + 2\langle y, w\rangle - \frac{2\langle g'(x,y)(v,w), g(x,\,y)\rangle}{\|g(x,y)\|} \right\|, \\ &\leq 2c(\|v\|+\|w\|) + 2\|g'(x,y)(v,w)\| \leq 6c(\|v\|+\|w\|), \\ |\varphi_3'(x,y)(v,w)| &\leq \|f'(x,y)(v,w)\| + \|p'(x,y)(v,w)\| \cdot \|h(x,y)\| + \|h'(x,y)(v,w)\| \\ &\leq M_1(\|v\|+\|w\|) \end{aligned}$$

where $M_1 = 2(|r|+1) + 4\rho^{-1}c^2(|r|+1)$. By the mean-value theorem, for any given $(x, y) \in U$, there exists $(\bar{x}, \bar{y}) \in U$ on the line segment joining (a, b) to (x, y) such that

$$|\varphi_3(x,y) - \varphi_3(a,b)| = |\varphi_3'(\bar{x},\bar{y})(x-a, y-b)| \leq M_1(\|x-a\| + \|y-b\|).$$

We claim that $\varphi_3(a, b) = 0$. To see this, from Lemma 3.2, $|\alpha| = \|a'\|$, $|\beta| = \|b'\|$ and $\alpha\beta = \langle a', b'\rangle$, which implies that $\|\alpha a' + \beta b'\| = \alpha^2 + \beta^2$, and

$$\begin{aligned} \varphi_3(a,b) &= \alpha + r\beta - \frac{1}{\alpha^2+\beta^2}(\alpha\|a'\|^2 + (r\alpha+\beta)\langle a', b'\rangle + r\beta\|b'\|^2) \\ &= \alpha + r\beta - \frac{1}{\alpha^2+\beta^2}(\alpha^3 + r\alpha^2\beta + \alpha\beta^2 + r\beta^3) = 0. \end{aligned}$$

This claim implies that

$$|\varphi_3(x,\,y)| \leq M_1(\|x-a\| + \|y-b\|) \quad \text{for any } (x,\,y) \in U. \tag{21}$$

In addition, noting that $\tau_1(a, b) = 0$ and applying the Mean Value Theorem to τ_1,

$$\sqrt{\tau_1(x,y)} \leq M_2 \cdot \sqrt{\|x-a\| + \|y-b\|} \quad \text{for any } (x,y) \in U, \tag{22}$$

where $M_2 = \sqrt{6c}$. Now from equations (20)–(22) it follows that, for any $(x, y) \in U$,

$$|\varphi_1(x,y) - \varphi_2(a,b)| = |\varphi_1(x,y)| \leq M_1 M_2(\|x-a\| + \|y-b\|)^{3/2},$$

which says that φ_1 is F-differentiable at (a, b) with $\varphi_1'(a, b)$ being the zero mapping in $\mathcal{L}(\mathcal{H} \times \mathcal{H}, \mathbb{R})$. So, ψ_r is F-differentiable at (a, b) with $\psi_r'(a, b) = \varphi_2'(a, b)$.

(b) From part (a), we know that $\psi_r'(x, y) = \varphi_2'(x, y)$. To compute $\varphi_2'(x, y)$, we write

$$\varphi_4(x,y) := f(x,y) + \langle p(x,y), h(x,y)\rangle = \lambda + r\mu + \left\langle \frac{\lambda x' + \mu y'}{\|\lambda x' + \mu y'\|}, x' + ry' \right\rangle.$$

From the expression of $\varphi_2(x,y)$, it follows that $\varphi_2(x,y) = \sqrt{\tau_2(x,y)} \cdot \varphi_4(x,y)$. Hence,

$$\varphi_2'(x,y)(v,w) = \frac{\tau_2'(x,y)(v,w)}{2\sqrt{\tau_2(x,y)}} \cdot \varphi_4(x,y) + \sqrt{\tau_2(x,y)} \cdot \varphi_4'(x,y)(v,w) \tag{23}$$

for any $v, w \in \mathcal{H}$. By the expressions of $\varphi_4(x,y)$ and $\tau_2(x,y)$,

$$\begin{aligned} \varphi_4'(x,y)(v,w) &= f'(x,y)(v,w) + \langle p'(x,y)(v,w), h(x,y)\rangle + \langle p(x,y), h'(x,y)(v,w)\rangle, \\ \tau_2'(x,y)(v,w) &= 2\langle v,x\rangle + 2\langle w,y\rangle + 2\frac{\langle g'(x,y)(v,w), g(x,y)\rangle}{\|g(x,y)\|}. \end{aligned} \tag{24}$$

Since $|\lambda| = \|x'\|$, $|\mu| = \|y'\|$, $\|\lambda x' + \mu y'\| = \lambda^2 + \mu^2$, $\langle x', y'\rangle = \lambda\mu$ by Lemma 3.2, we have

$$g(x,y) = \lambda^2 + \mu^2, \quad \sqrt{\tau_2(x,y)} = 2\sqrt{\lambda^2+\mu^2} \quad \text{and} \quad \varphi_4(x,y) = 2(\lambda + r\mu).$$

Using these equalities and $\mu x' = \lambda y'$, it is not hard to calculate that for any v, $w \in \mathcal{H}$,

$$\begin{aligned} f'(x,y)(v,w) &= \langle v,e\rangle + r\langle w,e\rangle = \langle v + rw, e\rangle; \\ \langle p'(x,y)(v,w), h(x,y)\rangle &= \frac{\langle g'(x,y)(v,w), h(x,y)\rangle}{\|g(x,y)\|} \\ &\quad - \frac{\langle g'(x,y)(v,w), g(x,y)\rangle}{\|g(x,y)\|^3}\langle g(x,y), h(x,y)\rangle = 0; \\ \langle p(x,y), h'(x,y)(v,w)\rangle &= \left\langle \frac{g(x,y)}{\|g(x,y)\|}, v - \langle v,e\rangle e + rw - r\langle w,e\rangle e \right\rangle \\ &= (\lambda^2+\mu^2)^{-1}\langle \lambda x' + \mu y', v + rw\rangle; \\ \langle g'(x,y)(v,w), g(x,y)\rangle &= \langle \lambda v, \lambda x' + \mu y'\rangle + \langle v,e\rangle \cdot \langle x' - \lambda e, \lambda x' + \mu y'\rangle \\ &\quad + \langle \mu w, \lambda x' + \mu y'\rangle + \langle w,e\rangle \cdot \langle y' - \mu e, \lambda x' + \mu y'\rangle \\ &= \langle v, \lambda^2 x' + \lambda\mu y'\rangle + \lambda\langle v,e\rangle(\lambda^2+\mu^2) \\ &\quad + \langle w, \lambda\mu x' + \mu^2 y'\rangle + \mu\langle w,e\rangle(\lambda^2+\mu^2) \\ &= (\lambda^2+\mu^2)\langle v,x'\rangle + \lambda\langle v,e\rangle(\lambda^2+\mu^2) \\ &\quad + (\lambda^2+\mu^2)\langle w,y'\rangle + \mu\langle w,e\rangle(\lambda^2+\mu^2) \\ &= (\lambda^2+\mu^2)(\langle v,x\rangle + \langle w,y\rangle). \end{aligned}$$

Combining the last three equations with equation (24), it follows that

$$\begin{aligned} \tau_2'(x,y)(v,w) &= 4\langle v,x\rangle + 4\langle w,y\rangle; \\ \sqrt{\tau_2(x,y)}\varphi_4'(x,y)(v,w) &= 2\sqrt{\lambda^2+\mu^2}\left\{\langle v + rw, e\rangle + \frac{\langle \lambda x' + \mu y', v + rw\rangle}{\lambda^2+\mu^2}\right\} \\ &= 2\sqrt{\lambda^2+\mu^2}\left\langle \frac{\lambda x' + \mu y'}{\lambda^2+\mu^2} + e, v + rw\right\rangle \\ &= 2\left\langle \left(x^2+y^2\right)^{1/2}, v + rw\right\rangle, \end{aligned}$$

where the last equality holds since $\tau(x,y)=\tau_2(x,y)$. This together with (23) yields that

$$\begin{aligned}
\varphi_2'(x,y)(v,w) &= \frac{2(\lambda+r\mu)}{\sqrt{\lambda^2+\mu^2}}\cdot(\langle v,x\rangle+\langle w,y\rangle)+2\left\langle \left(x^2+y^2\right)^{1/2}, v+rw\right\rangle,\\
\|\varphi_2'(x,y)(v,w)\| &\leq 2(1+|r|)(\|v\|\cdot\|x\|+\|w\|\cdot\|y\|)+2\sqrt{\|x\|^2+\|y\|^2}\cdot\|v+rw\|\\
&\leq 4(1+|r|)\sqrt{\|x\|^2+\|y\|^2}\cdot\sqrt{\|v\|^2+\|w\|^2}.
\end{aligned}$$

Together with $\varphi_r'(x,y)=\varphi_2'(x,y)$, we obtain the desired results. $\square$

Next we use Lemmas 4.1 and 4.2 to establish the F-differentiability of Ψ_t, and present the explicit formula for the differential of Ψ_t. Note that, for any given point $(x,y)\in\mathcal{H}\times\mathcal{H}$, the differential $\Psi_t'(x,y)$ induces two continuous linear mappings in $\mathcal{L}(\mathcal{H},\mathbb{R})$, which are $v\longmapsto\Psi_t'(x,y)(v,0)$ and $w\longmapsto\Psi_t'(x,y)(0,w)$ for $v,w\in\mathcal{H}$, called the partial derivatives of Ψ_t at (x,y) w.r.t. x and y, respectively. It is well known that for any given $l\in\mathcal{L}(\mathcal{H},\mathbb{R})$ there is a unique point $a\in\mathcal{H}$ such that $l(v)=\langle a,v\rangle$ for all $v\in\mathcal{H}$. We let $D_1\Psi_t(x,y)\in\mathcal{H}$ and $D_2\Psi_t(x,y)\in\mathcal{H}$ be such that

$$\Psi_t'(x,y)(v,0)=\langle D_1\Psi_t(x,y),v\rangle \quad\text{and}\quad \Psi_t'(x,y)(0,w)=\langle D_2\Psi_t(x,y),w\rangle$$

for all $v,w\in\mathcal{H}$. By identifying $D_1\Psi_t(x,y)$ with the mapping $v\longmapsto\Psi_t'(x,y)(v,0)$, we shall call $D_1\Psi_t(x,y)$ the partial derivative of Ψ_t at (x,y) w.r.t. x. Similarly, $D_2\Psi_t(x,y)$ is called the partial derivative of Ψ_t at (x,y) w.r.t. y.

Theorem 4.1 *The function Ψ_t with $0<t<2$ is F-differentiable on $\mathcal{H}\times\mathcal{H}$. Also,*

(a) *If $(x,y)=(0,0)$, then $D_1\Psi_t(x,y)=D_2\Psi_t(x,y)=0\in\mathcal{H}$.*

(b) *If $x=x'+\lambda e\in\mathcal{H}$ and $y=y'+\mu e\in\mathcal{H}$ with $x^2+y^2\in\partial\mathbb{K}\setminus\{0\}$, then*

$$\begin{aligned}
D_1\Psi_t(x,y) &= 2\left(\frac{\lambda+(t-1)\mu}{\tau}-1\right)\Phi_t(x,y)\\
D_2\Psi_t(x,y) &= 2\left(\frac{(t-1)\lambda+\mu}{\tau}-1\right)\Phi_t(x,y)
\end{aligned}\tag{25}$$

with $\tau=\sqrt{(\lambda-\mu)^2+2t\lambda\mu}$.

(c) *If $(x,y)\in\mathcal{H}\times\mathcal{H}$ with $x^2+y^2\in\mathrm{int}\mathbb{K}$, then*

$$\begin{aligned}
D_1\Psi_t(x,y) &= 2\left[(x+(t-1)y)\bullet L_z^{-1}\Phi_t(x,y)-\Phi_t(x,y)\right]\\
D_2\Psi_t(x,y) &= 2\left[((t-1)x+y)\bullet L_z^{-1}\Phi_t(x,y)-\Phi_t(x,y)\right]
\end{aligned}\tag{26}$$

with $z=\left[(x-y)^2+2t(x\bullet y)\right]^{1/2}$ and L_z^{-1} defined as in equation (12).

Proof. For $0 < t < 2$, we consider the mapping $S_t : \mathcal{H} \times \mathcal{H} \to \mathcal{H} \times \mathcal{H}$ defined by

$$S_t(x, y) := \left(x + (t-1)y, \sqrt{t(2-t)}y\right).$$

Note that $S_t \in \mathrm{GL}(\mathcal{H} \times \mathcal{H})$ with the inverse given by

$$S_t^{-1}(x, y) = \left(x + \frac{1-t}{\sqrt{t(2-t)}}y, \frac{y}{\sqrt{t(2-t)}}\right) \quad \text{for any } x,\, y \in \mathcal{H},$$

and

$$\Phi_t \circ S_t^{-1}(x, y) = \left(x^2 + y^2\right)^{1/2} - \left(x + \sqrt{2t^{-1} - 1}\; y\right). \tag{27}$$

Therefore, every Φ_t is continuously F-differentiable on the open set

$$\begin{aligned} \Omega &:= \left\{(x, y) \in \mathcal{H} \times \mathcal{H} \mid (x-y)^2 + 2t(x \bullet y) \in \mathrm{int}\mathbb{K}\right\} \\ &= \left\{(x, y) \in \mathcal{H} \times \mathcal{H} \mid x^2 + y^2 \in \mathrm{int}\mathbb{K}\right\}, \end{aligned} \tag{28}$$

if the mapping $(x, y) \mapsto (x^2 + y^2)^{1/2}$ is continuously F-differentiable on Ω, where the second equality in (28) is due to Lemma 3.2. Since the mapping $x \longmapsto x^2$ is continuously F-differentiable on $\mathcal{H}$, the mapping $(x, y) \longmapsto x^2 + y^2$ is continuously F-differentiable on $\mathcal{H} \times \mathcal{H}$. By Lemma 4.1 (a), $(x, y) \mapsto (x^2 + y^2)^{1/2}$ is then continuously F-differentiable on Ω. It remains to show that Ψ_t is F-differentiable on

$$\begin{aligned} \partial\Omega &= \left\{(x, y) \in \mathcal{H} \times \mathcal{H} \mid (x-y)^2 + 2t(x \bullet y) \in \partial\mathbb{K}\right\} \\ &= \left\{(x, y) \in \mathcal{H} \times \mathcal{H} \mid x^2 + y^2 \in \partial\mathbb{K}\right\}. \end{aligned} \tag{29}$$

From equation (27), for any $x,\, y \in \mathcal{H}$,

$$\begin{aligned} \Psi_t \circ S_t^{-1}(x, y) &= \left\|\Phi_t \circ S_t^{-1}(x, y)\right\|^2 = \left\|(x^2 + y^2)^{1/2} - (x + \sqrt{2t^{-1} - 1}\; y)\right\|^2, \\ &= \|x\|^2 + \|y\|^2 - \psi_r(x, y) + \|x + ry\|^2 \end{aligned}$$

with $r = \sqrt{2t^{-1} - 1}$. Notice that the mapping $x \mapsto \|x\|^2$ is continuously differentiable, whereas by Lemma 4.2 (a) the mapping $\psi_r(x, y)$ is F-differentiable at every point of $\partial\Omega$. Therefore, Ψ_t is F-differentiable on the set $\partial\Omega$.

(a) By Lemma 4.2, $\Psi_t'(0, 0)$ is the zero mapping in $\mathcal{L}(\mathcal{H} \times \mathcal{H}, \mathbb{R})$, which implies that

$$D_1\Psi_t(0, 0) \;=\; D_2\Psi_t(0, 0) = 0 \in \mathcal{H}.$$

We next compute the partial derivatives of Ψ_t at points in $\mathcal{H} \times \mathcal{H} \setminus \{(0, 0)\}$. From the definition of Ψ_t, it follows that for any $x, y \in \mathcal{H}$,

$$\Psi_t(x, y) = \|x + (t-1)y\|^2 + t(2-t)\|y\|^2 - \psi_r \circ S_t(x, y) + \|x + y\|^2$$

with $r = \sqrt{2t^{-1} - 1}$. Therefore, for all $v, w \in \mathcal{H}$, we have

$$\begin{aligned} \Psi_t'(x,y)(v,w) &= 2\langle x + (t-1)y, v + (t-1)w\rangle + 2t(2-t)\langle y, w\rangle \\ &\quad -\psi_r'(S_t(x,y)) \circ S_t(v,w) + 2\langle x + y, v + w\rangle, \end{aligned}$$

which in turn implies that

$$\Psi_t'(x,y)(v,0) = 2\langle 2x + ty, v\rangle - \psi_r'(S_t(x,y)) \circ S_t(v,0), \tag{30}$$

$$\Psi_t'(x,y)(0,w) = 2\langle tx + 2y, w\rangle - \psi_r'(S_t(x,y)) \circ S_t(0,w). \tag{31}$$

(b) Let $x = x' + \lambda e \in \mathcal{H}$ and $y = y' + \mu e \in \mathcal{H}$ with $x^2 + y^2 \in \partial\mathbb{K} \setminus \{0\}$, and write

$$a = x + (t-1)y,\ b = \sqrt{t(2-t)}y,\ c = (t-1)x + y,\ \text{and } z = (a^2 + b^2)^{1/2}.$$

Note that $z \in \partial\mathbb{K} \setminus \{0\}$ and $\Phi_t(x,y) = z - (x+y)$. It follows from Lemma 4.2 (b) that

$$\psi_r'(a,b)(v,w) = \frac{2(\alpha + r\beta)}{\sqrt{\alpha^2 + \beta^2}} \cdot (\langle a, v\rangle + \langle b, w\rangle) + 2\langle z, v + rw\rangle$$

for all $v, w \in \mathcal{H}$, where $\alpha = \langle a, e\rangle = \lambda + (t-1)\mu$ and $\beta = \langle b, e\rangle = \sqrt{t(2-t)}\mu$. Since

$$\alpha + r\beta = \lambda + \mu,\ \alpha^2 + \beta^2 = \tau^2 \text{ and } v + (t-1)w + r\sqrt{t(2-t)}w = v + w,$$

we have

$$\begin{aligned} \psi_r'(S_t(x,y)) \circ S_t(v,w) &= \psi_r'(a,b)(v + (t-1)w, \sqrt{t(2-t)}w) \\ &= \frac{2(\lambda + \mu)}{\tau} \cdot (\langle a, v\rangle + \langle c, w\rangle) + 2\langle z, v + w\rangle. \end{aligned}$$

This means that

$$\begin{aligned} \psi_r'(S_t(x,y)) \circ S_t(v,0) &= \frac{2(\lambda + \mu)}{\tau} \cdot \langle a, v\rangle + 2\langle z, v\rangle, \\ \psi_r'(S_t(x,y)) \circ S_t(0,w) &= \frac{2(\lambda + \mu)}{\tau} \cdot \langle c, w\rangle + 2\langle z, w\rangle. \end{aligned}$$

Using equations (30) and (31), it then follows that

$$\begin{aligned} \frac{1}{2}D_1\Psi_t(x,y) &= 2x + ty - \frac{\lambda + \mu}{\tau} \cdot a - z = \left(1 - \frac{\lambda + \mu}{\tau}\right) a - \Phi_t(x,y) \\ \frac{1}{2}D_2\Psi_t(x,y) &= tx + 2y - \frac{\lambda + \mu}{\tau} \cdot c - z = \left(1 - \frac{\lambda + \mu}{\tau}\right) c - \Phi_t(x,y). \end{aligned}$$

Now to obtain the two equalities in (25), it suffices to prove that

$$\begin{aligned} \frac{\lambda + (t-1)\mu}{\tau}\ \Phi_t(x,y) &= \left(1 - \frac{\lambda + \mu}{\tau}\right) a, \\ \frac{(t-1)\lambda + \mu}{\tau}\ \Phi_t(x,y) &= \left(1 - \frac{\lambda + \mu}{\tau}\right) c. \end{aligned} \tag{32}$$

Since $\alpha = \langle a, e\rangle$ and $\tau = \langle z, e\rangle$, we write $a = a' + \alpha e$ and $z = z' + \tau e$ with

$$a' = x' + (t-1)y' \quad \text{and} \quad \zeta = \frac{(\lambda + (t-1)\mu)x' + (\lambda(t-1)+\mu)y'}{\tau}.$$

Noting that $\Phi_t(x, y) = z - (x + y)$, we readily have

$$\frac{\lambda + (t-1)\mu}{\tau}\langle \Phi_t(x, y), e\rangle = \frac{\alpha}{\tau}(\tau - \lambda - \mu) = \left(1 - \frac{\lambda+\mu}{\tau}\right)\alpha.$$

On the other hand, since $\lambda y' = \mu x'$ by Lemma 3.2, it is not hard to verify that

$$\begin{aligned}\lambda a' &= \alpha x', \quad \mu a' = \alpha y', \quad \frac{\alpha}{\tau}(x' + y') = \frac{\lambda+\mu}{\tau}a',\\ \frac{\alpha}{\tau}z' &= \frac{\lambda + (t-1)\mu}{\tau}\cdot\frac{\lambda a'}{\tau} + \frac{\lambda(t-1)+\mu}{\tau}\cdot\frac{\mu a'}{\tau} = a',\end{aligned}$$

and consequently

$$\frac{\alpha}{\tau}(z' - x' - y') = \left(1 - \frac{\lambda+\mu}{\tau}\right)a'.$$

The two sides show that the first equality in (32) holds. Using the similar arguments, we can prove that the second equality in (32) also holds.

(c) Write $q(x, y) = x^2 + y^2$. Then, by the definition of $\sigma(x)$ given by Lemma 4.1,

$$\psi_r(x, y) = \langle \sigma(q(x, y)), x + y\rangle.$$

Applying the Chain Rule of differential, we have for any $v, w \in \mathcal{H}$

$$\psi_r'(S_t(x, y)) \circ S_t(v, w) = 2\langle \sigma'(z^2) \circ q'(S_t(x, y)) \circ S_t(v, w), x + y\rangle + 2\langle z, v + w\rangle,$$

and consequently,

$$\begin{aligned}\psi_r'(S_t(x, y)) \circ S_t(v, 0) &= 2\langle \sigma'(z^2) \circ q'(S_t(x, y))(v, 0),\ x + y\rangle + 2\langle z, v\rangle\\ \psi_r'(S_t(x, y)) \circ S_t(0, w) &= 2\langle \sigma'(z^2) \circ q'(S_t(x, y))((t-1)w, \sqrt{t(2-t)}w),\ x + y\rangle\\ &\quad +2\langle z, w\rangle.\end{aligned}$$

Noting that $q'(x, y)(v, w) = 2(x \bullet v) + 2(y \bullet w)$ and $\sigma'(z^2)$ is self-adjoint, we have

$$\langle \sigma'(z^2) \circ q'(S_t(x, y))(v, 0),\ x + y\rangle = 2\langle a \bullet v, \sigma'(z^2)(x + y)\rangle = 2\langle v, a \bullet \sigma'(z^2)(x + y)\rangle$$

and

$$\begin{aligned}&\left\langle \sigma'(z^2) \circ q'(S_t(x, y))((t-1)w, \sqrt{t(2-t)}w), x + y\right\rangle\\ =\ & 2(t-1)\left\langle w, b \bullet \sigma'(z^2)(x + y)\right\rangle + 2t(2-t)\left\langle w,\ b \bullet \sigma'(z^2)(x + y)\right\rangle\\ =\ & 2\left\langle w, b \bullet \sigma'(z^2)(x + y)\right\rangle.\end{aligned}$$

Together with equations (30)–(31), it then follows that

$$\begin{aligned}
\frac{1}{2}D_1\Psi_t(x,y) &= 2x+ty-2a\bullet\sigma'(z^2)(x+y)-z,\\
\frac{1}{2}D_2\Psi_t(x,y) &= tx+2y-2b\bullet\sigma'(z^2)(x+y)-z.
\end{aligned}$$

From Lemma 4.1, we have $2\sigma'(z^2)v=L_z^{-1}v$ for all $v\in\mathcal{H}$. Therefore,

$$\begin{aligned}
\frac{1}{2}D_1\Psi_t(x,y) &= (x+(t-1)y)-\Phi_t(x,y)-a\bullet L_z^{-1}(x+y)\\
&= a\bullet L_z^{-1}z-\Phi_t(x,y)-a\bullet L_z^{-1}(x+y)\\
&= a\bullet L_z^{-1}\Phi_t(x,y)-\Phi_t(x,y),
\end{aligned}$$

where the second equality is using $L_z^{-1}z=e$. This proves the first equality in (26). Similar, we can obtain the second equality in (26). The proof is complete. $\Box$

In what follows, we investigate the continuity of the differential Ψ_t'. From the proof of Theorem 4.1, we see that, to establish the continuity of Ψ_t' in $\mathcal{H}\times\mathcal{H}$, it suffices to show that the differential of ψ_r is continuous at every point $(a,b)\in\partial\Omega$. The following proposition shows that ψ_r' is continuous at $\partial\Omega\backslash\{(0,0)\}$.

Proposition 4.1 *Let ψ_r be defined as in Lemma 4.2. Then its F-differential is continuous at every point $(a,b)\in\partial\Omega\backslash\{(0,0)\}$.*

Proof. We shall use notations given in the proof of Lemma 4.2. Recall that

$$\psi_r(x,y)=\varphi_1(x,y)+\varphi_2(x,y)\quad \text{for any } (x,\,y)\in U,$$

and φ_2 is continuously F-differentiable on U, particularly at (a,b). To prove the continuity of F-differential of φ_1 at (a,b), we recall that $\varphi_1'(x,y)=0\in\mathcal{L}(\mathcal{H}\times\mathcal{H},\mathbb{R})$ whenever $(x,y)\in U$ with $x^2+y^2\in\partial\mathbb{K}$. Hence, for any $(x,y)\in U$ with $x^2+y^2\in\mathrm{int}\mathbb{K}$,

$$\varphi_1'(x,y)(v,w)-\varphi_1'(a,b)(v,w)=\frac{2\tau_1'(x,y)(v,w)}{\sqrt{\tau_1(x,y)}}\varphi_3(x,y)+\sqrt{\tau_1(x,y)}\varphi_3'(x,y)(v,w).$$

From the proof of Lemma 4.2, we know that for any $(x,y)\in U$ and any $v,w\in\mathcal{H}$,

$$\|\tau_1'(x,y)(v,w)\|\le 6c(\|v\|+\|w\|)\quad\text{and}\quad \|\varphi_3'(x,y)(v,w)\|\le M_1(\|v\|+\|w\|),$$

where $M_1>0$ and $c>0$ are constants. In addition, from the expression of $\varphi_3(x,y)$ and Lemma 3.3, it follows that for any $(x,y)\in U$ with $x^2+y^2\in\mathrm{int}\mathbb{K}$,

$$\begin{aligned}
|\varphi_3(x,y)| &= \left|\lambda+\gamma\mu-\left\langle\frac{\lambda x'+\mu y'}{\|\lambda x'+\mu y'\|},x'+ry'\right\rangle\right|\\
&\le \left|\lambda-\left\langle\frac{\lambda x'+\mu y'}{\|\lambda x'+\mu y'\|},x'\right\rangle\right|+|r|\left|\mu-\left\langle\frac{\lambda x'+\mu y'}{\|\lambda x'+\mu y'\|},y'\right\rangle\right|\\
&\le (1+|r|)\cdot\tau_1(x,y).
\end{aligned}$$

Now, for any $(x,y)\in U$ with $x^2+y^2\in \mathrm{int}\mathbb{K}$ and for any $v,w\in\mathcal{H}$, we have

$$|\varphi_1'(x,y)(v,w)-\varphi_1'(a,b)(v,w)|\le \big[12c(1+|r|)+M_1\big]\sqrt{\tau_1(x,y)}\cdot(\|v\|+\|w\|),$$

which in turn implies that

$$\|\varphi_1'(x,y)-\varphi_1'(a,b)\|\le 2\big[12c(1+|r|)+M_1\big]\sqrt{\tau_1(x,y)}\ \text{ for any } (x,y)\in U.$$

Since $\tau_1(a,b)=0$, we have $\|\varphi_1'(x,y)-\varphi_1'(a,b)\|\to 0$ as $(x,y)\to(a,b)$. $\square$

To prove the continuity of the differential ψ_r' at $(0,0)$, we need the following lemma which establishes the boundedness of the differentials of $(x,y)\longmapsto (x^2+y^2)^{1/2}$ on Ω.

Lemma 4.3 *Let $\sigma:\mathbb{K}\to\mathbb{K}$ be given as in Lemma 4.1 and $q(x,y):=x^2+y^2$ for any $x,y\in\mathcal{H}$. Then $\hat\sigma=\sigma\circ q$ is continuously F-differentiable on Ω, and moreover, there is a constant $C_1>0$ such that $\|\hat\sigma'(x,y)\|\le C_1$ for all $(x,y)\in\Omega$.*

Proof. Since σ is continuously F-differentiable on $\mathrm{int}\mathbb{K}$, and q is continuously F-differentiable on $\mathcal{H}\times\mathcal{H}$, it follows that $\hat\sigma$ is continuously F-differentiable on Ω. In the following, we prove that the F-differential of $\hat\sigma$ is bounded on Ω. For any point $(x,y)\in\Omega$, we write $x=x'+\lambda e$ and $y=y'+\mu e$. Then, for all $v,w\in\mathcal{H}$,

$$q'(x,y)(v,w)=2\big[\langle v,\,e\rangle x'+\lambda v+\langle v,\,x'\rangle e+\langle w,\,e\rangle y'+\mu w+\langle w,\,y'\rangle e\big].$$

Now, applying Lemma 4.1 (a) yields that

$$\begin{aligned}\hat\sigma'(x,y)(v,w)&=\sigma'(x^2+y^2)\circ q'(x,y)(v,w)\\&=\frac{\sqrt{\tau_1\tau_2}}{2\tau}\cdot\Big\langle\big(x^2+y^2\big)^{-1/2},q'(x,y)(v,w)\Big\rangle\big(x^2+y^2\big)^{-1/2}\\&\quad+\frac{1}{2\tau}\cdot\big[q'(x,y)(v,w)-\langle q'(x,y)(v,w),e\rangle e\big],\end{aligned}\tag{33}$$

where $\tau=\tau(x,y)$ and $\tau_j=\tau_j(x,y)$ for $j=1,2$ with $\tau(x,y)$ and $\tau_j(x,y)$ given as in the proof of Lemma 4.2. Using a direct computation and noting that $\sqrt{\|x\|^2+\|y\|^2}\le\sqrt{2}\tau$,

$$\begin{aligned}\frac{1}{2\tau}\|q'(x,y)(v,w)-\langle q'(x,y)(v,w),e\rangle e\|&\le\frac{2}{\tau}\sqrt{\|x\|^2+\|y\|^2}\cdot\sqrt{\|v\|^2+\|w\|^2}\\&\le 2\sqrt{2}\sqrt{\|v\|^2+\|w\|^2}.\end{aligned}\tag{34}$$

By writing $z=\frac{\lambda x'+\mu y'}{\|\lambda x'+\mu y'\|}$ and using equations (10) and (19), it follows that

$$\begin{aligned}\big(x^2+y^2\big)^{-1/2}&=\frac{1}{\sqrt{\tau_1\tau_2}}\left(\frac{-\lambda x'-\mu y'}{\tau}+\tau e\right)\\&=\frac{1}{\sqrt{\tau_1\tau_2}}\left(\frac{\sqrt{\tau_1}-\sqrt{\tau_2}}{2}z+\frac{\sqrt{\tau_2}+\sqrt{\tau_1}}{2}e\right).\end{aligned}$$

This together with the expression of $q'(x,y)(v,w)$ implies that

$$\begin{aligned}
&\left\langle \left(x^2+y^2\right)^{-1/2}, q'(x,y)(v,w)\right\rangle \\
= \;& (\tau_1\tau_2)^{-1/2}\left(\sqrt{\tau_1}-\sqrt{\tau_2}\right)\left[\langle v,e\rangle\langle z,x'\rangle+\lambda\langle z,\,v\rangle+\langle w,\,e\rangle\,\langle z,\,y'\rangle+\mu\langle z,\,w\rangle\right] \\
&+(\tau_1\tau_2)^{-1/2}(\sqrt{\tau_2}+\sqrt{\tau_1})\left[\lambda\langle v,\,e\rangle+\langle v,\,x'\rangle+\mu\langle w,\,e\rangle+\langle w,\,y'\rangle\right] \\
= \;& \tau_2^{-1/2}\big[(\lambda+\langle z,x'\rangle)\langle v,e\rangle+\langle x'+\lambda z,v\rangle+(\mu+\langle z,y'\rangle)\langle w,e\rangle+\langle \mu z+y',w\rangle\big] \\
&+\tau_1^{-1/2}\big[(\lambda-\langle z,x'\rangle)\langle v,e\rangle+\langle x'-\lambda z,v\rangle+(\mu-\langle z,y'\rangle)\langle w,e\rangle+\langle y'-\mu z,w\rangle\big].
\end{aligned}$$

Noting that

$$\begin{aligned}
\|x'+\lambda z\|+\|y'+\mu z\| \;&\le\; \sqrt{2}(\|x\|+\|y\|)\;\le\; 2\sqrt{\tau_2}, \\
|\lambda+\langle z,x'\rangle|+|\mu+\langle z,y'\rangle| \;&\le\; \sqrt{2}(\|x\|+\|y\|)\;\le 2\sqrt{\tau_2},
\end{aligned}$$

it follows that

$$\frac{1}{\sqrt{\tau_2}}\|(\lambda+\langle z,x'\rangle)\langle v,e\rangle+\langle x'+\lambda z,v\rangle+(\mu+\langle z,y'\rangle)\langle w,e\rangle+\langle \mu z+y',w\rangle\|\le 4(\|v\|+\|w\|).$$

On the other hand, applying Lemma 3.3, we have

$$\frac{1}{\sqrt{\tau_1}}\|(\lambda-\langle z,x'\rangle)\langle v,e\rangle+\langle x'-\lambda z,v\rangle+(\mu-\langle z,y'\rangle)\langle w,e\rangle+\langle y'-\mu z,w\rangle\|\le 2(\|v\|+\|w\|).$$

Therefore,

$$\left|\langle (x^2+y^2)^{-1/2},\, q'(x,\,y)(v,\,w)\rangle\right|\le 6(\|v\|+\|w\|)\le 12\sqrt{\|v\|^2+\|w\|^2}. \tag{35}$$

In addition, since

$$\left\|\left(x^2+y^2\right)^{-1/2}\right\|^2=\langle (x^2+y^2)^{-1},\,e\rangle=\frac{\|x\|^2+\|y\|^2}{\tau_1\tau_2}\le\frac{2\tau^2}{\tau_1\tau_2},$$

we also have

$$\left\|\frac{\sqrt{\tau_1\tau_2}}{2\tau}\left(x^2+y^2\right)^{-1//2}\right\|\le\frac{1}{\sqrt{2}}. \tag{36}$$

Now combining the inequalities (34)–(36) with equation (33) leads to

$$\|\hat{\sigma}'(x,y)(v,w)\|\le 8\sqrt{2}\sqrt{\|v\|^2+\|w\|^2}$$

for all $(x,y)\in\Omega$ and $v,w\in\mathcal{H}$. Therefore, $\|\hat{\sigma}'(x,y)\|\le 8\sqrt{2}$ for all $(x,y)\in\Omega$. $\square$

Proposition 4.2 *Let ψ_r be the mapping defined as in Lemma 4.2. Then, there is a constant $C>0$, independent of r, such that*

$$\|\psi_r'(x,\,y)\|\le C(1+|r|)\sqrt{\|x\|^2+\|y\|^2}\quad \textit{for all } x,\,y\in\mathcal{H}.$$

Consequently, the F-differential of ψ_r is continuous at $(0,\,0)\in\mathcal{H}\times\mathcal{H}$.

Proof. From Lemma 4.2(b), we have $\psi_r'(0,0)=0$. So, it suffices to prove the inequality given in the theorem for $(x,y)\in\mathcal{H}\times\mathcal{H}\setminus\{(0,0)\}$. Let $\hat{\sigma}$ be given as in Lemma 4.3. Then, from the definition of ψ_r, it follows that for any $(x,y)\in\Omega$ and for $v,w\in\mathcal{H}$,

$$\psi_r'(x,y)(v,w)=2\langle\hat{\sigma}'(x,y)(v,w),x+ry\rangle+2\langle\hat{\sigma}(x,y),v+rw\rangle.$$

By Lemma 4.3, there is a constant $C_1>0$ such that for any $(x,y)\in\Omega$ and $v,w\in\mathcal{H}$

$$2|\langle\hat{\sigma}'(x,y)(v,w),x+ry\rangle|\le 2C_1(1+|r|)\sqrt{\|v\|^2+\|w\|^2}\sqrt{\|x\|^2+\|y\|^2}.$$

In addition, from the definition of $\hat{\sigma}(x,y)$, we also have

$$2|\langle\widehat{\sigma}(x,y),v+rw\rangle|\le 2(1+|r|)\sqrt{\|x\|^2+\|y\|^2}\cdot\sqrt{\|v\|^2+\|w\|^2}.$$

The last three equations show that, for $(x,\,y)\in\Omega$,

$$\|\psi_r'(x,y)\|\le 2(C_1+1)(1+|r|)\sqrt{\|x\|^2+\|y\|^2}.$$

The inequality together with Lemma 4.1 imply that for all $(x,y)\in\mathcal{H}\times H$,

$$\|\psi_r'(x,y)\|\le C(1+|r|)\sqrt{\|x\|^2+\|y\|^2},$$

where $C=2\max\{2,C_1+1\}$ is independent of r. $\square$

From Theorem 4.1, Prop.4.1 and Prop.4.2, we readily obtain the following result.

Theorem 4.2 *The function Ψ_t with $0<t<2$ is smooth everywhere on $\mathcal{H}\times\mathcal{H}$.*

5 Stationary point conditions

From the previous discussions, we learn that the complementarity problem $\mathrm{CP}(\mathbb{K},T)$ in the Hilbert space $\mathcal{H}$ with the continuously F-differentiable mapping $T:\mathcal{H}\to\mathcal{H}$ can be transformed into an unconstrained smooth minimization problem

$$\min_{x\in\mathcal{H}}\ f(x):=\Psi_t(x,T(x))\ \text{ with }\ 0<t<2. \tag{37}$$

However, when applying minimization algorithms for (37), we can only expect to obtain stationary points of (37). Thus, it is natural to ask under what conditions each stationary point of the minimization problem (37) is a solution to the problem $\mathrm{CP}(\mathbb{K},T)$. To achieve this goal, we first establish some favorable properties for the differential $\Psi_t'(x,y)$. This needs the following key lemma which generalizes the result of [9, Prop. 3.4] to $\mathcal{H}$.

Lemma 5.1 *For any $x, y \in \mathcal{H}$ and $z \succ_{\mathbb{K}} 0$, the following implications hold*

$$z^2 \succ_{\mathbb{K}} x^2 + y^2 \implies L_z^2 - L_x^2 - L_y^2 \succ 0, \tag{38}$$
$$z^2 \succ_{\mathbb{K}} x^2 \implies z \succ_{\mathbb{K}} x. \tag{39}$$

Moreover, the implications (38) and (39) remain true when "$\succ$" is replaced by "$\succeq$".

Proof. Similar to [9], we first prove (38) for the case where $z = (x^2 + y^2 + \delta e)^{1/2}$ for some $\delta > 0$. Fix any $x, y \in \mathcal{H}$ and any $\delta > 0$. Let $z = (x^2 + y^2 + \delta e)^{1/2}$. It suffices to prove that for any nonzero vector h in $\mathcal{H}$,

$$\begin{aligned} 0 < \left\langle h, (L_z^2 - L_x^2 - L_y^2)h \right\rangle &= \langle L_z h, L_z h\rangle - \langle L_z x, L_z x\rangle - \langle L_z y, L_z y\rangle \\ &= \|z \bullet h\|^2 - \|x \bullet h\|^2 - \|y \bullet h\|^2. \end{aligned}$$

Let $x = x' + \lambda e$, $y = y' + \mu e$, $z = z' + \nu e$ and $h = h' + \xi e$. We calculate that

$$\begin{aligned} \|z \bullet h\|^2 - \|x \bullet h\|^2 - \|y \bullet h\|^2 &= \xi^2\|z\|^2 + \nu^2\|h'\|^2 + 4\xi\nu\langle z', h'\rangle + \langle z', h'\rangle^2 \\ &\quad - \left[\xi^2\|x\|^2 + \lambda^2\|h'\|^2 + 4\xi\lambda\langle x', h'\rangle + \langle x', h'\rangle^2\right] \\ &\quad - \left[\xi^2\|y\|^2 + \mu^2\|h'\|^2 + 4\xi\mu\langle y', h'\rangle + \langle y', h'\rangle^2\right] \\ &= \xi^2\left[\|z\|^2 - \|x\|^2 - \|y\|^2\right] + \left[\nu^2 - \lambda^2 - \mu^2\right]\|h'\|^2 \\ &\quad + \left(\langle z', h'\rangle^2 - \langle x', h'\rangle^2 - \langle y', h'\rangle^2\right) \\ &\quad + 4\left(\xi\nu\langle z', h'\rangle - \xi\lambda\langle x', h'\rangle - \xi\mu\langle y', h'\rangle\right). \end{aligned} \tag{40}$$

From the expression of $z = (x^2 + y^2 + \delta e)^{1/2}$, it is not hard to obtain that

$$z' = \tau^{-1}(\lambda x' + \mu y') \quad \text{and} \quad \nu = \tau,$$

where

$$\tau = \sqrt{\frac{\|x\|^2 + \|y\|^2 + \delta e + \sqrt{(\|x\|^2 + \|y\|^2 + \delta e)^2 - 4\|\lambda x' + \mu y'\|^2}}{2}}.$$

Substituting the expression of z' above into $\langle z', h'\rangle$ and using $\nu = \tau$ yields that

$$\xi\nu\langle z', h'\rangle - \xi\lambda\langle x', h'\rangle - \xi\mu\langle y', h'\rangle = 0 \tag{41}$$

and

$$\begin{aligned} &[\nu^2 - \lambda^2 - \mu^2]\|h'\|^2 + \langle z', h'\rangle^2 - \langle x', h'\rangle^2 - \langle y', h'\rangle^2 \\ = \; &[\nu^2 - \lambda^2 - \mu^2]\|h'\|^2 + \frac{(\lambda^2 + \mu^2 - \tau^2)(\langle x', h'\rangle^2 + \langle y', h'\rangle^2)}{\tau^2} - \frac{\langle h', \mu x' - \lambda y'\rangle^2}{\tau^2} \\ = \; &\frac{(\tau^2 - \lambda^2 - \mu^2)(\tau^2\|h'\|^2 - \langle x', h'\rangle^2 - \langle y', h'\rangle^2)}{\tau^2} - \frac{\langle h', \mu x' - \lambda y'\rangle^2}{\tau^2}. \end{aligned} \tag{42}$$

Now combining equations (41)–(42) with equation (40) yields that

$$\|z \bullet h\|^2 - \|x \bullet h\|^2 - \|y \bullet h\|^2 = \xi^2 \left[\|z\|^2 - \|x\|^2 - \|y\|^2\right] - \frac{\langle h', \mu x' - \lambda y'\rangle^2}{\tau^2} + \frac{(\tau^2 - \lambda^2 - \mu^2)(\tau^2\|h'\|^2 - \langle x', h'\rangle^2 - \langle y', h'\rangle^2)}{\tau^2}. \tag{43}$$

Notice that $\|z\|^2 = \|x\|^2 + \|y\|^2 + \delta > \|x\|^2 + \|y\|^2$ is equivalent to

$$\frac{\|\lambda x' + \mu y'\|^2}{\tau^2} + \tau^2 > \|x'\|^2 + \lambda^2 + \|y'\|^2 + \mu^2.$$

Multiplying the two sides of the last inequality with τ^2, and then adding $\mu^2\|x'\|^2+\lambda^2\|y'\|^2$ to the both sides of the inequality, we obtain

$$(\lambda^2 + \mu^2)(\|x'\|^2 + \|y'\|^2) + \tau^4 - \tau^2(\|x'\|^2 + \lambda^2 + \|y'\|^2 + \mu^2) > \|\mu x' - \lambda y'\|^2$$

or

$$(\tau^2 - \|x'\|^2 - \|y'\|^2)(\tau^2 - \lambda^2 - \mu^2) > \|\mu x' - \lambda y'\|^2. \tag{44}$$

This means that both $\tau^2 - \|x'\|^2 - \|y'\|^2$ and $\tau^2 - \lambda^2 - \mu^2$ are positive or negative. If both are negative, then we would have $\|x'\|^2 + \|y'\|^2 \geq \tau^2$ and $\lambda^2 + \mu^2 \geq \tau^2$, which leads to a contradiction that $\|x\|^2 + \|y\|^2 \geq 2\tau^2 \geq \|x\|^2 + \|y\|^2 + \delta$. Consequently,

$$\tau^2 - \|x'\|^2 - \|y'\|^2 > 0 \quad \text{and} \quad \tau^2 - \lambda^2 - \mu^2 > 0.$$

Together with (43) and (44) and $\|z\|^2 > \|x\|^2 + \|y\|^2$, it follows that

$$\begin{aligned} & \|z \bullet h\|^2 - \|x \bullet h\|^2 - \|y \bullet h\|^2 \\ \geq \; & \frac{(\tau^2 - \lambda^2 - \mu^2)(\tau^2 - \|x'\|^2 - \|y'\|^2)\|h'\|^2}{\tau^2} - \frac{\|h'\|^2\|\lambda x' - \mu y'\|^2}{\tau^2} > 0. \end{aligned}$$

So, (38) holds for any $x, y \in \mathcal{H}$ and z of the form $z = (x^2 + y^2 + \delta e)^{1/2}$ for some $\delta > 0$. In view of Lemma 2.2, the rest arguments are same as those of [9, Prop. 3.4]. □

Lemma 5.2 *Let Ψ_t be given in (15) with $0 < t < 2$. Then, for any $x, y \in \mathcal{H}$,*

(a) $\langle D_1\Psi_t(x, y), D_2\Psi_t(x, y)\rangle \geq 0$ *with equality holding if and only if* $\Phi_t(x, y) = 0$;

(b) $\langle D_1\Psi_t(x, y), x\rangle + \langle D_2\Psi_t(x, y), y\rangle = 2\Psi_t(x, y)$;

(c) $D_1\Psi_t(x, y) = 0 \Longleftrightarrow D_2\Psi_t(x, y) = 0 \Longleftrightarrow \Psi_t(x, y) = 0$.

Proof. (a) We proceed the arguments by three cases shown as below.

Case (a.1): $(x, y) = (0, 0)$. Since $D_1\Psi_t(0,0) = D_2\Psi_t(0,0) = 0$, the result is true.

Case (a.2): $x^2 + y^2 \in \partial\mathbb{K}\backslash\{0\}$. Let $x = x' + \lambda e$ and $y = y' + \mu e$. By Theorem 4.1,

$$\langle D_1\Psi_t(x,y), D_2\Psi_t(x,y)\rangle = 4\left(\frac{\lambda + (t-1)\mu}{\tau} - 1\right)\left(\frac{\mu + (t-1)\lambda}{\tau} - 1\right)\Psi_t(x,y)$$

where $\tau = \sqrt{(\lambda-\mu)^2 + 2t\lambda\mu}$. Noting that τ can be rewritten as

$$\tau = \sqrt{(\lambda + (t-1)\mu)^2 + t(2-t)\mu^2} = \sqrt{(\mu + (t-1)\lambda)^2 + t(2-t)\lambda^2},$$

and λ and μ can not be zero simultaneously by Lemma 3.2, it follows that

$$\left(\frac{\lambda + (t-1)\mu}{\tau} - 1\right)\left(\frac{\mu + (t-1)\lambda}{\tau} - 1\right) > 0.$$

Hence, $\langle D_1\Psi_t(x,y), D_2\Psi_t(x,y)\rangle \geq 0$, and the equality holds if and only if $\Phi_t(x,y) = 0$.

Case (a.3): $x^2 + y^2 \in \mathrm{int}\mathbb{K}$. By Theorem 4.1 and the definition of L_x, we have

$$\begin{aligned}
&\langle D_1\Psi_t(x,y), D_2\Psi_t(x,y)\rangle \\
=\ &\left\langle \left(L_{x+(t-1)y}L_z^{-1} - I\right)\Phi_t(x,y), \left(L_{y+(t-1)x}L_z^{-1} - I\right)\Phi_t(x,y)\right\rangle \\
=\ &\left\langle \left(L_{x+(t-1)y} - L_z\right)L_z^{-1}\Phi_t(x,y), \left(L_{y+(t-1)x} - L_z\right)L_z^{-1}\Phi_t(x,y)\right\rangle \\
=\ &\left\langle L_z^{-1}\Phi_t(x,y), \left(L_z - L_{x+(t-1)y}\right)\left(L_z - L_{y+(t-1)x}\right)L_z^{-1}\Phi_t(x,y)\right\rangle
\end{aligned} \tag{45}$$

where $z = [(x-y)^2 + 2t(x \bullet y)]^{1/2}$ and $I \in \mathcal{L}(\mathcal{H})$ is an identity mapping. From elementary calculation, we obtain that

$$\begin{aligned}
&\left(L_z - L_{x+(t-1)y}\right)\left(L_z - L_{y+(t-1)x}\right) + \left(L_z - L_{y+(t-1)x}\right)\left(L_z - L_{x+(t-1)y}\right) \\
=\ &t\Big(L_z - L_x - L_y\Big)^2 + \left(L_z^2 - L_{x+(t-1)y}^2 - L_{\sqrt{t(2-t)}\ y}^2\right).
\end{aligned} \tag{46}$$

Since $x^2 + y^2 \in \mathrm{int}\mathbb{K}$, from Lemma 3.2 we get $z \in \mathrm{int}\mathbb{K}$. Noting that

$$z^2 - (x + (t-1)y)^2 - \left(\sqrt{t(2-t)}y\right)^2 = 0,$$

we have $L_z^2 - L_{x+(t-1)y}^2 - L_{\sqrt{t(2-t)}y}^2 \succeq 0$ by Lemma 5.1. Along with (45) and (46),

$$\langle D_1\Psi_t(x,y), D_2\Psi_t(x,y)\rangle \geq \frac{t}{2}\left\|(L_z - L_x - L_y)L_z^{-1}\Phi_t(x,y)\right\|^2 \geq 0,$$

which in turn implies that

$$\begin{aligned}
\langle D_1\Psi_t(x,y), D_2\Psi_t(x,y)\rangle = 0 \quad &\Longleftrightarrow \quad (L_z - L_x - L_y)L_z^{-1}\Phi_t(x,y) = 0 \\
&\Longleftrightarrow \quad L_{z-x-y}L_z^{-1}\Phi_t(x,y) = 0 \\
&\Longleftrightarrow \quad \Phi_t(x,y) \bullet (L_z^{-1}\Phi_t(x,y)) = 0.
\end{aligned}$$

Since $x \bullet y = 0$ implies $\langle x, y\rangle = 0$, the last equivalence means that

$$\langle D_1\Psi_t(x,y), D_2\Psi_t(x,y)\rangle = 0 \Longrightarrow \langle \Phi_t(x,y), L_z^{-1}\Phi_t(x,y)\rangle = 0 \Longrightarrow \Phi_t(x,y) = 0$$

where the last implication is due to $z \in \text{int}\mathbb{K}$ and Lemma 2.2. Conversely, if $\Phi_t(x,y) = 0$, $\langle D_1\Psi_t(x,y), D_2\Psi_t(x,y)\rangle = 0$ follows directly from (45). This proves part (a).

(b) It suffices to prove the assertion for $(x,y) \in \mathcal{H} \times H \setminus \{(0,0)\}$. If $x^2 + y^2 \in \partial\mathbb{K} \setminus \{0\}$, then using the formula (25), we obtain

$$\begin{aligned} & \langle D_1\Psi_t(x,y), x\rangle + \langle D_2\Psi_t(x,y), y\rangle \\ = \;& 2\left\langle \Phi_t(x,y), \left(\frac{\lambda + (t-1)\mu}{\tau} - 1\right)x + \left(\frac{(t-1)\lambda + \mu}{\tau} - 1\right)y\right\rangle \\ = \;& 2\left\langle \Phi_t(x,y), z - (x+y)\right\rangle = 2\Psi_t(x,y). \end{aligned}$$

When $x^2 + y^2 \in \text{int}\mathbb{K}$, from the formula (26), it follows that

$$\begin{aligned} & \langle D_1\Psi_t(x,y), x\rangle + \langle D_2\Psi_t(x,y), y\rangle \\ = \;& 2\langle L_z^{-1}\Phi_t(x,y), (x + (t-1)y) \bullet x + (y + (t-1)x) \bullet y\rangle - 2\langle \Phi_t(x,y), x+y\rangle \\ = \;& 2\langle L_z^{-1}\Phi_t(x,y), z^2\rangle - 2\langle \Phi_t(x,y), x+y\rangle \\ = \;& 2\langle \Phi_t(x,y), z - (x+y)\rangle = 2\Psi_t(x,y). \end{aligned}$$

(c) The result is direct by part (a) and the expression of Ψ_t' given by Theorem 4.1. $\square$

Now we are in a position to establish the main result of this section by Lemma 5.2.

Theorem 5.1 *Let $T : \mathcal{H} \to \mathcal{H}$ be a given continuously F-differentiable mapping and $f(x) = \Psi_t(x, T(x))$ with $0 < t < 2$. If T is monotone, then for every $x \in \mathcal{H}$, either (i) $f(x) = 0$ or (ii) $f'(x) \neq 0$ and $\langle d(x), f'(x)\rangle < 0$ with $d(x) = -D_2\Psi_t(x, T(x))$.*

Proof. Fix any $x \in \mathcal{H}$. From Theorem 4.2 and the continuous F-differentiability of T, it follows that $f : \mathcal{H} \to \mathbb{R}_+$ is continuously F-differentiable on $\mathcal{H}$. By the chain rule of differential, we have for any $v \in \mathcal{H}$,

$$\begin{aligned} f'(x)v &= \Psi_t'(x, T(x))(v, 0) + \Psi_t'(x, T(x))\,(0, T'(x)v) \\ &= \langle D_1\Psi_t(x, T(x)), v\rangle + \langle D_2\Psi_t(x, T(x)), T'(x)v\rangle, \end{aligned}$$

which means that

$$f'(x) = D_1\Psi_t(x, T(x)) + (T'(x))^T D_2\Psi_t(x, T(x)).$$

Suppose that $f'(x) = 0$. Then the last equation implies that

$$\langle D_1\Psi_t(x, T(x)), D_2\Psi_t(x, T(x))\rangle = -\langle D_2\Psi_t(x, T(x)), T'(x) D_2\Psi_t(x, T(x))\rangle.$$

Since T is continuously F-differentiable and monotone, the right hand side is nonpositive, and consequently, $\langle D_1\Psi_t(x,T(x)), D_2\Psi_t(x,T(x))\rangle \le 0$. Together with Lemma 5.2, it then follows that $f(x) = \Psi_t(x,T(x)) = 0$.

Suppose that $f'(x) \neq 0$. Then, from the expression of $d(x)$, it follows that

$$\begin{aligned}\langle d(x), f'(x)\rangle &= -\langle D_2\Psi_t(x,T(x)), D_1\Psi_t(x,T(x)) + (T'(x))^T D_2\Psi_t(x,T(x))\rangle \\ &= -\langle D_2\Psi_t(x,T(x)), D_1\Psi_t(x,T(x))\rangle \\ &\quad -\left\langle D_2\Psi_t(x,T(x)), (T'(x))^T D_2\Psi_t(x,T(x))\right\rangle \\ &\le -\langle D_2\Psi_t(x,T(x)), D_1\Psi_t(x,T(x))\rangle .\end{aligned}$$

where the first inequality is using the monotonicity of T. By Lemma 5.2 (a), the right hand side is nonpositive and equals zero if and only if $\Psi_t(x,T(x)) = 0$, i.e., x is a solution of the minimization problem (37). However, the latter can not be true since $f'(x) \neq 0$, and consequently, $\langle d(x), f'(x)\rangle < 0$. The proof is completed. $\Box$

Theorem 5.1 states that if $x \in \mathcal{H}$ is not a solution of $\mathrm{CP}(\mathbb{K},T)$, then we can always find a descent direction $d(x)$ at this point. Based on this, an iterative descent algorithm can be designed for the self-dual conic complementarity problem $\mathrm{CP}(\mathbb{K},T)$.

6 Conclusions

We have developed a merit function method for the infinitely dimensional SOCCP $\mathrm{CP}(\mathbb{K},T)$ by extending Kanzow and Kleinmichel NCP-function to the Hilbert space. We believe that the merit functions given in this paper will be useful in other contexts, and further research work will be given to the specific applications of the merit function method. Using the techniques in this paper, other well-known merit functions, for example, the Yamashita-Fukushima merit function [15] can be also extended analogously. Specifically, we can define the Yamashita-Fukushima merit function in Hilbert space as

$$\psi_{\rm YF}(x,y) := \psi_0(\langle x,y\rangle) + \psi_{\rm FB}(x,y) \quad \forall x,y \in \mathcal{H}\times\mathcal{H},$$

where $\psi_0 : \mathbb{R} \to \mathbb{R}_+$ is any smoothing function satisfying

$$\psi_0(t) = 0 \;\; \forall t \le 0 \quad \text{and} \quad \psi_0'(t) > 0 \;\; \forall t > 0.$$

In addition, although the spectral factorization of vectors is not used in the analysis of this paper, we want to point out that, by the Jordan product associated with $\mathbb{K}$, every $x = x' + \lambda e \in \mathcal{H}$ can be written as $x = \lambda_1(x)u_x^{(1)} + \lambda_2(x)u_x^{(2)}$ with

$$\lambda_j(x) = \lambda + (-1)^j\|x'\| \;\text{ and }\; u_x^{(j)} = \frac{1}{2}\left(e + (-1)^j\bar{x}'\right), \quad j = 1,2$$

where $\bar{x}' = \frac{x'}{\|x'\|}$ if $x' \neq 0$, and otherwise $\bar{x}'$ is an arbitrary unit vector in $\langle e \rangle^\perp$.

Acknowledgments The authors would like to thank the referees for their careful reading of the paper and helpful suggestions.

References

[1] F. Alizadeh and D. Goldfarb, *Second-order cone programming*, Mathematical Programming, vol. 95, pp. 3–51, 2003.

[2] E. D. Andersen, C. Roos, and T. Terlaky, *On implementing a primal-dual interior-point method for conic quadratic optimization*, Mathematical Programming, vol. 95, pp. 249–277, 2003.

[3] J-S. Chen, *Two classes of merit functions for second-order cone complementarity problem*, Mathematical Methods of Operations Research, vol. 64, pp. 495-519, 2006.

[4] J-S. Chen and S-H. Pan, *A one-parametric class of merit functions for the second-order cone complementarity problem*, Computational Optimization and Applications, doi:10.1007/s10589-008-9182-9, 2009.

[5] X.-D. Chen, D. Sun, and J. Sun, *Complementarity functions and numerical experiments for second-order cone complementarity problems*, Computational Optimization and Applications, vol. 25, pp. 39–56, 2003.

[6] J.-S. Chen and P. Tseng, *An unconstrained smooth minimization reformulation of the second-order cone complementarity problem*, Mathematical Programming, vol. 104, pp. 293–327, 2005.

[7] J. Faraut and A. Kor'anyi, *Analysis on symmetric cones*, Oxford Mathematical Monographs Oxford University Press, New York, 1994.

[8] A. Fischer, *A special Newton-type optimization method*, Optimization, vol. 24, pp. 269–284, 1992.

[9] M. Fukushima, Z-Q. Luo, and P. Tseng, *Smoothing functions for second-order cone complementarity problems*, SIAM Journal on Optimization, vol. 12, pp. 436–460, 2002.

[10] L. Faybusovich and T. Tsuchiya, *Primal-dual algorithms and infinite-dimensional Jordan algebras of finite rank*, Mathematical Programming, vol. 97, pp. 471–493, 2003.

[11] S. Hayashi, N. Yamashita, and M. Fukushima, *A combined smoothing and regularization method for monotone second-order cone complementarity problems*, SIAM Journal of Optimization, vol. 15, pp. 593–615, 2005.

[12] C. Kanzow and H. Kleinmichel, *A new class of semismooth Newton-type methods for nonlinear complementarity problems*, Computational Optimization and Applications, vol. 11, pp. 227-251, 1998.

[13] Z-Q. Luo and P. Tseng, *A new class of merit functions for the nonlinear complementarity problem*, In : Ferris, M.C., Pang, J-S., (eds), Complementarity and Variational Problems: State of the Art, SIAM Publications, Philadelphia, pp. 204-225, 1997.

[14] O. L. Mangasarian, *Equivalence of the complementarity problem to a system of nonlinear equations*, SIAM Journal on Applied Mathematics, vol. 31, pp. 89-92, 1976.

[15] N. Yamashita and M. Fukushima, *A new merit function and a descent method for semidefinite complementarity problems*, In : Fukushima, M., Qi, L., (eds), Reformulation - Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods, Kluwer Academic Publishers, Boston, pp. 405-420, 1999.

[16] C. Kanzow, I. Ferenczi and M. Fukushima, *Semismooth methods for linear and nonlinear second-order cone programs*, Technical Report, Department of Applied Mathematics and Physics, Kyoto University, 2006.

[17] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, *Application of second-order cone programming*, Linear Algebra and its Applications, vol. 284, pp. 193–228, 1998.

[18] R. D. C. Monteiro and T. Tsuchiya, *Polynomial convergence of primal-dual algorithms for the second-order cone programs based on the MZ-family of directions*, Mathematical Programming, vol. 88, pp. 61-83, 2000.

[19] S.-H. Pan and J.-S. Chen, *A damped Guass-Newton method for the second-order cone complementarity problem*, Applied Mathematics and Optimization, vol.59, pp. 293-318, 2009.

[20] D. Sun and J. Sun, *Strong semismoothness of Fischer-Burmeister SDC and SOC function*, Mathematical Programming, vol. 103, pp. 575–581, 2005.

[21] T. Tsuchiya, *A convergence analysis of the scaling-invariant primal-dual path-following algorithms for second-order cone programming*, Optimization Methods and Software, vol. 11, pp. 141–182, 1999.

[22] A. Yoshise, *Interior point trajectories and a homogenous model for nonlinear complementarity problems over symmetric cones*, SIAM Journal on Optimization, vol. 17, pp. 1129–1153, 2006.