Operations Research Letters, vol. 36, pp. 385-392, 2008.

The SC^1 property of the squared norm of the SOC Fischer-Burmeister function

Jein-Shan Chen [1]
Department of Mathematics
National Taiwan Normal University
Taipei, Taiwan 11677

Defeng Sun [2]
Department of Mathematics
National University of Singapore
Singapore 117543

Jie Sun [3]
Department of Decision Sciences
National University of Singapore
Singapore 117592

March 26, 2007
(revised May 25, 2007)

Abstract: We show that the gradient mapping of the squared norm of Fischer-Burmeister function is globally Lipschitz continuous and semismooth, which provide a theoretical basis for solving nonlinear second order cone complementarity problems via the conjugate gradient method and the semismooth Newton's method.

Key words. Second-order cone, merit function, spectral factorization, Lipschitz continuity, semismoothness.

AMS subject classifications. 26B05, 26B35, 90C33, 65K05

---

[1]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.

[2]E-mail: matsundf@nus.edu.sg

[3]E-mail: jsun@nus.edu.sg. This author's work is partially supported by Singapore-MIT Alliance.

1 Introduction

A popular approach to solving the nonlinear complementarity problem (NCP) is to reformulate it as the global minimization via a certain merit function over $\mathbb{R}^n$. For this approach to be effective, the choice of the merit function is crucial. A popular choice of the merit function is the squared norm of the Fischer-Burmeister (FB) function $\Psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ defined by

$$\Psi(a,b) \;:=\; \frac{1}{2}\sum_{i=1}^{n} |\phi(a_i,b_i)|^2, \tag{1}$$

for all $a = (a_1, \cdots, a_n)^T \in \mathbb{R}^n$ and $b = (b_1, \cdots, b_n)^T \in \mathbb{R}^n$. The aforementioned Fischer-Burmeister function is denoted by $\Phi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ whose i-th component function is $\Phi_i(a,b) = \phi(a_i,b_i)$ with $\phi : \mathbb{R}^2 \to \mathbb{R}$ given by

$$\phi(a_i,b_i) = \sqrt{a_i^2 + b_i^2} - a_i - b_i. \tag{2}$$

It is well-known that the FB function satisfies

$$\phi(a_i,b_i) = 0 \quad \Longleftrightarrow \quad a_i \geq 0, \quad b_i \geq 0, \quad a_i b_i = 0. \tag{3}$$

It has been shown that ϕ^2 is smooth (continuously differentiable) even though ϕ is not differentiable. This merit function and its analysis were subsequently extended by Tseng [12] to the semidefinite complementarity problem (SDCP) although only differentiability, not continuous differentiability, was established. In fact, the FB function for the SDCP is the matrix-valued function $\Phi : \mathcal{S}^n \times \mathcal{S}^n \to \mathcal{S}^n$ defined by

$$\Phi(X,Y) := (X^2 + Y^2)^{1/2} - (X + Y),$$

while the squared norm of the FB function for the SDCP is the function $\Psi : \mathcal{S}^n \times \mathcal{S}^n \to \mathbb{R}_+$ given by

$$\Psi(X,Y) := \frac{1}{2}\|\Phi(X,Y)\|^2,$$

where $\mathcal{S}^n$ denotes the set of real $n \times n$ symmetric matrices. The function Φ has been proved to be strongly semismooth everywhere [11]. More recently, the squared norm of the matrix-valued FB function Ψ was reported in [10] to be a smooth function and its gradient is Lipschitz continuous.

The *second-order cone* (SOC), also called the Lorentz cone, in $\mathbb{R}^n$ is defined as

$$\mathcal{K}^n := \{(x_1,x_2) \in \mathbb{R} \times \mathbb{R}^{n-1} \mid \|x_2\| \leq x_1\}, \tag{4}$$

where $\|\cdot\|$ denotes the Euclidean norm. By definition, $\mathcal{K}^1$ is the set of nonnegative reals $\mathbb{R}_+$. The second-order cone complementarity problem (SOCCP) which is to find $x, y \in \mathbb{R}^n$ satisfying

$$x = F(\zeta), \;\; y = G(\zeta), \;\; \langle x,y\rangle = 0, \quad x \in \mathcal{K}^n, \quad y \in \mathcal{K}^n, \tag{5}$$

where $\langle \cdot, \cdot \rangle$ is the Euclidean inner product and $F, G : \mathbb{R}^n \to \mathbb{R}^n$ are continuous (possibly nonlinear) functions. The FB function for the SOCCP is the vector-valued function $\phi_{\rm FB} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ defined by

$$\phi_{\rm FB}(x, y) := (x^2 + y^2)^{1/2} - (x + y), \tag{6}$$

and the squared norm of the FB function for the SOCCP is $\psi_{\rm FB} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ given by

$$\psi_{\rm FB}(x, y) := \frac{1}{2}\|\phi_{\rm FB}(x, y)\|^2. \tag{7}$$

Note that x^2 and y^2 in (6) mean $x \circ x$ and $y \circ y$, respectively ("$\circ$" is introduced in Sec. 2); and $x + y$ means the usual componentwise addition of vectors. It is known that $x^2 \in \mathcal{K}^n$ for all $x \in \mathbb{R}^n$. Moreover, if $x \in \mathcal{K}^n$ then there exists a unique vector in $\mathcal{K}^n$ denoted by $x^{1/2}$ such that $(x^{1/2})^2 = x^{1/2} \circ x^{1/2} = x$. Therefore, the FB function given as in (6) is well-defined for all $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$. Besides, it was shown in [5] that property (3) of ϕ can be extended to $\phi_{\rm FB}$. Thus, $\psi_{\rm FB}$ is a merit function for the SOCCP since the SOCCP can be expressed as an unconstrained minimization problem:

$$\min_{\zeta \in \mathbb{R}^n} \; f(\zeta) := \psi_{\rm FB}(F(\zeta), G(\zeta)). \tag{8}$$

Like in the NCP and the SDCP cases, $\psi_{\rm FB}$ is shown to be smooth, and when ∇F and $-\nabla G$ are column monotone, every stationary point of (8) solves SOCCP; see [2].

The last hurdle to cross in applying (8) to solve (5) is to show that the gradient of $\psi_{\rm FB}$ is sufficiently smooth to warrant the convergence of appropriate computational methods. In particular, we are concerned with the conjugate gradient methods and the semismooth Newton's methods [3]. The former methods generally require the Lipschitz continuity of the gradient ($f \in LC^1$ for short since $f : \mathbb{R}^n \to \mathbb{R}$ is said to be an LC^1 function if it is continuously differentiable and its gradient is locally Lipschitz continuous), while the latter require that the gradient is semismooth ($f \in SC^1$ for short since f is called an SC^1 function if it is continuously differentiable and its gradient is semismooth), in addition to being Lipschitz continuous.

The main purpose of this paper is to show that the gradient function of $\psi_{\rm FB}$ defined as in (7) is globally Lipschitz continuous and semismooth, which is an important property for superlinear convergence of semismooth Newton methods [9]. It should be noted that this result is not a direct implication from a similar result on function $\Psi(X, Y)$ recently published in [10]. Different analysis is necessary for the proof of Lipschitz continuity.

Throughout this paper, $\mathbb{R}^n$ denotes the space of n-dimensional real column vectors and the supscript "T" denotes transpose. For any differentiable function $f : \mathbb{R}^n \to \mathbb{R}$, $\nabla f(x)$ denotes the gradient of f at x. For any differentiable mapping $F = (F_1, ..., F_m)^T : \mathbb{R}^n \to \mathbb{R}^m$, $\nabla F(x) = [\nabla F_1(x) \;\cdots\; \nabla F_m(x)]$ is a $n \times m$ matrix denoting the transposed

Jacobian of F at x. For any symmetric matrices $A, B \in \mathbb{R}^{n\times n}$, we write $A \succeq B$ (respectively, $A \succ B$) to mean $A-B$ is positive semidefinite (respectively, positive definite). For nonnegative scalars α and β, we write $\alpha = O(\beta)$ to mean $\alpha \leq C\beta$, with C independent of α and β.

2 Preliminaries

For any $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, we define their *Jordan product* associated with $\mathcal{K}^n$ as

$$x \circ y \ := \ (\langle x, y\rangle,\ y_1 x_2 + x_1 y_2). \tag{9}$$

The identity element under this product is $e := (1, 0, \ldots, 0)^T \in \mathbb{R}^n$. We write x^2 to mean $x \circ x$ and write $x + y$ to mean the usual componentwise addition of vectors. It is known that $x^2 \in \mathcal{K}^n$ for all $x \in \mathbb{R}^n$. Moreover, if $x \in \mathcal{K}^n$, then there exists a unique vector in $\mathcal{K}^n$, denoted by $x^{1/2}$, such that $(x^{1/2})^2 = x^{1/2} \circ x^{1/2} = x$.

For any $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, we define a linear mapping from $\mathbb{R}^n$ to $\mathbb{R}^n$ as

$$\begin{array}{rccl} L_x : & \mathbb{R}^n & \longrightarrow & \mathbb{R}^n \\ & y & \longrightarrow & L_x y := \begin{bmatrix} x_1 & x_2^T \\ x_2 & x_1 I \end{bmatrix} y \ . \end{array}$$

It can be easily verified that $x \circ y = L_x y$, $\forall y \in \mathbb{R}^n$, and L_x is positive definite (and hence invertible) if and only if $x \in \text{int}(\mathcal{K}^n)$. However, $L_x^{-1} y \neq x^{-1} \circ y$, for some $x \in \text{int}(\mathcal{K}^n)$ and $y \in \mathbb{R}^n$, i.e., $L_x^{-1} \neq L_{x^{-1}}$.

In addition, any $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$ can be decomposed as

$$x = \lambda_1 u^{(1)} + \lambda_2 u^{(2)}, \tag{10}$$

where λ_1, λ_2 and $u^{(1)}$, $u^{(2)}$ are the *spectral values* and the associated *spectral vectors* of x, with respect to $\mathcal{K}^n$, given by

$$\lambda_i \ = \ x_1 + (-1)^i \|x_2\|, \tag{11}$$

$$u^{(i)} \ = \ \begin{cases} \frac{1}{2}\Big(1, (-1)^i \dfrac{x_2}{\|x_2\|}\Big), & \text{if } \ x_2 \neq 0, \\ \frac{1}{2}\left(1, (-1)^i w\right), & \text{if } \ x_2 = 0, \end{cases} \tag{12}$$

for $i = 1, 2$, with w being any vector in $\mathbb{R}^{n-1}$ satisfying $\|w\| = 1$.

The above spectral factorization of x, as well as x^2 and $x^{1/2}$ and the matrix L_x, have various interesting properties (cf. [5]). We list some properties that we will use later.

Property 2.1 *For any $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$ with spectral values λ_1, λ_2 and spectral vectors $u^{(1)}, u^{(2)}$, the following results hold.*

(a) $x^2 = \lambda_1^2 u^{(1)} + \lambda_2^2 u^{(2)} \in \mathcal{K}^n$.

(b) *If $x \in \mathcal{K}^n$, then $0 \le \lambda_1 \le \lambda_2$ and $x^{1/2} = \sqrt{\lambda_1}\, u^{(1)} + \sqrt{\lambda_2}\, u^{(2)}$.*

(c) *If $x \in \operatorname{int}(\mathcal{K}^n)$, then $0 < \lambda_1 \le \lambda_2$, $det(x) = \lambda_1\lambda_2$, and L_x is invertible with*

$$L_x^{-1} = \frac{1}{det(x)} \begin{bmatrix} x_1 & -x_2^T \\ -x_2 & \dfrac{det(x)}{x_1} I + \dfrac{1}{x_1} x_2 x_2^T \end{bmatrix}.$$

(d) *$x \circ y = L_x y$ for all $y \in \mathbb{R}^n$, and $L_x \succ 0$ if and only if $x \in \operatorname{int}(\mathcal{K}^n)$.*

For any function $f : \mathbb{R} \to \mathbb{R}$, the following vector-valued function associated with $\mathcal{K}^n$ ($n \ge 1$) was considered in [6, 7]

$$f^{\text{soc}}(x) = f(\lambda_1)u^{(1)} + f(\lambda_2)u^{(2)} \quad \forall x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}. \tag{13}$$

For a recent treatment, see [1, 5]. If f is defined only on a subset of $\mathbb{R}$, then f^{soc} is defined on the corresponding subset of $\mathbb{R}^n$.

Since we aim to prove that the merit function ψ_{FB} defined as in (7) has a Lipschitz continuous gradient, we now write down the gradient function of ψ_{FB} as below. Let $\phi_{\text{FB}}, \psi_{\text{FB}}$ be given by (6) and (7), respectively. Then, from [2, Prop. 1], we know that $\nabla_x \psi_{\text{FB}}(0,0) = \nabla_y \psi_{\text{FB}}(0,0) = 0$. If $(x,y) \ne (0,0)$ and $x^2 + y^2 \in \operatorname{int}(\mathcal{K}^n)$, then

$$\begin{aligned} \nabla_x \psi_{\text{FB}}(x,y) &= \left(L_x L^{-1}_{(x^2+y^2)^{1/2}} - I \right) \phi_{\text{FB}}(x,y), \\ \nabla_y \psi_{\text{FB}}(x,y) &= \left(L_y L^{-1}_{(x^2+y^2)^{1/2}} - I \right) \phi_{\text{FB}}(x,y). \end{aligned} \tag{14}$$

If $(x,y) \ne (0,0)$ and $x^2 + y^2 \notin \operatorname{int}(\mathcal{K}^n)$, then $x_1^2 + y_1^2 \ne 0$ and

$$\begin{aligned} \nabla_x \psi_{\text{FB}}(x,y) &= \left(\frac{x_1}{\sqrt{x_1^2 + y_1^2}} - 1 \right) \phi_{\text{FB}}(x,y), \\ \nabla_y \psi_{\text{FB}}(x,y) &= \left(\frac{y_1}{\sqrt{x_1^2 + y_1^2}} - 1 \right) \phi_{\text{FB}}(x,y). \end{aligned} \tag{15}$$

Next, we also state some important technical lemmas which will be used in proving our main results. Lemma 2.1 describes the behavior of (x,y) when $x^2 + y^2$ lies on the boundary of $\mathcal{K}^n$; and Lemma 2.2 measures how close $x^2 + y^2$ comes to the boundary of $\mathcal{K}^n$. Lemma 2.3 says the matrices appeared in the gradient function (14) of ψ_{FB} is uniformly bounded.

Lemma 2.1 *[2, Lemma 2] For any $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$ with $x^2 + y^2 \notin \operatorname{int}(\mathcal{K}^n)$, we have*

$$\begin{aligned} x_1^2 &= \|x_2\|^2, \\ y_1^2 &= \|y_2\|^2, \\ x_1 y_1 &= x_2^T y_2, \\ x_1 y_2 &= y_1 x_2. \end{aligned}$$

Lemma 2.2 *[2, Lemma 3] For any $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$ with $x_1 x_2 + y_1 y_2 \neq 0$, we have*

$$\left(x_1 - \frac{(x_1x_2 + y_1y_2)^T x_2}{\|x_1x_2 + y_1y_2\|}\right)^2 \leq \left\| x_2 - x_1 \frac{x_1x_2 + y_1y_2}{\|x_1x_2 + y_1y_2\|} \right\|^2 \leq \|x\|^2 + \|y\|^2 - 2\|x_1x_2 + y_1y_2\|.$$

Lemma 2.3 *[2, Lemma 4] There exists a scalar constant $C > 0$ such that $\|L_x L^{-1}_{(x^2+y^2)^{1/2}}\|_F \leq C$, $\|L_y L^{-1}_{(x^2+y^2)^{1/2}}\|_F \leq C$ for all $(x, y) \neq (0, 0)$ satisfying $x^2 + y^2 \in \operatorname{int}(\mathcal{K}^n)$. ($\|A\|_F$ denotes the Frobenius norm of $A \in \mathbb{R}^{n \times n}$.)*

3 Main results

In this section, we will present the proof that the gradient function of $\psi_{\rm FB}$ is Lipschitz continuous. In fact, we will argue that $\nabla \psi_{\rm FB}$ is differentiable everywhere except $(x, y) = (0, 0)$ with $\|\nabla^2 \psi_{\rm FB}(x, y)\|$ being uniformly bounded. Then, by applying the Mean-Value Theorem for vector-valued functions, we conclude that $\nabla_x \psi_{\rm FB}$ and $\nabla_y \psi_{\rm FB}$ are globally Lipschitz continuous. We need the following three important lemmas to prove our main results.

Lemma 3.1 *Let $\omega : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ be defined by $\omega(x, y) := u(x, y) \circ v(x, y)$, where $u, v : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ are differentiable mappings. Then, ω is differentiable and*

$$\begin{aligned} \nabla_x \omega(x, y) &= \nabla_x u(x, y) L_{v(x,y)} + \nabla_x v(x, y) L_{u(x,y)}, \\ \nabla_y \omega(x, y) &= \nabla_y u(x, y) L_{v(x,y)} + \nabla_y v(x, y) L_{u(x,y)}. \end{aligned} \tag{16}$$

In particular, when $\omega(x, y) = x \circ y$, there hold

$$\nabla_x \omega(x, y) = L_y, \quad \nabla_y \omega(x, y) = L_x;$$

and when $\omega(x, y) = x^2 \circ y^2$, there hold

$$\nabla_x \omega(x, y) = 2L_x L_{y^2}, \quad \nabla_y \omega(x, y) = 2L_y L_{x^2}.$$

Proof. This is the product rule associated with Jordan product. Its proof is straightforward, so we omit it. $\square$

Lemma 3.2 *For any $x, y \in \mathbb{R}^n$, let $z(x,y) := (x^2+y^2)^{1/2}$, $F(x,y) := L_x L^{-1}_{z(x,y)}(x+y)$, and $G(x,y) := L_y L^{-1}_{z(x,y)}(x+y)$. Then, we have*

(a) *z is differentiable at $(x,y) \neq (0,0) \in \mathbb{R}^n \times \mathbb{R}^n$ with $x^2+y^2 \in \mathrm{int}(\mathcal{K}^n)$. Moreover*

$$\nabla_x z(x,y) = L_x L^{-1}_{z(x,y)}, \quad \nabla_y z(x,y) = L_y L^{-1}_{z(x,y)}.$$

(b) *F, G are differentiable at $(x,y) \neq (0,0) \in \mathbb{R}^n \times \mathbb{R}^n$ with $x^2+y^2 \in \mathrm{int}(\mathcal{K}^n)$. Moreover, $\|\nabla F(x,y)\|$, $\|\nabla G(x,y)\|$ are uniformly bounded at such points.*

Proof. (a) That the function z is differentiable is an immediate consequence of [7]. See also [1, Prop. 4]. Since, $z^2(x,y) = x^2+y^2$, applying Lemma 3.1 yields

$$2\nabla_x z(x,y) L_{z(x,y)} = 2L_x, \quad 2\nabla_y z(x,y) L_{z(x,y)} = 2L_y.$$

Hence, the desired results follow.

(b) For symmetry, it is enough to show that F is differentiable at $(x,y) \neq (0,0)$ with $x^2+y^2 \in \mathrm{int}(\mathcal{K}^n)$ and that $\|\nabla_x F(x,y)\|$, $\|\nabla_y F(x,y)\|$ are uniformly bounded. It is clear that F is differentiable at such points. The key part is to show the uniform boundedness of $\|\nabla_x F(x,y)\|$, $\|\nabla_y F(x,y)\|$. Let λ_1, λ_2 be the spectral values of x^2+y^2, then

$$\begin{aligned} \lambda_1 &:= \|x\|^2 + \|y\|^2 - 2\|x_1x_2 + y_1y_2\|, \\ \lambda_2 &:= \|x\|^2 + \|y\|^2 + 2\|x_1x_2 + y_1y_2\|. \end{aligned}$$

Thus, by Property 2.1(b), $z(x,y) := (x^2+y^2)^{1/2}$ has the spectral values $\sqrt{\lambda_1}, \sqrt{\lambda_2}$ and

$$z(x,y) = (z_1, z_2) = \left(\frac{\sqrt{\lambda_1}+\sqrt{\lambda_2}}{2}, \frac{\sqrt{\lambda_2}-\sqrt{\lambda_1}}{2} w_2 \right), \tag{17}$$

where $w_2 := \dfrac{x_1x_2+y_1y_2}{\|x_1x_2+y_1y_2\|}$ if $x_1x_2+y_1y_2 \neq 0$ and otherwise w_2 is any vector in $\mathbb{R}^{n-1}$ satisfying $\|w_2\| = 1$.

Now, let $u := L^{-1}_{z(x,y)}(x+y)$. By applying Property 2.1(c), we compute u as below.

$$\begin{aligned} u &= L^{-1}_{z(x,y)}(x+y) \\ &= \frac{1}{det(z(x,y))} \begin{bmatrix} z_1 & -z_2^T \\ -z_2 & \frac{det(z(x,y))}{z_1} I + \frac{1}{z_1} z_2 z_2^T \end{bmatrix} \begin{bmatrix} x_1+y_1 \\ x_2+y_2 \end{bmatrix} \\ &= \frac{1}{det(z(x,y))} \begin{bmatrix} (x_1+y_1)z_1 - (x_2+y_2)^T z_2 \\ -(x_1+y_1)z_2 + \frac{det(z)}{z_1}(x_2+y_2) + \frac{(x_2+y_2)^T z_2}{z_1} z_2 \end{bmatrix} \\ &:= \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}. \end{aligned}$$

We notice that $F(x,y) = L_x L^{-1}_{z(x,y)}(x+y) = L_x u = x \circ u$, where the last equality is due to Property 2.1(d). Then, by applying Lemma 3.1, we obtain

$$\begin{aligned} \nabla_x F(x,y) &= L_u + \nabla_x u(x,y) L_x, \\ \nabla_y F(x,y) &= \nabla_y u(x,y) L_x \end{aligned} \tag{18}$$

To show that $\|\nabla_x F(x,y)\|$ is uniformly bounded, we shall verify that both $\|L_u\|$ and $\|\nabla_x u(x,y)L_x\|$ are uniformly bounded. We prove them as follows.

(i) To see $\|L_u\|$ is uniformly bounded, it is sufficient to argue that $|u_1|, \|u_2\|$ are both uniformly bounded. First, we argue that $|u_1|$ is uniformly bounded. From the above expression of u, we have

$$u_1 = \frac{1}{det(z(x,y))}(x_1 z_1 - x_2^T z_2) + \frac{1}{det(z(x,y))}(y_1 z_1 - y_2^T z_2).$$

Following the similar arguments as in [2, Lemma 4] yields

$$\begin{aligned} u_1 &= \frac{1}{det(z(x,y))}(x_1 z_1 - x_2^T z_2) + \frac{1}{det(z(x,y))}(y_1 z_1 - y_2^T z_2) \\ &= \left[O(1) + \frac{(x_1 - x_2^T w_2)}{2\sqrt{\lambda_1}} \right] + \left[O(1) + \frac{(y_1 - y_2^T w_2)}{2\sqrt{\lambda_1}} \right], \end{aligned}$$

where $O(1)$ denotes terms that are uniformly bounded with bound independent of (x,y). Moreover, by Lemma 2.2, if $x_1 x_2 + y_1 y_2 \neq 0$ then $|x_1 - x_2^T w_2| \leq \|x_2 - x_1 w_2\| \leq \sqrt{\lambda_1}$. If $x_1 x_2 + y_1 y_2 = 0$ then $\lambda_1 = \|x\|^2 + \|y\|^2$ so that by choosing w_2 to further satisfy $x_2^T w_2 = 0$ we obtain $|x_1 - x_2^T w_2| \leq \|x_2 - x_1 w_2\| \leq \|x\| \leq \sqrt{\lambda_1}$. Similarly, it can be verified that $|y_1 - y_2^T w_2| \leq \sqrt{\lambda_1}$. Thus, $|u_1|$ is uniformly bounded.

Secondly, we argue that $\|u_2\|$ is also uniformly bounded. Again, using the expression of u and following the similar arguments as in [2, Lemma 4], we obtain

$$\begin{aligned} u_2 &= \frac{1}{det(z(x,y))}\left[-x_1 z_2 + \frac{det(z(x,y))}{z_1} x_2 + \frac{x_2^T z_2}{z_1} z_2 \right] \\ &\quad + \frac{1}{det(z(x,y))}\left[-y_1 z_2 + \frac{det(z(x,y))}{z_1} y_2 + \frac{y_2^T z_2}{z_1} z_2 \right] \\ &= \left[O(1) - \frac{x_1 w_2}{2\sqrt{\lambda_1}} + \frac{\frac{\sqrt{\lambda_2}}{\sqrt{\lambda_1}}(x_2^T w_2)}{2(\sqrt{\lambda_1}+\sqrt{\lambda_2})} w_2 \right] + \left[O(1) - \frac{y_1 w_2}{2\sqrt{\lambda_1}} + \frac{\frac{\sqrt{\lambda_2}}{\sqrt{\lambda_1}}(y_2^T w_2)}{2(\sqrt{\lambda_1}+\sqrt{\lambda_2})} w_2 \right] \\ &= \left[O(1) - \frac{x_1 w_2}{2(\sqrt{\lambda_1}+\sqrt{\lambda_2})} - \frac{\sqrt{\lambda_2}(x_1 - x_2^T w_2)}{2(\sqrt{\lambda_1}+\sqrt{\lambda_2})\sqrt{\lambda_1}} w_2 \right] \\ &\quad + \left[O(1) - \frac{y_1 w_2}{2(\sqrt{\lambda_1}+\sqrt{\lambda_2})} - \frac{\sqrt{\lambda_2}(y_1 - y_2^T w_2)}{2(\sqrt{\lambda_1}+\sqrt{\lambda_2})\sqrt{\lambda_1}} w_2 \right]. \end{aligned}$$

Using the same explanations as above for u_1 yields that each term is uniformly bounded. Thus, $\|u_2\|$ is uniformly bounded. This together with $|u_1|$ being uniformly bounded implies that $\|\nabla_x F(x,y)\| = \|L_u\| = \left\| \begin{bmatrix} u_1 & u_2^T \\ u_2 & u_1 I \end{bmatrix} \right\|$ is also uniformly bounded.

(ii) Now, it comes to show that $\|\nabla_x u(x,y) L_x\|$ is uniformly bounded. From the definition of $u := L^{-1}_{z(x,y)}(x+y)$, we know that $z(x,y) \circ u = x+y$. Applying Lemma 3.1 gives

$$\nabla_x z(x,y) L_u + \nabla_x u(x,y) L_{z(x,y)} = I,$$

which leads to

$$\begin{aligned}
& \nabla_x u(x,y) L_{z(x,y)} = I - \nabla_x z(x,y) L_u = I - (L_x L^{-1}_{z(x,y)}) L_u \\
\Rightarrow \quad & \nabla_x u(x,y) = \Big(I - L_x L^{-1}_{z(x,y)} L_u \Big) L^{-1}_{z(x,y)} \\
\Rightarrow \quad & \nabla_x u(x,y) L_x = \Big(I - L_x L^{-1}_{z(x,y)} L_u \Big) L^{-1}_{z(x,y)} L_x \\
\Rightarrow \quad & \nabla_x u(x,y) L_x = L^{-1}_{z(x,y)} L_x - L_x L^{-1}_{z(x,y)} L_u L^{-1}_{z(x,y)} L_x \\
\Rightarrow \quad & \nabla_x u(x,y) L_x = (L_x L^{-1}_{z(x,y)})^T - (L_x L^{-1}_{z(x,y)}) L_u (L_x L^{-1}_{z(x,y)})^T.
\end{aligned}$$

Therefore,

$$\|\nabla_x u(x,y) L_x\| \leq \|(L_x L^{-1}_{z(x,y)})^T\| + \|L_x L^{-1}_{z(x,y)}\| \cdot \|L_u\| \cdot \|(L_x L^{-1}_{z(x,y)})^T\|.$$

By Lemma 2.3, $\|L_x L^{-1}_{z(x,y)}\|$ is uniformly bounded, so is $\|(L_x L^{-1}_{z(x,y)})^T\|$. This together with $\|L_u\|$ being uniformly bounded shown as above yields $\|\nabla_x u(x,y) L_x\|$ is uniformly bounded.

From (i) and (ii), we conclude that $\|\nabla_x F(x,y)\|$ is uniformly bounded. Similar arguments apply to $\|\nabla_y F(x,y)\|$; and hence, $\|\nabla F(x,y)\|$ is uniformly bounded. Thus, we complete the proof. □

Lemma 3.3 *Let $\psi_{\rm FB}$ be defined as (7). Then, $\nabla \psi_{\rm FB}$ is continuously differentiable everywhere except for $(x,y) = (0,0)$. Moreover, $\|\nabla^2 \psi_{\rm FB}(x,y)\|$ is uniformly bounded for all $(x,y) \neq (0,0)$.*

Proof. For any $(x,y) \in \mathbb{R}^n \times \mathbb{R}^n$, let $z := (x^2+y^2)^{1/2}$. We prove this lemma by considering the following two cases.

(i) Consider all points $(x,y) \neq (0,0)$ with $x^2+y^2 \in \text{int}(\mathcal{K}^n)$. Since

$$\begin{aligned}
\nabla_x \psi_{\rm FB}(x,y) &= \Big(L_x L_z^{-1} - I \Big) \phi_{\rm FB}(x,y) \\
&= x - L_x L_z^{-1}(x+y) - \phi_{\rm FB}(x,y), \\
\nabla_y \psi_{\rm FB}(x,y) &= \Big(L_y L_z^{-1} - I \Big) \phi_{\rm FB}(x,y) \\
&= y - L_y L_z^{-1}(x+y) - \phi_{\rm FB}(x,y),
\end{aligned}$$

we can compute $\nabla^2\psi_{\rm FB}(x,y)$ as follows:

$$\begin{aligned}
\nabla^2_{xx}\psi_{\rm FB}(x,y) &= I-\nabla_x\Big(L_xL_z^{-1}(x+y)\Big)-\Big(L_xL_z^{-1}-I\Big), \qquad (19)\\
\nabla^2_{xy}\psi_{\rm FB}(x,y) &= -\nabla_y\Big(L_xL_z^{-1}(x+y)\Big)-\Big(L_yL_z^{-1}-I\Big),\\
\nabla^2_{yx}\psi_{\rm FB}(x,y) &= -\nabla_x\Big(L_yL_z^{-1}(x+y)\Big)-\Big(L_xL_z^{-1}-I\Big),\\
\nabla^2_{yy}\psi_{\rm FB}(x,y) &= I-\nabla_y\Big(L_yL_z^{-1}(x+y)\Big)-\Big(L_yL_z^{-1}-I\Big).
\end{aligned}$$

The continuity of $\nabla^2\psi_{\rm FB}$ at (x,y) thus follows. It is easy to see that $\|L_xL_z^{-1}\|$, $\|L_yL_z^{-1}\|$ are uniformly bounded by Lemma 2.3 ($\|\cdot\|$ and $\|\cdot\|_F$ are equivalent in $\mathbb{R}^{n\times n}$). Let $F(x,y):=L_xL_z^{-1}(x+y)$ and $G(x,y):=L_yL_z^{-1}(x+y)$. By Lemma 3.2, we know that $\left\|\nabla_x\Big(L_xL_z^{-1}(x+y)\Big)\right\|=\|\nabla_xF(x,y)\|$ is uniformly bounded. Likewise, we have that $\left\|\nabla_y\Big(L_xL_z^{-1}(x+y)\Big)\right\|$, $\left\|\nabla_x\Big(L_yL_z^{-1}(x+y)\Big)\right\|$, $\left\|\nabla_y\Big(L_yL_z^{-1}(x+y)\Big)\right\|$ are all uniformly bounded. Thus, we can conclude that $\|\nabla^2_{xx}\psi_{\rm FB}(x,y)\|$, $\|\nabla^2_{xy}\psi_{\rm FB}(x,y)\|$, $\|\nabla^2_{yx}\psi_{\rm FB}(x,y)\|$, $\|\nabla^2_{yy}\psi_{\rm FB}(x,y)\|$ are all uniformly bounded which implies that $\|\nabla^2\psi_{\rm FB}(x,y)\|$ is also uniformly bounded.

(ii) Consider all points $(x,y)\neq(0,0)$ with $x^2+y^2\notin{\rm int}(\mathcal{K}^n)$. Since

$$\begin{aligned}
\nabla_x\psi_{\rm FB}(x,y) &= \left(\frac{x_1}{\sqrt{x_1^2+y_1^2}}-1\right)\phi_{\rm FB}(x,y)\\
&= x-\frac{x_1}{\sqrt{x_1^2+y_1^2}}(x+y)-\phi_{\rm FB}(x,y),\\
\nabla_y\psi_{\rm FB}(x,y) &= \left(\frac{y_1}{\sqrt{x_1^2+y_1^2}}-1\right)\phi_{\rm FB}(x,y)\\
&= y-\frac{y_1}{\sqrt{x_1^2+y_1^2}}(x+y)-\phi_{\rm FB}(x,y),
\end{aligned}$$

we can compute $\nabla^2\psi_{\rm FB}(x,y)$ as follows:

$$\begin{aligned}
\nabla^2_{xx}\psi_{\rm FB}(x,y) &= I-\left(\frac{x_1}{\sqrt{x_1^2+y_1^2}}I+\frac{x_1y_1^2+y_1^3}{(x_1^2+y_1^2)^{3/2}}\begin{bmatrix}1&0\\0&\mathbf{0}\end{bmatrix}\right)-\left(\frac{x_1}{\sqrt{x_1^2+y_1^2}}-1\right)I, \qquad (20)\\
\nabla^2_{xy}\psi_{\rm FB}(x,y) &= -\left(\frac{x_1}{\sqrt{x_1^2+y_1^2}}I-\frac{x_1^2y_1+x_1y_1^2}{(x_1^2+y_1^2)^{3/2}}\begin{bmatrix}1&0\\0&\mathbf{0}\end{bmatrix}\right)-\left(\frac{y_1}{\sqrt{x_1^2+y_1^2}}-1\right)I,\\
\nabla^2_{yx}\psi_{\rm FB}(x,y) &= -\left(\frac{y_1}{\sqrt{x_1^2+y_1^2}}I-\frac{x_1^2y_1+x_1y_1^2}{(x_1^2+y_1^2)^{3/2}}\begin{bmatrix}1&0\\0&\mathbf{0}\end{bmatrix}\right)-\left(\frac{x_1}{\sqrt{x_1^2+y_1^2}}-1\right)I,\\
\nabla^2_{yy}\psi_{\rm FB}(x,y) &= I-\left(\frac{y_1}{\sqrt{x_1^2+y_1^2}}I+\frac{x_1^3+x_1^2y_1}{(x_1^2+y_1^2)^{3/2}}\begin{bmatrix}1&0\\0&\mathbf{0}\end{bmatrix}\right)-\left(\frac{y_1}{\sqrt{x_1^2+y_1^2}}-1\right)I,
\end{aligned}$$

where $\mathbf{0}$ denotes the $(n-1)\times(n-1)$ zero matrix. Following the similar arguments as in case (3) of [2, Prop. 2], one can verify that $\nabla_{xx}\psi_{\text{FB}}$, $\nabla_{xy}\psi_{\text{FB}}$, $\nabla_{yx}\psi_{\text{FB}}$, and $\nabla_{yy}\psi_{\text{FB}}$ are continuous at (x,y) too in this case (the verifications may be very tedious). Here we provide an alternative approach to verify it. Let $(a,b)\neq(0,0)$ and $a^2+b^2\notin \text{int}(\mathcal{K}^n)$. We want to prove that

$$\nabla_{xx}\psi_{\text{FB}}(x,y)\to\nabla_{xx}\psi_{\text{FB}}(a,b), \quad \text{as} \quad (x,y)\to(a,b). \tag{21}$$

Due to the neighborhood of such (a,b), we have to consider two subcases: (1) $(x,y)\neq(0,0)$ with $x^2+y^2\in\text{int}(\mathcal{K}^n)$ and (2) $(x,y)\neq(0,0)$ with $x^2+y^2\notin\text{int}(\mathcal{K}^n)$. It is clear that (21) holds in subcase (2) because the formula given in (20) is continuous. In subcase (1), we have

$$\begin{aligned}\nabla_{xx}\psi_{\text{FB}}(x,y) &= I-\nabla_x\left(L_xL_z^{-1}(x+y)\right)-\left(L_xL_z^{-1}-I\right) \qquad (22)\\ &= I-\left[L_u+\left(L_xL_z^{-1}\right)^T-\left(L_xL_z^{-1}\right)(L_u)\left(L_xL_z^{-1}\right)^T\right]-\left(L_xL_z^{-1}-I\right).\end{aligned}$$

In view of (19), (20) and (22), it suffices to show the following three statements for (21) to be held in this subcase (1):

(a) $L_xL_z^{-1}\to\dfrac{a_1}{\sqrt{a_1^2+b_1^2}}\,I$, as $(x,y)\to(a,b)$.

(b) $L_u\to\dfrac{a_1+b_1}{\sqrt{a_1^2+b_1^2}}\,I$, as $(x,y)\to(a,b)$.

(c) $L_u-(L_xL_z^{-1})(L_u)(L_xL_z^{-1})^T\to\dfrac{a_1^2(a_1+b_1)}{(a_1^2+b_1^2)^{3/2}}\,I$, as $(x,y)\to(a,b)$.

First, we know from [2, Prop. 2] that there holds

$$L_xL_z^{-1}(x+y)\to\frac{a_1}{\sqrt{a_1^2+b_1^2}}\,(a+b) \quad \text{as} \quad (x,y)\to(a,b),$$

which implies $L_xL_z^{-1}\to\dfrac{a_1}{\sqrt{a_1^2+b_1^2}}\,I$, as $(x,y)\to(a,b)$ since both $(x+y)$ and $L_xL_z^{-1}$ are continuous and $(x+y)\to(a+b)$ when $(x,y)\to(a,b)$. Secondly, if we look into the entries of L_u and compare them with the entries of $L_xL_z^{-1}$ (see [2, eq. (27)]), then it is clear that $L_u\to\dfrac{a_1+b_1}{\sqrt{a_1^2+b_1^2}}\,I$, as $(x,y)\to(a,b)$. Finally, part(c) follows immediately from part (a) and (b). Thus, we complete the verifications of (21). The other cases can be argued similarly for $\nabla_{xy}\psi_{\text{FB}}$, $\nabla_{yx}\psi_{\text{FB}}$, and $\nabla_{yy}\psi_{\text{FB}}$. In addition, it is also clear that each term in the above expressions (20) is uniformly bounded. Thus, we obtain that $\nabla^2\psi_{\text{FB}}$ is continuously differentiable near (x,y) and $\|\nabla^2\psi_{\text{FB}}(x,y)\|$ is uniformly bounded. □

Theorem 3.1 *Let $\psi_{\rm FB}$ be defined as (7). Then, $\nabla\psi_{\rm FB}$ is globally Lipschitz continuous, i.e., there exists a constant C such that for all $(x,y),\ (a,b)\in {\rm I\!R}^n\times{\rm I\!R}^n$,*

$$\begin{array}{rcl} \|\nabla_x\psi_{\rm FB}(x,y)-\nabla_x\psi_{\rm FB}(a,b)\| & \leq & C\|(x,y)-(a,b)\|, \\ \|\nabla_y\psi_{\rm FB}(x,y)-\nabla_y\psi_{\rm FB}(a,b)\| & \leq & C\|(x,y)-(a,b)\| \end{array} \tag{23}$$

and is semismooth everywhere.

Proof. Because of symmetry, we only need to show that the first part of (23) holds. For any $(x,y)\in {\rm I\!R}^n\times{\rm I\!R}^n$, let $z:=(x^2+y^2)^{1/2}$.

(i) First, we prove that $\nabla_x\psi_{\rm FB}$ is Lipschitz continuous at $(0,0)$. We have to discuss three subcases for completing the proof of this part.
If $(x,y)=(0,0)$, it is obvious that (23) is satisfied.
If $(x,y)\neq(0,0)$ with $x^2+y^2\in{\rm int}(\mathcal{K}^n)$, then

$$\|\nabla_x\psi_{\rm FB}(x,y)-\nabla_x\psi_{\rm FB}(0,0)\| = \|\nabla_x\psi_{\rm FB}(x,y)\| = \|x-L_xL_z^{-1}(x+y)-\phi_{\rm FB}(x,y)\|.$$

It is already known that x and $\phi_{\rm FB}(x,y)$ are Lipschitz continuous (see [11, Cor. 3.3]). In addition, Theorem 3.2.4 of [8, pp. 70] says that the uniform boundedness of $\nabla\Big(L_xL_z^{-1}(x+y)\Big)$ (by Lemma 3.2) yields the Lipschitz continuity. Thus, (23) is satisfied for this subcase.
If $(x,y)\neq(0,0)$ with $x^2+y^2\notin{\rm int}(\mathcal{K}^n)$, then

$$\|\nabla_x\psi_{\rm FB}(x,y)-\nabla_x\psi_{\rm FB}(0,0)\| = \|\nabla_x\psi_{\rm FB}(x,y)\| = \left\|x-\frac{x_1}{\sqrt{x_1^2+y_1^2}}(x+y)-\phi_{\rm FB}(x,y)\right\|.$$

Since $\left|\dfrac{x_1}{\sqrt{x_1^2+y_1^2}}\right|\leq 1$ and both $(x+y)$, $\phi_{\rm FB}(x,y)$ are known Lipschitz continuous, the desired result follows.

(ii) Secondly, we prove that $\nabla_x\psi_{\rm FB}$ is Lipschitz continuous at $(a,b)\neq(0,0)$. Let $(x,y)\in {\rm I\!R}^n\times{\rm I\!R}^n$, we wish to show that (23) is satisfied. In fact, if the line segment $[(a,b),(x,y)]$ does not contain the origin, then we can write

$$\begin{array}{rl} & \|\nabla_x\psi_{\rm FB}(x,y)-\nabla_x\psi_{\rm FB}(a,b)\| \\ \leq & \left\|\int_0^1 \nabla^2\psi_{\rm FB}[(a,b)+t((x,y)-(a,b))]dt\right\| \\ \leq & C\|(x,y)-(a,b)\|, \end{array}$$

where the first inequality is from the Mean-Value Theorem (see [8, Theorem 3.2.3]), and the second inequality is by Lemma 3.3. On the other hand, if the line segment $[(a,b),(x,y)]$ contains the origin, we can construct a sequence $\{(x^k,y^k)\}$ converging to

(x, y) but for each k, the line segment $[(a, b), (x^k, y^k)]$ does not contain the origin and apply the above inequalities to get

$$\|\nabla_x \psi_{\text{FB}}(x^k, y^k) - \nabla_x \psi_{\text{FB}}(a, b)\| \leq C\|(x^k, y^k) - (a, b)\|,$$

which, by the continuity, implies

$$\|\nabla_x \psi_{\text{FB}}(x, y) - \nabla_x \psi_{\text{FB}}(a, b)\| \leq C\|(x, y) - (a, b)\|.$$

Thus, (23) is satisfied.

To complete the proof of this theorem, we only need to show that $\nabla \psi_{\text{FB}}$ is semismooth at the origin as, by Lemma 3.3, $\nabla \psi_{\text{FB}}$ is continuously differentiable near any $(0, 0) \neq (x, y) \in \mathbb{R}^n \times \mathbb{R}^n$. From (14) and (15), we know that for any $t \in \mathbb{R}_+$ and $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ we have

$$\nabla \psi_{\text{FB}}(tx, ty) = t \nabla \psi_{\text{FB}}(x, y)\,.$$

Thus, $\nabla \psi_{\text{FB}}$ is directionally differentiable at the origin and for any $(0, 0) \neq (x, y) \in \mathbb{R}^n \times \mathbb{R}^n$

$$\nabla^2 \psi_{\text{FB}}(x, y)(x, y) = (\nabla \psi_{\text{FB}})'((x, y); (x, y)) = \nabla \psi_{\text{FB}}(x, y)\,.$$

This means that for any $(0, 0) \neq (x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ converging to $(0, 0)$,

$$\nabla \psi_{\text{FB}}(x, y) - \nabla \psi_{\text{FB}}(0, 0) - \nabla^2 \psi_{\text{FB}}(x, y)(x, y) = \nabla \psi_{\text{FB}}(x, y) - 0 - \nabla \psi_{\text{FB}}(x, y) = 0\,,$$

which, together with the Lipschitz continuity of $\nabla \psi_{\text{FB}}$ and the directional differentiability of $\nabla \psi_{\text{FB}}$ at the origin ($\nabla \psi_{\text{FB}}$ is, however, not differentiable at the origin), shows that $\nabla \psi_{\text{FB}}(x, y)$ is (strongly) semismooth at the origin. The proof is completed. □

From Theorem 3.1, we immediately obtain that the function ψ_{FB} defined as in (7) is an SC^1 function as well as an LC^1 function.

References

[1] J.-S. Chen, X. Chen, and P. Tseng, *Analysis of nonsmooth vector-valued functions associated with second-order cones*, Mathematical Programming, vol. 101, pp. 95-117, 2004.

[2] 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.

[3] 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.

[4] U. Faraut and A. Korányi, *Analysis on Symmetric Cones*, Oxford Mathematical Monographs, Oxford University Press, New York, 1994.

[5] 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.

[6] M. Koecher, *The Minnesota Notes on Jordan Algebras and Their Applications*, edited and annotated by A. Brieg and S. Walcher, Springer, Berlin, 1999.

[7] A. Korányi, *Monotone functions on formally real Jordan algebras*, Mathematische Annalen, vol. 269, pp. 73–76, 1984.

[8] J. Ortega and W. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, SIAM Classics in Applied Mathematics, 2000.

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

[10] C.-K. Sim, J. Sun, and D. Ralph, *A note on the Lipschitz continuity of the gradient of the squared norm of the matrix-valued Fischer-Burmeister function*, Mathematical Programming, vol. 107, pp. 547-553, 2006.

[11] D. Sun and J. Sun, *Strong semismoothness of the Fischer-Burmeister SDC and SOC complementarity functions*, Mathematical Programming, vol. 103, pp. 575-581, 2005.

[12] P. Tseng, *Merit function for semidefinite complementarity problems*, Mathematical Programming, vol. 83, pp. 159-185, 1998.