PROPERTIES OF A CLASS OF NCP-FUNCTIONS AND A RELATED SEMISMOOTH NEWTON METHOD FOR COMPLEMENTARITY PROBLEMS

YUAN MIN LI AND XING TAO WANG
Department of Mathematics
Harbin Institute of Technology
No. 92, West Da-Zhi Street, Harbin 150001, P. R. China
*Corresponding author: xingtao@hit.edu.cn

Received October 2010; revised February 2011

ABSTRACT. In this paper, we aim to explore properties of a class of NCP-functions and investigate a related semismooth Newton method for complementarity problems. Some favorite properties about the class of NCP-functions and its merit function are discussed including strong semismoothness, continuous differentiability and the nonsingularity of the element in C-subdifferential. In particular, we present an exact expression of the generalized gradient for the NCP-function. The level boundedness of the merit function is discussed. Based on these results, we investigate a semismooth Newton method and give its convergence analysis. As an application, we use this method to solve the frictionless contact problem.

Keywords: Complementarity problem, Semismooth Newton method, Complementarity function

1. Introduction. The nonlinear complementarity problem (simplified by NCP) is to find a point $x \in \mathbb{R}^n$ such that

$$x \geq 0, \quad F(x) \geq 0, \quad \langle x, F(x) \rangle = 0,$$

where $F : \mathbb{R}^n \to \mathbb{R}^n$ is a continuously differentiable mapping. The concept of complementarity is synonymous with the notion of system equilibrium; as a consequence, complementarity problems have numerous applications in engineering and economic science [1, 2, 3, 4]. Realistic complementarity problems arise from contact mechanics, structural mechanics, nonlinear obstacle, traffic equilibrium and economic equilibrium problems. An exhaustive review of the theory, algorithms and applications of NCP is given in [3]. A function-based policy problem was converted into a parameter-based policy problem in [5].

There is a great deal of practical interest in developing robust and efficient algorithms for solving the NCPs, including merit function approaches [6, 7, 8], nonsmooth Newton methods [9, 10], smoothing methods [11, 12, 13, 14, 15, 16] and regularization methods [17, 18, 19]. All the above mentioned methods are to reformulate each problem as an equivalent equation system via NCP-functions and then use Newton method to solve the equation system. A function $\phi : \mathbb{R}^2 \to \mathbb{R}$ is called an NCP-function if it satisfies

$$\phi(a, b) = 0 \iff a \geq 0, \quad b \geq 0, \quad \langle a, b \rangle = 0.$$

Many NCP-functions have been proposed [10, 20, 21, 22]. Among them, the FB function is one of the most popular NCP-functions, which is defined by

$$\phi_{FB}(a, b) = a + b - \sqrt{a^2 + b^2}.$$
One of the main generalizations of the FB function is defined by [10]:

\[ \phi_{\lambda}(a, b) = a + b - \sqrt{a^2 + b^2 + (\lambda - 2)ab} \]

where \( \lambda \) is a fixed parameter such that \( \lambda \in (0, 4) \). It has been proved in [10] that the function given by (4) possesses a system of favorite properties, such as strong semismoothness, Lipschitz continuity and directional differentiability. However, it has limitations in dealing with some monotone complementarity problems since the global convergence of the semismooth Newton method based on this function needs serious condition.

In view of the above shortcoming, the authors recently have proposed a new class of complementarity functions over symmetric cones [22]. When considering the nonnegative cone, this class of complementarity functions reduces to the NCP-function:

\[ \phi(a, b) = a + b - \sqrt{a^2 + b^2 + (\tau_1 - 2)ab + \tau_2 a_+ b_+} \]

where \( \tau_1 \in (0, 4) \) and \( \tau_2 > 0 \) are arbitrary but fixed parameters. In the setting of Jordan algebra, the authors have considered the continuous differentiability of the smoothed function of (5) in [22]. Also, they have discussed the level boundedness of the complementarity function (5) over symmetric cones in [23]. However, they do not investigate its subdifferential theory in the setting of Jordan algebra. Here, we explore all the properties about subdifferential of the function \( \phi \) defined by (5) in the setting of \( \mathbb{R}^n \). The NCP-function \( \phi \) defined by (5) inherits all the favorite properties from \( \phi_{\lambda} \). In particular, we give an exact expression of the generalized gradient \( \partial \phi(a, b) \). In this expression, when \( \tau_2 = 0 \), the generalized gradient for \( \phi_{\lambda} \) can be obtained, which is more convenient and more exact in computation than the one given in [10]. Furthermore, we give the level boundedness property under much milder condition than the ones in [10, 23]. This property makes the function \( \phi \) defined by (5) take greater advantage than \( \phi_{\lambda} \) defined by (4) in guaranteeing the global convergence of the related semismooth Newton method.

With the above characterization of \( \phi \) defined by (5), the NCP is equivalent to the following system of nonsmooth equations

\[ \Phi(x) = 0, \]

where the equation operator is

\[ \Phi(x) = \left( \begin{array}{c} \phi(x_1, F_1(x)) \\ \vdots \\ \phi(x_n, F_n(x)) \end{array} \right). \]

A natural merit function \( \Psi : \mathbb{R}^n \to \mathbb{R}^+ \) is obtained for the complementarity problems

\[ \Psi(x) = \frac{1}{2} \Phi(x)^T \Phi(x) = \sum_{i=1}^{n} \varphi(x_i, F_i(x)), \]

where

\[ \varphi(a, b) = \frac{1}{2} \phi(a, b)^2. \]

Some favorite properties of the equation operator (7) and the corresponding merit function (8) are discussed, including strong semismoothness, continuous differentiability and the nonsingularity of the C-subdifferential. A procedure is provided to calculate an element of the C-subdifferential \( \partial_C \Phi(x) \). Based on the results, we investigate a semismooth Newton method to solve the nonsmooth Equation (6). The algorithm is similar to the one in [10] but with some nice properties that the complementarity functions have one more parameter; the condition to get the global convergence is the weakest; the C-subdifferential is more convenient to calculate. The proposed method has the global convergence and
quadratic convergence under suitable conditions. Instead of solving the original NCP (1), we can get the solution based on the method. As an application, we use this method to solve the frictionless contact problem.

2. Preliminaries. In this section, we review some basic concepts that will be used in the subsequent analysis. We begin with the concept of generalized Jacobian. Let \( G : \mathbb{R}^n \to \mathbb{R}^m \) be a locally Lipschitz continuous mapping. Then \( G \) is almost everywhere differentiable by Rademacher’s Theorem [24]. In this case, the generalized Jacobian \( \partial G(x) \) of \( G \) at \( x \) (in the Clarke sense) is defined as the convex hull of the B-subdifferential

\[
\partial_B G(x) := \left\{ V \in \mathbb{R}^{m \times n} \mid \exists \{x^k\} \subseteq D_G : \{x^k\} \to x \text{ and } G'(x^k) \to V \right\},
\]

where \( D_G \) is the set of all differentiable points of \( G \) and \( G'(z) = \left( \frac{\partial G(z)}{\partial z} \right)_{m \times n} \). In other words, \( \partial G(x) = \text{co} \partial_B G(x) \). If \( m = 1 \), we call \( \partial G(x) \) the generalized gradient of \( G \) at \( x \) in which the element is a row vector. The calculation of \( \partial G(x) \) is usually difficult in practice, and hence, Qi [25] proposed the concept C-subdifferential of \( G \):

\[
\partial^C G(x) = \partial^T G_1(x) \times \cdots \times \partial^T G_m(x)
\]

which is easier to compute than the generalized Jacobian \( \partial G(x) \). Here, the right-hand side denotes the set of matrices in \( \mathbb{R}^{n \times m} \) whose \( i \)-th column is given by the transpose of the generalized gradient of the \( i \)-th component function \( G_i \). By [24, Proposition 2.6.2], \( \partial^T G(x) \subseteq \partial^C G(x) \). We assume that the reader is familiar with the concepts of (strongly) semismooth functions, and can refer to [18]. We also need the definitions of P-property which can be found in [26].

**Definition 2.1.** A function \( F : \mathbb{R}^n \to \mathbb{R}^n \) has the

(a) \( P_0 \)-property if for every \( x \) and \( y \) in \( \mathbb{R}^n \) with \( x \neq y \), there exists an index \( i \) such that

\[
(x_i - y_i)[F_i(x) - F_i(y)] \geq 0.
\]

(b) \( P \)-property if for every \( x \) and \( y \) in \( \mathbb{R}^n \) with \( x \neq y \), there exists an index \( i \) such that

\[
(x_i - y_i)[F_i(x) - F_i(y)] > 0.
\]

It is obvious that if a continuously differentiable function has the \( P_0 \)-property, then its Jacobian matrix also has the \( P_0 \)-property. Furthermore, if a matrix has the \( P_0 \)-property (\( P \)-property), then all its principal minors are nonnegative (positive).

3. Properties of the NCP-Function. In this section, some properties of the NCP-function \( \phi \), the equation operator \( \Phi \) and the merit function \( \Psi \) are discussed.

**Proposition 3.1.** The function \( \phi \) defined by (5) satisfies the following properties:

(a) \( \phi(a, b) = 0 \iff (a, b) \in N_\phi = \{(a, b) \mid a \geq 0, b \geq 0, ab = 0\} \).

(b) The generalized gradient \( \partial \phi(a, b) \) is equal to the set of all \((v_a, v_b)\) such that

\[
(v_a, v_b) = \begin{cases} 
\left( 1 - \frac{a + \frac{\tau_2 b}{2}}{\sqrt{a^2 + b^2 + (\tau_2 - 2)ab}} \right) \xi + \frac{\tau_2 b}{\sqrt{a^2 + b^2 + (\tau_2 - 2)ab}} \eta, 1 - \frac{b + \frac{\tau_2 a}{2}}{\sqrt{a^2 + b^2 + (\tau_2 - 2)ab}} \xi + \frac{\tau_2 a}{\sqrt{a^2 + b^2 + (\tau_2 - 2)ab}} \eta \\
(1 - \xi, 1 - \frac{\tau_2 - 2}{2} \xi - \frac{\sqrt{\tau_2}}{2} \eta) & \text{if } (a, b) \neq (0, 0) \\
& \text{if } (a, b) = (0, 0)
\end{cases}
\]

where \((\xi, \eta)\) is any vector satisfying \( \| (\xi, \eta) \| \leq 1 \) and

\[
\partial z_+ = \begin{cases} 
1 & \text{if } z > 0 \\
[0, 1] & \text{if } z = 0 \\
0 & \text{if } z < 0
\end{cases}
\].
Theorem 2.178] that \( g \to \infty \) if one of the following conditions is satisfied:

\[
a^k \to \infty \quad \text{and} \quad b^k \to \infty; \quad a^k \to -\infty; \quad b^k \to -\infty.
\]

**Proof:** (a) It follows from [23, Theorem 3.1].

(b) Let \( \phi(a, b) = a + b - \phi(a, b) + \tau \phi(a, b) \), where

\[
\phi_1(a, b) = gh(a, b) = \sqrt{a^2 + b^2 + (\tau_1 - 2)ab}, \quad \phi_2(a, b) = a + b_+,
\]

where \( g(a, b) = \sqrt{a^2 + b^2} \) and \( h(a, b) = (a + \frac{\tau_1 - 2}{2} b, \frac{\sqrt{(\tau_1 - 2)}}{2} b) \). Now, we show that \( \phi_1 \) and \( \phi_2 \) are C-regular. Since \( g(a, b) = \sqrt{a^2 + b^2} = ||(a, b)|| \) is convex and continuous, then \( g \) is C-regular. Because \( g \) is locally Lipschitz, \( h \) is continuously differentiable and \( h \) maps every neighborhood of \( u \) onto a dense subset of a neighborhood of \( h(u) \), it follows from [27, Theorem 2.178] that \( \partial \phi_1(a, b) = \partial g(h(a, b))h'(a, b) \) and \( \phi_1 \) is C-regular from [27, Corollary 2.179]. Since \( z_+ \) is convex and continuous, then \( \phi_2 \) is C-regular. So, it follows from [27, Proposition 2.174] that

\[
\partial \phi(a, b) = (1, 1) - \partial g(h(a, b))h'(a, b) + \tau_2(a_++b_+, b_+d_{a_+}).
\]

Since

\[
\partial g(h(a, b)) = \begin{cases} h(a, b)/||h(a, b)|| & \text{if } (a, b) \neq (0, 0) \\ (\xi, \eta) & \text{if } (a, b) = (0, 0) \end{cases}
\]

and

\[
h'(a, b) = \begin{pmatrix} 1 & \frac{\tau_1 - 2}{2} \\ 0 & \frac{\sqrt{(\tau_1 - 2)}}{2} \end{pmatrix}
\]

where \( ||(\xi, \eta)|| \leq 1 \), then we obtain the expression (10).

(c) From [10, Lemma 2.2], we know \( a + b - \sqrt{a^2 + b^2 + (\tau_1 - 2)ab} \) is strongly semismooth. Moreover, it is easy to see that the plus function \( z \to z_+ \) is strongly semismooth. Since the composition, product, sum of two strongly semismooth functions is strongly semismooth, then \( \phi \) is strongly semismooth on \( \mathbb{R}^2 \).

(d) For any \( (a, b) \in \mathbb{R}^2 \), we have

\[
\phi'(a, b) = \begin{cases} (1 - \frac{a + \frac{\tau_1 - 2}{2} b}{\sqrt{a^2 + (\tau_1 - 2)ab}} + \tau_2 b, 1 - \frac{b + \frac{\tau_1 - 2}{2} a}{\sqrt{a^2 + (\tau_1 - 2)ab}} + \tau_2 a) & \text{if } a > 0, \ b > 0 \\ (1 - \frac{a + \frac{\tau_1 - 2}{2} b}{\sqrt{a^2 + (\tau_1 - 2)ab}}, 1 - \frac{b + \frac{\tau_1 - 2}{2} a}{\sqrt{a^2 + (\tau_1 - 2)ab}}) & \text{if } a < 0 \ b < 0 \end{cases}
\]

Then, \( \phi(a, b) \) is continuously differentiable on \( \mathbb{R}^2 \).

(e) Let \( \{(a^k, b^k)\} \subset \mathbb{R}^2 \) be a sequence. We consider the following three cases.

Suppose that \( a^k \to \infty \) and \( b^k \to \infty \). For sufficiently large \( k \), we have

\[
a^k + b^k - \sqrt{(a^k)^2 + (b^k)^2 + (\tau_1 - 2)a^k b^k} \geq a^k + b^k - |a^k + b^k| = 0.
\]

This together with \( a_+^{b_+} b_+ \to \infty \) implies the result.

Suppose that \( a^k \to -\infty \). Then, \( a_+^{b_+} b_+ \to 0 \) and

\[
a^k + b^k - \sqrt{(a^k)^2 + (b^k)^2 + (\tau_1 - 2)a^k b^k} \leq \frac{4 - \tau_1}{2} a^k \to -\infty,
\]

which implies the result.

For the case \( b^k \to -\infty \), a similar analysis to the case \( a^k \to -\infty \) yields the result.

Since \( \Phi \) is (strongly) semismooth if and only if all component functions are (strongly) semismooth, we obtain the following result as an immediate consequence of Proposition 3.1.
Algorithm 3.1. The function $\Phi$ defined by (7) satisfies the following properties:

(a) $\Phi$ is semismooth.

(b) $\Phi$ is strongly semismooth if $F$ is LC$^q$ (i.e., the Jacobian of $F$ is locally Lipschitz continuous).

Since the generalized Jacobian is difficult to describe, we characterize the C-subdifferential of $\Phi$ below.

**Proposition 3.2.** For any $x \in \mathbb{R}^n$, we have

$$\partial_c \Phi(x) \subseteq D_a(x) + D_b(x)F'(x),$$

where $D_a(x) = \text{diag}\{a_i(x)\}$ and $D_b(x) = \text{diag}\{b_i(x)\}$ are diagonal matrices with entries $(a_i(x), b_i(x)) \in \partial \phi(x_i, F_i(x))$, where $\partial \phi(x_i, F_i(x))$ denotes the set from (10) being replaced by $(x_i, F_i(x))$.

**Proof:** By the definition of C-subdifferential, we have

$$\partial_c \Phi(x) = \partial^T \Phi_1(x) \times \cdots \times \partial^T \Phi_n(x).$$

Let $H_i(x) = (x_i, F_i(x))$. Then, $\Phi_i(x) = \phi H_i(x)$. Since $\phi$ is locally Lipschitz and $H_i$ is continuously differentiable, from [27, Theorem 2.178], we get

$$\partial \Phi_i(x) \subseteq \partial \phi(x_i, F_i(x)) H'_i(x) = \partial \phi(x_i, F_i(x)) \left( e_i^T F'_i(x) \right),$$

where $\partial \phi(x_i, F_i(x))$ denotes the set from (10) being replaced by $(x_i, F_i(x))$. Hence,

$$\partial_c \Phi(x) = (\partial^T \Phi_1(x), \ldots, \partial^T \Phi_n(x))^T \subseteq D_a(x) + D_b(x)F'(x)$$

where $D_a(x) = \text{diag}\{a_i(x)\}$ and $D_b(x) = \text{diag}\{b_i(x)\}$ are diagonal matrices with entries $(a_i(x), b_i(x)) \in \partial \phi(x_i, F_i(x))$.

We now provide a procedure to calculate an element of the C-subdifferential $\partial_c \Phi(x)$ at any point $x \in \mathbb{R}^n$. For simplicity, we follow the notation $\phi_1$ and $\phi_2$ as (11).

**Algorithm 3.1.** (Procedure to evaluate an element $V \in \partial_c \Phi(x)$)

**Step 0:** Let $x \in \mathbb{R}^n$ be given and $V_i$ denote the $i$-th row of a matrix $V \in \mathbb{R}^{n \times n}$.

**Step 1:** Set $S_1 = \{ i \mid x_i = F_i(x) = 0 \}$ and $S_2 = \{ i \mid x_i > 0, \ F_i(x) > 0 \}$.

**Step 2:** Set $z \in \mathbb{R}^n$ such that $z_i = 0$ for $i \notin S_1$ and $z_i = 1$ for $i \in S_1$.

**Step 3:** For $i \in S_1$, set

$$V_i = \left( 1 - \frac{z_i + \frac{n-2}{2} F_i'(x) z_i}{\phi_1(z_i, F_i'(x))} \right) e_i^T + \left( 1 - \frac{F_i'(x) z_i + \frac{n-2}{2} z_i}{\phi_1(z_i, F_i'(x))} \right) F'_i(x).$$

**Step 4:** For $i \in S_2$, set

$$V_i = \left( 1 - \frac{x_i + \frac{n-2}{2} F_i(x)}{\phi_1(x_i, F_i(x))} \right) e_i^T + \left( 1 - \frac{F_i(x) + \frac{n-2}{2} x_i}{\phi_1(x_i, F_i(x))} + \tau_i x_i \right) F'_i(x).$$

**Step 5:** For $i \notin S_1 \cup S_2$, set

$$V_i = \left( 1 - \frac{x_i + \frac{n-2}{2} F_i(x)}{\phi_1(x_i, F_i(x))} \right) e_i^T + \left( 1 - \frac{F_i(x) + \frac{n-2}{2} x_i}{\phi_1(x_i, F_i(x))} \right) F'_i(x).$$

The next result shows that the matrix $V$ calculated by Algorithm 3.1 is indeed an element from the C-subdifferential $\partial_c \Phi(x)$. 
Proposition 3.3. The element \( V \) calculated by Algorithm 3.1 is an element of the C-subdifferential \( \partial_C \Phi(x) \).

**Proof:** By the definition of \( \partial_C \Phi(x) \), we only need to consider \( V_i \in \partial \Phi_i(x) \).

Suppose \( i \in S_1 \). The mapping \( \phi_2 \) is C-regular with \( \partial \phi_2(x_i, F_i(x)) = \{(0,0)\} \). Let \( y^k = x + \varepsilon^k z \), where \( \varepsilon^k \) is a sequence of positive real numbers converging to zero. Then, \( y^k \to x, y^k > 0 \) and \( F_i(y^k) = F_i(x) + \varepsilon^k F'_i(h^k)z \) with \( h^k \) on the line segment from \( x \) to \( y^k \). Hence, the mapping \( x_i + F_i(x) - \phi_1(x_i, F_i(x)) \) is differentiable at \( y^k \) with its Jacobian

\[
1 - \frac{y^k_i + \frac{n-2}{2} F'_i(h^k)z}{\phi_1(y^k_i, F_i(y^k))} e_i^T + 1 - \frac{F_i(y^k) + \frac{n-2}{2} y^k_i}{\phi_1(y^k_i, F_i(y^k))} F'_i(y^k) \]

\[
= 1 - \frac{\varepsilon^k z_i + \frac{n-2}{2} \varepsilon^k F'_i(h^k)z}{\phi_1(\varepsilon^k z_i, \varepsilon^k F'_i(h^k)z)} e_i^T + 1 - \frac{\varepsilon^k F'_i(h^k)z + \frac{n-2}{2} \varepsilon^k z_i}{\phi_1(\varepsilon^k z_i, \varepsilon^k F'_i(h^k)z)} F'_i(y^k) \]

\[
\to 1 - \frac{z_i + \frac{n-2}{2} F'_i(x)z}{\phi_1(z_i, F'_i(x)z)} e_i^T + 1 - \frac{F'_i(x)z + \frac{n-2}{2} z_i}{\phi_1(z_i, F'_i(x)z)} F'_i(x) \]

So, (13) belongs to \( \partial \Phi_i(x) \) when \( i \in S_1 \).

For the case \( i \in S_2 \), it is obvious that (14) belongs to \( \partial \Phi_i(x) \).

Suppose \( i \notin S_1 \cup S_2 \). If \( x_i < 0 \) or \( F_i(x) < 0 \), the mapping \( \phi_2 \) is C-regular with \( \partial \phi_2(x_i, F_i(x)) = \{(0,0)\} \) and the mapping \( x_i + F_i(x) - \phi_1(x_i, F_i(x)) \) is differentiable at \( x \). Then, we have

\[
\partial \Phi_i(x) = \left\{ \left( 1 - \frac{x_i + \frac{n-2}{2} F_i(x)}{\phi_1(x_i, F_i(x))} \right) e_i^T + \left( 1 - \frac{F_i(x) + \frac{n-2}{2} x_i}{\phi_1(x_i, F_i(x))} \right) F'_i(x) \right\}. 
\]

If \( x_i = 0 \) and \( F_i(x) > 0 \). We choose a sequence \( \{y^k\} \subset \mathbb{R}^n \) such that \( y^k = x - \varepsilon^k e_i \), where \( \varepsilon^k \) is a sequence of positive real numbers converging to zero. Then, \( y^k_i < 0 \) and \( F_i(y^k) > 0 \) for sufficiently large \( k \). So, \( \Phi_i \) is continuously differentiable at these points \( y^k \) with

\[
\Phi'_i(y^k) = \left( 1 - \frac{y^k_i + \frac{n-2}{2} F'_i(y^k)}{\phi_1(y^k_i, F_i(y^k))} \right) e_i^T + \left( 1 - \frac{F_i(y^k) + \frac{n-2}{2} y^k_i}{\phi_1(y^k_i, F_i(y^k))} \right) F'_i(y^k). \tag{16}
\]

Taking the limit of the above equality gives the desired expression of \( V_i \) in Step 5.

If \( x_i > 0 \) and \( F_i(x) = 0 \). If \( F'_i(x) = 0 \), by Proposition 3.2, \( \partial \Phi_i(x) \subseteq a_i(x)e_i^T \). Since \( a_i(x) \) is single valued, we have

\[
\partial \Phi_i(x) = \left\{ \left( 1 - \frac{x_i + \frac{n-2}{2} F_i(x)}{\phi_1(x_i, F_i(x))} \right) e_i^T \right\}. 
\]

So, the expression of \( V_i \) in Step 5 is correct. Now, we consider the case \( F'_i(x) \neq 0 \). Given this situation, we define another sequence \( \{y^k\} \subset \mathbb{R}^n \) by \( y^k = x - \varepsilon^k (F'_i(x))^T \) where \( \varepsilon^k \) is a sequence of positive real numbers converging to zero. Since \( F \) is continuously differentiable, there is a vector sequence \( h^k \) on the open line segment from \( x \) to \( y^k \) with \( F_i(y^k) = F_i(x) - \varepsilon^k F'_i(h^k)(F'_i(x))^T \). For sufficiently large \( k \), we have \( y^k_i > 0 \) and \( F_i(y^k) < 0 \). Then \( \Phi_i \) is continuously differentiable at these points \( y^k \) with \( \Phi'_i(y^k) \) being equal to the expression on the right hand side of (16). Taking the limit of (16), we get the desired expression of \( V_i \) in Step 5.

Proposition 3.4. The function \( \varphi(a, b) \) defined by (9) satisfies the following properties:

(a) \( \varphi(a, b) \) is continuously differentiable on \( \mathbb{R}^2 \) with \( \varphi'(a, b) = \phi(a, b) \cdot V \) for any \( V \in \mathbb{R}^2 \). 

\[ \partial \phi(a, b) \]

\( (b) \ \varphi(a, b) = 0 \iff \varphi'(a, b) = 0 \iff \frac{\partial \varphi}{\partial a}(a, b) = 0 \iff \frac{\partial \varphi}{\partial b}(a, b) = 0. \]

\( (c) \ \frac{\partial \varphi}{\partial a}(a, b) \cdot \frac{\partial \varphi}{\partial b}(a, b) \geq 0. \)

**Proof:** (a) From [27, Theorem 2.178], it holds that \( \partial \varphi(a, b) \subseteq \text{co}\{\phi(a, b) \cdot \partial \phi(a, b)\} \). By simple calculation, we obtain

\[
\phi(a, b) \cdot \partial \phi(a, b) = \begin{cases} 
\phi(a, b) \left( \frac{\phi_1(a,b) - a - \frac{\tau_1 - 2}{\phi_1(a,b)} b}{\phi_1(a,b)} \right) + \tau_2 b, & \text{if } a > 0, b > 0 \\
0, & \text{if } a < 0 \text{ or } b < 0 \\
0, & \text{if } a = 0, b \geq 0 \\
0, & \text{if } a \geq 0, b = 0
\end{cases}
\]

Then, \( \text{co}\{\phi(a, b) \cdot \partial \phi(a, b)\} \) is single valued and hence \( \varphi'(a, b) = \phi(a, b) \cdot V \) for any \( V \in \partial \phi(a, b) \). From Corollary to Proposition 2.2.4 in [24], we know \( \varphi(a, b) \) is continuously differentiable everywhere. By straightforward calculation, we obtain \( (b) \) and \( (c) \) from (17).

**Proposition 3.5.** The function \( \Psi(x) \) defined by (8) is continuously differentiable with \( \Psi'(x) = \Phi(x)^T \partial_C \Phi(x) \).

**Proof:** By known rules on the calculus of generalized Jacobian, it holds that \( \partial \Phi_t^2(x) \subseteq \text{co}\{2\Phi_t(x) \cdot \partial \Phi_t(x)\} \). Since \( \Phi_t(x) \cdot \partial \Phi_t(x) \) is single valued, then \( \Psi'(x) = \sum \Phi_t(x) \cdot \partial \Phi_t(x) = \Phi(x)^T \partial_C \Phi(x) \) is single valued and \( \Psi(x) \) is continuously differentiable.

4. **Semismooth Newton Method.** This section deals with a semismooth Newton method based on the complementarity function \( \Phi \) and the merit function \( \Psi \). The convergence of the algorithm is also discussed.

**Algorithm 4.1.** *(A semismooth Newton method)*

**Step 0:** *(Initialization)*

Let \( \beta \in (0, 1), \sigma \in (0, \frac{1}{2}), p > 2, \rho > 0 \) and \( \varepsilon \geq 0 \). Choose any \( x^0 \in \mathbb{R}^n \). Set \( k = 0 \).

**Step 1:** *(Termination Check)*

If \( \|\Psi'(x^k)\| \leq \varepsilon \), stop.

**Step 2:** *(Search Direction Calculation)*

Choose \( V^k \in \partial_C \Phi(x^k) \) and let \( d^k \in \mathbb{R}^n \) be a solution of the following linear system of equations:

\[
V^k d = -\Phi(x^k). \tag{18}
\]

If we cannot find a solution \( d^k \) or if the descent test

\[
\Psi'(x^k) d^k \leq -\rho\|d^k\|^P \tag{19}
\]

is not satisfied, set \( d^k = -\Psi'(x^k)^T \).

**Step 3:** *(Line Search)*

Let \( l_k \) be the smallest nonnegative integer \( l \) such that

\[
\Psi(x^k + \beta^l d^k) \leq \Psi(x^k) + \sigma \beta^l \Psi(x^k) d^k. \tag{20}
\]

**Step 4:** Set \( x^{k+1} = x^k + \beta^k d^k \) and \( k = k + 1 \). Go to Step 1.

The algorithm is similar to the one in [10] but with some advantages: (a) the complementarity function has one more parameter which can be more flexible; (b) the condition to get the global convergence is much milder than the uniform P-property which is used in [10]; (c) we choose an element in \( \partial_C \Phi(x) \) instead of taking the generalized Jacobian \( \partial \Phi(x) \) which is difficult to calculate.
Definition 4.1. Given a solution $x^*$ of NCP ($F$), let $M = F'(x^*)$. Then $x^*$ is called an $R$-regular solution if $M_{aa}$ is nonsingular and the Schur-complement $M_{ab} - M_{ba}M_{aa}^{-1}M_{ab}$ has the $P$-property, where

$$\alpha = \{i \mid x_i^* > 0, F_i(x^*) = 0\}, \quad \beta = \{i \mid x_i^* = 0, F_i(x^*) = 0\} \quad \text{and} \quad \gamma = \{i \mid x_i^* = 0, F_i(x^*) > 0\}.$$ 

Theorem 4.1. If $x^*$ is an $R$-regular solution of NCP($F$), then all elements in $\partial_C\Phi(x^*)$ are nonsingular.

Proof: Due to Proposition 3.2, any element $V \in \partial_C\Phi(x^*)$ can be written as $D_a(x^*) + D_b(x^*)F'(x^*)$ for some nonnegative diagonal matrices $D_a$ and $D_b$. Without loss of generality, let

$$D_a = \text{diag}\{D_{a,a}, D_{a,b}, D_{a,\gamma}\}, \quad D_b = \text{diag}\{D_{b,a}, D_{b,b}, D_{b,\gamma}\}$$

and

$$M = F'(x^*) = \begin{pmatrix} M_{aa} & M_{a\beta} & M_{a\gamma} \\ M_{b\alpha} & M_{bb} & M_{b\gamma} \\ M_{\gamma\alpha} & M_{\gamma\beta} & M_{\gamma\gamma} \end{pmatrix}$$

where $D_{a,a} = (D_a)_{aa}$, etc. Also, we know that

$$D_{a,a} = 0_a, \quad D_{a,b} = \text{diag}\{1 - u_i\}_{\beta}, \quad D_{a,\gamma} = (2 - \frac{\tau_1}{2})I_\gamma + \tau_2\text{diag}\{F_i(x^*)s_i\}_{\gamma} \quad \text{(21)}$$

and

$$D_{b,a} = (2 - \frac{\tau_2}{2})I_\alpha + \tau_2\text{diag}\{x_i^*w_i\}_{\alpha}, \quad D_{b,b} = \text{diag}\{1 - (\frac{\tau_1}{2} - 1)u_i - \frac{\sqrt{\tau_1(4-\tau_1)}}{2}v_i\}_{\beta}, \quad D_{b,\gamma} = 0_{\gamma} \quad \text{(22)}$$

where $\|(u_i, v_i)\| \leq 1$ for all $i \in \beta$, $s_i \in [0, 1]$ for all $i \in \gamma$, $w_i \in [0, 1]$ for all $i \in \alpha$. Let $q \in \mathbb{R}^n$ be an arbitrary vector with $(D_a + D_b)q = 0$. It can be rewritten as:

$$(D_{a,a} + D_{a,b}M_{aa})q_a + D_{a,b}M_{a\beta}q_{\beta} + D_{b,a}M_{a\gamma}q_{\gamma} = 0,$$

$$D_{b,b}M_{b\alpha}q_{a} + (D_{a,b} + D_{b,b}M_{b\beta})q_{\beta} + D_{b,\gamma}M_{\gamma\beta}q_{\gamma} = 0,$$

$$D_{b,\gamma}M_{\gamma\alpha}q_{a} + D_{b,\gamma}M_{\gamma\beta}q_{\beta} + (D_{a,\gamma} + D_{b,\gamma}M_{\gamma\gamma})q_{\gamma} = 0.$$ 

Taking into account (21) and (22), we have

$$D_{b,a}M_{aa}q_{a} + D_{b,a}M_{a\beta}q_{\beta} + D_{b,a}M_{a\gamma}q_{\gamma} = 0, \quad \text{(23)}$$

$$D_{b,b}M_{b\alpha}q_{a} + (D_{a,b} + D_{b,b}M_{b\beta})q_{\beta} + D_{b,\beta}M_{\beta\gamma}q_{\gamma} = 0, \quad \text{(24)}$$

$$D_{a,\gamma}q_{\gamma} = 0.$$ 

Since the diagonal matrix $D_{a,\gamma}$ is positive, we obtain $q_{\gamma} = 0$. Hence, Equations (23) and (24) reduce to

$$D_{b,a}M_{aa}q_{a} + D_{b,a}M_{a\beta}q_{\beta} = 0, \quad \text{(25)}$$

$$D_{b,b}M_{b\alpha}q_{a} + (D_{a,b} + D_{b,b}M_{b\beta})q_{\beta} = 0. \quad \text{(26)}$$

Due to the nonsingularity of $D_{b,a}$ and $M_{aa}$, we directly obtain from (25) that

$$q_{a} = -M_{aa}^{-1}M_{a\beta}q_{\beta} \quad \text{(27)}$$

Substituting (27) into (26) yields that

$$[D_{a,b} + D_{b,b}(M_{b\beta} - M_{b\alpha}M_{aa}^{-1}M_{a\beta})]q_{\beta} = 0.$$ 

Let $N = M_{b\beta} - M_{b\alpha}M_{aa}^{-1}M_{a\beta}$. Then, we have

$$(D_{a,b})_{i}(q_{\beta})_{i} = -(D_{b,b})_{i}(Nq_{\beta})_{i}, \quad i \in \beta.$$ 

Without loss of generality, suppose

$$(D_{a,b})_{j} = 0, \quad j = 0, 1, \cdots, \beta_0 \quad \text{and} \quad (D_{a,b})_{k} > 0, \quad k = \beta_0 + 1, \cdots, |\beta|.$$
where \( \beta_0 = 0 \) means \( D_{a,\beta} \) is nonsingular. Then, \((D_{b,\beta})_j = 2 - \frac{q_j}{2} > 0 \), and hence, \((Nq_\beta)_j = 0 \) for \( j = 0, 1, \cdots, \beta_0 \). Therefore, \((q_\beta)_j(Nq_\beta)_j = 0 \), \( j = 0, \cdots, \beta_0 \) and
\[
(q_\beta)_k(Nq_\beta)_k = -\frac{(D_{b,\beta})_k}{(D_{a,\beta})_k} \cdot (Nq_\beta)_k^2 \leq 0, \quad k = \beta_0 + 1, \cdots, |\beta|.
\]
So, \( q_\beta \cdot Nq_\beta \leq 0 \). Since \( N \) has the P-property, we have \( q_\alpha = 0 \). From (27), we get \( q_\alpha = 0 \). Hence, \( V \in \partial_c \Phi(x^*) \) is nonsingular.

If the Jacobian \( F'(x^*) \) has the P-property, then every principal minor of \( F'(x^*) \) is nonsingular and the Schur-complement of every principal minor has the P-property. So, we have the following corollary.

**Corollary 4.1.** Suppose that the Jacobian \( F'(x^*) \) has the P-property, then all the elements in \( \partial_c \Phi(x^*) \) are nonsingular.

In order to guarantee the global convergence of the proposed method, we consider the following condition. It turns out that this condition is sufficient.

**Condition 4.1.** For any sequence \( \{x^k\} \) such that
\[
\|x^k\| \to \infty, \quad [-x^k]^+ < \infty, \quad [-F(x^k)]_+ < \infty,
\]
it holds
\[
\max_i (x^k)^+_i (F_i(x^k))^+_i \to \infty.
\]

**Theorem 4.2.** If \( F \) satisfies Condition 4.1, then the level sets
\[
L(c) = \{x \in \mathbb{R}^n \mid \Psi(x) \leq c\}
\]
are bounded for any fixed \( c \geq 0 \).

**Proof:** Assume on the contrary that there exists an unbounded sequence \( \{x^k\} \subset L(c) \) for some \( c \geq 0 \) such that \( \Psi(x^k) \leq c \) for all \( k \geq 0 \). By Proposition 3.1, there is no index \( i \) such that \( x^k_i \to -\infty \) or \( F_i(x^k) \to -\infty \). Since \( F \) satisfies Condition 4.1, there is a fixed index \( j \) such that \((x^k)^+_j(F_j(x^k))^+_j \to \infty \) at least on a subsequence. However, this implies \( \Psi(x^k) \) is unbounded. In fact,
\[
\Psi(x^k) = \frac{1}{2} \sum_{i=1}^{n} \phi^2(x^k_i, F_i(x^k))
\]
\[
\geq \frac{1}{2} \sum_{i=1}^{n} \left\{ \phi^2(x^k_i, F_i(x^k)) + \tau_2^2 [(x^k_i)^+_i (F_i(x^k))^+_i]^2 + 2 \tau_2 \phi(x^k_i, F_i(x^k)) (x^k_i)^+_i (F_i(x^k))^+_i \right\}
\geq \frac{1}{2} \tau_2^2 [(x^k_j)^+_j (F_j(x^k))^+_j]^2
\]
since \( \phi(x^k_i, F_i(x^k)) (x^k_i)^+_i (F_i(x^k))^+_i \) is nonnegative. Then, it contradicts with \( \Psi(x^k) \leq c \).

To the best of our knowledge, Condition 4.1 is the weakest assumption to guarantee bounded level sets for NCPs. Indeed, all the uniform P-property, \( R_0 \)-property and monotone property with a strictly feasible point satisfy this condition. This condition is much milder than the ones in [10, 23].

**Theorem 4.3.** Assume that \( x^* \) is a stationary point of \( \Psi \) such that the Jacobian \( F'(x^*) \) has the \( P_0 \)-property. Then, \( x^* \) is a solution of \( NCP(F) \).
Proof: By Proposition 3.2, we know
\[ \Psi'(x) = \Phi(x)^T \partial_c \Phi(x) \subseteq \Phi(x)^T [D_a(x) + D_b(x)F'(x)]. \]
Since \( \Phi(x)^T [D_a(x) + D_b(x)F'(x)] \) is single valued, then \( \Psi'(x) = \Phi(x)^T [D_a(x) + D_b(x)F'(x)] \). Suppose \( \Psi'(x^*) = 0 \) which means that
\[ [D_a(x^*) + (F'(x^*))^T D_b(x^*)] \Phi(x^*) = 0. \] (28)
We want to show that \( \Phi(x^*) = 0 \). Suppose the contrary. Consider the vector \( D_a(x^*) \Phi(x^*) \). Then, \( (D_a(x^*) \Phi(x^*))_i \neq 0 \) iff \( \Phi_i(x^*) \neq 0 \). In fact, if \( \Phi_i(x^*) \neq 0 \), \( (D_a(x^*) \Phi(x^*))_i \) can be zero iff \( a_i(x^*) = 0 \). However, \( \Phi_i(x^*) \neq 0 \) means that one of the following situations occurs: (a) \( x_i^* \neq 0 \) and \( F_i(x^*) \neq 0 \); (b) \( x_i^* = 0 \) and \( F_i(x^*) < 0 \); (c) \( x_i^* < 0 \) and \( F_i(x^*) = 0 \). In every case, it is obvious that \( a_i(x^*) > 0 \). So, \( (D_a(x^*) \Phi(x^*))_i \neq 0 \).

Similar reasoning can be repeated for the vector \( D_b(x^*) \Phi(x^*) \). Then, it is easy to verify that if \( \Phi(x^*) \neq 0 \), then \( D_a(x^*) \Phi(x^*) \) and \( D_b(x^*) \Phi(x^*) \) are both different from zero and have their nonzero elements in the same positions; such nonzero elements have the same sign. However, then for (28) to hold it would be necessary for \( (F'(x^*))^T \) to revert the sign of all the nonzero elements of \( D_b(x^*) \Phi(x^*) \), which contradicts with the fact that the transpose of \( F'(x^*) \) has the \( P_0 \)-property since \( F'(x^*) \) has the \( P_0 \)-property.

The following lemma follows from the proof of [28, Proposition 3.1], which is useful in proving the quadratic convergence.

Lemma 4.1. Suppose \( G : R^n \rightarrow R^n \) is locally Lipschitzian. If all \( V \in \partial_C G(x) \) are nonsingular, then there is a neighborhood \( N(x) \) of \( x \) and a constant \( C \) such that for any \( y \in N(x) \) and any \( V \in \partial_C G(y) \), \( V \) is nonsingular and \( \|V^{-1}\| \leq C \).

Theorem 4.4. The following results hold for Algorithm 4.1:
(a) Any accumulation point is a stationary point of \( \Psi \). Furthermore, if \( F \) satisfies Condition 4.1, such an accumulation point exists.
(b) Let \( x^* \) be an accumulation point such that \( F'(x^*) \) has the \( P_0 \)-property. Then, \( x^* \) is a solution of NCP(\( F \)).
(c) If \( x^* \) is an \( R \)-regular solution of NCP(\( F \)), then the whole sequence generated by Algorithm 4.1 converges to \( x^* \), and the rate of convergence is \( Q \)-superlinear (\( Q \)-quadratic if \( F \) is an LC\(^4 \) function).

Proof: (a) Suppose \( x^k \rightarrow x^* \) but \( \Psi'(x^*) \neq 0 \). If \( d^k \) is a solution of \( V^kd = -\Phi(x^k) \), then
\[ \Psi(x^{k+1}) \leq \Psi(x^k) + \sigma \beta \Psi'(x^k)d^k = \Psi(x^k) - 2\sigma \beta \Psi(x^k) + (1 - 2\sigma \beta)^2 \Psi(x^k). \]
Taking the limit we have \( \Psi(x^*) = 0 \) and hence \( \Psi'(x^*) = 0 \) which is a contradiction. If \( d^k = -\Psi'(x^k) \), then
\[ \Psi(x^{k+1}) \leq \Psi(x^k) + \sigma \beta \Psi'(x^k)d^k = \Psi(x^k) - \sigma \beta \|\Psi'(x^k)\|^2. \]
Taking the limit we have \( \Psi'(x^*) = 0 \) which is a contradiction. Hence, any accumulation point is a stationary point of \( \Psi \). Furthermore, it follows from Theorem 4.2 that if \( F \) satisfies Condition 4.1, such an accumulation point exists.
(b) By (a), we know any accumulation point \( x^* \) is a stationary point of \( \Psi \). Then, from Theorem 4.3 that \( x^* \) is a solution of NCP(\( F \)) if \( F'(x^*) \) has the \( P_0 \)-property.
(c) If \( x^* \) is an \( R \)-regular solution of NCP(\( F \)), then the elements in \( \partial_\Phi(x^*) \subseteq \partial_c \Phi(x^*) \) are nonsingular by Theorem 4.1. From Theorem 3.1, \( \Phi \) is semismooth everywhere. Then \( x^* \) is a BD-regular solution of \( \Phi(x) = 0 \). So, it follows from [29, Theorem 11] that \( \{x^k\} \) converges to \( x^* \).

To consider the superlinear convergence, we need to show that eventually the direction is always the solution of (18). For sufficiently large \( k \), by Lemma 4.1, we have that any
$V^k \in \partial \phi(x^k)$ is nonsingular and $\| (V^k)^{-1} \| \leq C$ ($C$ is a constant). Hence, system (18) always admits a solution for $k$ sufficiently large.

Since $\Phi$ is semismooth, we have
\[
\| x^k + d^k - x^* \| = \| x^k - (V^k)^{-1} \Phi(x^k) - x^* \|
= \| - (V^k)^{-1} (\Phi(x^k) - \Phi(x^*) - V^k(x^k - x^*)) \|
= o(\| x^k - x^* \|).
\]
Then, by the Lipschitz continuity of $\Phi$, we get
\[
\| \Phi(x^k + d^k) - \Phi(x^*) \| = O(\| x^k + d^k - x^* \|) = o(\| x^k - x^* \|).
\]
Therefore, for all sufficiently large $k$, we have $x^{k+1} = x^k + d^k$ and
\[
\| x^{k+1} - x^* \| = o(\| x^k - x^* \|)
\]
proving the suplinear convergence. The quadratic convergence is similar.

5. Application to Frictionless Contact Problem. The frictionless contact problem of linear elastic bodies with small deformation is considered in this section. The problem can be approximated using the finite element (FE) technique. The FE-approximation is in the form of the following complementarity problem, where both the displacement vector $u \in \mathbb{R}^{n_d}$ and the contact pressure vector $p \in \mathbb{R}^{n_c}$ are treated as unknowns [30].

\[
K u - T^T p = f
\]
\[
-p_i \geq 0, \quad (i = 1, \ldots, n_c) \tag{30}
\]
\[
(-T u + h)_i \geq 0, \quad (i = 1, \ldots, n_c) \tag{31}
\]
\[
-p_i (-T u + h)_i = 0, \quad (i = 1, \ldots, n_c) \tag{32}
\]
where $K \in \mathbb{R}^{n_d \times n_d}$ is the stiffness matrix with the P-property, $T \in \mathbb{R}^{n_c \times n_d}$ is the constraint matrix, $f$ denotes the vector of external loads and $h$ denotes the initial gap vector.

The complementarity problem presented by Equations (29)-(32) can be rewritten as follows:
\[
\Phi(p) = \begin{pmatrix}
\phi(-p_1, F_1(-p)) \\
\vdots \\
\phi(-p_{n_c}, F_{n_c}(-p))
\end{pmatrix} = 0 \tag{33}
\]
where the complementarity function $\phi$ is defined by (5) and $F(-p) = h - TK^{-1}(f + T^T p)$. Since $K$ has the P-property, $F$ and $F'$ have the $P_0$-property. Then, we can use the above semismooth Newton method to solve the nonsmooth Equation (33). Instead of solving the original frictionless contact problem, we can obtain the solution by applying the semismooth Newton method.

6. Conclusions. In this paper, we investigate a class of NCP-functions which contains the famous penalized FB function as a special case. The class of NCP-function and its merit function is shown to enjoy some favorite properties, such as strong semismoothness, continuous differentiability and the nonsingularity of the element in $C$-subdifferential. In particular, we present an exact expression of the generalized gradient for the NCP-function. The merit function is shown to have the level boundedness property under mild condition. Based on these results, we present a semismooth Newton method to solve the NCP and this proposed method has some advantages. It is globally convergent and quadratically convergent under suitable conditions. Applying this method, we can solve the frictionless contact problem.
Acknowledgement. The authors are very grateful to the associate editor and the referees for their helpful and valuable suggestions. The research is supported partially by National Natural Science Foundation of China (Grant No. 10871056 and 10971150) and by Science Research Foundation in Harbin Institute of Technology (Grant No. HIT200708).

REFERENCES


