The Australian Journal of Mathematical Analysis and Applications

http://ajmaa.org

Volume 9, Issue 1, Article 11, pp. 1-12, 2012

EQUIVALENCE OF THE NONSMOOTH NONLINEAR COMPLEMENTARITY PROBLEMS TO UNCONSTRAINED MINIMIZATION

M. A. TAWHID AND J. L. GOFFIN

Received 1 February, 2010; accepted 21 June, 2010; published 20 March, 2012.

DEPARTMENT OF MATHEMATICS AND STATISTICS, SCHOOL OF ADVANCED TECHNOLOGIES AND MATHEMATICS, THOMPSON RIVERS UNIVERSITY, 900 MCGILL ROAD, PO BOX 3010, KAMLOOPS, BC V2C 5N3 CANADA

ALEXANDRIA UNIVERSITY AND EGYPT JAPAN UNIVERSITY OF SCIENCE AND TECHNOLOGY, ALEXANDRIA-EGYPT
mtawhid@tru.ca

FACULTY OF MANAGEMENT, MCGILL UNIVERSITY, 1001 SHERBROOKE STREET WEST, MONTREAL, QUEBEC, H3A 1G5 CANADA.
Jean-Louis.Goffin@McGill.ca

ABSTRACT. This paper deals with nonsmooth nonlinear complementarity problem, where the underlying functions are nonsmooth which admit the H-differentiability but not necessarily locally Lipschitzian or directionally differentiable. We consider a reformulation of the nonlinear complementarity problem as an unconstrained minimization problem. We describe H-differentials of the associated penalized Fischer-Burmeister and Kanzow and Kleinmichel merit functions. We show how, under appropriate $\mathbf{P}_0$, semimonotone ($\mathbf{E_0}$), $\mathbf{P}$, positive definite, and strictly semimonotone ($\mathbf{E}$) -conditions on an H-differential of f, finding local/global minimum of a merit function (or a 'stationary point' of a merit function) leads to a solution of the given nonlinear complementarity problem. Our results not only give new results but also unify/extend various similar results proved for C^1.

Key words and phrases: H-Differentiability, semismooth-functions, locally Lipschitzian, nonlinear complementarity problem, NCP function, merit function, unconstrained minimization.

2000 *Mathematics Subject Classification.* 90C33, 49J52.

ISSN (electronic): 1449-5910

The research of the first author is supported in part by the Natural Sciences and Engineering Research Council of Canada. The research of the second author is supported by the Natural Sciences and Engineering Research Council of Canada.

1. INTRODUCTION

Given a function $f : R^n \to R^n$. Consider the nonlinear complementarity problem, denoted by the NCP(f), which can be defined as

$$\text{find } \bar{x} \in R^n \text{ such that } \bar{x} \geq 0\,,\;\; f(\bar{x}) \geq 0 \;\text{ and }\; \langle f(\bar{x}), \bar{x}\rangle = 0.$$

This problem arises in many applications, e.g., in operations research, economic models, and engineering sciences (such as dynamic rigid- body model, nonlinear obstacle problems, traffic equilibrium problems, optimal control problems, taxation and subsidies, invariant capital stock, and spatial price equilibria), see [6], [14] for a more detail description. One of the popular approaches to solve the NCP is to reformulate it as an unconstrained minimization problem whose global minima are coincident with the solution of the NCP and the objective function of this unconstrained minimization problem is called a merit function for the NCP [4], [5], [8], [10], [16], [17], [19], [20], [30]. Most of the merit functions in these references based on the implicit Lagrangian function [4], [16], [20], [30], the square Fischer-Burmeister function [5], [10], [16], [17], [19], and for other NCP functions see, e.g., the survey paper [7].
In this paper we consider nonsmooth nonlinear complementarity problem NCP(f) when the underlying functions are nonsmooth which admit the H-differentiability but not necessarily locally Lipschitzian or directionally differentiable. By considering an NCP function $\Phi : R^n \to R^n$ associated with NCP(f) so that

$$\Phi(\bar{x}) = 0 \Leftrightarrow \;\bar{x} \text{ solves NCP}(f),$$

and the corresponding merit function

$$\Psi(x) := \sum_{i=1}^{n} \Phi_i(x). \tag{1.1}$$

We consider the following NCP functions:

(1)
$$\phi_1(a,b) := \frac{1}{2}\left[\phi_\beta(a,b)\right]^2 = \frac{1}{2}\left[a+b-\sqrt{(a-b)^2+\beta ab}\right]^2$$
where $\phi_1, \phi_\beta : R^2 \to R$. NCP function ϕ_β was proposed by Kanzow and Kleinmichel [18]
$$\phi_\beta(a,b) := a+b-\sqrt{(a-b)^2+\beta ab} \tag{1.2}$$
where β is a fixed parameter in $(0,4)$. We note that when $\beta = 2$, ϕ reduces to the Fischer-Burmeister function, while as $\beta \to 0$, ϕ_β becomes
$$\phi(a,b) := a+b-\sqrt{(a-b)^2}\;(= 2\min\{a,b\}).$$
Then the merit function associated to ϕ_1 at $\bar{x}$ is defined as in (1.1) where
$$\begin{aligned}\Phi_i(\bar{x}) &= \phi_1(\bar{x}_i, f_i(\bar{x})) = \tfrac{1}{2}\left[\phi_\beta(\bar{x}_i, f_i(\bar{x}))\right]^2\\ &:= (1/2)\left[\bar{x}_i + f_i(\bar{x}) - \sqrt{(\bar{x}_i - f_i(\bar{x}))^2 + \beta\bar{x}_i f_i(\bar{x})}\right]^2.\end{aligned} \tag{1.3}$$

(2)
$$\phi_2(a,b) = \frac{1}{2}\left[\phi_\lambda(a,b)\right]^2 := \frac{1}{2}\left[\lambda\phi_{FB}(a,b) + (1-\lambda)a_+ b_+\right]^2 \tag{1.4}$$
where $\phi_2, \phi_\lambda : R^2 \to R$. NCP function ϕ_λ is called the penalized Fischer-Burmeister function [1]
$$\phi_\lambda(a,b) := \lambda\phi_{FB}(a,b) + (1-\lambda)a_+ b_+ \tag{1.5}$$

where ϕ_{FB} is called Fischer-Burmeister function, $a_+ = \max\{0, a\}$ and $\lambda \in (0, 1)$ is a fixed parameter. Then its merit function associated to ϕ_2 at $\bar{x}$ is defined as in (1.1) where

$$\begin{aligned} \Phi_i(\bar{x}) &= \phi_2(\bar{x}_i, f_i(\bar{x})) = \tfrac{1}{2}\left[\phi_\lambda(\bar{x}_i, f_i(\bar{x}))\right]^2 \\ &:= \tfrac{1}{2}\left[\lambda\phi_{FB}(\bar{x}_i, f_i(\bar{x})) + (1-\lambda)\bar{x}_{i+}\, f_i(\bar{x})_+\right]^2. \end{aligned} \tag{1.6}$$

In this paper, we describe H-differentials of the square Kanzow and Kleinmichel function and the square penalized Fischer-Burmeister function, and their merit functions. Also, we show how, under appropriate $\mathbf{P}_0$, semimonotone ($\mathbf{E_0}$), $\mathbf{P}$, positive definite, and strictly semimonotone ($\mathbf{E}$) -conditions on an H-differential of f, finding local/global minimum of Ψ (or a 'stationary point' of Ψ) leads to a solution of the given nonlinear complementarity problem. Our results unify/extend various similar results proved in the literature for C^1 and semismooth functions [1], [18].

Our approach relies on the concepts of H-differentiability and H-differential of a function [13] because of the following reasons: H-differentiability implies continuity, any superset of an H-differential is an H-differential, and H-differentials enjoy simple sum, product, chain rules, a mean value theorem and a second order Taylor-like expansion, and inverse and implicit function theorems, see [11], [12], [13]; a H-differentiable function need not be locally Lipschitzian function or directionally differentiable; the Fréchet derivative of a differentiable function, the Clarke generalized Jacobian of a locally Lipschitzian function [2], the Bouligand differential of a semismooth function [23], and the C-differential of Qi [24] are particular instances of H-differential; moreover, the closure of the H-differential is an approximate Jacobian [15].

For some applications of H-differentiability to optimization problems, nonlinear complementarity problems and variational inequalities, see e.g. [29], [28].

2. PRELIMINARIES

Throughout this paper, we regard vectors in R^n as column vectors. We denote the inner-product between two vectors x and y in R^n by either $x^T y$ or $\langle x, y\rangle$. Vector inequalities are interpreted componentwise. For a matrix A, A_i denotes the ith row of A. For a differentiable function $f : R^n \to R^m$, $\nabla f(\bar{x})$ denotes the Jacobian matrix of f at $\bar{x}$.

Definition 2.1. A function $\phi : R^2 \to R$ is called an NCP function if

$$\phi(a, b) = 0 \Leftrightarrow ab = 0, a \geq 0,\, b \geq 0.$$

We call ϕ a nonnegative NCP function if $\phi(a, b) \geq 0$ on R^2. For the problem NCP(f), we define

$$\Phi(x) = \left[\ \phi(x_1, f_1(x)) \cdots \phi(x_i, f_i(x)) \cdots \phi(x_n, f_n(x))\ \right]^T \tag{2.1}$$

and, call $\Phi(x)$ an NCP function for NCP(f). We call Φ a nonnegative NCP function for NCP(f) if ϕ is nonnegative.

We need the following definitions from [3], [22].

Definition 2.2. A matrix $A \in R^{n\times n}$ is called
(a) $\mathbf{P_0}$ ($\mathbf{P}$) if $\forall x \in R^n, x \neq 0$, there exists i such that $x_i \neq 0$ and $x_i\,(Ax)_i \geq 0$ (> 0) or equivalently, every principle minor of A is nonnegative (respectively, positive).
(b) semimonotone ($\mathbf{E_0}$) (strictly semimonotone ($\mathbf{E}$))-matrix if

$$\forall x \in R^n_+,\ x \neq 0,\ \text{there exists } i \text{ such that } x_i\,(Ax)_i \geq 0\ (> 0).$$

Definition 2.3. For a function $f : R^n \to R^n$, we say that f is a
(i) *monotone* if

$$\langle f(x) - f(y), x - y\rangle \geq 0 \quad \text{for all } x, y \in R^n.$$

(ii) $\mathbf{P_0}(\mathbf{P})$-function if, for any $x \neq y$ in R^n,

$$\max_{\{i:x_i \neq y_i\}} (x-y)_i[f(x)-f(y)]_i \geq 0 \ (>0). \tag{2.2}$$

We note that every monotone (strictly monotone) function is a $\mathbf{P}_0(\mathbf{P})$-function.

The following result is from [22], [27].

Theorem 2.1. *Under each the following conditions, $f : R^n \to R^n$ is a $\mathbf{P}_0(\mathbf{P})$-function.*

(a) *f is Fréchet differentiable on R^n and for every $x \in R^n$, the Jacobian matrix $\nabla f(x)$ is a $\mathbf{P}_0(\mathbf{P})$-matrix.*

(b) *f is locally Lipschitzian on R^n and for every $x \in R^n$, the generalized Jacobian $\partial f(x)$ consists of $\mathbf{P}_0(\mathbf{P})$-matrices.*

(c) *f is semismooth on R^n (in particular, piecewise affine or piecewise smooth) and for every $x \in R^n$, the Bouligand subdifferential $\partial_B f(x)$ consists of $\mathbf{P}_0(\mathbf{P})$-matrices.*

(d) *f is H-differentiable on R^n and for every $x \in R^n$, an H-differential $T_f(x)$ consists of $\mathbf{P}_0(\mathbf{P})$-matrices.*

The following definition and examples from Gowda and Ravindran [13].

Definition 2.4. Given a function $f : \Omega \subseteq R^n \to R^m$ where Ω is an open set in R^n and $x^* \in \Omega$, we say that a nonempty subset $T(x^*)$ (also denoted by $T_f(x^*)$) of $R^{m\times n}$ is an *H-differential of* f at x^* if for every sequence $\{x^k\} \subseteq \Omega$ converging to x^*, there exist a subsequence $\{x^{k_j}\}$ and a matrix $A \in T(x^*)$ such that

$$f(x^{k_j}) - f(x^*) - A(x^{k_j} - x^*) = o(||x_j^k - x^*||). \tag{2.3}$$

We say that f is H-differentiable at x^* if f has an H-differential at x^*.

Remarks

As noted in [29], it is easily seen that if a function $f : \Omega \subseteq R^n \to R^m$ is H-differentiable at a point $\bar{x}$, then there exist a constant $L > 0$ and a neighbourhood $B(\bar{x}, \delta)$ of $\bar{x}$ with

$$||f(x) - f(\bar{x})|| \leq L||x - \bar{x}||, \ \forall x \in B(\bar{x}, \delta). \tag{2.4}$$

Conversely, if condition (2.4) holds, then $T(\bar{x}) := R^{m\times n}$ can be taken as an H-differential of f at $\bar{x}$. We thus have, in (2.4), an alternate description of H-differentiability. But, as we see in the sequel, it is the identification of an appropriate H-differential that becomes important and relevant. Clearly any function locally Lipschitzian at $\bar{x}$ will satisfy (2.4). For real valued functions, condition (2.4) is known as the 'calmness' of f at $\bar{x}$. This concept has been well studied in the literature of nonsmooth analysis (see [26], Chapter 8).

Example 2.2. Let $f : R^n \to R^m$ be Fréchet differentiable at $x^* \in R^n$ with Fréchet derivative matrix (= Jacobian matrix derivative) $\{\nabla f(x^*)\}$ such that

$$f(x) - f(x^*) - \nabla f(x^*)(x - x^*) = o(||x - x^*||).$$

Then f is H-differentiable with $\{\nabla f(x^*)\}$ as an H-differential.

Example 2.3. Let $f : \Omega \subseteq R^n \to R^m$ be locally Lipschitzian at each point of an open set Ω. For $x^* \in \Omega$, define the Bouligand subdifferential of f at x^* by

$$\partial_B f(x^*) = \{\lim \nabla f(x^k) : x^k \to x^*, x^k \in \Omega_f\}$$

where Ω_f is the set of all points in Ω where f is Fréchet differentiable. Then, the (Clarke) generalized Jacobian [2]

$$\partial f(x^*) = co\partial_B f(x^*)$$

is an H-differential of f at x^*.

Example 2.4. Consider a locally Lipschitzian function $f : \Omega \subseteq R^n \to R^m$ that is semismooth at $x^* \in \Omega$ [21], [23], [25]. This means for any sequence $x^k \to x^*$, and for any $V_k \in \partial f(x^k)$,

$$f(x^k) - f(x^*) - V_k(x^k - x^*) = o(||x^k - x^*||).$$

Then the Bouligand subdifferential

$$\partial_B f(x^*) = \{\lim \nabla f(x^k) : x^k \to x^*, x^k \in \Omega_f\}.$$

is an H-differential of f at x^*. In particular, this holds if f is piecewise smooth, i.e., there exist continuously differentiable functions $f_j : R^n \to R^m$ such that

$$f(x) \in \{f_1(x), f_2(x), \ldots, f_J(x)\} \quad \forall x \in R^n.$$

Example 2.5. Let $f : R^n \to R^n$ be C-differentiable [24] in a neighborhood D of x^*. This means that there is a compact upper semicontinuous multivalued mapping $x \mapsto T(x)$ with $x \in D$ and $T(x) \subset R^{n\times n}$ satisfying the following condition at any $a \in D$: For any $V \in T(x)$,

$$f(x) - f(a) - V(x - a) = o(||x - a||).$$

Then, f is H-differentiable at x^* with $T(x^*)$ as an H-differential.

Remark 2.1. While the Fréchet derivative of a differentiable function, the Clarke generalized Jacobian of a locally Lipschitzian function [2], the Bouligand differential of a semismooth function [23], and the C-differential of a C-differentiable function [24] are particular instances of H-differential, the following simple example, taken from [11], shows that an H-differentiable function need not be locally Lipschitzian or directionally differentiable.

Example 2.6. Consider on R,

$$f(x) = x\sin(\frac{1}{x}) \text{ for } x \neq 0 \text{ and } f(0) = 0.$$

Then f is H-differentiable on R with

$$T(0) = [-1, 1] \text{ and } T(c) = \{\sin(\frac{1}{c}) - \frac{1}{c}cos(\frac{1}{c})\} \text{ for } c \neq 0.$$

We note that f is not locally Lipschitzian around zero. We also see that f is neither Fréchet differentiable or directionally differentiable.

3. H-DIFFERENTIALS OF SOME NCP/MERIT FUNCTIONS

In this section, we consider an NCP function Φ corresponding to NCP(f) and its merit function $\Psi := \sum_{i=1}^n \Phi_i$.

Theorem 3.1. *Suppose that Φ is H-differentiable at $\bar{x}$ with $T_\Phi(\bar{x})$ as an H-differential. Then $\Psi := \sum_{i=1}^n \Phi_i$ is H-differentiable at $\bar{x}$ with an H-differential given by*

$$T_\Psi(\bar{x}) = \{e^T B : B \in T_\Phi(\bar{x})\}.$$

Proof. To describe an H-differential of Ψ, let $\theta(x) = x_1 + \cdots + x_n$. Then $\Psi = \theta \circ \Phi$. Now by the chain rule for H-differentiability, we have $T_\Psi(\bar{x}) = (T_\theta \circ T_\Phi)(\bar{x})$ as an H-differential of Ψ at $\bar{x}$. Since $T_\theta(\bar{x}) = \{e^T\}$ where e is the vector of ones in R^n, we have

$$T_\Psi(\bar{x}) = \{e^T B : B \in T_\Phi(\bar{x})\}.$$

This completes the proof. ∎

Now we describe the H-differentials of the merit functions associated to square Kanzow and Kleinmichel function and square penalized Fischer-Burmeister function.

Example 3.2. Suppose that $f : R^n \to R^n$ has an H-differential $T(\bar{x})$ at $\bar{x} \in R^n$. Consider the associated square Kanzow and Kleinmichel function

$$\Phi(\bar{x}) := (1/2)\left[\bar{x} + f(\bar{x}) - \sqrt{(\bar{x} - f(\bar{x}))^2 + \beta\bar{x}\, f(\bar{x})}\right]^2, \tag{3.1}$$

where all the operations are performed componentwise. Let

$$J(\bar{x}) = \{i : f_i(\bar{x}) = 0 = \bar{x}_i\}.$$

Then the H-differential of Φ in (3.1) is given by

$$T_\Phi(\bar{x}) = \{VA + W : \ (A, V, W, d) \in \Gamma\},$$

where Γ is the set of all quadruples (A, V, W, d) with $A \in T(\bar{x})$, $||d|| = 1$, $V = diag(v_i)$ $W = diag(w_i)$ are diagonal matrices with

$$v_i = \begin{cases} \phi_\beta(\bar{x}_i, f_i(\bar{x}))\left[1 - \frac{-2(\bar{x}_i - f_i(\bar{x})) + \beta\bar{x}_i}{2\sqrt{(\bar{x}_i - f_i(\bar{x}))^2 + \beta\bar{x}_i f_i(\bar{x})}}\right] & \text{when } i \notin J(\bar{x}) \\ \phi_\beta(d_i, A_i d)\left[1 - \frac{-2(d_i - A_i d) + \beta d_i}{2\sqrt{(d_i - A_i d)^2 + \beta d_i (A_i d)}}\right] & \text{when } i \in J(\bar{x}) \text{ and } [(d_i - A_i d)^2 + \beta d_i (A_i d) > 0 \\ \text{arbitrary} & \text{when } i \in J(\bar{x}) \text{ and } (d_i - A_i d)^2 + \beta d_i (A_i d) = 0, \end{cases}$$

(3.2)

$$w_i = \begin{cases} \phi_\beta(\bar{x}_i, f_i(\bar{x}))\left[1 - \frac{2(\bar{x}_i - f_i(\bar{x})) + \beta f_i(\bar{x})}{2\sqrt{(\bar{x}_i - f_i(\bar{x}))^2 + \beta\bar{x}_i f_i(\bar{x})}}\right] & \text{when } i \notin J(\bar{x}) \\ \phi_\beta(d_i, A_i d)\left[1 - \frac{2(d_i - A_i d) + \beta A_i d}{2\sqrt{(d_i - A_i d)^2 + \beta d_i (A_i d)}}\right] & \text{when } i \in J(\bar{x}) \text{ and } (d_i - A_i d)^2 + \beta d_i (A_i d) > 0 \\ \text{arbitrary} & \text{when } i \in J(\bar{x}) \text{ and } (d_i - A_i d)^2 + \beta d_i (A_i d) = 0. \end{cases}$$

We can describe the H-differential of Φ in a way similar to the calculation and analysis of Examples 5-7 in [29]. By Theorem 3.1, the H-differential $T_\Psi(\bar{x})$ of $\Psi(\bar{x})$ consists of all vectors of the form $v^T A + w^T$ with $A \in T(\bar{x})$, v and w are columns vectors with entries defined by (3.2).

Example 3.3. Suppose that $f : R^n \to R^n$ has an H-differential $T(\bar{x})$ at $\bar{x} \in R^n$. Consider the associated square penalized Fischer-Burmeister function

$$\Phi(\bar{x}) := \frac{1}{2}\left[\lambda\phi_{FB}(\bar{x}, f(\bar{x})) + (1 - \lambda)\bar{x}_+ f(\bar{x})_+\right]^2. \tag{3.3}$$

where ϕ_{FB} is called Fischer-Burmeister function, $a_+ = \max\{0, a\}$ and $\lambda \in (0, 1)$ is a fixed parameter, and all the operations are performed componentwise. Let

$$J(\bar{x}) = \{i : f_i(\bar{x}) = 0 = \bar{x}_i\} \text{ and } K(\bar{x}) = \{i : \bar{x}_i > 0, \ f_i(\bar{x}) > 0\}.$$

For Φ in (3.3), a straightforward calculation shows that an H-differential is given by

$$T_\Phi(\bar{x}) = \{VA + W : \ (A, V, W, d) \in \Gamma\},$$

where Γ is the set of all quadruples (A, V, W, d) with $A \in T(\bar{x})$, $||d|| = 1$, $V = diag(v_i)$ and $W = diag(w_i)$ are diagonal matrices with

$$
v_i = \begin{cases} \phi_\lambda(\bar{x}_i, f_i(\bar{x})) \left[\lambda\left(1 - \frac{f_i(\bar{x})}{\sqrt{\bar{x}_i^2 + f_i(\bar{x})^2}}\right) + (1-\lambda)\bar{x}_i\right] & \text{when } i \in K(\bar{x}) \\ \phi_\lambda(d_i, A_i d) \left[\lambda\left(1 - \frac{A_i d}{\sqrt{d_i^2 + (A_i d)^2}}\right)\right] & \text{when } i \in J(\bar{x}) \text{ and } d_i^2 + (A_i d)^2 > 0 \\ \phi_\lambda(\bar{x}_i, f_i(\bar{x})) \left[\lambda\left(1 - \frac{f_i(\bar{x})}{\sqrt{\bar{x}_i^2 + f_i(\bar{x})^2}}\right)\right] & \text{when } i \notin J(\bar{x}) \cup K(\bar{x}) \\ \text{arbitrary} & \text{when } i \in J(\bar{x}) \text{ and } d_i^2 + (A_i d)^2 = 0, \end{cases}
$$

(3.4)

$$
w_i = \begin{cases} \phi_\lambda(\bar{x}_i, f_i(\bar{x})) \left[\lambda\left(1 - \frac{\bar{x}_i}{\sqrt{\bar{x}_i^2 + f_i(\bar{x})^2}}\right) + (1-\lambda)f_i(\bar{x})\right] & \text{when } i \in K(\bar{x}) \\ \phi_\lambda(d_i, A_i d) \left[\lambda\left(1 - \frac{d_i}{\sqrt{d_i^2 + (A_i d)^2}}\right)\right] & \text{when } i \in J(\bar{x}) \text{ and } d_i^2 + (A_i d)^2 > 0 \\ \phi_\lambda(\bar{x}_i, f_i(\bar{x})) \left[\lambda\left(1 - \frac{\bar{x}_i}{\sqrt{\bar{x}_i^2 + f_i(\bar{x})^2}}\right)\right] & \text{when } i \notin J(\bar{x}) \cup K(\bar{x}) \\ \text{arbitrary} & \text{when } i \in J(\bar{x}) \text{ and } d_i^2 + (A_i d)^2 = 0. \end{cases}
$$

The above calculation relies on the observation that the following is an H-differential of the one variable function $z \mapsto z_+$ at any $\bar{z}$:

$$
\Delta(\bar{z}) = \begin{cases} \{1\} & \text{if } \bar{z} > 0 \\ \{0,1\} & \text{if } \bar{z} = 0 \\ \{0\} & \text{if } \bar{z} < 0. \end{cases}
$$

Using Theorem 3.1, the H-differential $T_\Psi(\bar{x})$ of $\Psi(\bar{x})$ consists of all vectors of the form $v^T A + w^T$ with $A \in T(\bar{x})$, v and w are columns vectors with entries defined by (3.4).

4. MINIMIZING THE MERIT FUNCTION

For a given H-differentiable function $f : R^n \to R^n$, consider the associated NCP function Φ (as in Examples 3.2-3.3) and the corresponding merit function $\Psi := \sum_{i=1}^n \Phi_i$. It should be recalled that

$$
\bar{x} \text{ solves NCP}(f) \;\Leftrightarrow \Phi(\bar{x}) = 0 \Leftrightarrow \Psi(\bar{x}) = 0.
$$

The following lemma will be needed in our results. The proof is similar to Lemma 3.1 in [10].

Lemma 4.1. *Suppose that $f : R^n \to R^n$ is H-differentiable at $\bar{x}$ with $T(\bar{x})$ as an H-differential. Suppose that Φ is defined as in Examples 3.2-3.3, H-differentiable with an H-differential $T_\Phi(\bar{x})$ is given by*

$$
T_\Phi(\bar{x}) = \{VA + W : \; A \in T(\bar{x}), V = diag(v_i) \text{ and } W = diag(w_i)\}, \tag{4.1}
$$

and Ψ is H-differentiable with an H-differential $T_\Psi(\bar{x})$. Then Φ is nonnegative and the following properties hold:

$$
\left.\begin{array}{rl} \textbf{(i)} & \bar{x} \textit{ solves NCP(f)} \;\Leftrightarrow \Phi(\bar{x}) = 0. \\ \textbf{(ii)} & \textit{For } i \in \{1, \dots, n\},\; v_i\, w_i \geq 0. \\ \textbf{(iii)} & \textit{For } i \in \{1, \dots, n\},\; \Phi_i(\bar{x}) = 0 \Leftrightarrow (v_i, w_i) = (0, 0). \\ \textbf{(iv)} & \textit{If } 0 \in T_\Psi(\bar{x}), \textit{then } \Phi(\bar{x}) = 0 \Leftrightarrow v = 0. \end{array}\right\} \tag{4.2}
$$

In the subsequent subsections, we show that under appropriate conditions on an H-differential of f, a vector $\bar{x}$ is a solution of the NCP(f) if and only if zero belongs to $T_\Psi(\bar{x})$.

4.1. Under P_0-conditions.

Theorem 4.2. *Suppose that $f : R^n \to R^n$ is H-differentiable at $\bar{x}$ with an H-differential $T(\bar{x})$. Assume that Φ is defined as in Examples **??**-3.3. Suppose that $\Psi := \sum_{i=1}^n \Phi_i$ is H-differentiable at $\bar{x}$ with an H-differential given by*

$$T_\Psi(\bar{x}) = \{v^T A + w^T : (A, v, w) \in \Omega, \text{ with } v_i\, w_i > 0 \text{ whenever } \Phi_i(\bar{x}) \neq 0\}$$

where Ω is the set all triples (A, v, w) with $A \in T(\bar{x})$, v and w vectors in R^n satisfying properties (iii) and (iv) in (4.2).

Further suppose that $T(\bar{x})$ consists of $\mathbf{P}_0$-matrices. Then

$$\Phi(\bar{x}) = 0 \Leftrightarrow 0 \in T_\Psi(\bar{x}).$$

Proof. Suppose that $0 \in T_\Psi(\bar{x})$, so that for some $v^T A + w^T \in T_\Psi(\bar{x})$,

$$0 = v^T A + w^T$$

yielding $A^T v + w = 0$. Note that for any index i, $\Phi_i(\bar{x}) \neq 0 \Leftrightarrow\ v_i \neq 0$ (by property (iv) in (4.2) and $v_i w_i > 0$ when $\Phi_i(\bar{x}) \neq 0$) in which case $v_i(A^T v)_i = -v_i\, w_i < 0$ contradicting the $\mathbf{P}_0$-property of A. We conclude that $\Phi(\bar{x}) = 0$. Conversely, suppose that $\Phi(\bar{x}) = 0$. Then by property (iii) in (4.2) and the description of $T_\Psi(\bar{x})$, we have $T_\Psi(\bar{x}) = \{0\}$. ∎

As consequences of the above theorem, we state the results for the square penalized Fischer-Burmeister function for simplicity. However, it is possible to state a general result for any NCP function satisfying the assumptions of Theorem 4.2.

Corollary 4.3. *Let $f : R^n \to R^n$ be differentiable and $\Phi(x)$ be the square penalized Fischer-Burmeister function and $\Psi := \sum_{i=1}^n \Phi_i$. If f is $\mathbf{P_0}$-function, then $\bar{x}$ is a local minimizer to Ψ if and only if $\bar{x}$ solves NCP(f).*

In view of Example 2.4, if f is locally Lipschitzian with $T(\bar{x}) = \partial f(\bar{x})$, the above theorem reduces to the following result.

Corollary 4.4. *Let $f : R^n \to R^n$ be locally Lipschitzian. Let Φ be the square penalized Fischer-Burmeister function and $\Psi := \sum_{i=1}^n \Phi_i$. Further suppose that $\partial f(\bar{x})$ consists of $\mathbf{P_0}$-matrices.*
Then

$$\Psi(\bar{x}) = 0 \Leftrightarrow 0 \in \partial\Psi(\bar{x}).$$

Proof. In fact, by taking $T_f(x) = \partial f(x)$ in Theorem 4.2 and noting $\partial\Psi(x) \subseteq T_\Psi(x)$ for all x, we have the proof. ∎

4.2. Under semimonotone ($\mathbf{E_0}$)-conditions.

Theorem 4.5. *Suppose that $f : R^n \to R^n$ is H-differentiable at $\bar{x}$ with an H-differential $T(\bar{x})$. Assume that Φ is defined as in Examples 3.2-3.3. Suppose that $\Psi := \sum_{i=1}^n \Phi_i(\bar{x})$ is H-differentiable at $\bar{x}$ with an H-differential given by*

$$T_\Psi(\bar{x}) = \{v^T A + w^T : (A, v, w) \in \Omega, \text{ with } v_i > 0,\ w_i > 0 \text{ whenever } \Phi(\bar{x})_i \neq 0\}.\}$$

where Ω is the set all triples (A, v, w) with $A \in T(\bar{x})$, v and w vectors in R^n satisfying properties (iii) and (iv) in (4.2).Further suppose that $T(\bar{x})$ consists of semimonotone ($\mathbf{E_0}$)-matrices. Then

$$\Phi(\bar{x}) = 0 \Leftrightarrow 0 \in T_\Psi(\bar{x}).$$

Proof. Suppose that $\Phi(\bar{x}) = 0$. Then by property (iii) in (4.2)and the description of $T_\Psi(\bar{x})$, we have $T_\Psi(\bar{x}) = \{0\}$. Conversely, suppose that $0 \in T_\Psi(\bar{x})$. Then $v^T A + w^T = 0 \Rightarrow A^T v + w = 0$. We claim that $\Phi(\bar{x}) = 0$. Suppose, if possible, $\Phi(\bar{x}) \neq 0$. Then by property (iv) in (4.2), $v \neq 0$. Since $T(\bar{x})$ consists of $\mathbf{E_0}$-matrices and $A \in T(\bar{x})$, there exists an index i such that $0 \neq \Phi_i, 0 \neq v_i > 0$ and $0 \leq v_i(Av)_i$. By the fact , $v_i\, w_i > 0$, we have $0 \leq v_i(Av)_i = -v_i\, w_i < 0$ which is a contradiction. Hence $\Phi(\bar{x}) = 0$. ∎

Remark 4.1. We note that in Examples 3.2-3.3 if $\bar{x}$ is a strictly feasible point of NCP(f), i.e., $\bar{x}_i > 0$ and $f(\bar{x}_i) > 0$ for all $i \in \{1, \dots, n\}$, then we have $v_i > 0$ and $w_i > 0$.

4.3. **Under P-conditions.**

Theorem 4.6. *Suppose that $f : R^n \to R^n$ is H-differentiable at $\bar{x}$ with an H-differential $T(\bar{x})$. Assume that Φ is defined as in Examples 3.2-3.3. Suppose that $\Psi := \sum_{i=1}^n \Phi_i(\bar{x})$ is H-differentiable at $\bar{x}$ with an H-differential given by*

$$T_\Psi(\bar{x}) = \{v^T A + w^T : (A, v, w) \in \Omega\}$$

where Ω is the set all triples (A, v, w) with $A \in T(\bar{x})$, v and w vectors in R^n satisfying properties (ii), (iii), and (iv) in (4.2).

Further suppose that $T(\bar{x})$ consists of $\mathbf{P}$-matrices. Then

$$\Phi(\bar{x}) = 0 \Leftrightarrow 0 \in T_\Psi(\bar{x}).$$

Proof. To see this, suppose that $0 \in T_\Psi(\bar{x})$. Then $v^T A + w^T = 0 \Rightarrow A^T v + w = 0$. We claim that $\Phi(\bar{x}) = 0$. Suppose, if possible, $\Phi(\bar{x}) \neq 0$. Then by property (iv) in (4.2), $v \neq 0$. Since $T(\bar{x})$ consists of $\mathbf{P}$-matrices and $A \in T(\bar{x})$, there exists an index i such that $v_i \neq 0$ and $0 < v_i(Av)_i$. By property (ii) in (4.2), $v_i\, w_i \geq 0$. But $0 < v_i(Av)_i = -v_i\, w_i \leq 0$ which is a contradiction. Hence $\Phi(\bar{x}) = 0$. Conversely, suppose that $\Phi(\bar{x}) = 0$. Then by property (iii) in (4.2) and the description of $T_\Psi(\bar{x})$, we have $T_\Psi(\bar{x}) = \{0\}$. ∎

4.4. **Under positive-definite-conditions.**

Theorem 4.7. *Suppose that $f : R^n \to R^n$ is H-differentiable at $\bar{x}$ with an H-differential $T(\bar{x})$. Assume that Φ is defined as in Examples 3.2-3.3. Suppose that $\Psi := \sum_{i=1}^n \Phi_i(\bar{x})$ is H-differentiable at $\bar{x}$ with an H-differential given by*

$$T_\Psi(\bar{x}) = \{v^T A + w^T : (A, v, w) \in \Omega\}$$

where Ω is the set all triples (A, v, w) with $A \in T(\bar{x})$, v and w vectors in R^n satisfying properties (ii), (iii), and (iv) in (4.2).

Further suppose that $T(\bar{x})$ consists of positive-definite matrices. Then

$$\Phi(\bar{x}) = 0 \Leftrightarrow 0 \in T_\Psi(\bar{x}).$$

Proof. Suppose that $\Phi(\bar{x}) = 0$. Then by property (iii) in (4.2) and the description of $T_\Psi(\bar{x})$, we have $T_\Psi(\bar{x}) = \{0\}$. Conversely, suppose that $0 \in T_\Psi(\bar{x})$. Then $v^T A + w^T = 0 \Rightarrow A^T v + w = 0$. We claim that $\Phi(\bar{x}) = 0$. Suppose, if possible, $\Phi(\bar{x}) \neq 0$. Then by property (iv) in (4.2), $v \neq 0$. Since $T(\bar{x})$ consists of positive definite matrices and $A \in T(\bar{x})$,

$0 < \langle v, Av \rangle$. By property (ii) in (4.2), $\langle v, w \rangle \geq 0$. But $0 < \langle v, Av \rangle = -\langle v, w \rangle \leq 0$ which is a contradiction. Hence $\Phi(\bar{x}) = 0$. ∎

Remark 4.2. Since every positive definite matrix is also a $\mathbf{P}$-matrix, the proof of Theorem 4.7 follows from Theorem 4.6. However, we gave a general proof of Theorem 4.7.

4.5. **Under strictly semi-monotone (E)-conditions.**

Theorem 4.8. *Suppose that $f : R^n \to R^n$ is H-differentiable at $\bar{x}$ with an H-differential $T(\bar{x})$. Assume that Φ is defined as in Examples 3.2-3.3. Suppose that $\Psi := \sum_{i=1}^{n} \Phi_i(\bar{x})$ is H-differentiable at $\bar{x}$ with an H-differential given by*

$$T_\Psi(\bar{x}) = \{v^T A + w^T : (A, v, w) \in \Omega, \text{ with } v_i > 0,\ w_i \geq 0 \text{ whenever } \Phi(\bar{x})_i \neq 0\}.\}$$

where Ω is the set all triples (A, v, w) with $A \in T(\bar{x})$, v and w vectors in R^n satisfying properties (iii) and (iv) in (4.2).

Further suppose that $T(\bar{x})$ consists of $\mathbf{E}$*-matrices.Then*

$$\Phi(\bar{x}) = 0 \Leftrightarrow 0 \in T_\Psi(\bar{x}).$$

Proof. Suppose that $0 \in T_\Psi(\bar{x})$. Then $v^T A + w^T = 0 \Rightarrow A^T v + w = 0$. We claim that $\Phi(\bar{x}) = 0$. Suppose, if possible, $\Phi(\bar{x}) \neq 0$. Then by property (iv) in (4.2), $v \neq 0$.

Since $T(\bar{x})$ consists of $\mathbf{E}$-matrices and $A \in T(\bar{x})$, there exists an index i such that $0 < v_i(Av)_i$. By property (ii) in (4.2), $v_i w_i \geq 0$. But $0 < v_i(Av)_i = -v_i w_i \leq 0$ which is a contradiction. Hence $\Phi(\bar{x}) = 0$. Conversely, suppose $\Phi(\bar{x}) = 0$. Then by property (iii) in (4.2) and the description of $T_\Psi(\bar{x})$, we have $T_\Psi(\bar{x}) = \{0\}$. ∎

Remark 4.3. Note that in Examples 3.2-3.3 if $\bar{x}$ is a feasible point of NCP(f), i.e., $\bar{x}_i \geq 0$ and $f(\bar{x}_i) \geq 0$ for all $i \in \{1, \ldots, n\}$, then we have $v_i > 0$ and $w_i \geq 0$.

Remark 4.4. We note that we can use the so-called derivative-free descent method which does not need any explicit derivative of the function f involved in NCP to solve nonsmooth NCP. When f is C^1, Yamashita and Fukushima [30] and Geiger and Kanzow [10] proposed a descent method for minimizing the unconstrained minimization which does not require to compute the derivative of f and Ψ. Fischer [8] obtained similar results when f is locally Lipschitzian and the authors in [9] proved the global convergence when the underlying functions admit approximated Jacobians. Since H-differentiability implies continuity [13], we beleive that the algorithm in [9] will be applicable to our nonsmooth NCP when NCP function is based on Fischer-Burmeister function.

Concluding Remarks

We considered a nonlinear complementarity problem corresponding to H-differentiable functions, with an associated NCP function Φ and a merit function $\Psi(\bar{x}) := \sum_{i=1}^{n} \Phi_i(\bar{x})$. First, we described the H-differential of the squared Kanzow and Kleinmichel function, the squared penalized Fischer-Burmeister function and their merit functions. We gave conditions under which every global/local minimum or a stationary point of Ψ is a solution of NCP(f).

For nonlinear complementarity problem based on the squared Kanzow and Kleinmichel function/the squared penalized Fischer-Burmeister function, our results not only give new results but also recover/extend various results stated for nonlinear complementarity problem when the underlying functions are continuously differentiable (locally Lipschitzian, semismooth, C-differentiable) functions. Moreover, we considered NCP function on the basis of the squared penalized Fischer-Burmeister function and Φ in Example 3.3 which appeared to be new. Our results are applicable to any nonnegative NCP functions satisfying Lemma 4.1 in the paper and we can state a very general result for any nonnegative NCP function satisfying Lemma 4.1, but for simplicity, we consider the squared Kanzow and Kleinmichel function and the squared penalized Fischer-Burmeister function. The NCP functions in this paper are called unrestricted NCP functions. The restricted NCP functions (subject to some constraints such as $x \geq 0$) and their merit functions will be considered by the author as a future work.

We note here that similar methodologies under H-differentiability can be carried out for the following NCP functions:

(1)
$$\phi_1(a,b) := \frac{1}{2}\min^2\{a,b\}.$$

(2)
$$\phi_2(a,b) := \frac{1}{2}[(ab)^2 + \min^2\{0,a\} + \min^2\{0,b\}].$$

(3)
$$\phi_3(a,b) := a\,b + \frac{1}{2\alpha}\left[\max^2\{0, a-\alpha b\} - x^2 + \max^2\{0, b-\alpha a\} - b^2\right],$$

where $\alpha > 1$ is any fixed parameter.

(4)
$$\phi_4(a,b) := \frac{1}{2}[a + b - \sqrt{a^2+b^2}]^2.$$

REFERENCES

[1] B. CHEN, X. CHEN, and C. KANZOW, A penalized Fischer-Burmeister NCP-function: theoretical investigation and numerical results, *Mathematical Programming*, Vol. **88**, pp. 211-216, 2000.

[2] F. H. CLARKE, *Optimization and Nonsmooth Analysis*, Wiley, New York, 1983 (reprinted by SIAM , Philadelphia, PA, 1990).

[3] R. W. COTTLE, J.-S. PANG and R. E. STONE, *The Linear Complementarity Problem*, Academic Press, Boston, 1992.

[4] F. FACCHINEI, andC. KANZOW, On unconstrained and constrained stationary points of the implicit Lagrangian, *Journal of Optimization Theory and Applications*, Vol. **92**, pp. 99-115, 1997.

[5] F. FACCHINEI, and J. SOARES, A new merit function for nonlinear complementarity problems and related algorithm, *SIAM Journal on Optimization*, Vol. **7**, pp. 225-247, 1997.

[6] M. C. FERRIS and J.-S. PANG, Engineering and economic applications of complementarity problems, *SIAM Review. 39*, pp. 669-713, 1997.

[7] A. FISCHER, *An NCP-Function and Its Use for the Solution of Complementarity Problems, Recent Advances in Nonsmooth Optimization*, Edited by D.Z. Du, L.Qi, and R.S. Womersley, World Scientific Publishers, Singapore, Republic of Singapore, pp. 88-105, 1995.

[8] A. FISCHER, Solution of monotone complementarity problems with locally Lipschitzian functions, *Mathematical Programming*, Vol. **76**, pp. 513-532, 1997.

[9] A. FISCHER, V. JEYAKUMAR and D. T. LUC, Solution point characterizations and convergence analysis of a descent algorithm for nonsmooth continuous complementarity problems, *Journal of Optimization Theory and Applications*, Vol. **110**, pp. 493-513, 2001.

[10] C. GEIGER, and C. KANZOW, On the resolution of monotone complementarity problems, *Computational Optimization*, Vol. **5**, pp. 155-173, 1996.

[11] M. S. GOWDA, *A Note on H-Differentiable Functions*, Department of Mathematics and Statistics, University of Maryland, Baltimore County, Baltimore, Maryland 21250, November, 1998.

[12] M. S. GOWDA, Inverse and implicit function theorems for H-differentiable and semismooth functions, *Optimization Methods and Software*, Vol. **19**, pp. 443-461, 2004.

[13] M. S. GOWDA, and G. RAVINDRAN, Algebraic univalence theorems for nonsmooth functions, *Journal of Mathematical Analysis and Applications*, Vol. **252**, pp. 917-935, 2000.

[14] P. T. HARKER, and J.-S. PANG, Finite dimension variational inequality and nonlinear complementarity problems: a survey of theory, algorithms and applications, *Mathematical Programming*, Vol. **48,** pp. 161-220, 1990.

[15] J. JEYAKUMAR and D. T. LUC, Approximate Jacobians matrices for nonsmooth continuous maps and C^1-optimization, *SIAM Journal on Control and Optimization,* Vol. **36**, pp. 1815-1832, 1998.

[16] H. JIANG, Unconstrained minimization approaches to nonlinear complementarity problems, *Journal of Global Optimization,* Vol. **9**, pp. 169-181, 1996.

[17] C. KANZOW, Nonlinear complementarity as unconstrained optimization, *Journal of Optimization Theory and Applications*, Vol. **88**, pp. 139-155, 1996.

[18] 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.

[19] T. D. LUCA, F. FACCHINEI and C. KANZOW, A semismooth equation approach to the solution of nonlinear complementarity problems, *Mathematical Programming*, Vol. **75**, pp. 407-439, 1996.

[20] O. L. MANGASARIAN and M. V. SOLODOV, Nonlinear complementarity as unconstrained and constrained minimization, *Mathematical Programming*, Vol. **62**, pp. 277-297, 1993.

[21] R. MIFFLIN, Semismooth and semiconvex functions in constrained optimization, *SIAM Journal on Control and Optimization,* Vol. **15**, pp. 952-972, 1977.

[22] J. J. MORÉ and W. C. RHEINBOLDT, On P- and S- functions and related classes of N-dimensional nonlinear mappings, *Linear Algebra and its Applications*, Vol. **6**, pp. 45-68, 1973.

[23] L. QI, Convergence analysis of some algorithms for solving nonsmooth equations, *Mathematics of Operations Research*, Vol. **18**, pp. 227-244, 1993.

[24] L. QI, *C-Differentiability, C-Differential Operators and Generalized Newton Methods*, Research Report, School of Mathematics, The University of New South Wales, Sydney, New South Wales 2052, Australia, January 1996.

[25] L. QI and J. SUN, A nonsmooth version of Newton's method, *Mathematical Programming,* Vol. **58**, pp. 353-367, 1993.

[26] R. T. ROCKAFELLAR and R J.-B. WETS, *Variational Analysis*, Grundlehren der Mathematischen Wissenschaften, 317, Springer-Verlag, Berlin, Germany, 1998.

[27] Y. SONG, M. S. GOWDA and G. RAVINDRAN, On characterizations of $\mathbf{P}$- and $\mathbf{P}_0$-properties in nonsmooth functions, *Mathematics of Operations Research*, Vol. **25**, pp. 400-408, 2000.

[28] N. A. TAWHID, On the local uniqueness of solutions of variational inequalities under H-differentiability, *Journal of Optimization Theory and Applications*, Vol. **113**, pp. 149-164, 2002.

[29] M. A. TAWHID, and M. S. GOWDA, On two applications of H-differentiability to optimization and complementarity problems, *Computational Optimization and Applications*, Vol. **17**, pp. 279-299, 2000.

[30] N. YAMASHITA and M. FUKUSHIMA, On stationary points of the implicit lagrangian for nonlinear complementarity Problems, *Journal of Optimization Theory and Applications*, Vol. **84**, pp. 653-663, 1995.