Indag. Mathem., N.S., 14 (1), 81–101 March 31, 2003

Constrained second order optimization on non-archimedean fields

by K. Shamseddine and V. Zeidan†

Department of Mathematics, Michigan State University, East Lansing, MI 48824, USA
e-mail: khodr@math.msu.edu
Department of Mathematics, Michigan State University, East Lansing, MI 48824, USA
e-mail: zeidan@math.msu.edu

Communicated by Prof. T.A. Springer at the meeting of November 25, 2002

ABSTRACT

Constrained optimization on non-Archimedean fields is presented. We formalize the notion of a tangent plane to the surface defined by the constraints making use of an implicit function Theorem similar to its real counterpart. Then we derive necessary and sufficient conditions of second order for the existence of a local minimizer of a function subject to a set of equality and inequality constraints, based on a concept of continuity and differentiability that is stronger than the conventional one.

1. INTRODUCTION

In this paper, optimization over equalities and inequalities on non-Archimedean fields will be considered. We first review some basic terminology and facts about non-Archimedean fields. So let K be a totally ordered non-Archimedean field extension of $\mathbb{R}$. We introduce the following terminology.

Definition 1.1 ($\sim, \approx, \ll, H, \lambda$). For $x, y \in K$, we say $x \sim y$ if there exist $n, m \in \mathbb{N}$ such that $n|x| > |y|$ and $m|y| > |x|$; for nonnegative $x, y \in K$, we say that x is infinitely smaller than y and write $x \ll y$ if $nx < y$ for all $n \in \mathbb{N}$, and we say that x is infinitely small if $x \ll 1$ and x is finite if $x \sim 1$; finally, we say that x is ap-

1991 *Mathematics Subject Classification.* 26E30, 12J25, 11D88, 78M50, 80M50.
Key words and phrases. Non-Archimedean calculus, constraints, optimization, local extrema.
†Research supported by the National Science Foundation under grant # DMS-0072598.

proximately equal to y and write $x \approx y$ if $x \sim y$ and $|x - y| \ll |x|$. We also set $\lambda(x) = [x]$, the class of x under the equivalence relation $\sim$.

The set of equivalence classes H (under the relation $\sim$) is naturally endowed with an addition via $[x] + [y] = [x \cdot y]$ and an order via $[x] < [y]$ if $|y| \ll |x|$ (or $|x| \gg |y|$), both of which are readily checked to be well-defined. It follows that $(H, +, <)$ is a totally ordered group, often referred to as the Hahn group or skeleton group, whose neutral element is the class [1]. It follows from the above that the projection λ from K to H is a valuation.

The Theorem of Hahn [5] provides a complete classification of non-Archimedean extensions of $\mathbb{R}$ in terms of their skeleton groups. In fact, invoking the axiom of choice it is shown that the elements of any such field K can be written as formal power series over its skeleton group H with real coefficients, and the set of appearing 'exponents' forms a well-ordered subset of H. The coefficient of the qth power in the Hahn representation of a given x will be denoted by $x[q]$, and the number d will be defined by $d[1] = 1$ and $d[q] = 0$ for $q \neq 1$. It is easy to check that $0 < d^q \ll 1$ if and only if $q > 0$, and $d^q \gg 1$ if and only if $q < 0$; moreover, $x \approx x[\lambda(x)]d^{\lambda(x)}$ for all $x \neq 0$.

From general properties of formal power series fields [10,12], it follows that if H is divisible then K is algebraically closed; that is, every polynomial of odd degree over K has at least one root in K. For a general overview of the algebraic properties of formal power series fields, we refer to the comprehensive overview by Ribenboim [13], and for an overview of the related valuation theory the book by Krull [6]. A thorough and complete treatment of ordered structures can also be found in [11].

Throughout this paper, $\mathcal{N}$ will denote any totally ordered non-Archimedean field extension of $\mathbb{R}$ that is complete in the order topology and whose skeleton group H is Archimedean; i.e. a subgroup of $\mathbb{R}$. The smallest such field is the field of the formal Laurent series whose skeleton group is $\mathbb{Z}$; and the smallest such field that is also algebraically closed is the Levi-Civita field $\mathcal{R}$, first introduced in [7,8]. In this case $H = \mathbb{Q}$, and for any element $x \in \mathcal{R}$, the set of exponents in the Hahn representation of x is a left-finite subset of $\mathbb{Q}$, i.e. below any rational bound r there are only finitely many exponents.

The Levi-Civita field $\mathcal{R}$ is of particular interest because of its practical usefulness. Since the supports of the elements of $\mathcal{R}$ are left-finite, it is possible to represent these numbers on a computer [1]. Having infinitely small numbers, the errors in classical numerical methods can be made infinitely small and hence irrelevant in all practical applications. One such application is the computation of derivatives of real functions representable on a computer [15], where both the accuracy of formula manipulators and the speed of classical numerical methods are achieved.

In the light of such practical usefulness of infinitely small numbers, it is natural to study optimization questions on non-Archimedean fields with the hope to extend the methods mentioned in the previous paragraph to find local extrema of functions as closely as allowed by machine precision.

The following example shows that continuity or even differentiability of a function on a closed and bounded subset of $\mathcal{N}$ or $\mathcal{N}^n$ do not necessarily entail that the function assumes a maximum or a minimum on the set.

Example 1.2. Let $f : [-1, 1] \to \mathcal{N}$ be given by $f(x) = x - x[0]$.

Then f is continuous on $[-1, 1]$ and differentiable on $(-1, 1)$ with $f'(x) = 1$ for all $x \in (-1, 1)$. In fact, for all $y \neq x$ in $[-1, 1]$ satisfying $|y - x| \ll 1$, we have that $(y - x)[0] = 0$ and hence

$$\begin{aligned} \frac{f(y) - f(x)}{y - x} &= \frac{y - y[0] - x + x[0]}{y - x} = \frac{y - x - (y[0] - x[0])}{y - x} \\ &= \frac{y - x - (y - x)[0]}{y - x} = \frac{y - x}{y - x} = 1, \end{aligned}$$

which shows that f is differentiable at x for all $x \in (-1, 1)$ with $f'(x) = 1$. However, f assumes neither a maximum nor a minimum on $[-1, 1]$. The set $f([-1, 1])$ is bounded above by any positive real number and below by any negative real number; but it has neither a least upper bound nor a greatest lower bound.

Also, contrary to the real case, the following example shows that a function that is $2k$-times differentiable on an open interval (a, b) containing the point x_0, with $f^{(j)}(x_0) = 0$ for all $j \in \{1, \ldots, 2k - 1\}$ and $f^{(2k)}(x_0) \neq 0$, need not have a local extremum at x_0.

Example 1.3. Let $g : (-1, 1) \to \mathcal{N}$ be given by $g(x)[q] = x[q/3]$ and let $f : (-1, 1) \to \mathcal{N}$ be given by $f(x) = g(x) - x^4$. Then g is infinitely often differentiable on $(-1, 1)$ with $g^{(j)}(x) = 0$ for all $j \in \mathbb{N}$ and for all $x \in (-1, 1)$ [16]. It follows that f is four times differentiable on $(-1, 1)$ with $f'(0) = f''(0) = f'''(0) = 0$ and $f^{(4)}(0) = -24$. Now let $x \in (-1, 1)$ be such that $0 < |x| \ll 1$. Then $g(x) \approx x[\lambda(x)]d^{3\lambda(x)}$ and $x^4 \approx (x[\lambda(x)])^4 d^{4\lambda(x)} \ll |x[\lambda(x)]| d^{3\lambda(x)}$. Thus, $f(x) \approx x[\lambda(x)]d^{3\lambda(x)}$. It follows that $f(x) > f(0) = 0$ if $0 < x \ll 1$ and $f(x) < f(0)$ if $0 < -x \ll 1$; and hence f has no local extremum at 0.

The difficulties presented in Example 1.2 and Example 1.3 are due to the total disconnectedness of the field in the order topology [14] and makes the study of optimization more involved than in the real case. Thus, a stronger smoothness criterion is needed to study optimization on $\mathcal{N}$. In [16], we considered unconstrained one-dimensional optimization on the field $\mathcal{N}$, using the notion of continuity and differentiability based on the derivate concept [2]. In this paper, we generalize the concepts of derivate continuity and differentiability to higher dimensions. Then we use that to derive necessary and sufficient conditions of second order for the existence of a minimum of an $\mathcal{N}$-valued function on $\mathcal{N}^n$ subject to constraints.

We will consider the problem of minimizing a function $f : \mathcal{N}^n \to \mathcal{N}$, subject to the following set of constraints:

$$(1.1)\qquad \begin{cases} h_1(\vec{x}) = 0 \\ \quad\vdots \\ h_m(\vec{x}) = 0 \end{cases} \text{and} \begin{cases} g_1(\vec{x}) \leq 0 \\ \quad\vdots \\ g_p(\vec{x}) \leq 0 \end{cases},$$

where all the functions in Equation (1.1) are from $\mathcal{N}^n$ to $\mathcal{N}$. A point $\vec{x}_0 \in \mathcal{N}^n$ will be said to be a feasible point if it satisfies the constraints in Equation (1.1). Before deriving necessary and sufficient conditions for a feasible point $\vec{x}_0$ to be a local minimizer of f, we first review the concept of derivate differentiability [2,14] and extend the concept to higher dimensions.

2. DERIVATE CONTINUITY AND DIFFERENTIABILITY

In this section, we review the definitions of derivate continuity and differentiability in one dimension, as well as some related results that are useful for our purposes here, and we refer the interested reader to [2,14] for a more detailed study. Then we generalize these notions of continuity and differentiability to higher dimensions.

Definition 2.1. Let $D \subset \mathcal{N}$ be open and let $f : D \to \mathcal{N}$. Then we say that f is derivate continuous on D if there exists $M \in \mathcal{N}$, called a Lipschitz constant of f on D, such that

$$\left| \frac{f(y) - f(x)}{y - x} \right| \leq M \text{ for all } x \neq y \text{ in } D.$$

It follows immediately from Definition 2.1 that if $f : D \to \mathcal{N}$ is derivate continuous on D then f is uniformly continuous (in the conventional sense) on D.

Remark 2.2. It is clear that the concept of derivate continuity in Definition 2.1 coincides with that of uniform Lipschitz continuity when restricted to $\mathbb{R}$. We chose to call it derivate continuity here so that, after having defined derivate differentiability in Definition 2.3 and higher order derivate differentiability in Definition 2.5, we can think of derivate continuity as derivate differentiability of 'order zero', just as is the case for continuity in $\mathbb{R}$.

Definition 2.3. Let $D \subset \mathcal{N}$ be open, let $f : D \to \mathcal{N}$ be derivate continuous on D, and let I_D denote the identity function on D. Then we say that f is derivate differentiable on D if for all $x \in D$, the function $\frac{f - f(x)}{I_D - x} : D \setminus \{x\} \to \mathcal{N}$ is derivate continuous on $D \setminus \{x\}$. In this case, the unique continuation of $\frac{f - f(x)}{I_D - x}$ to D (see [14]) will be called the first derivate function (or simply the derivate function) of f at x and will be denoted by $F_{1,x}$; moreover, the function value $F_{1,x}(x)$ will be called the derivative of f at x and will be denoted by $f'(x)$.

It follows immediately from Definition 2.3 that if $f : D \to \mathcal{N}$ is derivate differentiable then f is differentiable in the conventional sense; moreover, the two derivatives at any given point of D agree. The following result provides a useful tool for checking the derivate differentiability of functions; we refer the interested reader to [14,16] for its Proof.

Theorem 2.4. *Let $D \subset \mathcal{N}$ be open and let $f : D \to \mathcal{N}$ be derivate continuous on D. Suppose there exists $M \in \mathcal{N}$ and there exists a function $g : D \to \mathcal{N}$ such that*

$$\left|\frac{f(y)-f(x)}{y-x} - g(x)\right| \leq M|y-x| \text{ for all } y \neq x \text{ in } D.$$

Then f is derivate differentiable on D, with derivative $f' = g$.

Definition 2.5 (n-times Derivate Differentiability). Let $D \subset \mathcal{N}$ be open, and let $f : D \to \mathcal{N}$. Let $n \geq 2$ be given in $\mathbb{N}$. Then we define n-times derivate differentiability of f on D inductively as follows: Having defined $(n-1)$-times derivate differentiability, we say that f is n-times derivate differentiable on D if f is $(n-1)$-times derivate differentiable on D and for all $x \in D$, the $(n-1)$st derivate function $F_{n-1,x}$ is derivate differentiable on D. For all $x \in D$, the derivate function $F_{n,x}$ of $F_{n-1,x}$ at x will be called the nth derivate function of f at x, and the number $f^{(n)}(x) = n!F'_{n-1,x}(x)$ will be called the nth derivative of f at x and denoted by $f^{(n)}(x)$.

One of the most useful consequences of the derivate differentiability concept is that it gives rise to a Taylor formula with remainder while the conventional (topological) differentiability does not; see [2,14]. We only state the result here and refer the reader to [2,14] for its Proof. We also note that, as an immediate result of Theorem 2.6, we obtain local expandability in Taylor series around $x_0 \in D$ of a given function that is infinitely often derivate differentiable on D [2,14].

Theorem 2.6 *(Taylor Formula with Remainder). Let $D \subset \mathcal{N}$ be open and let $f : D \to \mathcal{N}$ be n-times derivate differentiable on D. Let $x \in D$ be given, let $F_{n,x}$ be the nth order derivate function of f at x, and let $M_{n,x}$ be a Lipschitz constant of $F_{n,x}$ on D. Then for all $y \in D$, we have that*

$$f(y) = f(x) + \sum_{j=1}^{n} \frac{f^{(j)}(x)}{j!}(y-x)^j + r_{n+1}(x,y)(y-x)^{n+1},$$

with $\lambda(r_{n+1}(x,y)) \geq \lambda(M_{n,x})$.

Now we generalize the concepts of derivate continuity and derivate differentiability to functions of many variables. In the following, column vectors in $\mathcal{N}^n$ will be denoted by $\vec{x}, \vec{y}, \ldots$; and row vectors by $\vec{x}^T, \vec{y}^T, \ldots$.

Definition 2.7. Let $D \subset \mathcal{N}^n$ be open, let $f : D \to \mathcal{N}$, and let $\vec{u}$ be a unit vector. For each $\vec{x} \in D$, let

$$D_{\vec{x},\vec{u}} = \{t \in \mathcal{N} : \vec{x} + t\vec{u} \in D\}$$

and define $\phi_{\vec{x},\vec{u}} : D_{\vec{x},\vec{u}} \to \mathcal{N}$ by

$$\phi_{\vec{x},\vec{u}}(t) = f(\vec{x} + t\vec{u}).$$

Then we say that f is derivate differentiable on D in the direction of $\vec{u}$ if $\phi_{\vec{x},\vec{u}}$ is derivate differentiable on $D_{\vec{x},\vec{u}}$ for all $\vec{x}$ in D. Moreover, the derivative $\phi'_{\vec{x},\vec{u}}(0)$ will be called the directional derivative of f at $\vec{x}$ in the $\vec{u}$ direction and will be denoted by $\partial_{\vec{u}} f(\vec{x})$.

Definition 2.8 (Partial Derivatives). Let $D \subset \mathcal{N}^n$ be open, let $f : D \to \mathcal{N}$ and let $\{\vec{e}_1, \ldots, \vec{e}_n\}$ denote the standard orthonormal basis of $\mathcal{N}^n$. Then the partial derivatives of f are defined as the directional derivatives of f in the directions $\vec{e}_1, \ldots, \vec{e}_n$, if these exist. If the partial derivative in the direction $\vec{e}_k$ exists, we will denote it by $\partial_k f$. The gradient of f, denoted by ∇f, is defined to be the row vector whose components are the (first order) partial derivatives of f.

Definition 2.9. Let $D \subset \mathcal{N}^n$ be open, let $f : D \to \mathcal{N}$ and let $q \in \mathbb{N}$ be given. Then we say that f is C^q on D if all the partial derivatives of order smaller than or equal to q exist and are continuous on D (in the derivate sense).

Theorem 2.10. *Let $D \subset \mathcal{N}^n$ be open, let $f : D \to \mathcal{N}$ be C^1 on D and let $\vec{x} \in D$ be given. Then there exist $\delta, M > 0$ in $\mathcal{N}$ such that $B_\delta(\vec{x}) = \{\vec{z} \in \mathcal{N}^n : |\vec{z} - \vec{x}| < \delta\} \subset D$, and $|f(\vec{y}) - f(\vec{x}) - \nabla f(\vec{x})(\vec{y} - \vec{x})| \le M|\vec{y} - \vec{x}|^2$ for all $\vec{y} \in B_\delta(\vec{x})$.*

Proof. We use induction on the number of variables n. The result is true for $n = 1$ by definition of derivate differentiability in one dimension. Assume it is true for $n = k$ and we show that it is true for $n = k + 1$. So let $f : D \subset \mathcal{N}^{k+1} \to \mathcal{N}$ be C^1, and let $\vec{x} \in D$ be given. Since D is open, there exists $\delta_0 > 0$ in $\mathcal{N}$ such that $B_{\delta_0}^{(k+1)}(\vec{x}) \subset D$. Write $\vec{x} = (x_1, \ldots, x_k, x_{k+1})$, let $\vec{\zeta} = (x_1, \ldots, x_k) \in \mathcal{N}^k$ and let

$$D_1 = \left\{\vec{z} = (z_1, \ldots, z_k) \in \mathcal{N}^k : (z_1, \ldots, z_k, x_{k+1}) \in B_{\delta_0}^{(k+1)}(\vec{x})\right\}.$$

Then $D_1 = B_{\delta_0}^{(k)}(\vec{\zeta})$; and hence D_1 is open in $\mathcal{N}^k$. Define $\psi : D_1 \to \mathcal{N}$ by

$$\psi(z_1, \ldots, z_k) = f(z_1, \ldots, z_k, x_{k+1}).$$

Then ψ is C^1 on D_1 and the partial derivatives of ψ at $\vec{z}$ are given by

$$\partial_j \psi(z_1, \ldots, z_k) = \partial_j f(z_1, \ldots, z_k, x_{k+1}) \text{ for all } j \in \{1, \ldots, k\}.$$

Thus, by the induction hypothesis, there exist $\delta_1, M_1 > 0$ in $\mathcal{N}$ such that $B_{\delta_1}^{(k)}(\vec{\zeta}) \subset D_1$ (i.e. $\delta_1 \le \delta_0$) and

$$\left|\psi(\vec{\eta}) - \psi(\vec{\zeta}) - \nabla\psi(\vec{\zeta})(\vec{\eta} - \vec{\zeta})\right| \le M_1|\vec{\eta} - \vec{\zeta}|^2 \text{ for all } \vec{\eta} \in B_{\delta_1}^{(k)}(\vec{\zeta}).$$

Let $\vec{y} \in B_{\delta_1}^{(k+1)}(\vec{x})$ and let $\vec{w} = (y_1, \ldots, y_k)$. Then $\vec{w} \in B_{\delta_1}^{(k)}(\vec{\zeta})$; and hence

$$
\begin{aligned}
&|f(\vec{y}) - f(\vec{x}) - \nabla f(\vec{x})(\vec{y} - \vec{x})| \\
&\leq |f(\vec{y}) - \psi(\vec{w}) - \partial_{k+1} f(\vec{x})(y_{k+1} - x_{k+1})| \\
&\quad + \left| \psi(\vec{w}) - \psi(\vec{\zeta}) - \sum_{j=1}^{k} \partial_j f(\vec{x})(y_j - x_j) \right| \\
&= |f(\vec{w}, y_{k+1}) - f(\vec{w}, x_{k+1}) - \partial_{k+1} f(\vec{x})(y_{k+1} - x_{k+1})| \\
&\quad + \left| \psi(\vec{w}) - \psi(\vec{\zeta}) - \nabla\psi(\vec{\zeta})(\vec{w} - \vec{\zeta}) \right| \\
&\leq |f(\vec{w}, y_{k+1}) - f(\vec{w}, x_{k+1}) - \partial_{k+1} f(\vec{w}, x_{k+1})(y_{k+1} - x_{k+1})| \\
&\quad + | \partial_{k+1} f(\vec{w}, x_{k+1}) - \partial_{k+1} f(\vec{x})| \, |y_{k+1} - x_{k+1}| \\
&\quad + \left| \psi(\vec{w}) - \psi(\vec{\zeta}) - \nabla\psi(\vec{\zeta})(\vec{w} - \vec{\zeta}) \right| \\
&\leq M_2 |y_{k+1} - x_{k+1}|^2 + M_3 |\vec{w} - \vec{\zeta}| \, |y_{k+1} - x_{k+1}| + M_1 |\vec{w} - \vec{\zeta}|^2
\end{aligned}
$$

for some constants $M_2, M_3 \in \mathcal{N}$, which do not depend on $\vec{y}$ (nor on $\vec{x}$), since $\partial_{k+1} f$ is derivate continuous on D. Let $M = \max\{3M_1, 3M_2, 3M_3\}$. Then

$$
\begin{aligned}
|f(\vec{y}) - f(\vec{x}) - \nabla f(\vec{x})(\vec{y} - \vec{x})| &\leq M_2 |\vec{y} - \vec{x}|^2 + M_3 |\vec{y} - \vec{x}|^2 + M_1 |\vec{y} - \vec{x}|^2 \\
&\leq M |\vec{y} - \vec{x}|^2.
\end{aligned}
$$

Thus, the result is true for $n = k + 1$; and hence it is true for all n. □

As an immediate consequence of Theorem 2.10, we obtain the following result.

Corollary 2.11. *Let $D \subset \mathcal{N}^n$ be open, let $\vec{x}_0 \in D$ be given and let $f : D \to \mathcal{N}$ be C^1 on D. Then there exist $M, \delta > 0$ in $\mathcal{N}$ such that $B_\delta(\vec{x}_0) \subset D$ and, for all $\vec{x} \in B_\delta(\vec{x}_0)$, we have that*

$$
f(\vec{x}) = f(\vec{x}_0) + \nabla f(\vec{x}_0)(\vec{x} - \vec{x}_0) + R_2(\vec{x}_0, \vec{x}),
$$

where $|R_2(\vec{x}_0, \vec{x})| \leq M |\vec{x} - \vec{x}_0|^2$.

Lemma 2.12. *Let $D \subset \mathcal{N}^n$ be open, and let $f : D \to \mathcal{N}$ be C^1 on D. Fix $\vec{x}_0 \in D$, let $\vec{v} \in \mathcal{N}^n$ be a fixed vector and let $D_1 \subset \mathcal{N}$ be given by $D_1 = \{t \in \mathcal{N} : \vec{x}_0 + t\vec{v} \in D\}$. Define $\phi : D_1 \to \mathcal{N}$ by $\phi(t) = f(\vec{x}_0 + t\vec{v})$. Then ϕ is differentiable at 0 (in the conventional sense) with derivative $\phi'(0) = \nabla f(\vec{x}_0)\vec{v}$. Moreover, there exists $\delta > 0$ in $\mathcal{N}$ such that $B_\delta(0) \subset D_1$ and*

$$
\left| \frac{\phi(t) - \phi(0)}{t} - \phi'(0) \right| \leq M |\vec{v}|^2 |t| \text{ for all } t \in B_\delta(0) \setminus \{0\}.
$$

Proof. By Theorem 2.10, there exist $\delta_0, M > 0$ in $\mathcal{N}$ such that $B_{\delta_0}(\vec{x}_0) \subset D$ and

$$
|f(\vec{x}) - f(\vec{x}_0) - \nabla f(\vec{x}_0)(\vec{x} - \vec{x}_0)| \leq M |\vec{x} - \vec{x}_0|^2 \text{ for all } \vec{x} \in B_{\delta_0}(\vec{x}_0).
$$

Let $\delta = \delta_0 / |\vec{v}|$. Then, since $B_{\delta_0}(\vec{x}_0) \subset D$, we obtain that $B_\delta(0) \subset D_1$. Now let $t \in B_\delta(0) \setminus \{0\}$ be given. Then $\vec{x}_0 + t\vec{v} \in B_{\delta_0}(\vec{x}_0)$; and hence

$$|f(\vec{x}_0 + t\vec{v}) - f(\vec{x}_0) - \nabla f(\vec{x}_0)\vec{v}t| \le M|\vec{v}|^2 t^2.$$

Thus,

$$\left|\frac{\phi(t) - \phi(0)}{t} - \nabla f(\vec{x}_0)\vec{v}\right| \le M|\vec{v}|^2|t|,$$

which finishes the Proof of the Lemma. □

By repeated application of Corollary 2.11 and of Lemma 2.12 to the function ϕ and its derivatives, we obtain a Taylor formula with remainder, similar to the corresponding one in real calculus.

Corollary 2.13 *[Taylor Formula for Functions of Several Variables]. Let $D \subset \mathcal{N}^n$ be open, let $\vec{x}_0 \in D$ be given and let $f : D \to \mathcal{N}$ be C^q on D. Then there exist $M, \delta > 0$ in $\mathcal{N}$ such that $B_\delta(\vec{x}_0) \subset D$ and, for all $\vec{x} \in B_\delta(\vec{x}_0)$, we have that*

$$\begin{aligned} f(\vec{x}) = & f(\vec{x}_0) + \sum_{j=1}^{q} \left(\frac{1}{j!} \sum_{l_1,\ldots,l_j=1}^{n} \left(\partial_{l_1} \cdots \partial_{l_j} f(\vec{x}_0) \pi_{k=1}^{j} \left(x_{l_k} - x_{0,l_k} \right) \right) \right) \\ & + R_{q+1}(\vec{x}_0, \vec{x}), \end{aligned}$$

where $\left|R_{q+1}(\vec{x}_0, \vec{x})\right| \le M|\vec{x} - \vec{x}_0|^{q+1}$.

3. TANGENT PLANE

We start this section by first stating, without Proofs, the inverse function Theorem and the implicit function Theorem. The Proofs are very similar to those of the respective real ones: rather than taking $c \in \mathbb{R}$ such that $0 < c < 1$ as in [4] pp. 140–149, we use $c \in \mathcal{N}$ such that $0 < c \ll 1$ wherever the convergence of the sequence (c^n) is needed in the Proofs. Also, we can replace conventional differentiability with derivate differentiability without having to change the essence of the Proofs. The details of the last two statements are left as an exercise for the interested reader.

Notation 3.1. Let $m, n \in \mathbb{N}$ be given, let $D \subset \mathcal{N}^n$ be open and let $\vec{f} : D \to \mathcal{N}^m$ be C^1 on D. Then for $\vec{x} \in D$, the $m \times n$ matrix of the partial derivatives of the components of f will be denoted by $\mathcal{M}\vec{f}(\vec{x})$; that is

$$\mathcal{M}\vec{f}(\vec{x}) = \begin{pmatrix} \partial_1 f_1(\vec{x}) & \partial_2 f_1(\vec{x}) & \cdots & \partial_n f_1(\vec{x}) \\ \partial_1 f_2(\vec{x}) & \partial_2 f_2(\vec{x}) & \cdots & \partial_n f_2(\vec{x}) \\ \vdots & \vdots & \vdots & \vdots \\ \partial_1 f_m(\vec{x}) & \partial_2 f_m(\vec{x}) & \cdots & \partial_n f_m(\vec{x}) \end{pmatrix}.$$

Moreover, if $m = n$ then the determinant of $\mathcal{M}\vec{f}(\vec{x})$ will be called the Jacobian of $\vec{f}$ at $\vec{x}$ and denoted by $\mathcal{J}\vec{f}(\vec{x})$.

Theorem 3.2 *(Inverse Function Theorem). Let $n, q \in \mathbb{N}$ be given, let $D \in \mathcal{N}^n$ be* *open and let $\vec{f} : D \to \mathcal{N}^n$ be C^q. Let $\vec{x}_0 \in D$ be such that $\mathcal{J}\vec{f}(\vec{x}_0) \neq 0$. Then there exists an open set Ω containing $\vec{x}_0$ such that*

- $\Omega \subset D$,
- $\vec{f}|_\Omega$ *is one-to-one,*
- $\vec{f}(\Omega)$ *is open,*
- $\mathcal{J}\vec{f}(\vec{x}) \neq 0$ *for all* $\vec{x} \in \Omega$,
- *The inverse* $\vec{g}$ *of* $\vec{f}|_\Omega$ *is* C^q *on* $\vec{f}(\Omega)$,
- $\mathcal{M}\vec{g}(\vec{y}) = [\mathcal{M}\vec{f}(\vec{x})]^{-1}$ *if* $\vec{y} = \vec{f}(\vec{x})$ *and* $\vec{x} \in \Omega$.

Theorem 3.3 *[Implicit Function Theorem]. Let $r, m, q \in \mathbb{N}$ be given, let $D_r \subset \mathcal{N}^r$ and $D_m \subset \mathcal{N}^m$ be open and let $\vec{F} : D_r \times D_m \to \mathcal{N}^m$ be C^q. Let $(\vec{x}_0, \vec{z}_0) \in D_r \times D_m$ be such that $\vec{F}(\vec{x}_0, \vec{z}_0) = \vec{0}$ and $\mathcal{J}_z\vec{F}(\vec{x}_0, \vec{z}_0) \neq 0$. Then there exist neighborhoods Ω_r and Ω_m of $\vec{x}_0$ and $\vec{z}_0$, respectively and there exists a function $\vec{\phi} : \Omega_r \to \Omega_m$ that is C^q on Ω_r such that $\vec{F}\left(\vec{x}, \vec{\phi}(\vec{x})\right) = \vec{0}$ for all $\vec{x} \in \Omega_r$ and $\vec{\phi}(\vec{x}_0) = \vec{z}_0$.*

Now let's go back to our original problem stated at the end of the introduction. Let S be the hypersurface in $\mathcal{N}^n$ defined by

$$(3.1) \qquad \begin{cases} h_1(\vec{x}) = 0 \\ \qquad \vdots \\ h_m(\vec{x}) = 0 \end{cases}$$

and assume that $h_1, \ldots, h_m$ are C^1.

Definition 3.4. Let S be the surface defined by Equation (3.1) and let $\vec{x}_0 \in S$ be given. Then the tangent plane at $\vec{x}_0$ to S, denoted by $\mathcal{T}(\vec{x}_0)$, is the collection of the derivatives $\vec{x}'(t_0)$, where $\vec{x}(t) \in S$ for t in some open interval (a, b) in $\mathcal{N}$, $\vec{x}(t)$ is derivate differentiable on (a, b) and $\vec{x}(t_0) = \vec{x}_0$ for some $t_0 \in (a, b)$.

Definition 3.5. Let S be as in Definition 3.4, and let $\vec{x}_0 \in S$ be given. Then we say that $\vec{x}_0$ is a regular point of S if $\nabla h_1(\vec{x}_0), \ldots, \nabla h_m(\vec{x}_0)$ are linearly independent. That is, if $a_1, \ldots, a_m \in \mathcal{N}$ and if $a_1\nabla h_1(\vec{x}_0) + \cdots + a_m\nabla h_m(\vec{x}_0) = \vec{0}$ then $a_1 = \cdots = a_m = 0$.

Remark 3.6. We define matrix addition and multiplication in $\mathcal{N}$ the same way we do in $\mathbb{R}$; and hence we have the same criteria for the invertibility of a given matrix and the same procedures to compute the inverse. Thus, an $n \times n$ matrix $\mathcal{M}$ over $\mathcal{N}$ is invertible if and only if its n column vectors are linearly independent, if and only if its n row vectors are linearly independent in $\mathcal{N}$.

A simple expression for the tangent plane to S at a regular point $\vec{x}_0$ is given by the following theorem.

Theorem 3.7. *Let S be the surface defined by Equation (3.1), and let $\vec{x}_0$ be a regular point of S. Then the tangent plane at $\vec{x}_0$ to S is given by*

$$\mathcal{T}(\vec{x}_0) = \{\vec{y} \in \mathcal{N}^n : \nabla h_j(\vec{x}_0)\vec{y} = 0 \textit{ for all } j = 1, \ldots, m\}.$$

Proof. Let $\vec{y} \in \mathcal{T}(\vec{x}_0)$ be given. Then there exists a derivate differentiable curve $\vec{x}(t)$, $a < t < b$, in S such that $\vec{x}_0 = \vec{x}(t_0)$ and $\vec{y} = \vec{x}'(t_0)$ for some $t_0 \in (a,b)$. Thus, $h_j(\vec{x}(t)) = 0$ for all $t \in (a,b)$ and for all $j = 1, \ldots, m$. Since h_j is C^1 in the derivate sense, we obtain that h_j is C^1 in the conventional sense and the partial derivatives of h_j at any given point (in the derivate sense and in the conventional sense) agree; similarly, $\vec{x}$ is differentiable (in the conventional sense) on (a,b) and the derivatives at any $t \in (a,b)$ in the derivate sense and in the conventional sense agree. Hence, applying the chain rule, we obtain that $h_j(\vec{x}(t))$ is differentiable (in the conventional sense) at t for all $t \in (a,b)$ with $h_j'(\vec{x}(t)) = \nabla h_j(\vec{x}(t))\vec{x}'(t) = 0$ for all $t \in (a,b)$. In particular, $\nabla h_j(\vec{x}(t_0))\vec{x}'(t_0) = 0$; and this is true for all $j = 1, \ldots, m$. Therefore, $\nabla h_j(\vec{x}_0)\vec{y} = 0$ for all $j = 1, \ldots, m$.

Conversely, let $\vec{y} \in \mathcal{N}^n$ be such that $\nabla h_j(\vec{x}_0)\vec{y} = 0$ for all $j = 1, \ldots, m$. Define $\vec{F} : \mathcal{N} \times \mathcal{N}^m \to \mathcal{N}^m$ by

$$\vec{F}(t, \vec{z}) = \vec{h}\left(\vec{x}_0 + t\vec{y} + \left(\mathcal{M}\vec{h}(\vec{x}_0)\right)^T \vec{z}\right),$$

and consider now the equation $\vec{F}(t,\vec{z}) = \vec{0}$. Then, $\vec{F}(0,\vec{0}) = \vec{0}$ since $\vec{x}_0$ is feasible. Also,

$$F_z(0,\vec{0}) = \mathcal{M}\vec{h}(\vec{x}_0)\left(\mathcal{M}\vec{h}(\vec{x}_0)\right)^T$$

which is a nonsingular $m \times m$ matrix. This is so since the row vectors in $\mathcal{M}\vec{h}(\vec{x}_0)$, namely $\nabla h_1(\vec{x}_0), \ldots, \nabla h_m(\vec{x}_0)$, are linearly independent. Thus, by the implicit function theorem, there exists a C^1 function $\vec{z}(t)$ defined on some open interval $(-a,a) \subset \mathcal{N}$ such that

$$\vec{F}(t, \vec{z}(t)) = \vec{0} \text{ for all } t \in (-a,a) \text{ and } \vec{z}(0) = \vec{0}.$$

Define the function $\vec{x} : (-a,a) \to \mathcal{N}^n$ by $\vec{x}(t) = \vec{x}_0 + t\vec{y} + \left(\mathcal{M}\vec{h}(\vec{x}_0)\right)^T \vec{z}(t)$. Then $\vec{x}(t) \in S$ for all $t \in (-a,a)$, $\vec{x}(0) = \vec{x}_0$ and $\vec{x}'(0) = \vec{y} + \left(\mathcal{M}\vec{h}(\vec{x}_0)\right)^T \vec{z}'(0)$. But from $\vec{h}(\vec{x}(t)) = \vec{0}$ for all $t \in (-a,a)$, we obtain that

$$\vec{h}'(\vec{x}(t))\big|_{t=0} = \vec{0}; \text{ and hence } \mathcal{M}\vec{h}(\vec{x}_0)\left(\vec{y} + \left(\mathcal{M}\vec{h}(\vec{x}_0)\right)^T \vec{z}'(0)\right) = \vec{0}.$$

Since $\mathcal{M}\vec{h}(\vec{x}_0)\vec{y} = \vec{0}$ and since $\mathcal{M}\vec{h}(\vec{x}_0)\left(\mathcal{M}\vec{h}(\vec{x}_0)\right)^T$ is invertible, we obtain that $\vec{z}'(0) = \vec{0}$. Thus, $\vec{x}'(0) = \vec{y}$ and hence $\vec{y} \in \mathcal{T}(\vec{x}_0)$. □

4. NECESSARY AND SUFFICIENT OPTIMALITY CONDITIONS

In this section, we derive necessary and sufficient, second order, optimality conditions for a local minimum of a function subject to a set of constraints. We start with the following definition.

Definition 4.1. Let $\vec{x}_0$ be a feasible point for the constraints in Equation (1.1) and let $I(\vec{x}_0) = \{l \in \{1, \ldots, p\} : g_l(\vec{x}_0) = 0\}$. Then we say that $\vec{x}_0$ is regular for the constraints if $\{\nabla h_j(\vec{x}_0) : j = 1, \ldots, m; \nabla g_l(\vec{x}_0) : l \in I(\vec{x}_0)\}$ forms a linearly independent subset of vectors in $\mathcal{N}^n$.

The following theorem provides *necessary conditions* of second order for a local minimizer $\vec{x}_0$ of a function f subject to the constraints in Equation (1.1). The result is a generalization of the corresponding real result [9,3] and the Proof is similar to that of the latter; but one essential difference is the form of the remainder formula as in Equations (4.4), (4.5) and (4.6). In the real case, the remainder term is related to the second derivative at some intermediate point, while here that is not the case. However, the concept of derivate differentiability puts a bound on the remainder term; and this is instrumental to prove the desired result.

Theorem 4.2. *Suppose that f, $\{h_j\}_{j=1}^m$, $\{g_l\}_{l=1}^p$ are C^2 on some open set $D \subset \mathcal{N}^n$ containing the point $\vec{x}_0$ and that $\vec{x}_0$ is a regular point for the constraints in Equation (1.1). If $\vec{x}_0$ is a local minimizer for f under the given constraints, then there exist $\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_p \in \mathcal{N}$ such that*

(i) $\beta_l \geq 0$ *for all* $l \in \{1, \ldots, p\}$,

(ii) $\beta_l g_l(\vec{x}_0) = 0$ *for all* $l \in \{1, \ldots, p\}$,

(iii) $\nabla f(\vec{x}_0) + \sum_{j=1}^m \alpha_j \nabla h_j(\vec{x}_0) + \sum_{l=1}^p \beta_l \nabla g_l(\vec{x}_0) = \vec{0}$, *and*

(iv) $\vec{y}^T \left(\nabla^2 f(\vec{x}_0) + \sum_{j=1}^m \alpha_j \nabla^2 h_j(\vec{x}_0) + \sum_{l=1}^p \beta_l \nabla^2 g_l(\vec{x}_0) \right) \vec{y} \geq 0$ *for all* $\vec{y} \in \mathcal{N}^n$ *satisfying* $\nabla h_j(\vec{x}_0)\vec{y} = 0$ *for all* $j \in \{1, \ldots, m\}$, $\nabla g_l(\vec{x}_0)\vec{y} = 0$ *for all* $l \in L = \{k \in I(\vec{x}_0) : \beta_k > 0\}$ *and* $\nabla g_l(\vec{x}_0)\vec{y} \leq 0$ *for all* $l \in I(\vec{x}_0) \setminus L$.

Proof. Since $\vec{x}_0$ is a local minimizer for f over the constraints in Equation (1.1) and since, for $l \notin I(\vec{x}_0)$, $g_l(\vec{x}_0) < 0$, there exists $\epsilon > 0$ in $\mathcal{N}$ such that $\vec{x}_0$ is a minimum point for f in $B_\epsilon(\vec{x}_0)$ over the constraints $\vec{h}(\vec{x}) = \vec{0}$ and $g_l(\vec{x}) = 0$ for $l \in I(\vec{x}_0)$. That is, $\vec{x}_0$ is a solution for

$$\min f(\vec{x}) : \vec{h}(\vec{x}) = \vec{0}, g_l(\vec{x}) = 0 \text{ for } l \in I(\vec{x}_0), \vec{x} \in B_\epsilon(\vec{x}_0). \tag{4.1}$$

Since $\vec{x}_0$ is regular for the constraints in Equation (1.1), this is equivalent to saying that $\vec{x}_0$ is regular for the constraints in Equation (4.1). Thus, by Theorem 3.7, we have that the tangent plane to the constraint set S defined by Equation (4.1) is

$$\mathcal{P}(\vec{x}_0) = \left\{ \vec{y} \in \mathcal{N}^n : \mathcal{M}\vec{h}(\vec{x}_0)\vec{y} = \vec{0} \text{ and } \nabla g_l(\vec{x}_0)\vec{y} = 0 \text{ for } l \in I(\vec{x}_0) \right\}.$$

Let $\vec{y} \in \mathcal{P}(\vec{x}_0)$. Then there exists a derivate differentiable curve $\vec{x}(t)$, $a < t < b$, with $\vec{x}(t) \in S$ for all $t \in (a, b)$, and $\vec{x}'(t_0) = \vec{y}$ and $\vec{x}(t_0) = \vec{x}_0$ for some $t_0 \in (a, b)$. Then

$$f(\vec{x}(t)) \geq f(\vec{x}(t_0)) = f(\vec{x}_0) \text{ for all } t \in (a, b).$$

Hence, using a result in [16] about local minima, we have that

(4.2) $\quad f'(\vec{x}(t)))|_{t=t_0} = 0 = \nabla f(\vec{x}_0)\vec{x}'(t_0) = \nabla f(\vec{x}_0)\vec{y}$ and

(4.3) $\quad f''(\vec{x}(t)))|_{t=t_0} \geq 0.$

Equation (4.2) yields that

$$\nabla f(\vec{x}_0) \in \mathcal{P}(\vec{x}_0)^{\perp} = \{\vec{z} \in \mathcal{N}^n : \vec{z} \cdot \vec{w} = 0 \text{ for all } \vec{w} \in \mathcal{P}(\vec{x}_0)\}.$$

It follows that $\nabla f(\vec{x}_0) \in N(\vec{x}_0)$, where

$$N(\vec{x}_0) = \left\{ \vec{z} \in \mathcal{N}^n : \vec{z} = -\sum_{j=1}^{m} \alpha_j \nabla h_j(\vec{x}_0) - \sum_{l \in I(\vec{x}_0)} \beta_l \nabla g_l(\vec{x}_0) \text{ with } \alpha_j, \beta_l \in \mathcal{N} \right\}.$$

Therefore, there exist $\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_p \in \mathcal{N}$ such that $\beta_l = 0$ for $l \notin I(\vec{x}_0)$ and $\nabla f(\vec{x}_0) + \sum_{j=1}^{m} \alpha_j \nabla h_j(\vec{x}_0) + \sum_{l=1}^{p} \beta_l \nabla g_l(\vec{x}_0) = \vec{0}$. Clearly, $\beta_l g_l(\vec{x}_0) = 0$ for all $l \in \{1, \ldots, p\}$. Hence (ii) and (iii) hold.

To prove (i), we need to show that $\beta_l \geq 0$ for all $l \in I(\vec{x}_0)$. Suppose that $\beta_{l_0} < 0$ for some $l_0 \in I(\vec{x}_0)$. If $\mathcal{P}_{l_0}(\vec{x}_0)$ is the tangent plane to

$$S_{l_0}(\vec{x}_0) = \{\vec{x} \in D : h_j(\vec{x}) = 0 \text{ for all } j; g_l(\vec{x}) = 0 \text{ for all } l \in I(\vec{x}_0) \text{ and } l \neq l_0\}$$

then the regularity of $\vec{x}_0$ for the constraints in Equation (1.1) yields the existence of some $\vec{y} \in \mathcal{P}_{l_0}(\vec{x}_0)$ such that $\nabla g_{l_0}(\vec{x}_0)\vec{y} < 0$. By Theorem 3.7 applied to S_{l_0}, it follows that there exists a derivate differentiable curve $\vec{x}(t)$, $a < t < b$, in S_{l_0} with $\vec{x}'(t_0) = \vec{y}$ and $\vec{x}(t_0) = \vec{x}_0$. Thus,

$$\begin{aligned} f'(\vec{x}(t))|_{t=t_0} &= \nabla f(\vec{x}_0)\vec{y} \\ &= -\sum_{j=1}^{m} \alpha_j \nabla h_j \vec{y} - \sum_{l=1}^{p} \beta_l \nabla g_l(\vec{x}_0)\vec{y} \\ &= -\beta_{l_0} \nabla g_{l_0}(\vec{x}_0)\vec{y} < 0. \end{aligned}$$

Since g_{l_0} is C^1 on D and since $\vec{x}(t)$ is derivate differentiable on (a, b), we have that

(4.4) $\quad g_{l_0}(\vec{x}(t)) = g_{l_0}(\vec{x}_0) + \nabla g_{l_0}(\vec{x}_0)(\vec{x}(t) - \vec{x}_0) + S_2(\vec{x}_0, \vec{x}(t))(\vec{x}(t) - \vec{x}_0)^2,$

and

(4.5)
$$\begin{aligned} \vec{x}(t) &= \vec{x}_0 + \vec{x}'(t_0)(t - t_0) + \vec{s}_2(t_0, t)(t - t_0)^2 \\ &= \vec{x}_0 + (t - t_0)\vec{y} + \vec{s}_2(t_0, t)(t - t_0)^2, \end{aligned}$$

where $S_2(\vec{x}_0, \vec{x}(t))$ and $\vec{s}_2(t_0, t)$ are bounded on D and (a, b), respectively. Substituting for $\vec{x}(t) - \vec{x}_0$ from Equation (4.5) into Equation (4.4), we readily obtain that

(4.6) $\quad g_{l_0}(\vec{x}(t)) = g_{l_0}(\vec{x}_0) + (t - t_0)\nabla g_{l_0}(\vec{x}_0)\vec{y} + r_2(t_0, t)(t - t_0)^2,$

where $|r_2(t_0, t)|$ is bounded. Since $\nabla g_{l_0}(\vec{x}_0)\vec{y} < 0$, then for t close enough to t_0 and for $t - t_0 > 0$, we get $g_{l_0}(\vec{x}(t)) \leq 0$. Thus, $\vec{x}(t)$ is feasible for the constraints in Equation (1.1) for t near t_0 and $t > t_0$. Hence from $f'(\vec{x}(t))|_{t=t_0} < 0$, we obtain a contradiction with the optimality of $\vec{x}_0$ for the constraints in Equation (1.1). Therefore, $\beta_l \geq 0$ for all $l \in I(\vec{x}_0)$; and hence $\beta_l \geq 0$ for all $l \in \{1, \ldots, p\}$.

Finally, to prove (iv), let $\vec{y} \in \mathcal{N}^n$ be such that $\nabla h_j(\vec{x}_0)\vec{y} = 0$ for all $j \in \{1, \ldots, m\}$, $\nabla g_l(\vec{x}_0)\vec{y} = 0$ for all $l \in L$ and $\nabla g_l(\vec{x}_0)\vec{y} \leq 0$ for all $l \in I(\vec{x}_0) \setminus L$. Let $I_1 = \{l \in I(\vec{x}_0) : \nabla g_l(\vec{x}_0)\vec{y} = 0\}$. Then $\vec{y}$ is in the tangent plane to the constraint set

$$S_{I_1} = \left\{\vec{x} \in D : h_j(\vec{x}) = 0 \text{ for all } j \in \{1, \ldots, m\} \text{ and } g_l(\vec{x}) = 0 \text{ for all } l \in I_1\right\}.$$

Again, by Theorem 3.7 applied to S_{I_1}, it follows that there exists a derivate differentiable curve $\vec{x}(t)$, $a < t < b$, in S_{I_1} with $\vec{x}'(t_0) = \vec{y}$ and $\vec{x}(t_0) = \vec{x}_0$.

Now let $l \in I(\vec{x}_0) \setminus I_1$ be given. Then for all $t \in (a, b)$, we obtain that

$$\begin{aligned} g_l(\vec{x}(t)) &= g_l(\vec{x}_0) + \nabla g_l(\vec{x}_0)(\vec{x}(t) - \vec{x}_0) + \frac{1}{2}(\vec{x}(t) - \vec{x}_0)^T \nabla^2 g_l(\vec{x}_0)(\vec{x}(t) - \vec{x}_0) \\ &\quad + r_{3,l}(\vec{x}_0, \vec{x}(t)) \\ &= (t - t_0)\nabla g_l(\vec{x}_0)\vec{y} + \frac{1}{2}(t - t_0)^2 \vec{y}^T \nabla^2 g_l(\vec{x}_0)\vec{y} + R_{3,l}(t_0, t), \end{aligned}$$

where $|R_{3,l}(t_0, t)| \leq M_{3,l}|t - t_0|^3$ for some constant $M_{3,l} \geq 0$ in $\mathcal{N}$, and where use has been made of the fact that g_l is C^2 at $\vec{x}_0$ and $\vec{x}(t)$ is derivate differentiable at t_0. Since $\nabla g_l(\vec{x}_0)\vec{y} < 0$, we obtain that $g_l(\vec{x}(t)) < 0$ for all $t > t_0$ in (a, b) sufficiently close to t_0. Thus, for all such t, $\vec{x}(t)$ is a feasible point for the constraints and hence

$$f(\vec{x}_0) \leq f(\vec{x}(t)) = f(\vec{x}_0) + (t - t_0)\nabla f(\vec{x}_0)\vec{y} + \frac{1}{2}(t - t_0)^2 \vec{y}^T \nabla^2 f(\vec{x}_0)\vec{y} + R_{3,f}(t_0, t),$$

where

$$|R_{3,f}(t_0, t)| \leq M_{3,f}|t - t_0|^3 = M_{3,f}(t - t_0)^3$$

for some constant $M_{3,f} \geq 0$ in $\mathcal{N}$. Thus, for all $t > t_0$ in $\mathcal{N}$, sufficiently close to t_0, we have that

$$0 \leq (t - t_0)\nabla f(\vec{x}_0)\vec{y} + \frac{1}{2}(t - t_0)^2 \vec{y}^T \nabla^2 f(\vec{x}_0)\vec{y} + R_{3,f}(t_0, t) \tag{4.7}$$

$$\begin{aligned} 0 = \alpha_j h_j(\vec{x}(t)) &= \alpha_j (t - t_0)\nabla h_j(\vec{x}_0)\vec{y} + \frac{1}{2}\alpha_j (t - t_0)^2 \vec{y}^T \nabla^2 h_j(\vec{x}_0)\vec{y} \\ &\quad + R_{3,h_j,\alpha_j}(t_0, t) \text{ for } j = 1, \ldots, m \end{aligned} \tag{4.8}$$

$$\begin{aligned} 0 = \beta_l g_l(\vec{x}(t)) &= \beta_l (t - t_0)\nabla g_l(\vec{x}_0)\vec{y} + \frac{1}{2}\beta_l (t - t_0)^2 \vec{y}^T \nabla^2 g_l(\vec{x}_0)\vec{y} \\ &\quad + R_{3,g_l,\beta_l}(t_0, t) \text{ for } l = 1, \ldots, p, \end{aligned} \tag{4.9}$$

where

$$\max_{j=1,\ldots,m; l=1,\ldots,p}\left\{|R_{3,f}(t_0, t)|, |R_{3,h_j,\alpha_j}(t_0, t)|, |R_{3,g_l,\beta_l}(t_0, t)|\right\} \leq M_0(t - t_0)^3,$$

for some constant $M_0 \in \mathcal{N}$.

Adding Equation (4.7), Equation (4.8) for $j = 1, \ldots, m$, and Equation (4.9) for $l = 1, \ldots, p$, and using (iii), we obtain that

$$
\begin{aligned}
(4.10)\qquad 0 \le \frac{1}{2}(t - t_0)^2 \vec{y}^T \left(\nabla^2 f(\vec{x}_0) + \sum_{j=1}^{m} \alpha_j \nabla^2 h_j(\vec{x}_0) + \sum_{l=1}^{p} \beta_l \nabla^2 g_l(\vec{x}_0) \right) \vec{y} \\
+ R_3(t_0, t),
\end{aligned}
$$

where

$$
R_3(t_0, t) = R_{3,f}(t_0, t) + \sum_{j=1}^{m} R_{3,h_j,\alpha_j}(t_0, t) + \sum_{l=1}^{p} R_{3,g_l,\beta_l}(t_0, t)
$$

and hence

$$
|R_3(t_0, t)| \le M(t - t_0)^3 \text{ for some constant } M \ge 0 \text{ in } \mathcal{N}.
$$

If $\vec{y}^T (\nabla^2 f(\vec{x}_0) + \sum_{j=1}^{m} \alpha_j \nabla^2 h_j(\vec{x}_0) + \sum_{l=1}^{p} \beta_l \nabla^2 g_l(\vec{x}_0)) \vec{y} < 0$, then Equation (4.10) would yield a contradiction for $0 < t - t_0 < 1/(2M)$. Thus,

$$
\vec{y}^T \left(\nabla^2 f(\vec{x}_0) + \sum_{j=1}^{m} \alpha_j \nabla^2 h_j(\vec{x}_0) + \sum_{l=1}^{p} \beta_l \nabla^2 g_l(\vec{x}_0) \right) \vec{y} \ge 0. \quad \square
$$

In the following Theorem, we present second order *sufficient conditions* for a feasible point $\vec{x}_0$ to be a local minimum of a function f subject to the constraints in Equation (1.1). It is a generalization of the real result [3] and reduces to it, when restricted to functions from $\mathbb{R}^n$ to $\mathbb{R}$. In fact, since ϵ in condition (iv) below is allowed to be infinitely small, the condition $|\nabla h_j(\vec{x}_0)\vec{y}| < \epsilon$ would reduce to $\nabla h_j(\vec{x}_0)\vec{y} = 0$, when restricted to $\mathbb{R}$. Similarly, one can readily see that the other conditions are mere generalizations of the corresponding real ones. However, the Proof is different than that of the real result since the supremum principle does not hold in $\mathcal{N}$.

Theorem 4.3. *Suppose that f, $\{h_j\}_{j=1}^{m}$, $\{g_l\}_{l=1}^{p}$ are C^2 on some open set $D \subset \mathcal{N}^n$ containing the point $\vec{x}_0$ and that $\vec{x}_0$ is a feasible point for the constraints in Equation (1.1) such that, for some $\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_p \in \mathcal{N}$ and for some $\epsilon, \gamma > 0$ in $\mathcal{N}$, we have that*

(i) *$\beta_l \ge 0$ for all $l \in \{1, \ldots, p\}$,*
(ii) *$\beta_l g_l(\vec{x}_0) = 0$ for all $l \in \{1, \ldots, p\}$,*
(iii) *$\nabla f(\vec{x}_0) + \sum_{j=1}^{m} \alpha_j \nabla h_j(\vec{x}_0) + \sum_{l=1}^{p} \beta_l \nabla g_l(\vec{x}_0) = \vec{0}$, and*
(iv) *$\vec{y}^T (\nabla^2 f(\vec{x}_0) + \sum_{j=1}^{m} \alpha_j \nabla^2 h_j(\vec{x}_0) + \sum_{l=1}^{p} \beta_l \nabla^2 g_l(\vec{x}_0)) \vec{y} \ge \gamma$ for all $\vec{y} \in \mathcal{N}^n$ satisfying $|\vec{y}| = 1$, $|\nabla h_j(\vec{x}_0)\vec{y}| < \epsilon$ for all $j \in \{1, \ldots, m\}$, $|\nabla g_l(\vec{x}_0)\vec{y}| < \epsilon$ for all $l \in L = \{k : \beta_k > 0\}$ and $\nabla g_l(\vec{x}_0)\vec{y} < \epsilon$ for all $l \in I(\vec{x}_0) \setminus L$, where $I(\vec{x}_0) = \{k : g_k(\vec{x}_0) = 0\}$.*

Then $\vec{x}_0$ is a strict local minimum for f under the constraints of Equation (1.1).

Proof. Since D is open, there exists $\delta_0 > 0$ in $\mathcal{N}$ such that $B_{\delta_0}(\vec{x}_0) \subset D$ and

Corollary 2.13 holds for f, h_j, g_l on $B_{\delta_0}(\vec{x}_0)$ for all $j \in \{1, \ldots, m\}$ and for all $l \in \{1, \ldots, p\}$. For all $\vec{x} \in B_{\delta_0}(\vec{x}_0)$, we have by Corollary 2.13 that

$$
\begin{aligned}
(4.11) \qquad f(\vec{x}) &= f(\vec{x}_0) + \nabla f(\vec{x}_0)(\vec{x} - \vec{x}_0) \\
&\quad + \frac{1}{2}(\vec{x} - \vec{x}_0)^T \nabla^2 f(\vec{x}_0)(\vec{x} - \vec{x}_0) + R_{3,f}(\vec{x}_0, \vec{x}) \\
(4.12) \qquad h_j(\vec{x}) &= h_j(\vec{x}_0) + \nabla h_j(\vec{x}_0)(\vec{x} - \vec{x}_0) \\
&\quad + \frac{1}{2}(\vec{x} - \vec{x}_0)^T \nabla^2 h_j(\vec{x}_0)(\vec{x} - \vec{x}_0) + R_{3,h_j}(\vec{x}_0, \vec{x}) \\
(4.13) \qquad g_l(\vec{x}) &= g_l(\vec{x}_0) + \nabla g_l(\vec{x}_0)(\vec{x} - \vec{x}_0) \\
&\quad + \frac{1}{2}(\vec{x} - \vec{x}_0)^T \nabla^2 g_l(\vec{x}_0)(\vec{x} - \vec{x}_0) + R_{3,g_l}(\vec{x}_0, \vec{x})
\end{aligned}
$$

for all $j \in \{1, \ldots, m\}$ and for all $l \in \{1, \ldots, p\}$, where

$$
\max_{\substack{1 \le j \le m \\ 1 \le l \le p}} \left\{ \left| R_{3,f}(\vec{x}_0, \vec{x}) \right|, \left| R_{3,h_j}(\vec{x}_0, \vec{x}) \right|, \left| R_{3,g_l}(\vec{x}_0, \vec{x}) \right| \right\} \le M_0 |\vec{x} - \vec{x}_0|^3
$$

for some constant $M_0 \in \mathcal{N}$.

Adding Equation (4.11), Equation (4.12) multiplied by α_j for all $j \in \{1, \ldots, m\}$, and Equation (4.13) multiplied by β_l for all $l \in \{1, \ldots, p\}$, we obtain that

$$
\begin{aligned}
f(\vec{x}) + \sum_{j=1}^{m} \alpha_j h_j(\vec{x}) + \sum_{l=1}^{p} \beta_l g_l(\vec{x}) &= f(\vec{x}_0) + \sum_{j=1}^{m} \alpha_j h_j(\vec{x}_0) + \sum_{l=1}^{p} \beta_l g_l(\vec{x}_0) \\
&\quad + \left(\nabla f(\vec{x}_0) + \sum_{j=1}^{m} \alpha_j \nabla h_j(\vec{x}_0) + \sum_{l=1}^{p} \beta_l \nabla g_l(\vec{x}_0) \right)(\vec{x} - \vec{x}_0) \\
(4.14) \qquad &\quad + \frac{1}{2}(\vec{x} - \vec{x}_0)^T \left(\nabla^2 f(\vec{x}_0) + \sum_{j=1}^{m} \alpha_j \nabla^2 h_j(\vec{x}_0) + \sum_{l=1}^{p} \beta_l \nabla^2 g_l(\vec{x}_0) \right)(\vec{x} - \vec{x}_0) \\
&\quad + R_3(\vec{x}_0, \vec{x}) \\
&= f(\vec{x}_0) + \frac{1}{2}(\vec{x} - \vec{x}_0)^T \left(\nabla^2 f(\vec{x}_0) + \sum_{j=1}^{m} \alpha_j \nabla^2 h_j(\vec{x}_0) + \sum_{l=1}^{p} \beta_l \nabla^2 g_l(\vec{x}_0) \right)(\vec{x} - \vec{x}_0) \\
&\quad + R_3(\vec{x}_0, \vec{x})
\end{aligned}
$$

where we have made use of (ii) and (iii) and the fact that $h_j(\vec{x}_0) = 0$ for all $j \in \{1, \ldots, m\}$, and where

$$
R_3(\vec{x}_0, \vec{x}) = R_{3,f}(\vec{x}_0, \vec{x}) + \sum_{j=1}^{m} \alpha_j R_{3,h_j}(\vec{x}_0, \vec{x}) + \sum_{l=1}^{p} \beta_l R_{3,g_l}(\vec{x}_0, \vec{x}).
$$

Thus, $|R_3(\vec{x}_0, \vec{x})| \le M |\vec{x} - \vec{x}_0|^3$, where

$$
(4.15) \qquad M = \left(1 + \sum_{j=1}^{m} |\alpha_j| + \sum_{l=1}^{p} |\beta_l| \right) M_0.
$$

Now suppose that $\vec{x}_0$ is not a strict local minimum for f under the constraints of

Equation (1.1) and let $t \in \mathbb{N}$ be such that $d^t < \delta_0$, where d is the infinitely small number defined in the introduction. Then for all $k \in \mathbb{N}$, there exists $\vec{y}_k \in B_{d^{t+k}}(\vec{x}_0) \setminus \{\vec{x}_0\} \subset B_{\delta_0}(\vec{x}_0)$ such that $\vec{y}_k$ is feasible and $f(\vec{y}_k) \leq f(\vec{x}_0)$. Thus, the sequence $\{\vec{y}_k\}_{k\in\mathbb{N}}$ converges to $\vec{x}_0$. For each $k \in \mathbb{N}$, write $\vec{y}_k = \vec{x}_0 + \delta_k \vec{s}_k$, where $|\vec{s}_k| = 1$ and $0 < \delta_k < d^{t+k} \ll d^k$. Then $\{\delta_k\}$ is a null sequence in $\mathcal{N}$.

For all $k \in \mathbb{N}$, we have that

$$h_j(\vec{y}_k) = h_j(\vec{x}_0) + \nabla h_j(\vec{x}_0)(\vec{y}_k - \vec{x}_0) + R_{2,h_j}(\vec{x}_0, \vec{y}_k) \text{ for all } j \in \{1, \ldots, m\},$$

where

$$\left|R_{2,h_j}(\vec{x}_0, \vec{y}_k)\right| \leq M_{2,h_j} |\vec{y}_k - \vec{x}_0|^2 = M_{2,h_j} \delta_k^2,$$

for some $M_{2,h_j} > 0$ in $\mathcal{N}$, by Corollary 2.11. Thus, for all $j \in \{1, \ldots, m\}$,

$$0 = 0 + \nabla h_j(\vec{x}_0)\delta_k \vec{s}_k + R_{2,h_j}(\vec{x}_0, \vec{y}_k),$$

or

$$\nabla h_j(\vec{x}_0)\vec{s}_k = -\frac{R_{2,h_j}(\vec{x}_0, \vec{y}_k)}{\delta_k}.$$

Thus,

$$\left|\nabla h_j(\vec{x}_0)\vec{s}_k\right| \leq \frac{\left|R_{2,h_j}(\vec{x}_0, \vec{y}_k)\right|}{\delta_k} \leq M_{2,h_j}\delta_k \underset{k\to\infty}{\longrightarrow} 0.$$

Hence, there exists $N_1 \in \mathbb{N}$ such that

$$\left|\nabla h_j(\vec{x}_0)\vec{s}_k\right| < \min\left\{\epsilon, \min_{q\in L}\left\{\frac{\epsilon\beta_q}{4m|\alpha_j|}\right\}\right\} \text{ for all } k \geq N_1 \text{ and for all } j \in \{1, \ldots, m\}.$$

Also, for all $k \in \mathbb{N}$, we have that

$$f(\vec{x}_0) \geq f(\vec{y}_k) = f(\vec{x}_0) + \nabla f(\vec{x}_0)(\vec{y}_k - \vec{x}_0) + R_{2,f}(\vec{x}_0, \vec{y}_k),$$

where

$$\left|R_{2,f}(\vec{x}_0, \vec{y}_k)\right| \leq M_{2,f}|\vec{y}_k - \vec{x}_0|^2 = M_{2,f}\delta_k^2,$$

for some $M_{2,f} > 0$ in $\mathcal{N}$, by Corollary 2.11. Thus,

$$\nabla f(\vec{x}_0)\delta_k\vec{s}_k + R_{2,f}(\vec{x}_0, \vec{y}_k) \leq 0, \text{ or } \nabla f(\vec{x}_0)\vec{s}_k \leq -\frac{R_{2,f}(\vec{x}_0, \vec{y}_k)}{\delta_k} \underset{k\to\infty}{\longrightarrow} 0.$$

Hence, there exists $N_2 \in \mathbb{N}$ such that

$$\nabla f(\vec{x}_0)\vec{s}_k < \min_{q\in L}\left\{\frac{\epsilon\beta_q}{4}\right\} \text{ for all } k \geq N_2.$$

Moreover, for all $l \in I(\vec{x}_0)$, we have for all $k \in \mathbb{N}$ that

$$0 \geq g_l(\vec{y}_k) = g_l(\vec{x}_0) + \nabla g_l(\vec{x}_0)\delta_k\vec{s}_k + R_{2,g_l}(\vec{x}_0, \vec{y}_k) = \nabla g_l(\vec{x}_0)\delta_k\vec{s}_k + R_{2,g_l}(\vec{x}_0, \vec{y}_k),$$

where

$$\left|R_{2,g_l}(\vec{x}_0, \vec{y}_k)\right| \leq M_{2,g_l}\delta_k^2,$$

for some $M_{2,g_l} > 0$ in $\mathcal{N}$. Thus,

$$\nabla g_l(\vec{x}_0)\vec{s}_k \leq -\frac{R_{2,g_l}(\vec{x}_0,\vec{y}_k)}{\delta_k} \underset{k\to\infty}{\longrightarrow} 0.$$

Hence, there exists $N_3 \in \mathbb{N}$ such that, for all $k \geq N_3$,

$$\nabla g_l(\vec{x}_0)\vec{s}_k < \min_{q\in L}\left\{\frac{\beta_q\epsilon}{4p\beta_l}\right\} \text{ for all } l \in L \text{ and}$$
$$\nabla g_l(\vec{x}_0)\vec{s}_k < \epsilon \text{ for all } l \in I(\vec{x}_0)\setminus L.$$

Let $N > \max\{N_1, N_2, N_3\}$ be such that $Md^N < \gamma/2$, where M is as in Equation (4.15). Then

$$\text{(4.16)} \quad \left|\nabla h_j(\vec{x}_0)\vec{s}_N\right| < \min\left\{\epsilon, \min_{q\in L}\left\{\frac{\epsilon\beta_q}{4m|\alpha_j|}\right\}\right\} \text{ for all } j \in \{1,\ldots,m\},$$

$$\text{(4.17)} \quad \nabla f(\vec{x}_0)\vec{s}_N < \min_{q\in L}\left\{\frac{\epsilon\beta_q}{4}\right\},$$

$$\text{(4.18)} \quad \nabla g_l(\vec{x}_0)\vec{s}_N < \min_{q\in L}\left\{\frac{\epsilon\beta_q}{4p\beta_l}\right\} \text{ for all } l \in L,$$

$$\text{(4.19)} \quad \nabla g_l(\vec{x}_0)\vec{s}_N < \epsilon \text{ for all } l \in I(\vec{x}_0)\setminus L.$$

Two cases are to be considered.

Case 1: Assume that $\nabla g_l(\vec{x}_0)\vec{s}_N > -\epsilon$ for all $l \in L$. Then it follows from Equation (4.18) that

$$|\nabla g_l(\vec{x}_0)\vec{s}_N| < \epsilon \text{ for all } l \in L.$$

Also, from Equation (4.16), we have that

$$\left|\nabla h_j(\vec{x}_0)\vec{s}_N\right| < \epsilon \text{ for all } j \in \{1,\ldots,m\}.$$

Thus condition (iv) of the theorem entails that

$$\text{(4.20)} \quad \vec{s}_N^T\left(\nabla^2 f(\vec{x}_0) + \sum_{j=1}^m \alpha_j\nabla^2 h_j(\vec{x}_0) + \sum_{l=1}^p \beta_l\nabla^2 g_l(\vec{x}_0)\right)\vec{s}_N \geq \gamma.$$

On the other hand, replacing $\vec{x}$ by $\vec{y}_N = \vec{x}_0 + \delta_N\vec{s}_N$ in Equations (4.14) and using the fact that $\vec{y}_N$ is feasible, we obtain that

$$\begin{aligned} f(\vec{x}_0) \geq f(\vec{y}_N) &= f(\vec{y}_N) + \sum_{j=1}^m \alpha_j h_j(\vec{y}_N) \\ &\geq f(\vec{y}_N) + \sum_{j=1}^m \alpha_j h_j(\vec{y}_N) + \sum_{l=1}^p \beta_l g_l(\vec{y}_N) \\ &= f(\vec{x}_0) + \frac{1}{2}\delta_N^2\vec{s}_N^T\left(\nabla^2 f(\vec{x}_0) + \sum_{j=1}^m \alpha_j\nabla^2 h_j(\vec{x}_0) + \sum_{l=1}^p \beta_l\nabla^2 g_l(\vec{x}_0)\right)\vec{s}_N \\ &\quad + R_3(\vec{x}_0,\vec{y}_N), \end{aligned}$$

from which we obtain that

$$\vec{s}_N^T\left(\nabla^2 f(\vec{x}_0)+\sum_{j=1}^{m}\alpha_j\nabla^2 h_j(\vec{x}_0)+\sum_{l=1}^{p}\beta_l\nabla^2 g_l(\vec{x}_0)\right)\vec{s}_N \leq -\frac{2R_3(\vec{x}_0,\vec{y}_N)}{\delta_N^2},$$

where

$$|R_3(\vec{x}_0,\vec{y}_N)| \leq M|\vec{y}_N-\vec{x}_0|^3 = M\delta_N^3.$$

Thus,

$$-\frac{2R_3(\vec{x}_0,\vec{y}_N)}{\delta_N^2} \leq 2M\delta_N < 2Md^N < \gamma.$$

It follows that

$$\vec{s}_N^T\left(\nabla^2 f(\vec{x}_0)+\sum_{j=1}^{m}\alpha_j\nabla^2 h_j(\vec{x}_0)+\sum_{l=1}^{p}\beta_l\nabla^2 g_l(\vec{x}_0)\right)\vec{s}_N < \gamma,$$

which contradicts Equation (4.20).

Case 2: Assume that $\nabla g_{l_0}(\vec{x}_0)\vec{s}_N < -\epsilon$ for some $l_0 \in L$ and let $L_0 = \{l \in L : \nabla g_{l_0}(\vec{x}_0)\vec{s}_N < 0\}$. Then, using Equation (4.17) and condition (iii) of the theorem, we have that

$$\begin{aligned}\frac{\epsilon\beta_{l_0}}{4} &> \nabla f(\vec{x}_0)\vec{s}_N\\ &= -\sum_{j=1}^{m}\alpha_j\nabla h_j(\vec{x}_0)\vec{s}_N-\sum_{l\in L\setminus L_0}\beta_l\nabla g_l(\vec{x}_0)\vec{s}_N-\sum_{l\in L_0}\beta_l\nabla g_l(\vec{x}_0)\vec{s}_N\\ &> -\sum_{j=1}^{m}\alpha_j\nabla h_j(\vec{x}_0)\vec{s}_N-\sum_{l\in L\setminus L_0}\beta_l\nabla g_l(\vec{x}_0)\vec{s}_N+\epsilon\beta_{l_0}.\end{aligned}\tag{4.21}$$

But, using Equation (4.16), we have that

$$\begin{aligned}\left|-\sum_{j=1}^{m}\alpha_j\nabla h_j(\vec{x}_0)\vec{s}_N\right| &\leq \sum_{j=1}^{m}|\alpha_j||\nabla h_j(\vec{x}_0)\vec{s}_N| < \sum_{j=1}^{m}|\alpha_j|\min_{q\in L}\left\{\frac{\epsilon\beta_q}{4m|\alpha_j|}\right\}\\ &\leq \sum_{j=1}^{m}|\alpha_j|\frac{\epsilon\beta_{l_0}}{4m|\alpha_j|}=\frac{\epsilon\beta_{l_0}}{4}.\end{aligned}$$

Hence

$$-\sum_{j=1}^{m}\alpha_j\nabla h_j(\vec{x}_0)\vec{s}_N > -\frac{\epsilon\beta_{l_0}}{4}.\tag{4.22}$$

Also, using Equation (4.18), we have that

$$\begin{aligned}\left|-\sum_{l\in L\setminus L_0}\beta_l\nabla g_l(\vec{x}_0)\vec{s}_N\right| &= \sum_{l\in L\setminus L_0}\beta_l\nabla g_l(\vec{x}_0)\vec{s}_N < \sum_{l\in L\setminus L_0}\beta_l\min_{q\in L}\left\{\frac{\epsilon\beta_q}{4p\beta_l}\right\}\\ &= \sum_{l\in L\setminus L_0}\min_{q\in L}\left\{\frac{\epsilon\beta_q}{4p}\right\} \leq \frac{\epsilon\beta_{l_0}}{4}.\end{aligned}$$

Hence

$$(4.23) \qquad -\sum_{l\in L\setminus L_0} \beta_l \nabla g_l(\vec{x}_0)\vec{s}_N > -\frac{\epsilon\beta_{l_0}}{4}.$$

Substituting Equation (4.22) and Equation (4.23) into Equation (4.21), we obtain that

$$\frac{\epsilon\beta_{l_0}}{4} > -\frac{\epsilon\beta_{l_0}}{4} - \frac{\epsilon\beta_{l_0}}{4} + \epsilon\beta_{l_0} = \frac{\epsilon\beta_{l_0}}{2},$$

a contradiction since $\beta_{l_0} > 0$. Thus, $\vec{x}_0$ is a strict local minimum for f over the constraints of Equation (1.1). $\square$

Example 4.4. Minimize

$$f(x_1, x_2, x_3) = dx_2 + dx_3 - x_1x_2 - x_1x_3 - x_2x_3,$$

subject to the constraints

$$(4.24) \qquad \begin{cases} x_1 + x_2 & = 2 + d \\ \quad\;\; x_1 & \leq 1 + d \\ \quad\;\; x_2 & \leq 2 \\ \quad\;\; x_3 & \leq 3 - d + 2d^2 + 5d^3, \end{cases}$$

where d is the infinitely small number defined in the introduction.

For the function f to have a local minimum at a regular point $\vec{x}_0 = (x_1, x_2, x_3)^T$ subject to the constraints in Equation (4.24), the necessary conditions of Theorem 4.2 must hold at $\vec{x}_0$. The first order conditions of that theorem entail that there must exist $\alpha, \beta_1, \beta_2, \beta_3 \in \mathcal{N}$ such that

$$(4.25) \qquad \begin{cases} \beta_l \geq 0 \text{ for } l = 1, 2, 3, \\ \beta_1(x_1 - 1 - d) = 0, \\ \beta_2(x_2 - 2) = 0, \\ \beta_3(x_3 - 3 + d - 2d^2 - 5d^3) = 0, \\ -x_2 - x_3 + \alpha + \beta_1 = 0, \\ d - x_1 - x_3 + \alpha + \beta_2 = 0, \\ d - x_1 - x_2 + \beta_3 = 0. \end{cases}$$

Using the constraints in Equation (4.24), a close inspection of the conditions in Equation (4.25) shows that those conditions are simultaneously satisfied only at

$$(4.26) \qquad \begin{cases} x_1 = 1 + d, \\ x_2 = 1, \\ x_3 = 3 - d + 2d^2 + 5d^3, \\ \alpha = 4 - d + 2d^2 + 5d^3, \\ \beta_1 = 0, \\ \beta_2 = 0, \\ \beta_3 = 2. \end{cases}$$

With $\vec{x}_0 = (1 + d, 1, 3 - d + 2d^2 + 5d^3)^T$, and using the notations of the Proof of Theorem 4.3, we have here $I(\vec{x}_0) = \{1, 3\}$ and $L = \{3\}$. Since $\nabla h(\vec{x}_0)$, $\nabla g_1(\vec{x}_0)$ and $\nabla g_3(\vec{x}_0)$ are linearly independent, the point $\vec{x}_0$ is regular for the constraints.

To show that $\vec{x}_0$ is indeed a strict local minimizer of f subject to the constraints in Equation (4.24), it remains to show that condition (iv) of Theorem 4.3 holds at $\vec{x}_0$ for the choices of $\alpha, \beta_1, \beta_2, \beta_3$ in Equation (4.26). Let $\epsilon = d$ and $\gamma = 1/2$. Then for all $\vec{y} \in \mathcal{N}^3$ satisfying $|\vec{y}| = 1$, $|\nabla h(\vec{x}_0)\vec{y}| < \epsilon$, $|\nabla g_3(\vec{x}_0)\vec{y}| < \epsilon$ and $\nabla g_1(\vec{x}_0)\vec{y} < \epsilon$, we have that

$$y_1^2 + y_2^2 + y_3^2 = 1, |y_1 + y_2| < d, |y_3| < d, \text{ and } y_1 < d.$$

It follows that

$$\begin{aligned}
&\vec{y}^T\left(\nabla^2 f(\vec{x}_0) + \alpha\nabla^2 h(\vec{x}_0) + \sum_{l=1}^{3} \beta_l \nabla^2 g_l(\vec{x}_0)\right)\vec{y} = \vec{y}^T\nabla^2 f(\vec{x}_0)\vec{y} \\
&= (y_1\ y_2\ y_3)\begin{pmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \\
&= -2y_1y_2 - 2y_1y_3 - 2y_2y_3 \\
&= -(y_1 + y_2 + y_3)^2 + (y_1^2 + y_2^2 + y_3^2) \\
&= 1 - (y_1 + y_2 + y_3)^2 \\
&= 1 - (y_1 + y_2)^2 - y_3^2 - 2y_3(y_1 + y_2) \\
&\geq 1 - |y_1 + y_2|^2 - |y_3|^2 - 2|y_3||y_1 + y_2| \\
&> 1 - d^2 - d^2 - 2d^2 = 1 - 4d^2 \\
&> \frac{1}{2} = \gamma.
\end{aligned}$$

Thus the conditions of Theorem 4.3 are satisfied at $\vec{x}_0$, and hence $\vec{x}_0$ is a strict local minimizer of f under the constraints in Equation (4.24).

Remark 4.5. In the example above, any infinitely small ϵ_0 can replace d and any positive real number γ_0 smaller than 1 can replace $1/2$ in showing that condition (iv) of Theorem 4.3 holds. This is so since $1 - \epsilon_0^2 > \gamma_0$ for all such ϵ_0 and γ_0.

5. ACKNOWLEDGMENTS

For their nice hospitality, the first author is thankful to the faculty and staff of the Center for Advanced Mathematical Sciences (CAMS) at the American University of Beirut (AUB), where much of this paper was written during a visit to CAMS for the whole Summer of 2001.

REFERENCES

[1] Berz, M. – Calculus and Numerics on Levi-Civita Fields. In Berz, M., C. Bischof, G. Corliss, and A. Griewank, editors. Computational Differentiation: Techniques, Applications, and Tools, pp. 19–35, Philadelphia, SIAM (1996).

[2] Berz, M. – Nonarchimedean analysis and rigorous computation. International Journal of Applied Mathematics **2**, 889–930 (2000).

[3] Fiacco, Anthony V. and Garth P. McCormick – Nonlinear Programming; Sequential Unconstrained Minimization Techniques. SIAM, Philadelphia (1990).
[4] Fleming, W. – Functions of Several Variables. Springer-Verlag, New York (1976).
[5] Hahn, H. – Über die nichtarchimedischen Größensysteme. Sitzungsbericht der Wiener Akademie der Wissenschaften Abt. 2a **117**, 601–655 (1907).
[6] Krull, Wolfgang – Allgemeine Bewertungstheorie. J. Reine Angew. Math. **167**, 160–196 (1932).
[7] Levi-Civita, Tullio – Sugli infiniti ed infinitesimi attuali quali elementi analitici. Atti Ist. Veneto di Sc., Lett. ed Art., 7a **4**, 1765 (1892).
[8] Levi-Civita, Tullio – Sui numeri transfiniti. Rend. Acc. Lincei, 5a **7**, 91,113 (1898).
[9] Luenberger, David G. – Linear and Nonlinear Programming. Addison-Wesley, Reading, Massachusetts, 2nd edition (1984).
[10] MacLane, S. – The universality of formal power series fields. Bulletin American Mathematical Society **45**, 888 (1939).
[11] Priess-Crampe, S. – Angeordnete Strukturen: Gruppen, Körper, projektive Ebenen. Springer, Berlin (1983).
[12] Rayner, F.J. – Algebraically Closed Fields Analogous to Fields of Puiseux Series. J. London Math. Soc. **8**, 504–506 (1974).
[13] Ribenboim, Paulo – Fields: Algebraically Closed and Others. Manuscripta Mathematica **75**, 115–150 (1992).
[14] Shamseddine, K. – New Elements of Analysis on the Levi-Civita Field. PhD thesis, Michigan State University, East Lansing, Michigan, USA, 1999. Also Michigan State University report MSUCL-1147.
[15] Shamseddine, K. and M. Berz – Exception Handling in Derivative Computation with Non-Archimedean Calculus. In: Berz, M., C. Bischof, G. Corliss and A. Griewank, editors. Computational Differentiation: Techniques, Applications, and Tools, pp. 37–51, Philadelphia. SIAM (1996).
[16] Shamseddine, K. and V. Zeidan – One-Dimensional Optimization on Non-Archimedean Fields. Journal of Nonlinear and Convex Analysis **2**, 351–361 (2001).

(Received 1 August 2002)