SURVEY

Real Analysis Exchange
Vol. 35(2), 2009/2010, pp. 311–342

Ivan Ginchev, University of Insubria, Department of Economics, Via Monte Generoso 71, 21100 Varese, Italy. email: ivan.ginchev@uninsubria.it

Davide La Torre, University of Milan, Department of Economics, Business and Statistics, Via Conservatorio 7, 20122 Milano, Italy. email: davide.latorre@unimi.it

Matteo Rocca, University of Insubria, Department of Economics, Via Monte Generoso 71, 21100 Varese, Italy. email: matteo.rocca@uninsubria.it

$C^{k,1}$ FUNCTIONS, CHARACTERIZATION, TAYLOR'S FORMULA AND OPTIMIZATION: A SURVEY

Abstract

The present paper is a survey on $C^{k,1}$ functions. Both theoretical and numerical results related to this class of nonsmooth functions are presented. We also give few new results and pose some open problems for further investigation.

1 Introduction.

Much of classical calculus is based on the notions and properties of gradient and differential; major subjects such as optimization and differential equations heavily involve the notion of derivative. As a result, to develop any kind of nonsmooth calculus this definition has to be replaced by a new one, trying to preserve, in some generalized form, most of the basic properties and results. This is what several authors have done; extending classical calculus by first modifying the concept of gradient. Sometime these new notions have been defined as generalized derivatives, gradients or differentials. Different schools

Mathematical Reviews subject classification: Primary: 26A03, 26A04, 26B05, 26B10; Secondary: 26A05

Key words: $C^{k,1}$ functions, generalized Taylor's formula, implicit function theorem, nonsmooth optimization

Received by the editors February 8, 2010

Communicated by: Paul D. Humke

introduced and constructed new tools and theories that allow treating several nonsmooth function classes (see, for instance, [21, 66, 68, 69, 86, 89, 76]). One of the most important such classes is the one of Lipschitz functions from $\mathbb{R}^n$ to $\mathbb{R}$; this class of functions has many important properties and many applications. They have been widely used in the literature (see for instance [38]). A smaller class, which is also very important is the collection of the so called $C^{1,1}$ functions, i.e. Fréchet differentiable functions with locally Lipschitz derivatives. This definition was introduced by Hiriart-Urruty and others ([40, 39]). Hiriart-Urruty also introduced the concept of a generalized Hessian matrix for $C^{1,1}$ functions and proved second order optimality conditions for nonlinear constrained problems. Many authors have highlighted relevant real applications in which second order differentiability of the involved data cannot be assumed but for which $C^{1,1}$ regularity holds. For instance, the extended linear-quadratic programming problem used in the context of stochastic programming and optimal control, even in the fully quadratic case, doesn't use a twice differentiable objective function; however these objective functions are differentiable and their derivatives are Lipschitzian. The augmented Lagrangian method of a twice smooth nonlinear programming problem is another example. On the other hand, from a computational perspective, the interest in $C^{1,1}$ functions comes from the fact that several numerical schemes need the Lipschitz property of their derivatives to be convergent. A natural generalization is the class of $C^{k,1}$ functions, with k being a positive integer. A function f from $\mathbb{R}^n$ (or from an open set in $\mathbb{R}^n$) to $\mathbb{R}$ is said to be $C^{k,1}$ if f is Fréchet differentiable up to order k with a locally Lipschitz derivative of order k. We say that f is $C^{k,1}$ near $x \in \mathbb{R}^n$ if there is an open neighborhood U of x such that f restricted to U is $C^{k,1}$. The class $C^{0,1}$ is defined as the class of locally Lipschitz functions. The class of $C^{k,1}$ functions has been considered by several authors in the literature; for instance, Luc [63], considering the class of $C^{k,1}$ functions, introduced the notion of a generalized k-th differential, extended Taylor's formula, proved higher order optimality conditions, and provided characterizations of generalized convex functions. Many other generalized derivatives have been introduced to deal with problems involving $C^{k,1}$ data. Among them are notions due to Dini-Hadamard [13], Peano and Riemann [1, 2, 22, 51, 70, 72], Michel and Penot [67], and Cominetti and Correa [14].

We now list three examples to show the importance of this class of nonsmooth functions in real applied problems.

Example 1.1. *Many problems in science and engineering (see, for instance, [76] and the references therein) can be formulated in terms of a nonsmooth* *semi-infinite optimization problem such as the following:*

$$\textit{Minimize } f(x) \tag{1.1}$$

$$\textit{subject to } \max_{t\in[a,b]} \phi_j(x,t) \leq 0,\ j = 1,\dots,l \tag{1.2}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ *is* $C^{1,1}$ *and* $\phi_j : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ *is* C^2, $j = 1,\dots,$. *A possible approach for solving this kind of problem is to convert it into equality constraints*

$$h_j(x) = \int_a^b \left[\max\left\{\phi_j(x,y), 0\right\}\right]^2 dt = 0,\ j = 1,\dots,l. \tag{1.3}$$

Since ϕ_j *is* C^2, *it is easy see that the function* h_j *is* $C^{1,1}$ *with gradient*

$$\nabla h_j(x) = 2\int_a^b \max\{\phi_j(x,t), 0\}\nabla\phi_j(x,t)dt,\ j = 1,\dots,l. \tag{1.4}$$

Example 1.2. *Consider the following minimization problem: Minimize* $f(x)$ *over all* $x \in \mathbb{R}^N$ *such that* $g_1(x) \leq 0,\dots,g_m(x) \leq 0$. *Letting* r *denote a positive parameter, the augmented Lagrangian* L_r *(see [76] and references therein) is defined on* $\mathbb{R}^n \times \mathbb{R}^m$ *as*

$$L_r(x,y) = f(x) + \frac{1}{4r}\sum_{i=1}^m \{[y_i + 2rg_i(x)]^+\}^2 - y_i^2. \tag{1.5}$$

Then $L_r(x,\cdot)$ *is concave and* $L_r(\cdot,y)$ *is convex whenever the minimization problem is a convex minimization problem. By replacing* $y = 0$ *we have*

$$L_r(x,0) = f(x) + r\sum_{i=1}^m [g_i^+(x)]^2 \tag{1.6}$$

which is the ordinary penalized version of the minimization problem. L_r *is differentiable everywhere on* $\mathbb{R}^n \times \mathbb{R}^m$ *with*

$$\nabla_x L_r(x,y) = \nabla f(x) + \sum_{j=1}^m [y_j + 2rg_j(x)]^+ \nabla g_j(x) \tag{1.7}$$

$$\frac{\partial L_r}{\partial y_i}(x,y) = \max\left\{g_i(x), -\frac{y_i}{2r}\right\}, i = 1\dots m. \tag{1.8}$$

Example 1.3. *Several formulations of applied problems involve strongly discontinuous data. For instance, for a given vector $x \in \mathbb{R}^n$ the step function* $\mathrm{supp}(x)$*, which counts the number of positive components of x, can be found in portfolio optimization, data classification and neural networks (see [49, 58]). One possible approach for dealing with this kind of function consists of approximating the problem by a nonlinear continuous function which can be easily proved to be a $C^{1,1}$ function. This allows approximating a strongly discontinuous optimization problem by a $C^{1,1}$ optimization problem for which a well established numerical scheme is available.*

Many results have also been extended by dealing with vector $C^{k,1}$ functions, set valued $C^{k,1}$ functions, multi-objective and set valued $C^{k,1}$ optimization problems. Among them are the papers [15, 29, 30, 31, 32, 33, 34, 36, 37, 49, 50, 60, 62]. The present paper is a survey of $C^{k,1}$ functions, mostly concentrating on the case $k \geq 1$. We shall also give a few new results and pose some problems for further investigation. The paper is organized as follows. Section 2 is devoted to a characterization of $C^{k,1}$ functions; in Section 3 a generalized Taylor's formula is provided; in Section 4 optimality conditions for both unconstrained and constrained problems are presented as well as some numerical techniques; while Section 5 concludes with some final remarks and several applications of $C^{k,1}$ functions to boundary value problems.

2 Characterization Of $C^{k,1}$ Functions.

In this section we provide a characterization of $C^{k,1}$ functions through k-th discretized differences. Our approach is mainly based on the papers [51, 52, 53]. There are other different characterizations in the literature provided by several authors and using different notions of generalized derivatives. We give a short overview of one of them at the end of this section. The main advantage of using discretized differences consists of having a characterization which doesn't require any kind of extra regularity assumption on f or its derivatives.

Let $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ be a function defined on an open set Ω. For such a function let

$$\Theta^k f(x,t,w) = \sum_{i=0}^{k} (-1)^{k-i} \binom{k}{i} f(x + itw - \frac{1}{2}ktw). \tag{2.1}$$

where $w \in \mathbb{R}^n$ and $t \in \mathbb{R}$. Similarly one can introduce

$$\theta^k f(x,t,w) = \sum_{i=0}^{k} (-1)^{k-i} \binom{k}{i} f(x + itw). \tag{2.2}$$

Let us note that whenever $t = 0$ or $w = 0$ we have $\Theta^k f(x, t, w) = \theta^k f(x, t, w) = 0$. It is also easy to see that $\theta^k f(x, t, w) = \Theta^k f(x + \frac{k}{2}tw, t, w)$. Through discretized differences one can also introduction a notion of generalized convexity, as stated in the following definition.

Definition 2.1. *[10, 27] A continuous function $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ is said to be locally k-convex at $x_0 \in \Omega$ when*

$$\frac{\Theta^{k+1} f(x, t, w)}{t^{k+1}} \geq 0, \tag{2.3}$$

$\forall x$ in a neighborhood U of x_0, $\forall w \in \mathbb{R}^n$ and $\forall t$ such that $x \pm \frac{k}{2}tw \in \Omega$.

When $k = 1$ the previous definition reduces to that for convex functions. Now suppose that $f : \Omega \to \mathbb{R}$ be a function such that $\Theta^{k+1}(x, t, w)/t^{k+1} \geq M$, for each x in a neighborhood U of x_0, t in a neighborhood V of 0 and $w \in S^1$. If $M \geq 0$, then f is obviously k-convex. If $M < 0$, let

$$p(x) = p(x_1, \ldots, x_n) = \sum_{j=0}^{k+1} \sum_{i_1+i_2+\ldots+i_n=j} c_{i_1,\ldots,i_n} x_1^{i_1} \cdots x_n^{i_n} \tag{2.4}$$

be a polynomial of degree at most $k+1$ in the variables $x_1, \ldots x_n$. It is known that, letting $w = (w_1, \ldots, w_n)$,

$$\frac{\Theta^{k+1} p(x, t, w)}{t^{k+1}} = \sum_{i_1+i_2+\ldots+i_n=k+1} c_{i_1,\ldots,i_n} w_1^{i_1} \cdots w_n^{i_n} \tag{2.5}$$

so that one can always choose the coefficients of the polynomial so that

$$\inf_{w \in S^1} \frac{\Theta^{k+1} p(x, t, w)}{t^{k+1}} \geq -M \tag{2.6}$$

for every x and t, and hence the function $f(x) + p(x)$ is locally k-convex at x_0. The following result provides a characterization of $C^{k,1}$ functions.

Theorem 2.2. *[53] Assume that the function $f : \Omega \to \mathbb{R}$ is bounded on a neighborhood of the point $x_0 \in \Omega$. Then f is of class $C^{k,1}$ at x_0 if and only if there exist neighborhoods U of x_0 and $V \subset \mathbb{R}$ of 0 such that $\frac{\Theta^{k+1} f(x,t,w)}{t^{k+1}}$ is bounded on $U \times V \backslash \{0\}$ uniformly with respect to $w \in S^1$.*

Corollary 2.3. *[53] Assume that the function f is bounded on a neighborhood of x_0. Then f is of class $C^{k,1}$ at x_0 if and only if there exist neighborhoods U of x_0 and V of 0 such that $\frac{\theta^{k+1} f(x,t,w)}{t^{k+1}}$ is bounded on $U \times V \backslash \{0\}$, uniformly with respect to $w \in S^1$.*

Through the notions of discretized differences it is possible to introduce the notions of Riemann derivatives. The k-th Riemann derivative of f at a point $x \in \Omega$ in the direction w is defined as

$$D_R^k f(x,w) = \lim_{t \to 0} \frac{\Theta^k f(x,t,w)}{t^k}, \tag{2.7}$$

if this limit exists. If the existence of the limit is replaced by the limsup or liminf one can introduce the notion of upper and lower Riemann derivatives as follows:

$$\overline{D}_R^k f(x,w) = \limsup_{t \to 0} \frac{\Theta^k f(x,t,w)}{t^k}, \tag{2.8}$$

and

$$\underline{D}_R^k f(x,w) = \liminf_{t \to 0} \frac{\Theta^k f(x,t,w)}{t^k}. \tag{2.9}$$

The corresponding k-th Riemann-type derivative is defined as

$$d_R^k f(x,w) = \lim_{t \to 0} \frac{\theta^k f(x,t,w)}{t^k}. \tag{2.10}$$

For a survey on Riemann derivatives and their relationships with other definitions of generalized derivatives one can see for instance [1, 2, 11, 22, 27, 26, 70]. The previous characterization can be now reformulated in terms of Riemann derivatives, as stated in the following corollary.

Corollary 2.4. *[53] Assume that f is continuous on a neighborhood of the point x_0. Then f is of class $C^{k,1}$ at x_0 if and only if both $\overline{D}_{k+1} f(x)$ and $\underline{D}_{k+1} f(x)$ are bounded on a neighborhood of x_0.*

Previous results extend the elementary condition which relates the Lipschitzian condition on $f^{(k)}$ and the boundedness of $f^{(k+1)}$. We have generalized this relation without requiring any differentiability hypothesis and linking the existence and the Lipschitz behaviour of $f^{(k)}$ to the boundedness of $\frac{\Theta^{k+1} f(x,t,w)}{t^{k+1}}$ or of the upper and lower Riemann derivatives. Similar conditions can be expressed in terms of $\underline{d}^{k+1} f$ and $\overline{d}^{k+1} f$. Further extensions of these results can be obtained for the class of $C^{k,\alpha}$ functions that is the set of all functions for which $f^{(k)}$ exists in a neighborhood of x_0 and $f^{(k)}$ is locally Hölderian of degree α at x_0. In particular, for this class of functions, it can be proved that the boundedness of certain discretized differences is related to higher order smooothness and quasi-smoothness conditions (see [11, 19, 20, 23]).

We are going to conclude this section by presenting an interesting characterization of the class of $C^{1,1}$ function, which has been provided by Ioffe in [46] and based on the notion of Clarke and Dini generalized derivatives. Before presenting the main result, we need to introduce the following definitions and terminology. For a locally Lipschitz function f,

$$f^o(x,w) = \limsup_{y \to x, t \to 0^+} \frac{f(y+tw) - f(x)}{t} \tag{2.11}$$

is the Clarke directional derivative of f at x along the direction w ([13]). We say that f is Clarke regular at x if

$$f^o(x,w) = d^- f(x,w) = \liminf_{t \to 0^+} \frac{f(x+tw) - f(x)}{t} \tag{2.12}$$

for all w. The quantity $d^- f(x,w)$ is called the lower Dini directional derivative of f at x along w. The upper Dini directional derivative $d^+ f(x,w)$ is defined in the same way with lim inf replaced by lim sup. Finally, recall that

$$\partial^- f(x) = \{x^* : \langle x^*, w\rangle \leq d^- f(x,w), \forall w\} \tag{2.13}$$

is the Dini subdifferential of f at x. We are ready to present the following result.

Theorem 2.5. *[46] Let f be a real-valued function that is defined and locally Lipschitz on an open subset Ω of $\mathbb{R}^n$. Then the following properties are equivalent:*

1. *f is $C^{1,1}$ on Ω;*
2. *f is Clarke regular on Ω and $f^o(\cdot,\cdot)$ satisfies the Lipschitz condition in a neighborhood of every $(x,0) \in \Omega \times \mathbb{R}^n$;*
3. *$\partial^- f(x)$ is nonempty for any $x \in \Omega$ and locally Lipschitz (as a set-valued mapping from $\mathbb{R}^n$ into $\mathbb{R}^n$);*
4. *$d^- f(x,w)$ satisfies the Lipschitz condition in a neighborhood of every $(x,0) \in \Omega \times \mathbb{R}^n$;*

We remark that other characterizations of $C^{1,1}$ functions that involve different notions of second order generalized derivatives can be found in [14, 40, 67, 71].

3 Taylor's Formula and the Implicit Function Theorem.

In this section, following D. T. Luc [63] and using Clarke's generalized Jacobian [13], a Taylor formula for a $C^{k,1}$ function $f : \mathbb{R}^n \to \mathbb{R}$ is presented. This result is applied to derive calculus rules for the generalized Hessian in the Implicit Function Theorem with $C^{k,1}$ functions and second-order characterization of quasiconvex functions.

3.1 Taylor's Formula.

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a C^{k+1} function. The classical Taylor Theorem states that for every pair of points $a, b \in \mathbb{R}^n$, there is a point c in the open interval $c \in (a, b)$ such that

$$f(b)-f(a) = \sum_{i=1}^{k} \frac{1}{i!} D^i f(a)(b-a, \dots, b-a) + \frac{1}{(k+1)!} D^{k+1} f(c)(b-a, \dots, b-a) .$$

where $D^i f(a) : \mathbb{R}^n \times \mathbb{R}^n \times \dots \mathbb{R}^n (\text{i times}) \to \mathbb{R}$ denotes the multilinear mapping defined on $(\mathbb{R}^n)^i$ which represents the differential of f of order i at the point a. Theorem 3.1 below generalizes Taylor's Theorem for $C^{k,1}$ functions. It applies the notion of a subgradient of higher order as defined in the sequel. For every $f \in C^{k,1}$ by Rademacher's Theorem, its k-th order derivative $D^k f$ is differentiable almost everywhere. (For the exact formulation and the proof of Rademacher's Theorem see e.g. [38].) The generalized Jacobian of $D^k f$ at $x \in \mathbb{R}^n$ in Clarke's sense [13], denoted by $\bar{J} D^k f(x)$, is defined as the convex hull of all $(n^k \times n)$-matrices obtained as the limit of a sequence of the form $J D^k f(x_i)$ where $\{x_i\}_{i=1}^{\infty}$ converges to x and the classical Jacobian matrix $J D^k f(x_i)$ of $D^k f$ at x_i exists. The $(k+1)$-th order subdifferential of f at x is defined as the set $\partial^{k+1} f(x) := \bar{J} D^k f(x)$. The elements of this set are called the $(k+1)$-th order subgradients of f at x. They can be considered as multilinear functions on the space $\mathbb{R}^n \times \dots \times \mathbb{R}^n$ $(k+1$ times). When $k = 1$ the set $\partial f(x) := \partial^1 f(x) = \bar{J} f(x)$ is the Clarke subdifferential of f at x and its elements are the Clarke subgradients of f at x. The space of the $(n^k \times n)$-matrices $A = (a_{ij})$, $i = 1, \dots n^k$, $j = 1, \dots, n$, is endowed with the norm $\|A\| = \left(\sum_{i=1}^{n^k} \sum_{j=1}^{n} a_{ij}^2 \right)^{1/2}$.

Theorem 3.1 (Luc [63]). *Let f be a $C^{k,1}$ function from $\mathbb{R}^n$ to $\mathbb{R}$ and let a, b be two arbitrary points in $\mathbb{R}^n$. Then there exist a point $c \in (a, b)$ and a*

$(k+1)$-th order subgradient A_b of f at c such that

$$f(b)-f(a) = \sum_{i=1}^{k} \frac{1}{i!} D^i f(a)(b-a,\ldots,b-a)+\frac{1}{(k+1)!} A_b(b-a,\ldots,b-a)\,. \quad (3.1)$$

Moreover, there exists a neighborhood U of a and a positive K such that $\|A_b\| \leq K$ for all $b \in U$.

For $k = 0$ this theorem gives:

Corollary 3.2 (Lebourg Mean Value Theorem [59])**.** *Let f be a $C^{0,1}$ function from $\mathbb{R}^n$ to $\mathbb{R}$ and let a, b be two arbitrary points in $\mathbb{R}^n$. Then there exist a point $c \in (a,\, b)$ and a subgradient A_b of f at c such that*

$$f(b) - f(a) = A_b(b-a)\,.$$

Moreover, there exists a neighborhood U of a and a positive K such that $\|A_b\| \leq K$ for all $b \in U$.

For references, other generalizations and applications of the Mean Value Theorem we refer to [64].

3.2 The Implicit Function Theorem.

The Implicit Function Theorem plays an important role in analysis. For $C^{k,1}$ functions the following result holds.

Theorem 3.3 (The Implicit Function Theorem)**.** *Let $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ be a $C^{k,1}$ function with the property that $f(x_0, y_0) = 0$ and suppose that every matrix of $\pi_y \partial f(x_0, y_0)$ is invertible. Then there exists an open neighborhood U of x_0 and an open neighborhood V of y_0 such that for any $x \in U$ there exists a unique $g(x) \in V$ with the property that $f(x, g(x)) = 0$. The function $g : U \to V$ satisfies $g(x_0) = y_0$ and $g \in C^{k,1}$.*

Here π_y stands for the projection operator on the y-space, in this case $\mathbb{R}^m$. When $k = 0$ this theorem is established in [13]. When $k \geq 1$, then f is also C^1, the condition "every matrix of $\pi_y \partial f(x_0, y_0)$ is invertible" can be written in the form $\nabla_y f(x_0, y_0) \neq 0$. The assertion with $g \in C^1$ is the classical Implicit Function Theorem. The fact that $g \in C^{k,1}$ is established in [63].

Since the function g in Theorem 3.3 is $C^{k,1}$, an important problem is to give a formula for $\partial^{k+1} f(x_0)$. Such formulas when $k = 0$ and $k = 1$ are established in [63].

a) When $k = 0$, then

$$\partial g(x_0) \subseteq -\left(\partial_y f(x_0, y_0)\right)^{-1} \partial_x f(x_0, y_0)\,.$$

b) When $k = 1$, then the following partition for every $(n+m) \times (n+m)$ matrix H is used.

$$H = \begin{pmatrix} H_{xx} & H_{xy} \\ H_{yx} & H_{yy} \end{pmatrix},$$

where the dimensions of the submatrices H_{xx}, H_{xy}, H_{yx}, H_{yy} are $n \times n$, $n \times m$, $m \times n$, $m \times m$ respectively. With this partition

$$\begin{aligned}\partial^2 g(x_0) \subseteq -\left(\partial_y f(x_0, y_0)\right)^{-1} \{H_{xx} + (Dg(x_0))^\top H_{yx} + H_{xy} Dg(x_0) \\ + H_{yy} (Dg(x_0))^\top D(g(x_0)) \mid H \in \partial f(x_0, g(x_0))\}\,,\end{aligned}$$

where $(\dots)^\top$ denotes the transpose of the matrix in the parentheses.

3.3 Quasiconvexity.

We recall that a function f from $\mathbb{R}^n$ to $\mathbb{R}$ is said to be quasiconvex if for every $x,\, y \in \mathbb{R}^n$ and every $\lambda \in (0,\, 1)$ one has $f(\lambda x + (1-\lambda) y) \le \max(f(x),\, f(y))$. Conditions for quasiconvexity of $C^{1,1}$ functions are obtained in [63] as an application of the results from the previous two subsection. The necessary conditions in Theorem 3.4 is derived on the basis of the Taylor's formula (3.1) while the proof of the sufficient conditions in Theorem 3.5 uses also the Implicit Function Theorem 3.3. See also [16] for additional information on this. The formulations use the following notations (with $k = 1$).

$$D_+^{k+1} f(x)(u) = \sup\{A(u, \dots, u) \mid A \in \partial^{k+1} f(x)\}\,, \tag{3.2}$$

$$D_-^{k+1} f(x)(u) = \inf\{A(u, \dots, u) \mid A \in \partial^{k+1} f(x)\}\,. \tag{3.3}$$

Theorem 3.4 (Quasiconvexity, Necessary condition, [63])**.** *Let f be a quasiconvex $C^{1,1}$ function. Then for every $x,\, u \in \mathbb{R}^n$, $Df(x)(u) = 0$ implies $D_+^2 f(x)(u) \ge 0$.*

The following theorem generalizes from C^2 to $C^{1,1}$ functions a result of Crouzeix [16].

Theorem 3.5 (Quasiconvexity, Sufficient condition, [63])**.** *Let the $C^{1,1}$ function f satisfy the following conditions for every $x,\, u \in \mathbb{R}^n$, $u \ne 0$:*

$$Df(x)(u) = 0 \quad \textit{implies} \quad D_-^2 f(x)(u) \ge 0\,, \tag{3.4}$$

$$\textit{and} \quad Df(x) = 0 \quad \textit{implies} \quad D_-^2 f(x)(u) > 0\,. \tag{3.5}$$

Then f is quasiconvex.

Under the hypotheses of Theorem 3.5 the function f is pseudoconvex, (see the Remark in [63, page 668]). Recall that the function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be pseudoconvex if for all $x,\ y \in \mathbb{R}^n$, $f(y) < f(x)$ implies $Df(x)(y - x) \leq 0$.

The pseudoconvex functions generalize the convex functions and obey some similar properties; for instance every local minimum of a pseudoconvex function is a global minimum. A larger class with similar properties is the class of pseudoconvex functions of order k([35]). A natural question, left here as an open problem, is whether Theorem 3.5 admits an appropriate generalization with respect to the class of $C^{k,1}$ pseudoconvex functions of order k.

4 Optimization.

In this section we deal with optimality conditions for problems with $C^{k,1}$ functions. Initially following Luc [63] we present optimality conditions based on Taylor's formula from the previous section. Thereafter we prove Taylor's formula with an integral representation of the remainder and on its basis establish optimality conditions for $C^{k,1}$ functions using Dini directional derivatives of order $k + 1$. We show that this new result implies that of [63]. We follow an approach similar to that of Ginchev, Guerraggio, Rocca [30], [29], [31], [34] for vector optimization problems with $C^{1,1}$ data (and that of [28], [32], [29] for vector optimization problems with $C^{0,1}$ data). In the present paper for simplicity we restrict our discussion to scalar problems.

The importance of the locally Lipschitz, that is $C^{0,1}$, optimization both from theoretical and practical points of view is well known. Actually it gave birth to an intensively studied, vast area of mathematics referred to recently as variational analysis, see [76], [68], [69]. Optimization problems with $C^{1,1}$ data were considered initially by Hiriart-Urruty, Strodiot, and Hien Nguen [40]. Since then such problems, both scalar and vector, have been frequently studied. For instance second-order conditions for $C^{1,1}$ scalar problems are studied in [17, 25, 43, 44, 45, 48, 54, 55, 56, 57, 88, 87, 91, 92] and for $C^{1,1}$ vector problems in [36, 37, 49, 50, 60, 61, 62]. Here is a short explanation about the choice of the Dini derivatives in the optimality conditions. In [26] optimality conditions are proposed in terms of the so called Hadamard derivatives which work with quite arbitrary functions. The Hadamard derivatives are however inconsistent with the classical derivatives in the case of differentiable functions. That is why in [30] Dini derivatives were considered instead of Hadamard ones, which on the one hand have similar structure to Hadamard derivatives, while on the other hand are consistent with the classical derivatives. However it is shown in [30] that second-order conditions designed like the ones in [26] but applying Dini derivatives do not work for arbitrary functions. The question is posed to find a class of functions $\mathcal{F}$ for which these conditions work. It is shown there that an appropriate choice is the class $\mathcal{F} = C^{1,1}$. Recently Bednařík and Pastor [4] introduced a more general class of functions called ℓ-stable functions, which also could be an appropriate choice. In our opinion the class of ℓ-stable functions deserves a separate study from the point of view of real analysis. It is shown in [3, 4, 5, 6, 7] that the optimality conditions for scalar problems with ℓ-stable functions not only generalize those of [30], but also the ones from [8], [14] and [3].

4.1 Some Optimality Conditions.

Taylor's formula from Section 3 is applied to derive the following optimality conditions.

Theorem 4.1 (Necessary conditions, [63]). *Let $x_0 \in \mathbb{R}^n$ be a local minimum of f with the property that $D^i f(x_0) = 0$ for $i = 1, \ldots, k$. Then $D_+^{k+1} f(x_0)(u) \geq 0$ for all $u \in \mathbb{R}^n$. In particular, if k is even, then $0 \in \partial^{k+1} f(x_0)(u, \ldots, u)$ for all $u \in \mathbb{R}^n$.*

Theorem 4.2 (Sufficient conditions, [63]). *Let $x_0 \in \mathbb{R}^n$ be a point with the property that $D^i f(x_0) = 0$ for $i = 1, \ldots, k$, and $D_-^{k+1} f(x_0)(u) > 0$ for all $u \in \mathbb{R}^n$, $u \neq 0$. Then x_0 is a local strict minimum of f.*

4.2 Taylor's Formula With An Integral Form Of The Rest.

We establish a variant of Taylor's formula.

Theorem 4.3. *Let f be a $C^{k,1}$ function from $\mathbb{R}^n$ to $\mathbb{R}$ and let a, b be two arbitrary points in $\mathbb{R}^n$. Then*

$$f(b) - f(a) = \sum_{i=1}^{k} \frac{1}{i!} D^i f(a)(b-a, \ldots, b-a) + R_{k+1}\,,$$

where

$$\begin{aligned} R_{k+1} = {} & \frac{1}{(k-1)!} \int_0^1 (1-s)^{k-1} \big(D^k f(a + s(b-a))(b-a, \ldots, b-a) \\ & - D^k f(a))(b-a, \ldots, b-a) \big)\, ds \qquad (4.1) \\ = {} & \frac{1}{k!} \int_0^1 (1-s)^k D^{k+1} f(a + s(b-a))(b-a, \ldots, b-a)\, ds\,. \end{aligned}$$

PROOF. Define the function

$$g(t) = f(a + t(b-a)) - f(a) - \sum_{i=1}^{k} \frac{t^i}{i!} D^i(a)(b-a,\dots,b-a)\,.$$

Then $g(0) = 0$ and

$$\begin{aligned} g(1) &= f(b) - f(a) - \sum_{i=1}^{k} \frac{1}{i!} D^i(a)(b-a,\dots,b-a) \\ &= \int_0^1 \int_0^{t_1} \dots \int_0^{t_{k-1}} \big(D^k f(a + s(b-a))(b-a,\dots,b-a) \\ &\qquad - D^k f(a)(b-a,\dots,b-a)\big)\, ds \\ &= \frac{1}{(k-1)!} \int_0^1 (1-s)^{k-1} \big(D^k f(a+s(b-a))(b-a,\dots,b-a) \\ &\qquad - D^k f(a))(b-a,\dots,b-a)\big)\, ds \\ &= \frac{1}{(k-1)!} \int_0^1 (1-s)^{k-1} \int_0^s D^{k+1} f(a + \tau(b-a))(b-a,\dots,b-a)\, d\tau\, ds \\ &= \frac{1}{k!} \int_0^1 (1-s)^k D^{k+1} f(a + s(b-a))(b-a,\dots,b-a)\, ds\,. \end{aligned}$$

To obtain the next to last expression we have used that $D^k f(a+s(b-a))(b-a,\dots,b-a)$ is absolutely continuous (since Lipschitz) in s. To obtain the last expression we have applied the Fubini Theorem. $\square$

The second integral in the representation (4.1) can be used to obtain the inclusion

$$\begin{aligned} &\frac{1}{k!} \int_0^1 (1-s)^k D^{k+1} f(a + s(b-a))(b-a,\dots,b-a)\, ds \\ &\qquad \in \frac{1}{(k+1)!} \partial^{k+1} f(c)(b-a,\dots,b-a) \end{aligned} \tag{4.2}$$

for some $c \in [a,\, b]$, which shows, roughly speaking, that Theorem 4.3 is a refinement of Theorem 3.1. Actually this result shows that in general the remainder term in Theorem 4.3 gives a more precise estimation, with the only exception that in (4.2), and hence in the representation of the remainder in Theorem 4.3, the point c possibly takes as values a or b, the extremes points of $[a,\, b]$, while in Theorem 4.3 this is excluded. In the case $f \in C^{k+1}$ the inclusion (4.2) turns into equality and the result is a simple application of the Integral Mean Value Theorem. As Theorem 4.4 below shows, the integral representation of the rest from Theorem 4.3 is more convenient for estimations than the form involving an intermediate point and generalized gradients from Theorem 3.1. We skip the proof of (4.2) in general, since on the one hand this result won't be used in the sequel, and on the other hand we don't wish to enter more deeply into the theory of Clarke generalized gradients. The first integral in the representation (4.1) is used to obtain the estimation in Theorem 4.4 below. For x, $w \in \mathbb{R}^n$ and $t \in \mathbb{R}$ (we are interested further in the case $t > 0$) we introduce the notation:

$$\Delta^{k+1} f(x_0, t, w) = \frac{(k+1)!}{t^{k+1}} \left(f(x+tw) - f(x) - \sum_{i=1}^{k} \frac{t^i}{i!} D^i f(x)(w, \ldots, w) \right).$$

In the sequel we also use the notation $S := \{u \in \mathbb{R}^n \mid \|u\| = 1\}$. For $x_0 \in \mathbb{R}^n$ and $r >$ we put also $B(x_0, r) = \{x \in \mathbb{R}^n \mid \|x - x_0\| < r\}$.

Theorem 4.4. *Let f be $C^{k,1}$ on $B(x_0, r)$, where $x_0 \in \mathbb{R}^n$, $r > 0$, with $D^k f(x_0)$ being Lipschitz with constant λ. Then for all u, $v \in S$, and all t with $0 < t < r$,*

$$|\Delta^{k+1} f(x_0, t, u) - \Delta^{k+1} f(x_0, t, v)| \leq (k+1)\, \lambda\, \|u - v\|. \tag{4.3}$$

For $v = 0$ we get

$$|\Delta^{k+1} f(x_0, t, u)| \leq (k+1)\, \lambda\, \|u\|. \tag{4.4}$$

PROOF. The estimation (4.3) follows from the following chain:

$$|\Delta^{k+1} f(x_0, t, u) - \Delta^{k+1} f(x_0, t, v)|$$

$$= \frac{(k+1)!}{t^{k+1}} \cdot \frac{1}{(k-1)!} \cdot \Big| \int_0^1 (1-s)^{k-1} \big(D^k f(x_0 + stu) - D^k f(x_0)\big)(tu, \ldots, tu)\, ds$$

$$- \int_0^1 (1-s)^{k-1} \big(D^k f(x_0 + stv) - D^k f(x_0)\big)(tv, \ldots, tv)\, ds \Big|$$

$$= \frac{(k+1)k}{t} \cdot \Big| \int_0^1 (1-s)^{k-1} \Big(\big((D^k f(x_0 + stu) - D^k f(x_0))(u, \ldots, u)\big)$$

$$-(D^k f(x_0 + stv) - D^k f(x_0))(u, \ldots, u)\big)$$

$$+(D^k f(x_0 + stv) - D^k f(x_0))(u - v, u, \ldots, u)$$

$$+(D^k f(x_0 + stv) - D^k f(x_0))(v, u - v, \ldots, u) + \ \ldots$$

$$+(D^k f(x_0+stv)-D^k f(x_0))(v,v,\dots,u-v)\Big)\,ds\,\Big|$$

$$\le (k+1)k\,\lambda \int_0^1 (1-s)^{k-1}(s+\dots+s)\,ds\ \|u-v\|$$

$$=(k+1)^2 k\,\lambda \int_0^1 (1-s)^{k-1}s\,ds\ \|u-v\| = (k+1)\,\lambda\ \|u-v\|\,.$$

For the last equality we applied the integral $\int_0^1 (1-s)^{k-1}s\,ds = \frac{1}{k(k+1)}$. Now (4.4) follows immediately from $\Delta^{k+1}f(x_0,t,0)=0$ and (4.3). □

4.3 Unconstrained Problems.

In this section we deal with the problem of finding the local minimum of a $C^{k,1}$ function $f:\mathbb{R}^n\to\mathbb{R}$. (All results are also valid for functions f defined on an open subset of $\mathbb{R}^n$.) We denote this problem by

$$\min\, f(x)\,. \tag{4.5}$$

Our goal is to discuss optimality conditions for this problem. In the sequel we make use of the $(k+1)$-th order lower Dini directional derivative of f, defined at the point $x\in\mathbb{R}^n$ in direction $u\in\mathbb{R}^n$ by

$$f_-^{(k+1)}(x,u)=\liminf_{t\to 0^+}\Delta^{k+1}f(x,t,u)\,. \tag{4.6}$$

A suitable notation for this derivative could be also $D_-^{k+1}f(x)(u,\dots,u)$; we avoid it however because $D_-^{k+1}f(x)(u,\dots,u)$ has been used earlier in another sense. Moreover, to unify the notation we also put

$$f_-^{(i)}(x,u):=D^i f(x)(u,\dots,u),\quad i=1,\dots,k\,.$$

Note that a formula similar to (4.6) is valid (with $k+1$ substituted by i and $\Delta^i f(x,t,u)$ defined similarly to $\Delta^{k+1}f(x,t,u)$).

The following conditions are important in the forthcoming discussion.

$\mathbb{N}_-^{(i)}(x,u)$: $\quad f_-^{(j)}(x,u)=0$, $j=1,\dots,i-1$, implies $f_-^{(i)}(x,u)\ge 0$.
$\mathbb{S}_-^{(i)}(x,u)$: $\quad f_-^{(j)}(x,u)=0$, $j=1,\dots,i-1$, implies $f_-^{(i)}(x,u)>0$.

Clearly, condition $\mathbb{N}_-^{(1)}(x,u)$ is simply $f_-^{(1)}(x,u)\ge 0$, and similarly condition $\mathbb{S}_-^{(1)}(x,u)$ is $f_-^{(1)}(x,u)>0$.

Theorem 4.5 (Necessary conditions, [26])**.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a $C^{k,1}$ function near x_0 for which x_0 is a local minimum. Then for all $u \in S$ conditions $\mathbb{N}_-^{(i)}(x_0, u)$, $i = 1, \ldots, k+1$, hold.*

The strict minima are in some sense more restrictive than the usual ones. We will say that x_0 is a strict minimum of order $k+1$ of f (or of problem (4.5)) if there is $r > 0$ and $a > 0$ such that

$$f(x) > f(x_0) + a\,\|x - x_0\|^{k+1} \quad \text{for all} \quad x \in B(x_0, r)\,. \tag{4.7}$$

Obviously, each strict minimum of order $k+1$ is a strict minimum, and moreover a minimum of f.

Theorem 4.6 (Sufficient conditions, Strict minima)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a $C^{k,1}$ function near x_0. Let $x_0 \in \mathbb{R}^n$ be such that for each $u \in S$ there exist $\bar\imath = \bar\imath(u) \in \{1, \ldots, k+1\}$, $\bar\delta = \bar\delta(u) > 0$ and $\bar t = \bar t(u) > 0$ for which*

a) Condition $\mathbb{S}_-^{(\bar\imath)}(x_0, u)$ is satisfied,

b) $\sum_{i=1}^{\bar\imath-1} \frac{t^i}{i!} f_-^{(i)}(x_0, v) \geq 0$ for $0 < t < \bar t$ and $v \in S \cap B(u, \bar\delta)$.

Then x_0 is a strict minimum of order $k+1$.

Proof. Since $f \in C^{k,1}$ near x_0, we can choose $\bar r > 0$ such that f is $C^{k,1}$ on $B(x_0, \bar r)$, and all $D^i f$, $i = 1, \ldots, k$, are Lipschitz with a constant $\lambda > 0$ on $B(x_0, \bar r)$.

Fix $u \in S$. Condition $\mathbb{S}_-^{(\bar\imath)}(x_0, u)$ gives

$$f_-^{(\bar\imath)}(x_0, u) = \liminf_{t \to 0^+} \frac{\bar\imath\,!}{t^{\bar\imath}}\,(f(x_0 + tu) - f(x_0)) > 0\,,$$

whence there exists $\varepsilon = \varepsilon(u) > 0$ and $r_1 = r_1(u)$, $0 < r_1 < \min(\bar r, 1)$, such that

$$f(x_0 + tu) - f(x_0) > \frac{t^{\bar\imath}}{\bar\imath\,!}\,\varepsilon \quad \text{for} \quad 0 < t < r_1\,.$$

From Theorem 4.4, with $k+1$ replaced by $\bar\imath$, for $0 < t < \bar r$ and $v \in S \cap B(u, \bar\delta)$

$$\begin{aligned} f(x_0 + tv) - f(x_0) &\geq f(x_0 + tu) - f(x_0) \\ + \sum_{i=1}^{\bar\imath-1} \frac{t^i}{i!}\, f_-^{(i)}(x_0, v) - \frac{t^{\bar\imath}}{\bar\imath\,!}\,\lambda\,\|v - u\| &\geq \frac{t^{\bar\imath}}{\bar\imath\,!}\,(\varepsilon - \bar\imath\,\lambda\,\|v - u\|\,)\,. \end{aligned}$$

Choose $\delta_1 = \min(\bar{\delta}, \varepsilon/(2\bar{\imath}\lambda))$. Then for $v \in S \cap B(u, \delta_1)$ and $0 < t < r_1$ we have

$$f(x_0 + tv) - f(x_0) \geq \frac{\varepsilon}{2\bar{\imath}!}\, t^{\bar{\imath}} \geq \frac{\varepsilon}{2(k+1)!}\, t^{k+1}\,.$$

Thus, we have shown that for every $u \in S$ there exist $r = r(u)$, a neighborhood $U = U(u)$ of u, and $a = a(u) > 0$, such that

$$f(x_0 + tv) \geq f(x_0) + a\, t^{k+1} \quad \text{for} \quad 0 < t < r \quad \text{and} \quad v \in S \cap U\,.$$

Because of the compactness of S, this implies that x_0 is a strict minimum of order $k+1$. □

Theorem 4.2 is a corollary from Theorem 4.6. Indeed, when $D^i f(x_0) = 0$, $i = 1, \ldots, k$, and $D_-^{k+1} f(x_0)(u) > 0$ for all $u \in \mathbb{R}^n \setminus \{0\}$, then conditions *a)* and *b)* are satisfied. Condition *a)* follows from

$$f_-^{(i)}(x_0, u) = D^i f(x_0)(u, \ldots, u) = 0, \quad i = 1, \ldots, k,$$

and, with regard to Theorem 4.3,

$$f_-^{(k+1)}(x_0, u) = \liminf_{t \to 0^+} \frac{(k+1)!}{t^{k+1}}\, (f(x_0 + tu) - f(x_0))$$
$$= \liminf_{t \to 0^+}\, (k+1) \int_0^1 (1-s)^k D^{k+1} f(x_0 + stu)(u, \ldots, u)\, ds > 0\,.$$

Here we have used that $D^{k+1} f(x) > 0$ for x in some neighborhood of x_0, a consequence of $D_-^{k+1} f(x_0)(u) > 0$ and the upper semi continuity of $D^{k+1} f$. In fact, we have shown that for all $u \in S$ condition $\mathbb{S}_-^{(k+1)}(x_0, u)$ is satisfied. Condition *b)* follows immediately from $f_-^{(i)}(x_0, u) = D^i f(x_0)(u, \ldots, u) = 0$, $i = 1, \ldots, k$. Then $\sum_{i=1}^k \frac{t^i}{i!} f_-^{(i)}(x_0, v) \equiv 0$.

Theorem 4.6 admits a converse with assumptions similar to Theorem 4.2. This fact is shown in the next theorem and stresses that the strict minimum of order $k+1$ is the appropriate notion of a solution for problem (4.5).

Theorem 4.7. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be $C^{k,1}$ near $x_0 \in \mathbb{R}^n$ and let $D^i f(x_0) = 0$, $i = 1, \ldots, k$. If x_0 is a strict minimum of order $k+1$ then for all $u \in S$ condition $\mathbb{S}_-^{(k+1)}(x_0, u)$ is satisfied.*

Proof. By the hypotheses there exist $a > 0$ and $\delta > 0$ such that

$$f(x_0 + tu) \geq f(x_0) + a\, t^{k+1} \quad \text{for} \quad 0 < t < \delta \quad \text{and} \quad u \in S\,,$$

$$f_-^{(i)}(x_0,u) = D^i f(x_0)(u,\dots,u) = 0, \quad i = 1,\dots,k\,.$$

Now

$$\begin{aligned}\Delta^{k+1} f(x_0,t,u) &= \frac{(k+1)!}{t^{k+1}}\left(f(x_0+tu) - f(x_0)\right)\\ &\geq \frac{(k+1)!}{t^{k+1}}\, a\, t^{k+1} = (k+1)!\, a\,,\end{aligned}$$

and

$$f_-^{(k+1)}(x_0,u) = \liminf_{t\to 0^+} \Delta^{k+1} f(x_0,t,u) \geq (k+1)!\, a > 0\,.$$

Therefore condition $\mathbb{S}_-^{(k+1)}(x_0,u)$ holds. □

The next two theorems are corollaries of Theorems 4.6 and 4.7 in the cases $k=0$ and $k=1$ respectively.

Theorem 4.8. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a $C^{0,1}$ function near $x_0 \in \mathbb{R}^n$. Let $f_-^{(1)}(x_0,u) > 0$ for all $u \in S$. Then x_0 is a strict minimum of order 1. Conversely, if x_0 is a strict minimum of order 1, then $f_-^{(1)}(x_0,u) > 0$ for all $u \in S$.*

Proof. One need observe that condition $\mathbb{S}_-^{(1)}(x_0,u)$ is $f_-^{(1)}(x_0,u) > 0$, and for $k=0$ condition $b)$ in Theorem 4.6 is missing. □

Theorem 4.9 ([30])**.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a $C^{1,1}$ function near $x_0 \in \mathbb{R}^n$. Let $D^1 f(x_0) = 0$ and $f_-^{(2)}(x_0,u) > 0$ for all $u \in S$. Then x_0 is a strict minimum of order 2. Conversely, if x_0 is a strict minimum of order 2, then $D^1 f(x_0) = 0$ and $f_-^{(2)}(x_0,u) > 0$ for all $u \in S$.*

Proof. Let $D^1 f(x_0) = 0$ and $f_-^{(2)}(x_0,u) > 0$ for all $u \in S$. This means that $\mathbb{S}_-^{(2)}(x_0,u)$ holds for all $u \in S$, which verifies condition $a)$ of Theorem 4.6. Condition $b)$ follows trivially from $t f_-^{(1)}(x_0,v) = t D^1 f(x_0)(v) = 0$. Conversely, if x_0 is a strict minimum of order 2, and hence a minimum, then necessarily $D^1 f(x_0) = 0$. The hypotheses of Theorem 4.7 are satisfied. Hence $\mathbb{S}_-^{(2)}(x_0,u)$ holds, which gives $f_-^{(2)}(x_0,u) > 0$ for all $u \in S$. □

4.4 Constrained Problems.

In this section we deal with the scalar constrained problem

$$\begin{aligned} \min\, f(x_1,\dots,x_n)\,, \\ g_i(x_1,\dots,x_n) \le 0, \quad i=1,\dots,p, \\ h_j(x_1,\dots,x_n) = 0, \quad j=1,\dots,q. \end{aligned} \tag{4.8}$$

under the assumption that all functions are $C^{k,1}$. (We say that the problem is with $C^{k,1}$ data.) This problem is the scalar counterpart of the vector optimization problem considered in [32] and [34] with the assumptions that the problem data are respectively $C^{0,1}$ and $C^{1,1}$ functions. We will adapt these results for the scalar problem considered here. Our approach is based on the Dini set-valued derivatives. Let $\Phi : \mathbb{R}^n \to \mathbb{R}^m$ be a $C^{k,1}$ function and $x^0 \in \mathbb{R}^n$. Then the $(k+1)$-th order Dini set-valued derivative of Φ at x in direction $u \in \mathbb{R}^n$ is defined by

$$\Phi_u^{(k+1)}(x^0) = \operatorname*{Limsup}_{t\to 0^+} \Delta^{k+1}\Phi(x,t,u)\,.$$

Here $\operatorname{Limsup}_{t\to+0}$ stands for the Painlevé-Kuratowski limit. In other words we have $y \in \Phi_u^{(k+1)}(x^0)$ if there exists a sequence $t_\nu \to 0^+$ such that $y = \lim_{\nu\to\infty} \Delta^{k+1}\Phi(x,t_\nu,u)$. Instead of $\Phi_u^{(1)}(x^0)$ and $\Phi_u^{(2)}(x^0)$ we will write $\Phi'_u(x^0)$ and $\Phi''_u(x^0)$ respectively. With problem (4.4) we relate the vector functions $g : \mathbb{R}^n \to \mathbb{R}^p$ whose components are the functions g_i, $i = 1,\dots,p$, and $h : \mathbb{R}^n \to \mathbb{R}^q$ whose components are the functions h_j, $j = 1,\dots,q$. We will suppose that $q \le n$ and will assume the hypotheses of the implicit function theorem that guarantee the equation $h(x) = 0$ to be solved with respect to q of the components of $x = (x_1,\dots,x_n)$. Substituting them in the objective function and in the inequality constraints, we transform the problem into one with only inequality constraints to which we can apply the optimality conditions from [28] (in the case $k = 0$) or [29] (in the case $k = 1$). In such a way we obtain Theorems 4.10 and 4.10 stated below.

In the sequel we will use the notation. $\mathbb{R}_+ = [0,+\infty)$, $\mathbb{R}_+^p = (\mathbb{R}_+)^p$. For $y \in \mathbb{R}_+$ let $\mathbb{R}_+[0] = \mathbb{R}_+$ and $\mathbb{R}_+[y] = \mathbb{R}$ if $y > 0$. For $y = (y_1,\dots,y_p) \in \mathbb{R}_+^p$ we put $\mathbb{R}_+^p[y] = \mathbb{R}_+[y_1] \times \cdots \times \mathbb{R}_+[y_p]$. Further, for $y \in \mathbb{R}_+$ we put $\mathbb{R}_+^*[0] = \mathbb{R}_+$, and $\mathbb{R}_+^*[y] = \{0\}$ if $y > 0$. Similarly, for $y = (y_1,\dots,y_p) \in \mathbb{R}_+^p$ let $\mathbb{R}_+^{p\,*}[y] = \mathbb{R}_+^*[y_1] \times \cdots \times \mathbb{R}_+^*[y_p]$. The scalar product of the vectors $z^i = (z_1^i,\dots,z_p^i) \in \mathbb{R}^p$, $i = 1,2$, is $\langle z^1, z^2\rangle = \sum_{i=1}^p z_i^1 z_i^2$. We apply upper indices for the points, as e. g. z^1 and z^2 here, or x^0 for the reference point, reserving the lower indices for their components. Recall that x^0 is a minimum point for problem (4.4) if there is $r > 0$ such that $f(x) \ge f(x^0)$ for all feasible $x \in B(x^0, r)$ (x is said to be feasible if it satisfies the constraints $g(x) \le 0$ and $h(x) = 0$). Similarly, we call x^0 a strict minimum point of order k if there is $r > 0$ and $a > 0$ such that inequality (4.7) for all feasible points $x \in B(x^0, r)$. For brevity we say *minimum* instead of *minimum point*, and *strict minimum* instead of *strict minimum point*.

We will consider the cases $k = 0$ and $k = 1$ separately.

Case $k = 0$**.**

Now the assumption is that f, g, h are $C^{0,1}$ functions. The goal is to establish first-order optimality conditions the feasible point $x^0 = (x_1^0, \dots, x_n^0)$ to be a minimum or strict minimum. Further we will suppose that the variable feasible point $x = (x_1, \dots, x_n)$ admits a representation $x = (x_a, x_b)$ where x_a unifies some $n - q$ components of x and x_b the remaining q components of x, such that

$\mathbb{A}^0(x^0)$: All the matrices in $\pi_b \, \partial h(x^0)$ are invertible.

Theorem 4.10 ([32])**.** *Consider problem (4.4) with f, g, h being $C^{0,1}$ functions near $x^0 \in \mathbb{R}^n$. Assume that x^0 is a feasible point satisfying condition $\mathbb{A}^0(x^0)$.*

(Necessary Conditions) *Let x^0 be a minimum of problem (4.4). Then for each $u \in S$*

$$\mathbb{N}_p^{0,1}(x^0, u) : (f,\, g,\, h)'_u(x^0) \cap \big(- \big(\operatorname{int} \mathbb{R}_+ \times \operatorname{int} \mathbb{R}_+^p[-g(x^0)] \times \{0\} \big) \big) = \emptyset\,.$$

(Sufficient Conditions) *Suppose that for each $u \in S$ the following condition is satisfied:*

$$\mathbb{S}_p^{0,1}(x^0, u) : (f,\, g,\, h)'_u(x^0) \cap \big(- \big(\mathbb{R}_+ \times \mathbb{R}_+^p[-g(x^0)] \times \{0\} \big) \big) = \emptyset\,.$$

Then x^0 is a strict minimum of order 1 of problem (4.4).

The primal form conditions $\mathbb{N}_p^{0,1}(x^0, u)$ and $\mathbb{S}_p^{0,1}(x^0, u)$ admit an equivalent dual form representation, respectively.

$$\mathbb{N}_d^{0,1}(x_0, u) : \left\{ \begin{array}{c} \forall (y^0,\, z^0,\, 0) \in (f,\, g,\, h)'_u(x^0) : \exists (\xi^0,\, \eta^0) \in \mathbb{R}_+ \times \mathbb{R}_+^{p\,*}[-g(x^0)] : \\ (\xi^0,\, \eta^0) \ne (0,\, 0) \quad \text{and} \quad \langle \xi^0,\, y^0 \rangle + \langle \eta^0,\, z^0 \rangle \ge 0\,. \end{array} \right.$$

$$\mathbb{S}_d^{0,1}(x^0, u) : \left\{ \begin{array}{c} \forall (y^0,\, z^0,\, 0) \in (f,\, g,\, h)'_u(x^0) : \exists (\xi^0,\, \eta^0) \in \mathbb{R}_+ \times \mathbb{R}_+^{p\,*}[-g(x^0)] : \\ (\xi^0,\, \eta^0) \ne (0,\, 0) \quad \text{and} \quad \langle \xi^0,\, y^0 \rangle + \langle \eta^0,\, z^0 \rangle > 0\,. \end{array} \right.$$

In this form one can recognize ξ^0, η^0 as Lagrange multipliers. The relation $\eta^0 = (\eta_1^0, \dots, \eta_p^0) \in \mathbb{R}_+^{p\,*}[-g(x^0)]$ is equivalent to $\eta_i^0\, g_i(x^0) = 0$, $i = 1, \dots, p$, which in the classical Karush-Kuhn-Tucker (KKT) theory is known as the complementary slackness condition. In contrast to the classical KKT theory the multipliers here depend on the direction, which gives some flexibility and generality. We emphasize, that optimality conditions in nonsmooth optimization are often designed on the basis of the smooth counterpart, the classical KKT conditions. Theorem 6.1.1 in [13, page 228] establishes conditions for problems with $C^{0,1}$ data of this type based on the Clarke generalized gradient. Let us stress that this theorem, as the usual results involving KKT conditions, is limited only to necessary optimality conditions, while Theorem 4.10 presents sufficient ones. In the case of necessary conditions, it is shown in [32] that Theorem 4.10 works in cases when [13] fails. The disadvantage of conditions based on Dini derivatives is that in general the Dini derivative does not admit convenient calculus rules, while the Clarke generalized gradient does. However, they are still effective when the functions considered present some regularity as it is shown in [33].

Case $k = 1$**.**

Now the assumption is that f, g, h are $C^{1,1}$ functions. The goal is to establish second-order optimality conditions for the feasible point $x^0 = (x_1^0, \dots, x_n^0)$ to be a minimum or strict minimum. For simplicity, the Jacobians of f, g and h at a point x will be denoted by $f'(x)$, $g'(x)$ and $h'(x)$ respectively. We will suppose that at the reference feasible point x^0

$\mathbb{A}^1(x^0)$: The Jacobian $h'(x^0)$ is of full range q.

Note that condition $\mathbb{A}^1(x^0)$ is only a more natural formulation of condition $\mathbb{A}^0(x^0)$ with regard to the present now case of a differentiable function h.

Theorem 4.11 ([34])**.** *Consider problem (4.4) with f, g, h being $C^{1,1}$ functions. Let x^0 be a feasible point and condition $\mathbb{A}^1(x^0)$ be satisfied.*

(Necessary Conditions) *Let x^0 be a minimum of problem (4.4). Then for all $u \in S$ both conditions $\mathbb{N}_p^{0,1}(x^0, u)$ and $\mathbb{N}_d^{1,1}(x^0, u)$ given below are satisfied.*

$$\mathbb{N}_d^{1,1}(x^0,u) : \left\{ \begin{array}{c} \textit{if } (f'(x^0)u,\, g'(x^0)u,\, h'(x^0)u) \\ \in -(\mathbb{R}_+ \times \mathbb{R}_+^p[-g(x^0)] \setminus \operatorname{int}\mathbb{R}_+ \times \operatorname{int}\mathbb{R}_+^p[-g(x^0)]) \times \{0\} \\ \textit{then } \forall\, (y^0, z^0, w^0) \in (f,\, g,\, h)''_u(x^0) : \exists\, (\xi^0, \eta^0, \zeta^0) : \\ (\xi^0, \eta^0) \in \mathbb{R}_+ \times \mathbb{R}_+^{p\,*}[-g(x^0)] \setminus \{(0,0)\}, \\ \langle \xi^0, f'(x^0)\, u\rangle + \langle \eta^0, g'(x^0)\, u\rangle = 0,\ \zeta^0 \in \mathbb{R}^q \textit{ satisfies (4.9)}, \\ \textit{and } \langle \xi^0,\, y^0\rangle + \langle \eta^0,\, z^0\rangle + \langle \zeta^0,\, w^0\rangle \ge 0\,. \end{array} \right.$$

(Sufficient Conditions) *Suppose that for each $u \in S$ one of the conditions $\mathbb{S}_p^{0,1}(x^0,u)$ and $\mathbb{S}_d^{1,1}(x^0,u)$ given below is satisfied.*

$$\mathbb{S}_d^{1,1}(x^0,u): \left\{ \begin{array}{c} (f'(x^0)u,\, g'(x^0)u,\, h'(x^0)u) \\ \in -(\mathbb{R}_+ \times \mathbb{R}_+^p[-g(x^0)] \setminus \operatorname{int}\mathbb{R}_+ \times \operatorname{int}\mathbb{R}_+^p[-g(x^0)]) \times \{0\} \\ \text{and } \forall\, (y^0, z^0, w^0) \in (f,\, g,\, h)''_u\,(x^0) : \exists\, (\xi^0, \eta^0, \zeta^0) : \\ (\xi^0, \eta^0) \in \mathbb{R}_+ \times \mathbb{R}_+^{p\,*}[-g(x^0)] \setminus \{(0,\,0)\}, \\ \langle \xi^0, f'(x^0)\,u\rangle + \langle \eta^0, g'(x^0)\,u\rangle = 0,\ \zeta^0 \in \mathbb{R}^q \text{ satisfies (4.9)}, \\ \text{and } \langle \xi^0,\, y^0\rangle + \langle \eta^0,\, z^0\rangle + \langle \zeta^0,\, w^0\rangle > 0\,. \end{array} \right.$$

Then x^0 is a strict minimum of order 2 of problem (4.4).

Both $\mathbb{N}_d^{1,1}(x^0,u)$ and $\mathbb{S}_d^{1,1}(x^0,u)$ refer to the following condition:

$$\begin{aligned} \langle \zeta^0,\, w^0\rangle = &-\langle \xi^0,\, f_{x_b}(x_a^0,x_b^0)\left(h_{x_b}^{-1}(x_a^0,x_b^0)\,w^0\right)\rangle \\ &-\langle \eta^0,\, g_{x_b}(x_a^0,x_b^0)\left(h_{x_b}^{-1}(x_a^0,x_b^0)\,w^0\right)\rangle\,. \end{aligned} \tag{4.9}$$

Here $x = (x_a, x_b)$ is a representation of x such that $\pi_{x_b}\, h'(x)$ is of full rank q. Let us mention that the existence of ζ^0 is guaranteed by the Riesz Theorem (saying that each linear functional $\ell(w^0)$ admits a representation $\langle \zeta^0, w^0\rangle = \ell(w^0)$ for some ζ^0).

With regard to the differentiability of the data, conditions $\mathbb{N}_p^{0,1}(x^0,u)$ and $\mathbb{S}_p^{0,1}(x^0,u)$ used in Theorem 4.11 admit a more convenient representation, noted below as $\mathbb{N}'$ and $\mathbb{S}'$ respectively.

$$\mathbb{N}' : (f'(x^0)u,\, g'(x^0)u,\, h'(x^0)u) \notin -(\operatorname{int}\mathbb{R}_+ \times \operatorname{int}\mathbb{R}_+^p[-g(x^0)]) \times \{0\}\,,$$

$$\mathbb{S}' : (f'(x^0)u,\, g'(x^0)u,\, h'(x^0)u) \notin -(\mathbb{R}_+ \times \mathbb{R}_+^p[-g(x^0)]) \times \{0\}\,.$$

In conclusion we remark that Theorems 4.10 and 4.11 establish first and second-order optimality conditions for the constrained problem (4.4) with $C^{0,1}$ and $C^{1,1}$ data respectively. The proof is based on the implicit function theorem for the equation $h(x_a, x_b) = 0$, excluding the equality constraints and the variables x_b, and reducing the problem in this way to the one considered in [28] or [29] respectively. In Section 4.5 the Taylor's formula of order 1 and 2 for the implicit function is presented with the remainder represented with a mean point and generalized gradients. Here for the exclusion process we needed rather the representation like in Theorem 4.3 with an integral form for the remainder. A similar problem for $k \geq 3$ is an open problem. Also an open problem, but of less importance, is establishing optimality conditions of order $k \geq 3$ for the constrained problem (4.4) with $C^{k,1}$ data.

4.5 Numerical Methods.

Several efficient numerical methods have been developed and implemented to solve nonsmooth optimization problems with $C^{1,1}$ functions (sometimes called LC^1 functions). Among them, the milestone is the one due to Qi and Sun (see [73]) which represents an extension of the classical Newton method for minimization of C^2 functions. See also [12], [18] and [85] for additional information. This method works under the assumption that the derivatives of the objective functions are semismooth. This is not an unusual hypothesis since, for instance, the derivative of the objective function of an extended linear-quadratic programming problem in the fully quadratic case is semismooth and the derivative of the augmented Lagrangian of a twice smooth nonlinear programming problem is also semismooth. The main difference between the classical Newton method and the one for $C^{1,1}$ functions is based on the choice of the second order differential. Since twice differentiability is not assumed, the Newton method for nonsmooth function at each step picks up one element belonging to the generalized Hessian in the sense of Clarke (see [13]). As already announced at the beginning of this section, convergence is proved under the extra hypothesis of semismoothness of the involved data. The generalized Newton method for $C^{1,1}$ problems reads as

$$x_{k+1} = x_k - V_k^{-1} \nabla f(x_k) \tag{4.10}$$

where $V_k \in \partial^2 f(x_k)$ is the generalized Hessian. The definition of semismoothness is the following.

Definition 4.12. *[73] ∇f is said to be semismooth at x if ∇f is locally Lipschitzian at x and*

$$\lim_{V \in \partial^2 f(x+th'), h' \to h, t \to 0} Vh' \tag{4.11}$$

exists for any h.

The local convergence of Newton's method for nonsmooth problems is stated in the following result.

Theorem 4.13. *[84] Suppose that f is a LC^1 function and ∇f is semismooth at x^*, x_k is sufficiently close to x^*, where x^* is a local minimum, $V \in \partial^2 f(x^*)$ is positive definite. Then the generalized Newton iteration 4.10 is well defined, and converges to x^* at a superlinear rate.*

There are results regarding the global convergence of this method and the analysis of the speed of convergence (see [74, 75, 84]).

5 Conclusions And Final Remarks.

We have presented a short survey on the theory of $C^{k,1}$ functions mainly concentrated on characterizations, generalized Taylor's formula and optimization of objects belonging to this class of nonsmooth functions. Before concluding this brief survey, we think it is worthwhile to give a quick look at how this class of functions has been used in other branches of nonsmooth calculus.

In fact, in the literature the class of $C^{k,1}$ functions has also been considered in several papers dealing with differential equations with initial conditions or boundary value problems. For instance, H.Shahgholian (see [79] and the references therein) proved $C^{1,1}$ regularity in semilinear elliptic problems, Y.Luo and A.Eberhad (see [65]) used $C^{1,1}$ approximation to prove comparison principles for viscosity solutions of curvature equations, and finally, M.Salo ([77]) considered the stable dependence of solutions to wave equations on metrics in the $C^{1,1}$ class. Another interesting application of $C^{1,1}$ functions for a new parametrix construction for the wave equation with variable coefficients is given in [80]. The same author shows the importance of regularity $C^{1,1}$ in considering L^2-L^q estimates for the spectral projections of an elliptic differential operator on a compact manifold ([81]). Other applications can be found in [9, 41, 47, 78, 82, 83, 90]. Without pretending to be exhaustive but just to give an idea of such applications, let us consider the main result obtained by H.Shahgholian in the above mentioned paper. He presents a simple proof of $C^{1,1}$ regularity of the solution u to the stationary reaction-diffusion equations

$$\nabla u = f(x, y). \tag{5.1}$$

Let B_1 the unit ball in $\mathbb{R}^n$, $W^{2,p}$ the classical Sobolev space and suppose that f be a function which satisfies

$$|f(x,t)| \leq M,\ |f(x,t) - f(y,t)| \leq M|x-y|,\ f'(x,t) \geq -M(weakly) \tag{5.2}$$

for all $x \in B_1$. A function $u \in W^{2,p}(B_1)$, $p > n$, is said to belong to the class $P := P(M, n)$ if u satisfies

- $\nabla u = f(x, y)$
- $\|u\|_{W^{2,p}} \leq M$

The main result in this paper states there exists a universal constant $C = C(M, n)$ such that for all $u \in P(M, n)$

$$|\nabla u(x) - \nabla u(y)| \leq C|x - y| \tag{5.3}$$

for all $x, y \in B_{1/2}$ or, in other words, that u is locally a $C^{1,1}$ function.

Acknowledgment. The authors wish to thank the referees for their constructive critique of the first draft and Professor Paul Humke for his support in this project.

References

[1] J. M. Ash, *Generalizations of the Riemann derivative*, Trans. Amer. Math. Soc., **126** (1967), 181–199.

[2] J. M. Ash, *Very generalized Riemann derivatives, generalized Riemann derivatives and associated summability methods*, Real Anal. Exchange, **11(1)** (1985), 10–29.

[3] D. Bednařík and K. Pastor, *Elimination of strict convergence in optimization*, SIAM J. Control Optim., **43** (2004), 1063–1077.

[4] D. Bednařík and K. Pastor, *On second-order conditions in unconstrained optimization*, Math. Program., Ser. A, **113(2)** (2008), 283–298.

[5] D. Bednařík and K. Pastor, *Differentiability properties of functions that are ℓ-stable at a point*, Nonlinear Anal., **69** (2008), 3128–3135.

[6] D. Bednařík and K. Pastor, *ℓ-stable functions are continuous*, Nonlinear Anal., **70** (2009), 2317–2324.

[7] D. Bednařík and K. Pastor, *Fréchet approach in second-order optimization*, Appl. Math. Letters, **22** (2009), 960–967.

[8] A. Ben-Tal and J. Zowe, *Necessary and sufficient optimality conditions for a class of nonsmooth minimization problems*, Math. Program., **24** (1982), 70–91.

[9] P. Bernard, *Existence of $C^{1,1}$ critical sub-solutions of the Hamilton-Jacobi equation on compact manifolds*, Ann. Sci. École Norm. Sup. (4), **40(3)** (2007), 445–452.

[10] P. S. Bullen, *A criterion for n-convexity*, Pac. J. Math., **36** (1971), 81–98.

[11] Z. Buczolich, M.J. Evans and P.D. Humke, *Approximate high order smooothness*, Acta Math. Hung., **61** (1993), 369–388.

[12] X. Chen, *Convergence of BFGS method for LC^1 convex constrained optimization*, SIAM J. Control Optim., **34** (1996), 2051–2063.

[13] F. H. Clarke, *Optimization and nonsmooth analysis*, Canadian Mathematical Society Series of Monographs and Advanced Texts, Wiley-Interscience Publications, John Wiley & Sons, Inc., New York, 1983.

[14] R. Cominetti and R. Correa, *A generalized second-order derivative in nonsmooth optimization*, SIAM J. Control Optim., **28** (1990), 789–809.

[15] C. Cusano, M. Fini and D. La Torre, *Characterization of convex vector functions and optimization*, J. Inequal. Pure Appl. Math., **5(4)** (2004), Article 101.

[16] J. P. Crouzeix, *On second-order conditions for quasiconvexity*, Math. Program., Ser. A, **18** (1980), 349–352.

[17] Q. M. Dong, *On second-order sufficient conditions for $C^{1,1}$ function minimization and a further discussion on the Hiriart-Urruty problem*, (Chinese, English summary), J. Systems Sci. Math. Sci. **8 (1)** (1988), 46–51.

[18] N. I. Duranovic-Milicic, *On an algorithm in $C^{1,1}$ optimization*, Filomat, **21(1)** (2007), 17–24.

[19] T. K. Dutta, *Generalised smooth functions*, Acta Math. Acad. Sci. Hungar., **40** (1982), 29–37.

[20] T. K. Dutta and N. Mukhopadhyay, *Generalised smooth functions II*, Acta. Math. Hung., **55** (1990), 47–56.

[21] A. Eberhard and R. Wenczel, *Some sufficient optimality conditions in nonsmooth analysis*, SIAM J. Optim., **20(1)** (2009), 251–296.

[22] M. J. Evans and C.E. Weil, *Peano derivatives: a survey*, Real Anal. Exchange, **7** (1981-82), 5–23.

[23] M. J. Evans, *High order smoothness*, Acta Math. Hung., **50** (1987), 17–20.

[24] P. G. Georgiev and N. Zlateva, *Second-order subdifferentials of $C^{1,1}$ functions and optimality conditions*, Set-Valued Anal., **4(2)** (1996), 101–117.

[25] P. G. Georgiev, *Second-order Clarke subdifferential of $C^{1,1}$ functions*, Nonlinear analysis and convex analysis, Yokohama Publ., Yokohama, 2007, 105–110.

[26] I. Ginchev, *Higher order optimality conditions in nonsmooth optimization*, Optimization, **51** (2002), 47–72.

[27] I. Ginchev, A. Guerraggio and M. Rocca, *Equivalence of (n+1)-th order Peano and usual derivatives for n-convex functions*, Real Anal. Exchange, **25** (2000), 145–156.

[28] I. Ginchev, A. Guerraggio and M. Rocca, *First-order conditions for $C^{0,1}$ constrained vector optimization*, Variational analysis and applications, 427–450, Nonconvex Optim. Appl., **79**, Springer, New York, 2005.

[29] I. Ginchev, A. Guerraggio and M. Rocca, *Second-order conditions in $C^{1,1}$ constrained vector optimization*, Math. Program., Ser. B, **104** (2005), 389–405.

[30] I. Ginchev, A. Guerraggio and M. Rocca, *From scalar to vector optimization*, Appl. Math., **51** (2006), 5–36.

[31] I. Ginchev, A. Guerraggio and M. Rocca, *Second-order conditions in $C^{1,1}$ vector optimization with inequality and equality constraints*, Recent Advances in Optimization, 29–44, Lecture Notes in Econom. and Math. Systems, **563**, Springer, Berlin, 2006.

[32] I. Ginchev, A. Guerraggio and M. Rocca, *Dini set-valued directional derivative in locally Lipschitz vector optimization*, J. Optim. Theory Appl., **143** (2009), 87–105.

[33] I. Ginchev, A. Guerraggio and M. Rocca, *Locally Lipschitz vector optimization with inequality and equality constraints*, Appl. Math., **55** (2010), 77–88.

[34] I. Ginchev, A. Guerraggio and M. Rocca, *Second-order Dini set-valued directional derivative in $C^{1,1}$ vector optimization*, Optim. Methods Softw., **25** (2010), 75–87.

[35] I. Ginchev and V.I. Ivanov, *Higher-order pseudoconvex functions*, Generalized convexity and related topics, 247–264, Lecture Notes in Econom. and Math. Systems, **583**, Springer, Berlin, 2007.

[36] A. Guerraggio and D. T. Luc, *Optimality conditions for $C^{1,1}$ vector optimization problems*, J. Optim. Theory Appl., **109 (3)** (2001), 615–629.

[37] A. Guerraggio and D. T. Luc, *Optimality conditions for $C^{1,1}$ constrained multiobjective problems*, J. Optim. Theory Appl., **116 (1)** (2003), 117–129.

[38] J. Heinonen, *Lectures on Lipschitz analysis*, Report, University of Jyväskylä Department of Mathematics and Statistics, **100**, University of Jyväskylä, Jyväskylä, 2005.

[39] J. B. Hiriart-Hurruty, *Contributions a la programmation mathematique: deterministe et stocastique*, Doctoral thesis, Univ. Clermont-Ferrand, 1977.

[40] J. B. Hiriart-Urruty, J. J. Strodiot and V. Hien Nguen, *Generalized Hessian matrix and second order optimality conditions for problems with $C^{1,1}$ data*, Appl. Math. Optim., **11** (1984), 169–180.

[41] G. Huisken, *$C^{1,1}$-regularity of solutions to variational inequalities*, Proc. Centre Math. Anal. Austral. Nat. Univ., Austral. Nat. Univ., Canberra, **5** (1984), 85–90.

[42] D. Klatte and R. Walzebok, *Sufficient optimality conditions for semi-infinite optimization problems as special $C^{1,1}$-problems*, (German summary), Wiss. Z. Pdagog. Hochsch. Halle/Kthen, **28(4)** (1990), 24–28.

[43] D. Klatte and K. Tammer, *On the second order sufficient conditions to perturbed $C^{1,1}$ optimization problems*, Optimization, **19** (1988), 169–180.

[44] D. Klatte, *Upper Lipschitz behavior of solutions to perturbed $C^{1,1}$ programs*, Math. Program., Ser. B, **88(2)** (2000), 285–311.

[45] B. Kummer, *An implicit-function theorem for $C^{0,1}$-equations and parametric $C^{1,1}$-optimization*, J. Math. Anal. Appl., **158(1)** (1991), 35–46.

[46] A. Ioffe and T. Milosz, *On a characterization of $C^{1,1}$ functions*, Kibernet. Sistem. Anal. **189** (3) (2002), 3–13. Translation in Cybernet. Systems Anal., **38(3)** (2002), 313–322.

[47] A. Ioffe and E. Schwartzman, *Parametric Morse lemmas for $C^{1,1}$-functions*, Recent developments in optimization theory and nonlinear analysis (Jerusalem, 1995), 139–147, Contemp. Math., **204**, Amer. Math. Soc., Providence, RI, 1997.

[48] D. La Torre, *Characterizations of $C^{k,1}$ functions, Taylor's formula and optimality conditions*, Doctoral thesis, University of Milan, 2001.

[49] D. La Torre, *On generalized derivatives for $C^{1,1}$ vector optimization problems*, J. Appl. Math., **7** (2003), 365–376.

[50] D. La Torre, *Necessary optimality conditions for nonsmooth vector optimization problems*, Math. Model. Analysis, **8(2)** (2003), 165–174.

[51] D. La Torre and M. Rocca, *$C^{k,1}$ functions and Riemann derivatives*, Real Anal. Exchange, **25(2)** (2000), 743–752.

[52] D. La Torre and M. Rocca, *Higher order smoothness conditions and differentiability of real functions*, Real Anal. Exchange, **26(2)** (2001), 657–667.

[53] D. La Torre D. and M. Rocca, *A characterization of $C^{k,1}$ functions*, Real Anal. Exchange, **27(2)** (2002), 515–534.

[54] D. La Torre and M. Rocca, *Remarks on second order generalized derivatives for differentiable functions with Lipschitzian Jacobian*, Appl. Math. E-Notes, **3** (2003), 130–137.

[55] D. La Torre and M. Rocca, *A survey on $C^{1,1}$ functions: theory, numerical methods and applications*, Rend. Circ. Mat. Palermo Suppl. ,**70(2)** (2002), 75–93.

[56] D. La Torre and M. Rocca, *On $C^{1,1}$ optimization problems*, J. Comput. Anal. Appl., **7(4)** (2005), 383–395.

[57] D. La Torre, M. Rocca, *$C^{1,1}$ functions and optimality conditions*, J. Concr. Appl. Math., **3(1)** (2005), 41–54.

[58] D. La Torre and C. Vercellis, *$C^{1,1}$ approximations of generalized support vector machines*, J. Concr. Appl. Math., **3(1)** (2005), 41–54.

[59] G. Lebourg, *Valeur moyenne pour gradient géneralizé*, C. R. Acad. Sci. Paris A, **281** (1975), 795–797.

[60] L. Liu, *The second-order conditions of nondominated solutions for $C^{1,1}$ generalized multiobjective mathematical programming*, J. Syst. Sci. Math. Sci., **4(2)** (1991), 128–138.

[61] L. Liu and M. Křížek, *The second-order optimality conditions for nonlinear mathematical programming with $C^{1,1}$ data*, Appl. Math., **42** (1997), 311–320.

[62] L. Liu, P. Neittaanmäki and M. Křížek, *Second-order optimality conditions for nondominated solutions of multiobjective programming with $C^{1,1}$ data*, Appl. Math., **45** (2000), 381–397.

[63] D. T. Luc, *Taylor's formula for $C^{k,1}$ functions*, SIAM J. Optimization, **5** (1995), 659–669.

[64] D. T. Luc, *A strong mean value theorem and applications*, Nonlinear Anal., **26** (1996), 915–923.

[65] Y. Luo and A. Eberhard, *An application of $C^{1,1}$ approximation to comparison principles for viscosity solutions of curvature equations*, Nonlinear Anal., **64(6)** (2006), 1236–1254.

[66] J. Marcinkiewicz and A. Zygmund, *On the differentiability of functions and summability of trigonometrical series*, Fund. Math., **26** (1936), 1–43.

[67] P. Michel and J. P. Penot, *Second-order moderate derivatives*, Nonlin. Anal., **22** (1994), 809–824.

[68] B. S. Mordukhovich, *Variational analysis and generalized differentiation. I. Basic theory*, Grundlehren der Mathematischen Wissenschaften, **330**, Springer-Verlag, Berlin, 2006.

[69] B. S. Mordukhovich, *Variational analysis and generalized differentiation. II. Applications*, Grundlehren der Mathematischen Wissenschaften, **331**, Springer-Verlag, Berlin, 2006.

[70] H. W. Oliver, *The exact Peano derivative*, Trans. Amer. Math.Soc., **76** (1954), 444–456.

[71] Z. Pales and V. Zeidan, *Very Generalized Hessian for $C^{1,1}$ functions in infinite-dimensional normed spaces*, Math. Programming, Ser. A, **74(1)** (1996), 59–78.

[72] G. Peano, *Sulla formula di Taylor*, Atti Accad. Sci. Torino, **27** (1891/92), 40–46. (Italian).

[73] L. Q. Qi and W.Sun, *A nonsmooth version of Newton's method*, Math. Programming, Ser. A, **58(3)** (1993), 353–368.

[74] L. Q. Qi, *Superlinearly convergent approximate Newton methods for LC^1 optimization problems*, Math. Programming, Ser. A, **64(3)** (1994), 277–294.

[75] L. Q. Qi, *LC^1 functions and LC^1 optimization*, Operations Research and Its Applications, World Publishing, Beijing (1996), 4-13.

[76] R. T. Rockafellar and R. J. B. Wets, *Variational analysis*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], **317**, Springer-Verlag, Berlin, 1998.

[77] M. Salo, *Stability for solutions of wave equations with $C^{1,1}$ coefficients*, Inverse Probl. Imaging, **1(3)** (2007), 537–556.

[78] M. E. Sarabi, *Taylor's expansion for $C^{1,1}$ functions in Asplund spaces*, Nonlinear Anal., **71(11)** (2009), 5707–5711.

[79] H. Shahgholian, *$C^{1,1}$ regularity in semilinear elliptic problems*, Comm. Pure Appl. Math., **56(2)** (2003), 278–281.

[80] H. F. Smith, *A parametrix construction for wave equations with $C^{1,1}$ coefficients*, Ann. Inst. Fourier, Grenoble, **48(3)** (1998), 797–835.

[81] H. F. Smith, *Spectral cluster estimates for $C^{1,1}$ metrics*, Amer. J. Math., **128(5)** (2006), 1069–1103.

[82] R. Song and Z. Vondracek, *Sharp bounds for Green functions and jumping functions of subordinate killed Brownian motions in bounded $C^{1,1}$ domains*, Electron. Comm. Probab., **9** (2004), 96–105.

[83] R. Song, *Estimates on the Dirichlet heat kernel of domains above the graphs of bounded $C^{1,1}$ functions*, Glas. Mat. Ser. III, **39(2)** (2004), 273–286.

[84] W. Sun, *Generalized Newton's method for LC^1 unconstrained optimization*, J.Comp. Math., **13(3)** (1995), 250–258.

[85] W. Sun, R. J. B. de Sampaio and J. Yuan, *Two algorithms for LC^1 unconstrained optimization*, J. Comp. Math., **18** (2000), 621–632.

[86] B. S. Thomson, *Symmetric properties of real functions*, Marcel Dekker, New York, 1993.

[87] X. Q. Yang and V. Jeyakumar, *Generalized second-order directional derivatives and optimization with $C^{1,1}$ functions*, Optimization, **26** (1992), 165–185.

[88] X. Q. Yang, *Second-order conditions in $C^{1,1}$ optimization with applications*, Numer. Funct. Anal. Optim., **14** (1993), 621–632.

[89] X. Q. Yang, *Generalized second-order derivatives and optimality conditions*, Nonlin. Anal., **23** (1994), 767–784.

[90] M. Wiegner, *The $C^{1,1}$-character of solutions of second order elliptic equations with gradient constraint*, Comm. Partial Differential Equations, **6(3)** (1981), 361–371.

[91] M. Zelen, *An example of a $C^{1,1}$ function, which is not a d.c. function*, Comment. Math. Univ. Carolin., **43(1)** (2002), 149–154.

[92] W. H. Zhao and Y. Gao, *Second-order optimality conditions for $C^{1,1}$ semidefinite programming*, Acta Math. Sci. Ser. A Chin. Ed., **28(1)** (2008), 155–163.