Control and Cybernetics

vol. **27** (1998) No. 1

Necessary optimality conditions for nonsmooth two-dimensional control systems described by Roesser's model

by

Marcin Studniarski

Faculty of Mathematics, University of Łódź,
ul. S. Banacha 22, 90-238 Łódź, Poland.
e-mail: marstud@imul.uni.lodz.pl

Abstract: Necessary optimality conditions for nonlinear nonsmooth two-dimensional discrete control systems are derived. Under additional convexity assumptions, these conditions can be obtained in the form of a maximum principle. For some special cases, sufficient optimality conditions are also presented.

Keywords: Two-dimensional discrete systems, optimal control, necessary conditions, generalized gradients.

1. Introduction

The aim of this paper is to obtain necessary conditions for nonlinear nonsmooth two-dimensional discrete control systems described by equations being a generalization of Roesser's model with variable coefficients (Kaczorek, 1985b).

A general theory of two-dimensional discrete systems (called also 2-D systems) has been presented in Kaczorek (1985a); see also references therein. However, the results concerning optimal control of such systems have appeared only in few papers. For instance, Kaczorek (1985a) gives a survey of results obtained for the minimum energy control problem for linear systems described by Roesser's model. Moreover, necessary optimality conditions (in the form of a maximum principle) for some classes of smooth nonlinear 2-D systems were obtained by Vasil'ev and Kirillova (1967) and Mansimov (1985).

In the present paper we apply a general method of obtaining necessary optimality conditions for discrete control problems with nondifferentiable data. This method, developed independently by Doležal (1982) and Studniarski (1982), consists in reducing the optimal control problem to some mathematical programming problem and then formulating necessary conditions in terms of the generalized gradients of Clarke (1975, 1983) and the partial generalized gradients introduced by Hiriart-Urruty (1979). The conditions obtained in this way can also be applied to the "classical" case where all the functions appearing in the problem are continuously differentiable. In this case, the generalized gradients reduce to usual gradients and all the inclusions occurring in the optimality conditions can be rewritten as equalities.

2. Definitions and preliminaries

We shall now recall some fundamental facts from the theory of generalized gradients of locally Lipschitzian functions, which can be found in the book of Clarke (1983). We shall consider finite-dimensional spaces only.

Let $f : \mathbf{R}^m \to \mathbf{R}$ be a locally Lipschitzian function (i.e. one that satisfies the Lipschitz condition in a neighbourhood of any point $x \in \mathbf{R}^m$). The generalized gradient of f at the point x is defined as follows:

$$\partial f(x) := \text{co}\,\{\lim_{\nu\to\infty} \nabla f(x^{(\nu)}) \mid x^{(\nu)} \to x, x^{(\nu)} \notin \Omega_f\}$$

where "co" denotes the convex hull, ∇f – the usual gradient, while Ω_f – the set of points $\mathbf{R}^m$ at which f fails to be differentiable (by Rademacher's theorem, Ω_f is of Lebesgue measure zero).

Given a convex function $g : \mathbf{R}^m \to \mathbf{R} \cup \{+\infty\}$, finite at the point x, we define the (one-sided) derivative of g at x in the direction y as follows:

$$g'(x;y) := \lim_{\lambda\downarrow 0} \lambda^{-1}(g(x+\lambda y) - g(x)). \tag{1}$$

By the *subdifferential* of g at x we mean the set

$$\widetilde{\partial} g(x) := \{u \in \mathbf{R}^m | \forall y \in \mathbf{R}^m, g(y) \geq g(x) + (y - x \mid u))\tag{2}$$

where $(\cdot \mid \cdot)$ denotes the usual inner product in $\mathbf{R}^m$. It is well known (Rockafellar, 1970) that $g'(x;y)$ exists for all $y \in \mathbf{R}^m$, and

$$\widetilde{\partial} g(x) := \{u \in \mathbf{R}^m | \forall y \in \mathbf{R}^m, g'(x;y) \geq (y \mid u)\}. \tag{3}$$

PROPOSITION 2.1 *(Clarke 1983). Let f, g be locally Lipschitzian functions on $\mathbf{R}^m$, and let $x \in \mathbf{R}^m$. Then*

a) *$\partial f(x)$ is a nonempty convex compact subset of $\mathbf{R}^m$;*

b) *$\partial(f+g)(x) \subset \partial f(x) + \partial g(x)$;*

c) *for any $\lambda \in \mathbf{R}$, $\partial(\lambda f)(x) = \lambda \partial f(x)$;*

d) *if f is continuously differentiable in a neighbourhood of x, then $\partial f(x) = \{\nabla f(x)\}$;*

e) *if f is convex, then $\partial f(x) = \widetilde{\partial} f(x)$.*

Let us now consider the Cartesian product $X = X_1 \times \cdots \times X_n$ where $X_i = \mathbf{R}^{m_i}$, $i = 1, \ldots, n$. Given a locally Lipschitzian function $f : X \to \mathbf{R}$, we define the *partial generalized gradient* $\partial_{x_i} f(x)$ of f with respect to the variable x_i at the point $x = (x_1, \ldots, x_n)$ as follows:

$$\partial_{x_i} f(x) := \mathrm{co}\{\lim_{\nu\to\infty} \nabla_{x_i} f(x^{(\nu)}) \mid x^{(\nu)} \to x, x^{(\nu)} \notin \Omega_f\}. \tag{4}$$

where $\nabla_{x_i} f(x^{(\nu)})$ stands for the usual partial gradient of f with respect to x_i at $x^{(\nu)}$. This definition was introduced by Hiriart-Urruty (1979) and differs from that adopted by Clarke (1983, p. 48); see remarks on p. 111 in Studniarski (1982).

The following proposition can easily be deduced from (4) and Proposition 2.1.

PROPOSITION 2.2 *(Studniarski, 1982). Let f, g be locally Lipschitzian functions on X, and let $x \in X$. Then, for each $i \in \{1, \ldots, n\}$,*

(a) $\partial_{x_i} f(x) = \mathrm{pr}_i(\partial f(x))$ *where* $\mathrm{pr}_i : X \to X_i$ *is defined by* $\mathrm{pr}_i(u) := u_i$ *for all* $u = (u_1, \ldots, u_n) \in X$;

(b) $\partial_{x_i}(f+g)(x) \subset \partial_{x_i} f(x) + \partial_{x_i} g(x)$;

(c) $\partial_{x_i}(\lambda f)(x) = \lambda \partial_{x_i} f(x)$ $(\lambda \in \mathbf{R})$;

(d) *if f does not depend on the variable x_i, then* $\partial_{x_i} f(x) = \{0\}$.

Using the partial generalized gradients, we can derive necessary optimality conditions for mathematical programming problems on the space X. We shall now formulate a result for the following particular problem:

$$\begin{aligned} &\text{minimize } f_0(x) \text{ subject to} \\ &f_1(x) = 0, \ldots, f_k(x) = 0, \\ &x \in A_1 \times \cdots \times A_n, \end{aligned} \tag{5}$$

where $f_j : X \to \mathbf{R}$, $j = 0, 1, \ldots, k$, are locally Lipschitzian, while the sets $A_i \subset X_i$, $i = 1, \ldots, n$, are closed and convex.

Given a convex set $A \subset \mathbf{R}^m$, let us denote by $N(x \mid A)$ the normal cone to A at the point $x \in A$ (Rockafellar, 1970), i.e.

$$N(x \mid A) := \{u \in \mathbf{R}^m | \forall a \in A, (a - x \mid u) \leq 0\}.$$

It is easy to verify that if $A = A_1 \times \cdots \times A_n \subset X$, then

$$\mathrm{pr}_i(N(x \mid A)) = N(x_i \mid A_i), \quad i = 1, \ldots, n. \tag{6}$$

The following theorem is a special case of Theorem 2.5 from Studniarski (1982). It can also be obtained from Theorem 6.1.1 of Clarke (1983) by using Proposition 2.2 (a) and formula (6).

THEOREM 2.1 *Suppose that $\overline{x} = (\overline{x}_1, \ldots, \overline{x}_n) \in X$ is a local minimum point in problem (5). Then there exist real numbers $\lambda_0 \geq 0, \lambda_1, \ldots, \lambda_k$, not all zero, such that*

$$0 \in \partial_{x_i}\left(\sum_{j=0}^{k} \lambda_j f_j\right)(\overline{x}) + N(\overline{x}_i \mid A_i), \ \ i = 1, \ldots, n.$$

3. Necessary optimality conditions in the general case

We shall consider a control system described by the equations

$$x^h_{i+1,j} = \varphi^h_{i,j}(x^h_{i,j}, x^v_{i,j}, u_{i,j}),$$
$$i = 0, 1, \ldots, r-1, \quad j = 0, 1, \ldots, s, \tag{7}$$
$$x^v_{i,j+1} = \varphi^v_{i,j}(x^h_{i,j}, x^v_{i,j}, u_{i,j}),$$
$$i = 0, 1, \ldots, r, \quad j = 0, 1, \ldots, s-1, \tag{8}$$

where
i is an integer-valued vertical coordinate,
j is an integer-valued horizontal coordinate,
$x^h_{i,j} \in \mathbf{R}^{n_1}$ is the horizontal state vector,
$x^v_{i,j} \in \mathbf{R}^{n_2}$ is the vertical state vector,
$u_{i,j} \in \mathbf{R}^m$ is the control vector,
r and s are fixed positive integers.
The boundary conditions for (7)-(8) are given by

$$\psi^h_j(x^h_{0,j}) = 0, \; j = 0, 1, \ldots, s, \tag{9}$$
$$\psi^v_i(x^v_{i,0}) = 0, \; i = 0, 1, \ldots, r. \tag{10}$$

We also assume the final condition

$$\psi(x^h_{r,s}, x^v_{r,s}) = 0 \tag{11}$$

and the following constraints on the control vectors:

$$u_{i,j} \in U_{i,j}, \quad (i,j) \in K, \tag{12}$$

where

$$K := (\{0, 1, \ldots, r\} \times \{0, 1, \ldots, s\}) \setminus \{(r,s)\}$$

and $U_{i,j}$ are given closed convex subsets of $\mathbf{R}^m$. Let us denote

$$Z := \mathbf{R}^{(r+1)(s+1)n_1} \times \mathbf{R}^{(r+1)(s+1)n_2} \times \mathbf{R}^{[(r+1)(s+1)-1]m}.$$

Elements of the space Z will be written down as triplets $z = (x^h, x^v, u)$ where

$$x^h = (x^h_{0,0}, \ldots, x^h_{0,s}, \ldots, x^h_{r,0}, \ldots, x^h_{r,s}),$$
$$x^v = (x^v_{0,0}, \ldots, x^v_{0,s}, \ldots, x^v_{r,0}, \ldots, x^v_{r,s}),$$
$$u = (u_{0,0}, \ldots, u_{0,s}, \ldots, u_{r-1,0}, \ldots, u_{r-1,s}, u_{r,0}, \ldots, u_{r,s-1}).$$

Let us now consider the cost functional $J : Z \to \mathbf{R}$ given by

$$J(z) := \sum_{(i,j)\in K} f_{i,j}(x^h_{i,j}, x^v_{i,j}, u_{i,j}) + f(x^h_{r,s}, x^v_{r,s}). \tag{13}$$

We assume that all the functions

$\varphi^h_{i,j}$: $\mathbf{R}^{n_1+n_2+m} \to \mathbf{R}^{n_1}$,
$\varphi^v_{i,j}$: $\mathbf{R}^{n_1+n_2+m} \to \mathbf{R}^{n_2}$,
ψ^h_j: $\mathbf{R}^{n_1} \to \mathbf{R}^{k_1}$,
ψ^v_i: $\mathbf{R}^{n_2} \to \mathbf{R}^{k_2}$,
ψ: $\mathbf{R}^{n_1+n_2} \to \mathbf{R}^{k}$,
$f_{i,j}$: $\mathbf{R}^{n_1+n_2+m} \to \mathbf{R}$,
f: $\mathbf{R}^{n_1+n_2} \to \mathbf{R}$,
appearing in (7)-(11) and (13), are locally Lipschitzian. In particular, if $\varphi^h_{i,j}$ and $\varphi^v_{i,j}$ are linear, equations (7)-(8) coincide with Roesser's model with variable coefficients, considered by Kaczorek (1985b).

We can now formulate the optimal control problem as follows

$$\text{minimize } J(z) \text{ over all } z \in Z$$

$$\text{satisfying (7)-(12).} \tag{14}$$

The necessary optimality conditions for this problem are stated in the following

THEOREM 3.1 *Suppose that $\overline{z} = (\overline{x}^h, \overline{x}^v, \overline{u})$ is a local minimum point for problem (14). Then there exist elements*

$$\lambda \geq 0,\ p^h_{i,j} \in \mathbf{R}^{n_1},\ i = 1, \ldots, r,\ j = 0, 1, \ldots, s,$$

$$p^v_{i,j} \in \mathbf{R}^{n_2},\ i = 0, 1, \ldots, r,\ j = 0, 1, \ldots, s, \tag{15}$$

$$w^h_j \in \mathbf{R}^{k_1},\ j = 0, 1, \ldots, s,$$

$$w^v_i \in \mathbf{R}^{k_2},\ i = 0, 1, \ldots, r,\ w \in \mathbf{R}^k,$$

not all zero, such that the functions $H_{i,j} : \mathbf{R}^{n_1+n_2+m} \to \mathbf{R}$ defined for $(i,j) \in K$ as follows:

$$H_{i,j} := (p^h_{i+1,j} \mid \varphi^h_{i,j}) + (p^v_{i,j+1} \mid \varphi^v_{i,j}) - \lambda f_{i,j},$$

$$i = 0, 1, \ldots, r-1,\ j = 0, 1, \ldots, s-1, \tag{16}$$

$$H_{r,j} := (p^v_{r,j+1} \mid \varphi^v_{r,j}) - \lambda f_{r,j},\ j = 0, 1, \ldots, s-1, \tag{17}$$

$$H_{i,s} := (p^h_{i+1,s} \mid \varphi^h_{i,s}) - \lambda f_{i,s}, i = 0, 1, \ldots, r-1, \tag{18}$$

(where, for instance, $(p^h_{i+1,j} \mid \varphi^h_{i,j})$ stands for the function $\mathbf{R}^{n_1+n_2+m} \ni y \to (p^h_{i+1,j} \mid \varphi^h_{i,j}(y)))$ satisfy the conditions

$$\partial_{x^h_{0,,j}} H_{0,j}(\overline{x}^h_{0,j}, \overline{x}^v_{0,j}, \overline{u}_{0,j}) \cap \partial(w^h_j \mid \psi^h_j)(\overline{x}^h_{0,j}) \neq \emptyset, \tag{19}$$

$$j = 0, 1, \ldots, s,$$

$$\partial_{x^v_{i,0}} H_{i,0}(\overline{x}^h_{i,0}, \overline{x}^v_{i,0}, \overline{u}_{i,0}) \cap \partial(w^v_i \mid \psi^v_i)(\overline{x}^v_{i,0}) \neq \emptyset, \tag{20}$$

$$i = 0, 1, \ldots, r,$$

$$p^h_{i,j} \in \partial_{x^h_{i,j}} H_{i,j}(\overline{x}^h_{i,j}, \overline{x}^v_{i,j}, \overline{u}_{i,j}),\ (i,j) \in K,\ i \neq 0, \tag{21}$$

$$p^v_{i,j} \in \partial_{x^v_{i,j}} H_{i,j}(\overline{x}^h_{i,j}, \overline{x}^v_{i,j}, \overline{u}_{i,j}),\quad (i,j) \in K,\quad j \neq 0, \tag{22}$$

$$p^h_{r,s} \in -\partial_{x^h_{r,s}}(\lambda f + (w \mid \psi))(\overline{x}^h_{r,s}, \overline{x}^v_{r,s}), \tag{23}$$

$$p^v_{r,s} \in -\partial_{x^v_{r,s}}(\lambda f + (w \mid \psi))(\overline{x}^h_{r,s}, \overline{x}^v_{r,s}), \tag{24}$$

$$\partial_{u_{i,j}} H_{i,j}(\overline{x}^h_{i,j}, \overline{x}^v_{i,j}, \overline{u}_{i,j}) \cap N(\overline{u}_{i,j} \mid U_{i,j}) \neq \emptyset,\ (i,j) \in K. \tag{25}$$

Proof. Let us introduce the notations

$$\widetilde{\varphi}^h_{i,j}(z) := x^h_{i+1,j} - \varphi^h_{i,j}(x^h_{i,j}, x^v_{i,j}, u_{i,j}),$$

$$i = 0, 1, \ldots, r-1, \quad j = 0, 1, \ldots, s,$$

$$\widetilde{\varphi}^v_{i,j}(z) := x^v_{i,j+1} - \varphi^v_{i,j}(x^h_{i,j}, x^v_{i,j}, u_{i,j}),$$

$$i = 0, 1, \ldots, r, \quad j = 0, 1, \ldots, s-1,$$

$$\widetilde{\psi}^h_j(z) := \psi^h_j(x^h_{0,j}), \quad j = 0, 1, \ldots, s,$$

$$\widetilde{\psi}^v_i(z) := \psi^v_i(x^v_{i,0}), \quad i = 0, 1, \ldots, r,$$

$$\widetilde{\psi}(z) := \psi(x^h_{r,s}, x^v_{r,s}),$$

$$A := \mathbf{R}^{(r+1)(s+1)n_1} \times \mathbf{R}^{(r+1)(s+1)n_2} \times (U_{0,0} \times \cdots \times U_{0,s} \times \cdots \times$$

$$\cdots \times U_{r-1,0} \times \cdots \times U_{r-1,s} \times U_{r,0} \times \cdots \times U_{r,s-1}). \tag{26}$$

Then problem (14) can be rewritten in the form

minimize $J(z)$ subject to

$$\widetilde{\varphi}^h_{i,j}(z) = 0,\ i = 0, 1, \ldots, r-1,\ j = 0, 1, \ldots, s,$$

$$\widetilde{\varphi}^v_{i,j}(z) = 0,\ i = 0, 1, \ldots, r,\ j = 0, 1, \ldots, s-1,$$

$$\widetilde{\psi}^h_j(z) = 0,\ j = 0, 1, \ldots, s, \tag{27}$$

$$\widetilde{\psi}^v_i(z) = 0,\ i = 0, 1, \ldots, r,$$

$$\widetilde{\psi}(z) = 0,$$

$$z \in A,$$

whence it can be seen at once that this is a special case of problem (5). It follows from Theorem 2.1 that there exist elements (15), not all zero, such that the function

$$L := L_1 + L_2 \tag{28}$$

where

$$L_1 := \lambda J + \sum_{\mu=0}^{r-1}\sum_{\nu=0}^{s}(p^{h}_{\mu+1,\nu} \mid \widetilde{\varphi}^{h}_{\mu,\nu}) + \sum_{\mu=0}^{r}\sum_{\nu=0}^{s-1}(p^{v}_{\mu,\nu+1} \mid \widetilde{\varphi}^{v}_{\mu,\nu}), \tag{29}$$

$$L_2 := \sum_{\nu=0}^{s}(w^h_{\nu} \mid \widetilde{\psi}^h_{\nu}) + \sum_{\mu=0}^{r}(w^v_{\mu} \mid \widetilde{\psi}^v_{\mu}) + (w \mid \widetilde{\psi}) \tag{30}$$

satisfies the conditions

$$0 \in \partial_{x^h_{i,j}} L(\overline{z}),\ i = 0, 1, \ldots, r,\ j = 0, 1, \ldots, s, \tag{31}$$

$$0 \in \partial_{x^v_{i,j}} L(\overline{z}),\ i = 0, 1, \ldots, r,\ j = 0, 1, \ldots, s, \tag{32}$$

$$0 \in \partial_{u_{i,j}} L(\overline{z}) + N(\overline{u}_{i,j} \mid U_{i,j}),\ (i, j) \in K. \tag{33}$$

It follows from Proposition 2.2 ((b) and (d)) that

$$\partial_{x^h_{0,j}} L_1(\overline{z}) \subset \partial_{x^h_{0,j}} (\lambda f_{0,j} - (p^h_{1,j} \mid \varphi^h_{0,j}) - (p^v_{0,j+1} \mid \varphi^v_{0,j}))(\overline{x}^h_{0,j}, \overline{x}^v_{0,j}, \overline{u}_{0,j}),$$

$$j = 0, 1, \ldots s-1,$$

$$\partial_{x^h_{0,s}} L_1(\overline{z}) \subset \partial_{x^h_{0,s}} (\lambda f_{0,s} - (p^h_{1,s} \mid \varphi^h_{0,s}))(\overline{x}^h_{0,s}, \overline{x}^v_{0,s}, \overline{u}_{0,s}),$$

$$\partial_{x^h_{0,j}} L_2(\overline{z}) \subset \partial(w^h_j \mid \psi^h_j)(\overline{x}^h_{0,j}),\ j = 0, 1, \ldots, s.$$

Hence, using (16) and (18), we find that condition (31) for $i = 0$ implies (19). Similarly, we have

$$\partial_{x^h_{i,j}} L_1(\overline{z}) \subset p^h_{i,j} + \partial_{x^h_{i,j}} (\lambda f_{i,j} - (p^h_{i+1,j} \mid \varphi^h_{i,j})$$

$$-(p^v_{i,j+1} \mid \varphi^v_{i,j}))\ (\overline{x}^h_{i,j}, \overline{x}^v_{i,j}, \overline{u}_{i,j}),$$

$$i = 1, \ldots, r-1,\ j = 0, 1, \ldots, s-1,$$

$$\partial_{x^h_{r,j}} L_1(\overline{z}) \subset p^h_{r,j} + \partial_{x^h_{r,j}} (\lambda f_{r,j} - (p^v_{r,j+1} \mid \varphi^v_{r,j}))(\overline{x}^h_{r,j}, \overline{x}^v_{r,j}, \overline{u}_{r,j}),$$

$$j = 0, 1, \ldots, s-1,$$

$$\partial_{x^h_{i,s}} L_1(\bar{z}) \subset p^h_{i,s} + \partial_{x^h_{i,s}}(\lambda f_{i,s} - (p^h_{i+1,s} \mid \varphi^h_{i,s}))(\bar{x}^h_{i,s}, \bar{x}^v_{i,s}, \bar{u}_{i,s}),$$

$$i = 1, \ldots, r-1,$$

$$\partial_{x^h_{i,j}} L_2(\bar{z}) = \{0\}, \ (i,j) \in K, \ i \neq 0.$$

Hence, from (16)-(18) it follows that condition (31) for $(i,j) \in K, i \neq 0$, implies (21). Next, we have

$$\partial_{x^h_{r,s}} L(\bar{z}) \subset p^h_{r,s} + \partial_{x^h_{r,s}}(\lambda f + (w \mid \psi))(\bar{x}^h_{r,s}, \bar{x}^v_{r,s}),$$

and so, (31) for $(i,j) = (r,s)$ implies (23). In a similar way, from (32) we deduce conditions (20), (22) and (24). Finally, (25) follows from (33).

4. The maximum principle

If a maximum principle formulation of the optimality conditions for problem (14) is required, additional convexity assumptions have to be imposed on the functions which describe the problem. The situation here is similar to that for one-dimensional systems (see Studniarski, 1982).

THEOREM 4.1 *Suppose that all the functions $\varphi^h_{i,j}$, $\varphi^v_{i,j}$ and $f_{i,j}$ appearing in (7), (8) and (13), respectively, satisfy the following conditions: for any $y_\nu \in \mathbf{R}^{n_\nu}$, $\nu = 1, 2$, the functions $\varphi^h_{i,j}(y_1, y_2, \cdot)$ and $\varphi^v_{i,j}(y_1, y_2, \cdot)$ are affine, while $f_{i,j}(y_1, y_2, \cdot)$ are convex. Then, condition (25) in Theorem 3.1 can be replaced by the following one:*

$$H_{i,j}(\bar{x}^h_{i,j}, \bar{x}^v_{i,j}, \bar{u}_{i,j}) = \max_{u \in U_{i,j}} H_{i,j}(\bar{x}^h_{i,j}, \bar{x}^v_{i,j}, u), \quad (i,j) \in K. \tag{34}$$

Proof. The proof is quite analogous to that of Theorem 4.5 of Studniarski (1982). We repeat it for the reader's convenience.

It follows from our assumptions that the functions $\overline{H}_{i,j} := -H_{i,j}(\bar{x}^h_{i,j}, \bar{x}^v_{i,j}, \cdot)$, $(i,j) \in K$, are convex. Using either Lemma 4.4 of Studniarski (1982) or Proposition 2.5.3 of Clarke (1983), we find that

$$\partial_{u_{i,j}}(-H_{i,j})(\bar{x}^h_{i,j}, \bar{x}^v_{i,j}, \bar{u}_{i,j}) = \widetilde{\partial}\overline{H}_{i,j}(\bar{u}_{i,j}), \ (i,j) \in K. \tag{35}$$

Hence, condition (25) is equivalent to

$$0 \in \widetilde{\partial}\overline{H}_{i,j}(\bar{u}_{i,j}) + N(\bar{u}_{i,j} \mid U_{i,j}), \ (i,j) \in K. \tag{36}$$

It is easy to check that $N(\bar{u}_{i,j} \mid U_{i,j}) = \widetilde{\partial}(\delta(\cdot \mid U_{i,j}))(\bar{u}_{i,j})$ where $\delta(\cdot \mid U_{i,j})$ is the indicator function of $U_{i,j}$:

$$\delta(u \mid U_{i,j}) := \begin{cases} 0 & \text{if } u \in U_{i,j}, \\ +\infty & \text{if } u \notin U_{i,j}. \end{cases} \tag{37}$$

Consequently, by the well-known theorem on the subdifferential of the sum of convex functions (Rockafellar, 1970), formula (36) gives

$$0 \in \widetilde{\partial}\overline{H}_{i,j} + \delta(\cdot \mid U_{i,j}))(\overline{u}_{i,j}), \quad (i,j) \in K.$$

In view of (2), this means that $\overline{H}_{i,j}$ attains at the point $\overline{u}_{i,j}$ its minimum on the set $U_{i,j}$, which is equivalent to (34).

We shall now deal with the special case when problem (14) is convex, that is, all the functions $\varphi^h_{i,j}, \varphi^v_{i,j}, \psi^h_j, \psi^v_i, \psi$ are affine, while $f_{i,j}$ and f are convex. Applying the theory of convex programming presented in Rockafellar (1970), we can prove that, under some additional hypotheses, Theorems 3.1 and 4.1 hold with $\lambda = 1$. Moreover, in some cases, the conditions given in these theorems are sufficient for optimality in problem (14). The details are given in the following two theorems.

THEOREM 4.2 *Suppose that problem (14) is convex and its optimal value (i.e. the infimum of $J(z)$ over all $z \in Z$ satisfying (7)-(12)) is greater than $-\infty$. Suppose that there exists a triplet $z = (x^h, x^v, u)$ satisfying (7)-(12) and such that for each $(i,j) \in K$, $u_{i,j}$ belongs to the relative interior of $U_{i,j}$. Then Theorems 3.1 and 4.1 are valid with $\lambda = 1$.*

Proof. Let $\overline{z} = (\overline{x}^h, \overline{x}^v, \overline{u})$ be a solution of (14). Applying Theorems 28.1 and 28.2 of Rockafellar (1970), we obtain that there exist elements (15), with $\lambda = 1$, such that $L + \delta(\cdot \mid A)$ attains its global minimum at the point $\overline{z}$ (here L and A are defined by (28) and (26), respectively, while $\delta(\cdot \mid A)$ is the indicator function, as in (37)). Hence

$$0 \in \widetilde{\partial}(L + \delta(\cdot \mid A))(\overline{z}) = \widetilde{\partial}L(\overline{z}) + N(\overline{z}|A).$$

From this, using Proposition 2.2 (a) and formula (6), we deduce (31)-(33) (note that, similarly as in (35), all partial generalized gradients of L at $\overline{z}$ coincide with partial subdifferentials). The remaining part of the proof is the same as in Theorems 3.1 and 4.1.

Let us now consider the Cartesian product $\mathbf{R}^{n_1} \times \mathbf{R}^{n_2} \times \mathbf{R}^m$ whose elements are written down as triplets (y_1, y_2, y_3). For a given convex function $g : \mathbf{R}^{n_1} \times \mathbf{R}^{n_2} \times \mathbf{R}^m \to \mathbf{R}$, we denote by $g'((\overline{y}_1, \overline{y}_2, \overline{y}_3); (y_1, y_2, y_3))$ the derivative of g at $(\overline{y}_1, \overline{y}_2, \overline{y}_3)$ in the direction (y_1, y_2, y_3) (see (1)). By $g'_{y_1}(\overline{y}_1, \overline{y}_2, \overline{y}_3; y_1)$ we denote the derivative of $g(\cdot, \overline{y}_2, \overline{y}_3)$ at $\overline{y}_1$ in the direction y_1. Analogously, we define $g'_{y_2}(\overline{y}_1, \overline{y}_2, \overline{y}_3; y_2)$ and $g'_{y_3}(\overline{y}_1, \overline{y}_2, \overline{y}_3; y_3)$.

THEOREM 4.3 *Suppose that problem (14) is convex, and that $\overline{z} = (\overline{x}^h, \overline{x}^v, \overline{u})$ satisfies the necessary conditions given in Theorem 3.1 (or 4.1) with $\lambda = 1$. Suppose, in addition, that all the functions $f_{i,j}$ and f satisfy the following conditions: for each $(y_1, y_2, y_3) \in \mathbf{R}^{n_1} \times \mathbf{R}^{n_2} \times \mathbf{R}^m$, we have*

$$(f_{i,j})'((\overline{x}^h_{i,j}, \overline{x}^v_{i,j}, \overline{u}_{i,j}); (y_1, y_2, y_3)) \geq$$

$$\geq \sum_{\nu=1}^{3} (f_{i,j})'_{y_\nu}(\overline{x}^h_{i,j}, \overline{x}^v_{i,j}, \overline{u}_{i,j}; y_\nu), \tag{38}$$

$$f'((\overline{x}^h_{r,s}, \overline{x}^v_{r,s}); (y_1, y_2)) \geq \sum_{\nu=1}^{2} f'_{y_\nu}(\overline{x}^h_{r,s}, \overline{x}^v_{r,s}; y_\nu). \tag{39}$$

Then $\overline{z}$ is an optimal solution of (14).

Proof. It suffices to show that $\overline{L} := L + \delta(\cdot \mid A)$ attains its global minimum at $\overline{z}$, which is equivalent to the following condition (see (2) and (3)):

$$\overline{L}'(\overline{z}; z) \geq 0 \text{ for all } z \in Z. \tag{40}$$

Let L_0 be the sum of all the summands occurring in (29) and (30) except for the first one; hence $L = J + L_0$. Since L_0 is affine, and thus differentiable, we have, for all $z = (x^h, x^v, u) \in Z$,

$$(L_0)'(\overline{z}; z) = \sum_{(i,j)\in S} (L_0)'_{x^h_{i,j}}(\overline{z}; x^h_{i,j})$$

$$+ \sum_{(i,j)\in S} (L_0)'_{x^v_{i,j}}(\overline{z}; x^v_{i,j}) + \sum_{(i,j)\in K} (L_0)'_{u_{i,j}}(\overline{z}; u_{i,j}) \tag{41}$$

where $S := K \cup \{(r,s)\}$. Next, from (13), (38) and (39) we obtain

$$J'(\overline{z}; z) \geq \sum_{(i,j)\in K} [(f_{i,j})'_{y_1}(\overline{x}^h_{i,j}, \overline{x}^v_{i,j}, \overline{u}_{i,j}; x^h_{i,j})$$

$$+(f_{i,j})'_{y_2}(\overline{x}^h_{i,j}, \overline{x}^v_{i,j}, \overline{u}_{i,j}; x^v_{i,j}) + (f_{i,j})'_{y_3}(\overline{x}^h_{i,j}, \overline{x}^v_{i,j}, \overline{u}_{i,j}; u_{i,j})]$$

$$+f'_{y_1}(\overline{x}^h_{r,s}, \overline{x}^v_{r,s}; x^h_{r,s}) + f'_{y_2}(\overline{x}^h_{r,s}, \overline{x}^v_{r,s}; x^v_{r,s}). \tag{42}$$

We also have

$$(\delta(\cdot \mid A))'(\overline{z}; z) = \sum_{(i,j)\in K} (\delta(\cdot \mid U_{i,j}))'(\overline{u}_{i,j}; u_{i,j}). \tag{43}$$

Conditions (41)-(43) mean that

$$\overline{L}'(\overline{z}; z) \geq$$

$$\sum_{(i,j)\in S} \overline{L}'_{x^h_{i,j}}(\overline{z}; x^h_{i,j}) + \sum_{(i,j)\in S} \overline{L}'_{x^v_{i,j}}(\overline{z}; x^v_{i,j}) + \sum_{(i,j)\in K} \overline{L}'_{u_{i,j}}(\overline{z}; u_{i,j}). \tag{44}$$

Since problem (14) is convex, it is easy to show that conditions (19)-(25) imply (31)-(33). This means that all the three summands on the right-hand side of (44) are nonnegative, and so, (40) holds.

REMARK 4.1 *It can easily be shown that if f is either Gâteaux differentiable at $(\overline{x}^h_{r,s}, \overline{x}^v_{r,s})$ or has the property that, for each fixed $y_2 \in \mathbf{R}^{n_2}$, the function $f'_{y_2}(\cdot, \overline{x}^v_{r,s}; y_2)$ is lower semicontinuous at $\overline{x}^h_{r,s}$, then f satisfies (39). However, there exist convex functions which do not satisfy (39); a simple example is given by $f : \mathbf{R}^2 \to \mathbf{R}$, $f(y_1, y_2) = \max\{y_1, y_2\}$. A similar remark can be made in regard to (38). See Studniarski (1991, Proposition 4.2) for details.*

References

CLARKE, F.H. (1975) Generalized gradients and applications. *Trans. Amer. Math. Soc.*, **205**, 247-262.

CLARKE, F.H. (1983) *Optimization and Nonsmooth Analysis.* New York, Wiley-Interscience.

DOLEŽAL., J. (1982) Non-smooth and non-convex problems in discrete optimal control. *Int J. Systems Sci.*, **13**, 9, 969-978.

HIRIART-URRUTY, J.-B. (1979) Refinements of necessary optimality conditions in nondifferentiable programming I. *Appl. Math. Optim.*, **5**, 1, 63-82.

KACZOREK, T. (1985A) *Two-dimensional linear systems.* Lecture Notes in Control and Information Sciences, **68**, Berlin, Springer-Verlag.

KACZOREK, T. (1985B) General response formula for Roesser's model with variable coefficients. *Bull. Polish Acad. Sci. Tech.* **33**, 7-8, 409-414.

MANSIMOV, K.B. (1985) Sufficient conditions of Krotov type in discrete two-parameter systems. *Automat. i Telemekh.*, **8**, 15-20 (in Russian).

ROCKAFELLAR, R.T. (1970) *Convex Analysis.* Princeton, New Jersey, Princeton University Press.

STUDNIARSKI, M. (1982) Necessary optimality conditions for a nonsmooth discrete control problem. *Control Cybernet.*, **11**, 3-4, 109-119.

STUDNIARSKI, M. (1991) The discrete maximum principle as a sufficient optimality condition. *Problems Control Inform. Theory*, **20**, 3, 179-186.

VASIL'EV, O.V., KIRILLOVA, F.M. (1967) Optimal processes in two-parameter discrete systems. *Dokl. Akad. Nauk SSSR*, **175**, 1, 17-19 (in Russian).